EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

ECB EZB EKT BCE EKP

WORKING PAPER NO. 86

RATIONAL EXPECTATIONS AND NEAR RATIONAL ALTERNATIVES: HOW BEST TO FORM EXPECTATIONS

BY M. BEEBY,
S. G. HALL
AND
S. B. HENRY

November 2001

EUROPEAN CENTRAL BANK

WORKING PAPER SERIES

WORKING PAPER NO. 86

RATIONAL EXPECTATIONS AND NEAR RATIONAL ALTERNATIVES: HOW BEST TO FORM EXPECTATIONS

BY M. BEEBY[1], S. G. HALL[1,2] AND S. B. HENRY[1]

November 2001

1 Centre for International Macroeconomics, Oxford University.

2 The Management School, Imperial College of Science Technology and Medicine.

This paper has been presented at the ECB workshop on 'Forecasting Techniques', September2001.

Address	**Kaiserstrasse 29**
	D-60311 Frankfurt am Main
	Germany
Postal address	**Postfach 16 03 19**
	D-60066 Frankfurt am Main
	Germany
Telephone	**+49 69 1344 0**
Internet	**http://www.ecb.int**
Fax	**+49 69 1344 6000**
Telex	**411 144 ecb d**

ISSN 1561-0810

Contents

Abstract

Learning rules are increasingly being used in macroeconomic models. However one criticism that has been levelled at this assumption is that the choice of variables for inclusion in the learning rule, and the actual specification of the learning rule itself, is arbitrary. In this paper we test how important the particular learning rule specification is by incorporating a battery of learning rules into a large-scale macro model. The model's dynamics are then compared to those from a version of the model simulated under rational expectations (RE). The results indicate that although there are large differences between the RE solution and each of the solutions under learning, differences amongst the learning rule solutions are minor.

JEL classification: C53, E43, F33
Keywords: Learning, Rational expectations, Bounded expectations, Kalman Filter

Non-technical summary

Since Lucas (1972, 1979) the dominant assumption in macroeconomics concerning the formation of agents' expectations has been that they are rational. That is, agents make no systematic errors in their forecasts (i.e. the forecast error is white noise); the expected future value of a series is equal to the value predicted by the economic model, subject to a shock. This is seen by some to be a highly restrictive assumption, requiring that agents not only use all their information optimally, but more importantly that they have all the information that is required to forecast. Requiring agents to know everything that drives a particular system places large demands on agents with the information requirements seeming to be too high. For instance, this assumption requires that agents know all the equations' functional forms, together with the processes that drive the shocks hitting the economy. As an alternative to the assumption of rational expectations, economists have begun experimenting with the hypothesis that agents learn about their economic environment over time. Under learning, agents do not know the true process that drives the expected values of a series so instead have a rule that they use to predict these future values. Each period as new information becomes available agents learn from their previous mistakes and improve their learning rule by updating the rule's parameter values. Marcet and Sargent (1989) have shown that so long as the variables that enter this rule are correlated with the variables that actually drive the process determining the expected values, then the learning model may converge to the equilibrium that would have been obtained under RE.

The attraction of learning then is that it allows agents to make mistakes in the short-run, but not in the long-run. However learning has itself been criticised since the learning rule that the modeller chooses for the agents is arbitrary; no theory exists for why agents may choose or prefer one rule instead of another. For even a medium size model the set of available rules appears large, since every possible permutation of the variables could conceivably be included as the learning rule. So a fundamental question is how does an agent select one rule to use from the very wide range of alternatives. Might some rules dominate others such that the agent would never choose the dominated rule when others are available? A second important set of issues relates to the model's dynamics under the various expectations assumptions, how do they differ as we change between learning rules, and how is the RE solution related to the solutions obtained under the many learning rules?

This paper begins by first attempting to understand how an agent might narrow down the set of rules that are available. The approach taken is first to ask how much of the information necessary for forecasting future values of a particular series can be obtained from a small subset of the remaining series and second, to what extent does the information contained in two series overlap for forecasting a third? Clearly, if two series contain the same information then there is no point in both series entering the learning rule. This will reduce the number of rules that the agent might choose to employ. It is desirable to obtain some measure of the amount of shared information in two series so that the number of potential rules can be eliminated in a systematic manner.

To disentangle the contribution of each series to the agent's information set, we follow the ideas set out in Collard and Juillard (2000). They build upon a recent contribution by Judd (1998), who argues that high order approximations for the expectation function may be needed when linearising stochastic optimal control models. Modellers have tended to limit themselves to first-order approximations because of the technical complexity thought to be required as one moves to a higher order. However Judd's method is considerably simply than earlier contributions and allows a second-order approximation to be easily computed. A second-order approximation becomes extremely useful in our context since it allows us to measure the extent to which the co-variation between two series may contribute to explain the variables of interest.

Having investigated a range of ways of selecting the learning rule and demonstrated that the learning rules gives very similar model properties over a very wide range of information sets we then investigate how these outcomes relate to the full RE outcome. We set up a weighted learning rule which combines the full RE expectation with the learning outcome and investigate the effect of moving from RE to learning. We find that a very small departure from RE moves the models properties very close to the range of learning solutions. This strongly suggests that RE is not representative of the models likely solution when agents are only weakly rational. A near RE solution such as learning then seems to be much more representative of the likely outcome for the model.

1 Introduction

Since Lucas (1971, 1972) the dominant assumption in macroeconomics concerning the formation of agents' expectations has been that they are rational. That is, agents make no systematic errors in their forecasts (i.e. the forecast error is white noise); the expected future value of a series is equal to the value predicted by the economic model, subject to a shock. This is seen by some to be a highly restrictive assumption, requiring that agents not only use all their information optimally, but more importantly that they have all the information that is required to forecast. Requiring agents to know everything that drives a particular system places large demands on agents with the information requirements seeming too high. For instance, this assumption requires that agents know all the equations' functional forms, together with the processes that drive the shocks hitting the economy. As an alternative to the assumption of rational expectations (RE), economists have begun experimenting with the hypothesis that agents learn about their economic environment over time. Under learning, agents do not know the true process that drives the expected values of a series so instead have a rule that they use to predict these future values. Each period as new information becomes available agents learn from their previous mistakes and improve their learning rule by updating the rule's parameter values. Marcet and Sargent (1989) have shown that so long as the variables that enter this rule are correlated with the variables that actually drive the process determining the expected values, then the learning model may converge to the equilibrium that would have be obtained under RE.

The attraction of learning then is that it allows agents to make mistakes in the short-run, but not in the long-run. However learning has itself been criticised since the learning rule that the modeller chooses for the agents is arbitrary; no theory exists for why agents may choose or prefer one rule over another. For even a medium size model the set of available rules appears large, since every possible permutation of the variables could conceivably be included as the learning rule? So a fundamental question is how does an agent select one rule to use from the very wide range of alternatives. Might some rules dominate others such that the agent would never choose the dominated rule when others are available? A second important set of issues is how different is the outcome of a model as we change between learning rules and how is the rational expectations(RE) solution related to the many learning rules. So we might find that rational expectations is similar to the many learning rules or we might find that the learning rules are all very similar but the rational model is qualitatively different from all of them. This would be an important result as it would suggest that the RE solution is not representative of a wide class of expectation formation procedures but that it has very special properties. So the question we are raising is does a small departure from full RE imply a big change in models properties or are these properties broadly similar for a range of near rational behaviour.

In this paper we first attempt to understand how the agent might narrow down the set of rules that are available. Our approach is to ask first, how much of the information necessary for forecasting future values of a particular series can be obtained from a small subset of the remaining series and second, to what extent does the information contained in two series overlap for forecasting a third? Clearly, if two series contain the same information then there is no point in both series entering a learning rule. This will reduce the number of learning rules that the agent might choose to employ. Here we want to obtain some measure of the amount of shared information in two series so that the number of potential rules can be eliminated in a systematic manner.

To disentangle the contribution of each series to the agent's information set, we follow the ideas set out in Collard and Juillard (2000). They build upon a recent contribution by Judd (1998), who argues that high order approximations for the expectation function may be needed when linearising stochastic optimal control models. Modellers have tended to limit themselves to first-order approximations because of the technical complexity thought to be required as one moves to a higher order. However Judd's method is considerably simply than earlier contributions and allows a second-order approximation to be easily computed. A second-order approximation becomes extremely useful in our context since it allows us to measure the extent to which the co-variation between two series may contribute to explain the variables of interest.

Having investigated a range of ways of selecting the learning rule and demonstrated that the learning rules give very similar model properties over a very wide range of information sets we then investigate how these outcomes relate to the full RE outcome. We set up a weighted learning rule which combines the full RE expectation with the learning outcome and investigate the effect of moving from full RE to full learning. We find that a very small departure from full RE moves the models properties very close to the range of learning solutions. This strongly suggests that RE is not representative of the models likely solution when agents are only weakly rational. A near RE solution such as learning then seems to be much more representative of the likely outcome for the model.

In the next section the theoretical framework is outlined. Section 3 discusses the ideas contained in Collard and Juillard, Section 4 presents the model with the results discussed in Section 5. In all of these results the benchmark against which the different learning rules are measured is the path that would have been obtained under RE. As an alternative, Section 6 compares the results to those from a very simple learning rule. Section 7 asks how much rationality is required in the learning rule for the model's properties to equal those from a model that assumes rational expectations. Section 8 concludes.

2 The Theoretical Framework

To gain some understanding of the contribution of learning to a model's solution and its dynamic behavior, the differences between learning and rational expectations can best be illustrated in a simple analytical model. Here, the cobweb model of Muth (1961) is used.

The cobweb model is a useful workhorse for studying the effects of learning and RE since supply is assumed to be a function of the expected price (plus additional exogenous terms), while demand is simply a function of the current price,

$$q_t^s = \alpha_1 E_{t-1} p_t + \alpha_2 w_{t-1} + u_t^s$$

$$q_t^d = \alpha_3 - \alpha_4 p_t + u_t^d$$

All parameters are positively signed. The first equation relates the quantity supplied of the good (q^s) to the previous periods expectation (E_{t-1}) of the price (p) this period, some additional exogenous terms (w) assumed to be white noise, plus an error term also assumed to be white noise (u^s). Quantity demanded of the good (q^d) is simply a function of the current price and a further white noise error.

The contribution of Muth (1961) was to assume that agents do not make systematic errors. They must therefore not only use the information they have efficiently but also use all available and relevant information. In the above context this requires that agents know both of the equations that drive prices and quantities. Assuming quantity supplied equals quantity demanded ($q^s = q^d$), the reduced form is given by

$$p_t = \boldsymbol{g} + \boldsymbol{b}_1 E_{t-1} p_t + \boldsymbol{b}_2 w_{t-1} + u_t \qquad (1)$$

where $\boldsymbol{g} = \boldsymbol{a}_0 / \boldsymbol{a}_4$, $\boldsymbol{b}_1 = -\boldsymbol{a}_1 / \boldsymbol{a}_4$, $\boldsymbol{b}_2 = -\boldsymbol{a}_2 / \boldsymbol{a}_4$, and $u = (u^d - u^s) / \boldsymbol{a}_4$ (note that â$_1$<0). Under RE it has to be the case that in equilibrium the expected price is the realised price ($E_{t-1}p_t=p_t$). Substituting this expression back into the reduced form (1) gives,

$$p_t = \bar{\boldsymbol{k}}_0 + \bar{\boldsymbol{k}}_1 w_{t-1} + \boldsymbol{e}_t$$

where $\bar{\boldsymbol{k}}_0 = \boldsymbol{g} / (1 - \boldsymbol{b}_1)$, $\bar{\boldsymbol{k}}_1 = \boldsymbol{g} / (1 - \boldsymbol{b}_2)$ and $\boldsymbol{e} = u / (1 - \boldsymbol{b}_2)$. Under RE then, $E_{t-1} p_t = \bar{\boldsymbol{k}}_0 + \bar{\boldsymbol{k}}_1 w_{t-1}$.

Incorporating learning into a model requires specifying the process by which agents actually learn. Here it is assumed that learning follows a statistical approach. Statistical learning is interesting since agents are treated as econometricians and like econometricians agents use past observations to make their best forecast of future realizations, updating their prediction rule depending upon the forecast error. Bray and Savin (1986) and Fourgeaud, Gourieroux and Pradel (1986) take the "cobweb" model and ask if agents would learn the RE solution over time. Suppose agents believe that prices evolve as follows,

$$p_t = \boldsymbol{k}_0 + \boldsymbol{k}_1 w_{t-1} + \boldsymbol{e}_t \qquad (2)$$

which is of the same form as the RE equilibrium, but where ê$_0$ and ê$_1$ are unknown. Agents are required to act like econometricians and estimate these parameters by regressing p_t on w_{t-1} and an intercept term using the historical data and a method of estimation such as least squares. Agents' forecasts would then be,

$$E_{t-1} p_t = \kappa_{0t-1} + \kappa_{1t-1} w_{t-1} \tag{3}$$

where the parameter estimates are ê$_{0t-1}$ and ê$_{1t-1}$, and are derived using standard least squares formulae,

$$\begin{bmatrix} \kappa_{0t-1} \\ \kappa_{1t-1} \end{bmatrix} = \left[\sum_{i=1}^{t-1} z_{i-1} z_{i-1}^{'} \right]^{-1} \left[\sum_{i=1}^{t-1} z_{i-1} p_i \right], \qquad \text{where } z_i=(1\ w_i) \tag{4}$$

Equations (1), (3) and (4) now form a fully specified dynamic system. The question asked by these authors is: will $(\kappa_{0t-1},\ \kappa_{1t-1})'$ $(\bar{\kappa}_{0t-1}, \bar{\kappa}_{1t-1})'$ as t ∞? They find that so long as â$_1$ is less than one, convergence is guaranteed. This condition always holds so long as the demand and supply curves have their usual slopes. The term E-stability is also useful here. An E-equilibrium exists if after repeatedly substituting into the actual process governing the model's evolution the perceived process, a fixed point is eventually reached where the expected process equals the actual process. In terms of the above system, an E-equilibrium exists so long as â$_1$ is less than one. The model is also said to be E-stable.

3 Which Learning Rule?

Specifying the particular learning rule used by agents can be an arbitrary exercise. The decision on which variables should be included is not straightforward even in small analytical models that have few variables. But it is much more difficult in the type of large-scale econometric models used by policymakers for forecasting and simulation exercises. These models may have over 500 variables and be highly non-linear. It is clearly impractical for all the variables to enter the learning rule, but how does the modeller decide which variables do enter, and how sensitive are the results to the chosen learning rule? Garratt and Hall (1997) provide an answer to the second question using the London Business School's econometric model. Expectations can enter these models in many areas, but it is in the exchange rate sector where they have been shown to have the largest effect. Typically expectations have been introduced via a forward-looking open arbitrage (nominal) exchange rate equation. Since it is a nominal variable, the easiest way to introduce a shock has been to shock an exogenous price, usually the world price of oil. In a small model it is straightforward to test if a stable equilibrium exists, the system's eigenvalues simply require computing. In large models it is more difficult since the eigenvalues can only be locally approximated. Garratt and Hall therefore concentrate on the evolution of the learning parameters. If the parameters cease changing then this is evidence they argue that the model has attained its equilibrium. Their results indicate that for four of the five rules they test (and where each rule is chosen in a totally arbitrary way) there is clear evidence of convergence, though for the final rule a longer simulation period would be required before convergence could be confirmed. Overall, it is argued that the large-scale model used is E-stable.

The authors then go on to ask how the dynamic paths of inflation and output are altered in response to the different rules. Again for four of the five rules the dynamic response of output is remarkably similar – at the end of the simulation they all predict that output is 3% lower following the oil price shock. The fifth rule does not appear to have converged. Garratt and Hall therefore conclude that where convergence does occur, the same end period values for output results for a range of learning rules. The length of the period required for learning seems to be in the order of five years. However, if the endpoints are similar the dynamic paths are not, so the form of the learning rule is important for the model's trajectory. Overall, Garratt and Hall show that dynamic responses, stability, and end-value solutions can be affected by the choice of learning rule. More significantly, their approach can be seen as offering a way of deciding between different forms of the general learning rule, by recourse to their stability and other dynamic properties when used in a macro model. In this sense, the paper is an illustration of one way to recover (identify) the form of the learning rule that is being used – to take an extreme example, a rule which is not E-stable, would be unusable eventually. That said, the procedure has limitations, the most obvious one being that several versions of the learning rule appear E-stable, thus it is difficult to choose between them on this criterion alone.

An entirely different approach, and the one used below, is to take the empirical evidence to establish the likely determinants of the variable concerned, and form an expectations rule based on this. We can refer to this as "estimating" the expectations rule, but on the understanding that the term is used as shorthand only since it is not an estimation approach. Instead, in the present context the essential idea is to establish (estimate) the quantitatively most important exogenous determinants of for example the expected exchange rate by the contribution these other series make to the evolution of the exchange rate.

Collard and Juillard (1999) set out the essential principles, although their applications are to small analytical models. Thus for the model

$$E_t(f(y_{t+1}, y_t, x_t, x_{t-1}, u_t, u_{t-1})) \qquad (5)$$

where x are the weakly exogenous variables, u are stochastic shocks, y the variables for which expectations are formed and for which there are a total of n equations. To obtain an expression for the formation of expectations two functions linking variables y and x to the shocks are postulated,

$$y_t = g(x_{t-1}, u_t) \qquad (6)$$

$$x_t = h(x_{t-1}, u_t) \qquad (7)$$

This enables us to write the expected value of y_{t+1} explicitly as

$$E_t y_{t+1} = E_t[g(h(x_{t-1}, u_t), u_{t+1})]. \qquad (8)$$

In other words the conditional expectation given by (8) can be summarised as

$$\overline{F}(x_{t-1}, u_t) = E_t[F(x_{t-1}, u_t, u_{t+1})] \qquad (9)$$

where the function F is equal to the function g(h). Equations (5-9) are the basic set of equations, but these are non-linear. They can be approximated however using Taylor expansions around a deterministic equilibrium, as Collard and Juillard show. Thus for the simplest, linear case the approximation for each of the equations in (9) is

$$\overline{F}_i(.) =_A E_t[F_i(\overline{x},0,0) + \sum_m \partial F_i / \partial x_m (x - \overline{x})^m + \sum_m \partial F_i / \partial u_m (u)^m] \qquad (10)$$

where the term in u_{t+1} is omitted as its expected value is assumed zero. Such linear approximations to the expectations function allow the model to be simulated. Our purpose in reviewing this technique is rather different though. It is to reveal the general relationship between the expectations function and the weakly exogenous variables x as shown in (10).

The approach we take using a large econometric model is in this spirit. First, every variable in the model is shocked and its effect on the value of the expected series we are interested in calculated. From this information one can determine which variables have the biggest effect on the future value of any particular series. Those variables that have the largest effect are then put into the learning rule. This information is comparable to the first-order term in the Collard and Juillard framework (10). However, it does not take into account any possible co-movement between the series included in the rule and it may be that two of the series contain essentially the same information (see below). To capture this co-movement two approaches are taken. First, the correlation matrix is calculated. This measures the extent to which any two series move together and would seem to capture the relationship expressed in the cross-derivatives expression in (10). If two series are highly correlated then the marginal contribution from including more than one of the two series in the rule is minor. Second, principal components analysis is used and those variables that are most correlated with the components are then included in the rule. Again, it would appear to make little sense including two variables in a rule that are both highly correlated with a particular principal component. Both these approaches might be expected to lead to a reduction in the size of the set of variables that are candidates for inclusion in the learning rule.

4 The Model

The above analysis is conducted on a medium sized econometric model of the aggregate Euro-11 economy (i.e. a model of those eleven countries that will be the inaugural members of European Monetary Union). Expectations can enter these models in many areas, but it is in the exchange rate sector where they have been shown to have the largest effect. Typically expectations have been introduced via a forward-looking open arbitrage real exchange rate equation,

$$e_t = \alpha E_{t-1} e_{t+1} + (1-\alpha) e_{t-1} + r_t \tag{11}$$

where e is the log of the real Euro/Dollar exchange rate and r is the real interest rate differential between short-term rates in Euroland and the United States (a proxy for the world interest rate). Expectations are assumed to take the following form,

$$E_{t-1} e_{t+1} = \gamma_{1t} + \sum_{j=1}^{k} \gamma_j Y_{t-j} \tag{12}$$

where there are k series included in the learning rule[1]. With time-varying parameters some process that can recursively update them needs formalizing. Here it is assumed that agents update their expectations each period by using the Kalman Filter, an extremely general tool in which the updating techniques used by other authors (e.g.

OLS) can be nested. To describe the conceptual framework take the following set of n equations as summarizing a large non-linear macro model,

$$Y_{it} = F_i(Y_{it}, X_{it}, E_t Y_{it+1}) \qquad \text{i=1,…,n, and t=1,…,T.} \tag{13}$$

where Y is a vector of current and lagged values of the n endogenous variables, X is a vector of exogenous variables and $E_t Y_{t+1}$ is the current period's expectation of next period's Y. It is assumed that agents have the following perceived law of motion for the endogenous variables,

$$E_t Y_{t+1} = D_t Z_t + \varepsilon_{1t} \qquad \varepsilon_{1t} \sim N(0,W) \tag{14}$$

where D is the time-varying matrix of parameters, Z is the information set consisting of all the current and lagged values of Y and X. The parameters evolve as,

$$D_t = D_{t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim N(0,Q) \tag{15}$$

It is the relative values of the Q and W matrices that determine the change in the parameters. Though in large-scale models, and unlike small analytical models, all error comes from the complexity of the model (ε_1), and not from stochastic terms in the model itself (ε_2). In Kalman Filter terminology, (14) is the measurement equation, (15) is the state equation and W and Q the matrices of hyper-parameters.

5 Empirical Results

In this section the empirical results are presented. First, each of the series in the model are shocked and their effect on real exchange rates calculated. Those series with the biggest effect are then included in the learning rule. Second, correlations between the exchange rate and the rest of the series are calculated for a moving window of four quarters; with those series that display the most consistent correlations included in the rule. Finally, principal components analysis is conducted, and those series that are most closely linked with each component are chosen to enter the rule.

Multi-Shocks

Each of the series (at every lag that they enter the model) are shocked and their effect on the expected value of next period's exchange rate calculated. Initially each of the series was shocked by a fixed per cent, as if a derivative was being calculated. But one drawback with this is that it doesn't take account of the extent to which some

[1] No currently dated variables are used to form expectations since including the current exchange rate would be close to assuming that the exchange rate is determined by itself.

series are more variable than others. As an example consider the behaviour of interest rates. They have varied relatively little so imposing a large change on them would have a huge impact upon the exchange rate. To compensate for this we take the size of the shock to be equal to the size of the series' historical deviation from trend. Table 1 presents the four best and four worst performing series.

The best performing series are nominal wage earnings and import prices, both their current values and their value lagged one quarter. One problem that may arise is the extent to which the current values of these series are known when the expectations are being formed, it is perhaps more usual to think of there being a time lag before the data become available. Similarly should a lagged dependent variable be included? Three possible combinations of these series for the learning rules are considered:

Rule 1: $$e_t = g_1 + g_{2t} e_{t-1} \tag{16}$$

Rule 2: $$e_t = g_1 + g_{2t} e_{t-1} + g_{3t} W_{t-1} \tag{17}$$

Rule 3: $$e_t = g_1 + g_{2t} e_{t-1} + g_{3t} W_{t-1} + g_{4t} PM_{t-1} \tag{18}$$

Note that each of the rules also includes a constant. To analyse the performance of these rules, government consumption in the model is shocked by one per cent in the initial period (1988 quarter 2) only. The resulting difference between the expected exchange rate and the actual exchange rate is then calculated, with the difference between these series adjusted to remove any negative signs. A time path is then created for the reported statistic (V) that measures this difference relative to next period's actual exchange rate,

$$V_t = V_{t-1} + \sqrt{\left(\frac{E_t e_{t+1} - e_{t+1}}{e_{t+1}} * 100 \right)^2}$$

This statistic captures the cumulative size of the forecast errors. Every simulation of the model then creates this variable so that it can be observed evolving over time. The value for V that is reported is its size at the end of the simulation period, here 2048 quarter four. Note that in a rational expectations model V_t is zero each period since the expected exchange rate can never deviate from the actual exchange rate. Clearly a "good" rule should be the one that has a small value for this term, relative to the other rules.

Kalman filter estimation requires choosing the initial value for the covariance matrix, plus the ratio of the error terms from the measurement and state equations. In a well-defined problem, estimation results are robust to the chosen initial values for the covariance matrix though not to the chosen value for the ratio of the error terms, as one might expect. Given the importance of this ratio, results are reported for a broad range of initial values, varying from very large numbers (i.e. 1000) to very small ones (i.e. 0.0001). The larger the number the more that the coefficients on the variables included in the rule are allowed to vary over time.

Figure 1 plots the cumulative forecast error for rules 1 to 3. The shape of the rules, and the magnitude of their deviation from the RE solution (i.e. 0) are similar in each case. For bigger values of the ratio of the error terms (termed CE in the figure), the resulting value for V is approximately invariant. All three rules then have a larger cumulative forecast error as the parameters on the variables in the learning rule are forced to vary less. As the ratio of error terms is further reduced the three learning rules move back towards convergence to the RE equilibrium[2]. It is surprising that the performance of these three rules is so similar. The rule that one might expect to perform best since it includes more information, Rule 3, does not perform any better than the other two rules. There is also some crossover between the rules, that is for certain values of the ratio of errors one rule has the smallest value of cumulative errors, and for other values for this ratio, a different rule performs better. The message seems to be that the simplest rule that includes only lagged exchange rates performs as well as the rules that include nominal wages and import prices. The marginal contribution of both of these series is extremely minor. This is true across a wide range of rules which are not reported here in detail given space constraints

To demonstrate the wide similarity in the model outcomes for different rules we will give details of three other rules which included only "unimportant" variables, that is those variables which, when shocked, had the least impact upon the real exchange rate.

Rule 4: $$e_t = g_1 + g_{2t} PGC_{t-1} \tag{19}$$

[2] We experimented with reducing the ratio of error terms even more but the rules appeared to diverge once this ratio is reduced even more.

Rule 5: $$e_t = g_1 + g_{2t} RSH_{t-2} \tag{20}$$

Rule 6: $$e_t = g_1 + g_{2t} C_{t-2} \tag{21}$$

Rule 4 includes only a constant and the price of government consumption (PGC), rule 5 has short-term interest rates (RSH) lagged two quarters and a constant, and rule 6 has a constant with consumption (C) also lagged two periods. Figure 2 plots their responses following the government expenditure shock. There are clearly large differences between these three rules. Rule 2, which includes only the short-term interest rate performs much worse than any of the other rules. Its cumulative forecast error term is larger than the corresponding term for any other term by an order of magnitude. Neither the magnitude nor shapes of the other two rules' cumulative error terms are markedly different from those obtained under the "important" rules (rules 1–3).

A comparison of rules 1 to 6 shows that differentiating the series on the basis of their effect on the real exchange rate does not lead to a clear difference in the performance of the rules. Three of the rules should perform much worse than the other three but it is only rule 5 that is noticeably different. There are two explanations for this performance. It could be that the Kalman filter is simply extremely effective at extracting the information necessary for forecasting a variable and that learning does indeed give rise to a very similar outcome over a very wide information set. It will therefore matter little which series is included in the learning rule, except for those variables that are totally unrelated to the real exchange rate. Alternatively, it could be that the criterion chosen to rank the series is not the optimal one. To explore this second issue alternative ranking methods are explored. First, a correlation matrix (between the exchange rate and the rest of the series) is constructed and then the series are ranked according to which series are most strongly correlated with the exchange rate. Second, principal component analysis is carried. Here the series chosen for inclusion in the learning rule are those variables that contain most of the information that is in these components. Both these methods have the added advantage that the inter-relationships between the series can be assessed. This enables the ideas in Collard and Juillard to be introduced.

Correlation Matrix

To capture the extent to which two series co-vary, a simple correlation matrix seems sufficient. But calculating the correlation between two series over the entire sample period is uninformative for two reasons. First, the degree of correlation can vary over time. Second, and perhaps more importantly, the information we are interested in is the usefulness of one series for forecasting another. For these reasons a rolling correlation was chosen. That is, for each quarter the correlation between the four previous values of a series with the current and three previous values of the real exchange rate was calculated. This resulted in a time series for each of the series' correlations and enabled us to calculate the standard deviations (SDs) for these series. The correlations were then ranked according to which had the lowest SD, since a smaller SD indicates that the series is consistently correlated with the exchange rate. Results for the best and worst performing series are shown in Table 2. The four best performing series are investment (IF), employment (ET), unemployment (UP) and the real wage (W/P), in that order. The four worst are the capital stock (K), the consumer price deflator (PC), consumption (C) and real personal net wealth (RPNW). A comparison of Tables 1 and 2 indicate that there is very little consistency between these two methods of choosing which series should enter the learning rules. Only consumption is a consistently poor performer. Figure 3 plots two learning rules[3], one that includes only investment (i.e. the best performing rule according to this correlation exercise) and one that includes only the capital stock (i.e. the rule that should perform the worst).

Rule 7: $$e_t = g_1 + g_{2t} IF_{t-1} \quad (22)$$

Rule 8: $$e_t = g_1 + g_{2t} K_{t-1} \quad (23)$$

The results indicate that the two rules in fact perform remarkably alike in terms of their shape and the size of their deviation from the RE equilibrium. Except for two of the chosen ratio of error terms, the "best" rule that includes investment also outperforms the "worst" rule, which includes the capital stock. These differences are minor however, suggesting that selecting series according to their correlation with the exchange rate utilising a moving window is also not a satisfactory criterion. Series that have little correlation with the exchange rate perform equally as well as those that are highly correlated.

[3] Again the series plotted are the deviations from the RE path for the real exchange rate.

Principal Components Analysis

One problem with including two series in a learning rule is not knowing the extent to which the series may contain essentially the same information. One advantage in assuming learning is that it assumes the agent may be limited in the amount of information available. Because of this it seems natural that the learning rule should be as simple as possible with a simple rule being preferred to a more complicated one. Including series that overlap in the amount of information that they contain for forecasting exchange rates is inefficient, and a situation the agent presumably wishes to avoid. This is the essence of Collard and Juillard's argument and motivates the use here of principle components analysis. The idea is to identify the principle components from all the series in the system, and include only those variables that most closely move with the components. Tables 3a and 3b show the results. Table 3a indicates that 58% of all the variation in each of the series can be explained by just two of the components. Given our requirement that the rules should be simple, it seems preferable to concentrate on selecting those two series that contain the information that is in these two components. Table 3b shows that employment and unemployment contain 94% of the information that is contained in the first principal component. This is a good illustration of the point that it is the second moment that we should be interested in, and not only the first. There seems little point in including both these series in any rule since, with a given workforce, they are perfectly negatively correlated with each other: as unemployment increases employment decreases, and vice versa. We therefore take the employment series as capturing the first component, with total imports being the series that most closely captures component two.

Rule 9: $$e_t = g_1 + g_{2t} ET_{t-1} + g_{3t} PM_{t-1} \tag{24}$$

Figure 4 plots the measure of the cumulative forecast errors from this rule, together with the best rule from the moving-window correlation exercise above. Despite these arguments for the use of principal components, there is little quantitative difference between the results from the two rules and from the rules graphed in Figures 3 and 4.

6 Using a Simple Learning Rule as the Benchmark

All of the above results assessed a given rule by comparing the exchange rate's path following a shock to the path that would have been obtained under RE. This is not necessarily the best way of measuring the efficiency of a rule. So here we ask how successful the rules are using a simple learning rule as the benchmark. The benchmark rule simply includes lagged real exchange rates (i.e. Rule 1). To assess which series should be included in the rules a correlation exercise is conducted, as above. The series are then ranked according to how consistently they track the exchange rate's path from Rule 1. The best performing series is relative wage costs (WCR).

Rule 10: $$e_t = \gamma_1 + \gamma_{2t} WCR_{t-1} \tag{25}$$

Figure 5 plots this rule and compares it to the simplified rule. As the figure shows, the two plots are virtually indistinguishable.

The conclusion from all these exercises seems to be very clear. For a very wide range of information sets the models properties are practically identical and quite different from the RE solution. We have to move to very poor rules to get any results, which are substantially different from the main cluster of results, and when this does occur the results are generally so poor that the model is clearly not converging in the long run to an expectations equilibria. So any rule which converges on an expectations equilibria seems to give very similar results.

7 Effect of Varying the Degree of Rationality

The natural next question that can be asked is how much rationality is needed in the learning rule for the results to converge to those from a model that assumes rational expectations. The above results show that varying the particular form of the learning rule has little effect on the model's properties. But it may be that there is a big divergence between the properties of the model obtained from the above simulations relative to those from an RE model. To test this the following rule for expectations was assumed,

$$E_t e_{t+1} = \alpha E_{REt} e_{t+1} + (1-\alpha) E_{LEt} e_{t+1} \tag{26}$$

where $E_{REt}e_{t+1}$ is the expected exchange rate assuming rational expectations, $E_{LEt}e_{t+1}$ the expected exchange rate assuming learning and where the parameter α is varied from 0 to 1. The assumed learning rule is Rule 3 (expected exchange rates are assumed to be determined by the actual exchange rate, earnings and import prices, all lagged one quarter) so clearly if α is set to zero then the model behaves exactly as described above and shown in Figure 1. Alternatively if α is set to one then the model behaves exactly as if all expectations are rational and the expectations error is zero, up to a white noise error term. Figure 6 presents the results from two experiments. In the first one α is set to 0.1 implying that there is very little rationality in the assumed form of the learning rule, in the second experiment α is set to 0.9, implying that it is rational expectations that are driving the results. The figure shows however that there are only minor differences between the two simulations. Remember, full RE would imply a straight line of zero in this graph, so just introducing 10% of learning behaviour into the expectations rule moves us almost all the way from the RE solution into the cluster of outcomes that all the learning results inhabit. This strongly suggests that there is a qualitative difference between the full RE outcome and any near RE solution. Any small departure from RE quickly shifts the models properties from the RE ones into something much more representative of the wide body of outcomes which may occur under many near rational learning schemes.

8 Conclusion

This paper set out to ask two basic questions; How should agents select the learning rule which they may use and does their choice have a substantial effect on the properties of an economic model. Second, does the full RE assumption lead to a reasonable approximation of the models properties under a wide range of near rational expectations formation procedures. If it does then the analytical attractiveness of this simple assumption undoubtedly suggests that RE is a sound basic assumption to employ in economic analysis. If it does not however then it throws serious dobt on the use of the full RE assumption in macroeconomic analysis.

Three methods for choosing the series that should enter the learning rule have been assessed. Except for the very worse rules, the results appear to be robust to the chosen form of the learning rule. One conclusion seems to be that learning is able to extract information from any series so well that the exact form of the rule is not important.

However while a very wide range of rules give almost identical answers they all differ substantially from the RE solution. To further check that this is not simply the result of the particular learning process being used we set op a weighted expectations scheme which allowed us to combine full RE expectations with a learning solution.

We demonstrated that very small departures from full RE (less than 10%) moved the models properties most of the way from the RE solution to the learning solution properties. This strongly suggests that RE is does not give rise to representative model properties and that if, in fact, real agents have only very small departures from rationality the RE assumption will not be a good approximation to their behaviour.

References

Bray, M. and Savin, N. (1986). Rational Expectations Equilibrium, Learning and Model Specification. *Econometrica*, 54: 1129-1160.

Collard, F. and Juillard, M. (2001). A Higher-Order Taylor Expansion Approach to Simulation of Stochastic Forward-Looking Models with an Application to a Non-Linear Phillips Curve. Forthcoming in *Computational Economics*.

Currie D.(1985)."Macroeconomic Policy Design and Control Theory- a failed partnership?" *Economic Journal*, 95: 285-306.

Evans, G.W. and Honkapohja, S. (1999). Learning Dynamics, Chapter 7 in *Handbook of Macroeconomics*. Taylor, J.B. and Woodford, M. (eds).

Fourgeaud, C., Gourieroux, C. and Pradel, J. (1986). Learning Procedures and Convergence to Rationality. *Econometrica*, 54: 845-868.

Garratt, A. and Hall, S.G. (1997). E-equilibria and Adaptive Expectations: Output and Inflation in the LBS model. *Journal of Economic Dynamics and Control*, 21: 1149-1171.

Judd, K. (1998). *Numerical Methods in Economics*. MIT Press.

Lucas, R.E. (1972). Expectations and the Neutrality of Money. *Journal of Economic Theory*, 4: 103-124.

Lucas, R. (1979). An Equilibrium Model of the Business Cycle. *Journal of Political Economy*, 83: 1113-1144.

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

Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, 29: 315-335.

Tables and Figures

Table 1: The Effect of a Shock to Eight Series on the Real Euro/Dollar Exchange Rate.

"Important" Series	Shock Size	Effect
W	0.16	14.88
W(-1)	0.16	13.48
PM	2.27	12.17
PM(-1)	2.27	8.38
"Unimportant" Series		
PGC(-1)	0.73	0
RSH(-2)	0.28	0.001
C(-2)	6.06	-0.001
ET	1.00	-0.003

Table 2: Best and Worst Performing Series According to Correlation with the Real Exchange Rate.

"Best" Series	"Worst" Series
IF	K
ET	PC
UP	C
W/P	RPNW

Tables 3a & 3b: Selecting Series According to Principal Components Analysis.

Component Number	Eigenvalue	Cumulative R^2
1	8.85	0.47
2	2.26	0.58
3	1.90	0.68
4	1.08	0.74
5	0.75	0.78
6	0.70	0.82
7	0.72	0.86
8	0.59	0.89
9	0.57	0.92
10	0.48	0.94

Factor Loadings				
Series	Component 1	Component 2	Component 3	Component 4
RRX	0.39	0.48	0.40	0.24
WCR	0.23	0.57	0.36	0.31
M	0.02	0.76	0.24	0.02
X	-0.67	-0.37	0.06	-0.16
PM	0.91	-0.13	0.08	-0.11
PPIN	0.91	-0.08	0.14	-0.16
WP	0.77	-0.36	-0.16	0.22
INF	-0.03	-0.15	0.81	-0.13
RSH	-0.29	-0.45	0.73	0.14
IF	-0.64	0.15	-0.13	0.48
ET	-0.94	-0.07	-0.02	0.02
UP	0.94	0.08	0.02	-0.02
NRR	-0.45	-0.57	0.42	0.15
UWC	0.89	-0.12	0.09	-0.05
GDP	-0.88	-0.15	-0.15	0.02
ER	0.91	-0.11	0.09	-0.10
C	-0.62	0.14	0.18	-0.02
K	-0.36	0.27	0.12	-0.69
RX	-0.77	0.36	0.16	-0.22

Figure 1: Three Learning Rules That Include the “Important” Series.

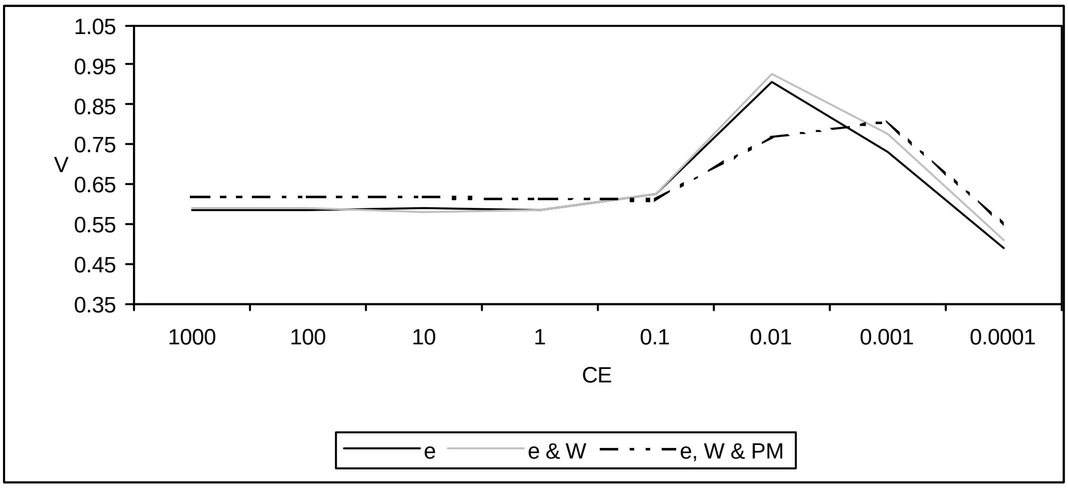


Figure 2: Three Learning Rules that Include the “Unimportant” Series.

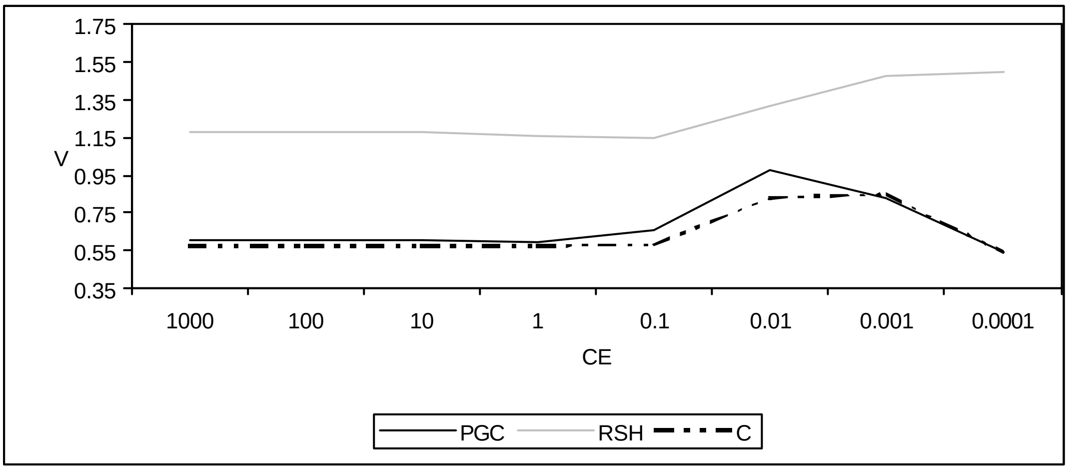


Figure 3: A Comparison of the “Best” (IF) and “Worst” (K) Rules According to the Correlation Exercises.

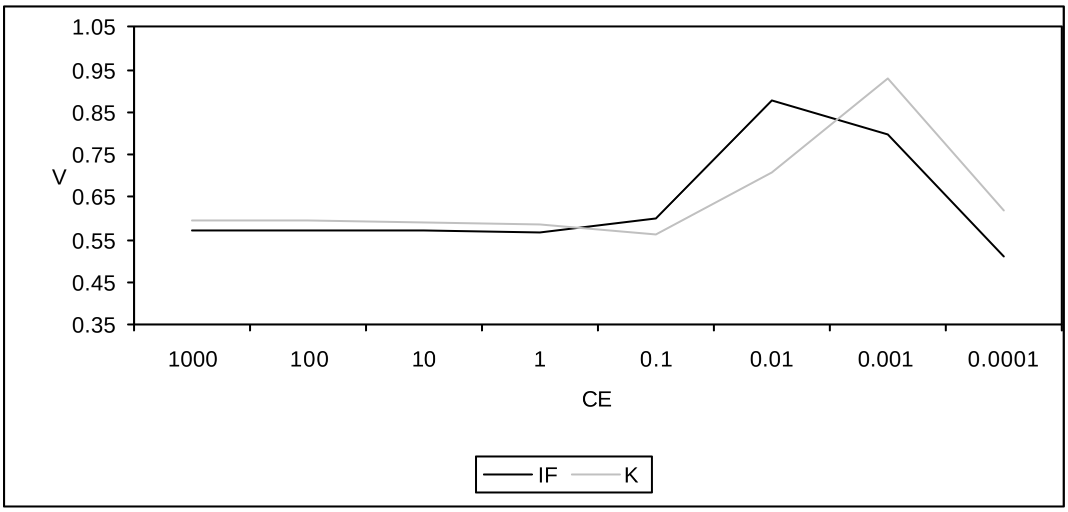


Figure 4: A Comparison of the "Best" Rule According to the Correlation Exercises with the "Best" Rule from the Principal Components Analysis.

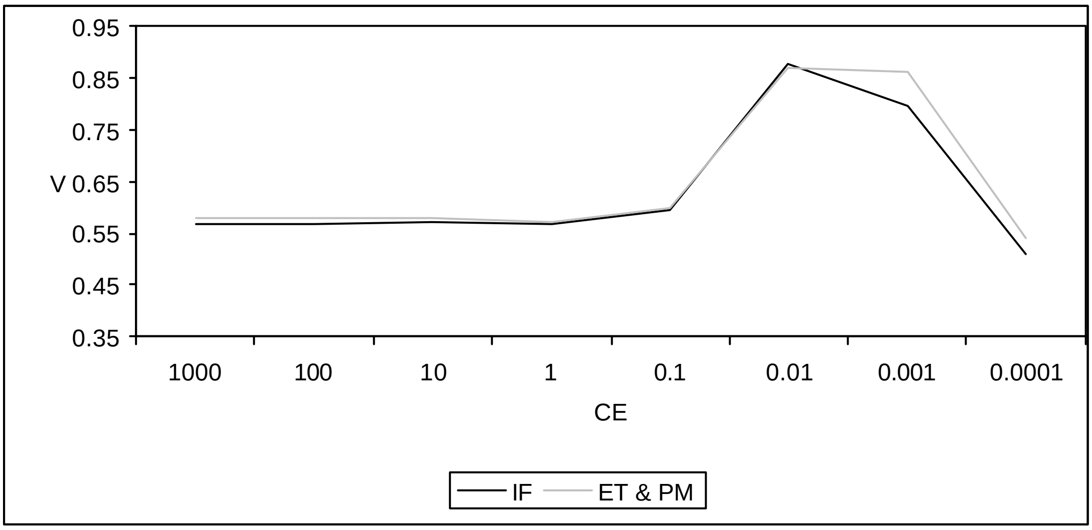


Figure 5: A Comparison of the "Best" Rule According to the Correlation Exercises (Assuming Learning) with a Simple Learning Rule.

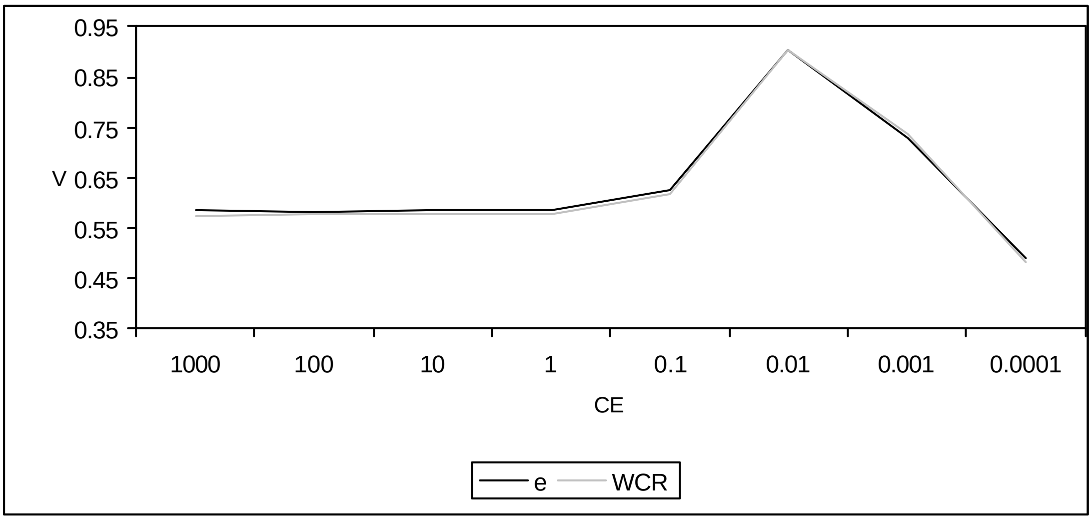
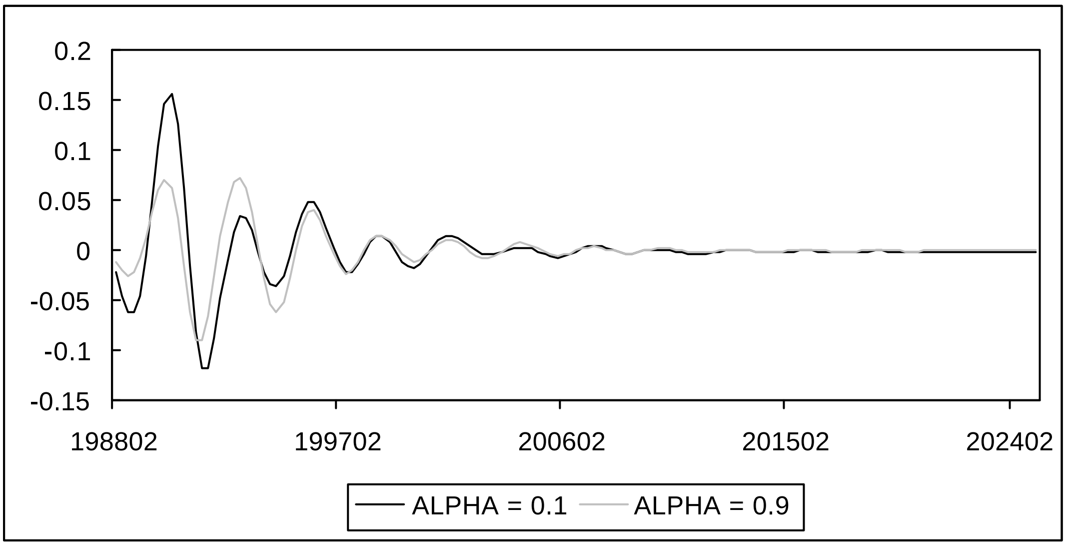


Figure 6: Varying the Degree of Rationality in the Learning Rule?



European Central Bank Working Paper Series

1 “A global hazard index for the world foreign exchange markets” by V. Brousseau and F. Scacciavillani, May 1999.

2 “What does the single monetary policy do? A SVAR benchmark for the European Central Bank” by C. Monticelli and O. Tristani, May 1999.

3 “Fiscal policy effectiveness and neutrality results in a non-Ricardian world” by C. Detken, May 1999.

4 “From the ERM to the euro: new evidence on economic and policy convergence among EU countries” by I. Angeloni and L. Dedola, May 1999.

5 “Core inflation: a review of some conceptual issues” by M. Wynne, May 1999.

6 “The demand for M3 in the euro area” by G. Coenen and J.-L. Vega, September 1999.

7 “A cross-country comparison of market structures in European banking” by O. de Bandt and E. P. Davis, September 1999.

8 “Inflation zone targeting” by A. Orphanides and V. Wieland, October 1999.

9 “Asymptotic confidence bands for the estimated autocovariance and autocorrelation functions of vector autoregressive models” by G. Coenen, January 2000.

10 “On the effectiveness of sterilized foreign exchange intervention” by R. Fatum, February 2000.

11 “Is the yield curve a useful information variable for the Eurosystem?” by J. M. Berk and P. van Bergeijk, February 2000.

12 “Indicator variables for optimal policy” by L. E. O. Svensson and M. Woodford, February 2000.

13 “Monetary policy with uncertain parameters” by U. Söderström, February 2000.

14 “Assessing nominal income rules for monetary policy with model and data uncertainty” by G. D. Rudebusch, February 2000.

15 “The quest for prosperity without inflation” by A. Orphanides, March 2000.

16 “Estimating the implied distribution of the future short term interest rate using the Longstaff-Schwartz model” by P. Hördahl, March 2000.

17 “Alternative measures of the NAIRU in the euro area: estimates and assessment” by S. Fabiani and R. Mestre, March 2000.

18 “House prices and the macroeconomy in Europe: Results from a structural VAR analysis” by M. Iacoviello, April 2000.

19 “The euro and international capital markets” by C. Detken and P. Hartmann, April 2000.

20 “Convergence of fiscal policies in the euro area” by O. De Bandt and F. P. Mongelli, May 2000.

21 “Firm size and monetary policy transmission: evidence from German business survey data” by M. Ehrmann, May 2000.

22 “Regulating access to international large value payment systems” by C. Holthausen and T. Rønde, June 2000.

23 “Escaping Nash inflation” by In-Koo Cho and T. J. Sargent, June 2000.

24 “What horizon for price stability” by F. Smets, July 2000.

25 “Caution and conservatism in the making of monetary policy” by P. Schellekens, July 2000.

26 “Which kind of transparency? On the need for clarity in monetary policy-making” by B. Winkler, August 2000.

27 “This is what the US leading indicators lead” by M. Camacho and G. Perez-Quiros, August 2000.

28 “Learning, uncertainty and central bank activism in an economy with strategic interactions” by M. Ellison and N. Valla, August 2000.

29 “The sources of unemployment fluctuations: an empirical application to the Italian case” by S. Fabiani, A. Locarno, G. Oneto and P. Sestito, September 2000.

30 “A small estimated euro area model with rational expectations and nominal rigidities” by G. Coenen and V. Wieland, September 2000.

31 “The disappearing tax base: Is foreign direct investment eroding corporate income taxes?” by R. Gropp and K. Kostial, September 2000.

32 “Can indeterminacy explain the short-run non-neutrality of money?” by F. De Fiore, September 2000.

33 “The information content of M3 for future inflation” by C. Trecroci and J. L. Vega, October 2000.

34 “Capital market development, corporate governance and the credibility of exchange rate pegs” by O. Castrén and T. Takalo, October 2000.

35 “Systemic risk: A survey” by O. De Bandt and P. Hartmann, November 2000.

36 “Measuring core inflation in the euro area” by C. Morana, November 2000.

37 “Business fixed investment: Evidence of a financial accelerator in Europe” by P. Vermeulen, November 2000.

38 "The optimal inflation tax when taxes are costly to collect" by F. De Fiore, November 2000.

39 "A money demand system for euro area M3" by C. Brand and N. Cassola, November 2000.

40 "Financial structure and the interest rate channel of ECB monetary policy" by B. Mojon, November 2000.

41 "Why adopt transparency? The publication of central bank forecasts" by P. M. Geraats, January 2001.

42 "An area-wide model (AWM) for the euro area" by G. Fagan, J. Henry and R. Mestre, January 2001.

43 "Sources of economic renewal: from the traditional firm to the knowledge firm" by D. R. Palenzuela, February 2001.

44 "The supply and demand for eurosystem deposits – The first 18 months" by U. Bindseil and F. Seitz, February 2001.

45 "Testing the Rank of the Hankel matrix: a statistical approach" by G. Camba-Mendez and G. Kapetanios, February 2001.

46 "A two-factor model of the German term structure of interest rates" by N. Cassola and J. B. Luís, February 2001.

47 "Deposit insurance and moral hazard: does the counterfactual matter?" by R. Gropp and J. Vesala, February 2001.

48 "Financial market integration in Europe: on the effects of EMU on stock markets" by M. Fratzscher, March 2001.

49 "Business cycle and monetary policy analysis in a structural sticky-price model of the euro area" by M. Casares, March 2001.

50 "Employment and productivity growth in service and manufacturing sectors in France, Germany and the US" by T. von Wachter, March 2001.

51 "The functional form of the demand for euro area M1" by L. Stracca, March 2001.

52 "Are the effects of monetary policy in the euro area greater in recessions than in booms?" by G. Peersman and F. Smets, March 2001.

53 "An evaluation of some measures of core inflation for the euro area" by J.-L. Vega and M. A. Wynne, April 2001.

54 "Assessment criteria for output gap estimates" by G. Camba-Méndez and D. R. Palenzuela, April 2001.

55 "Modelling the demand for loans to the private sector in the euro area" by A. Calza, G. Gartner and J. Sousa, April 2001.

56 “Stabilization policy in a two country model and the role of financial frictions” by E. Faia, April 2001.

57 “Model-based indicators of labour market rigidity” by S. Fabiani and D. Rodriguez-Palenzuela, April 2001.

58 “Business cycle asymmetries in stock returns: evidence from higher order moments and conditional densities” by G. Perez-Quiros and A. Timmermann, April 2001.

59 “Uncertain potential output: implications for monetary policy” by M. Ehrmann and F. Smets, April 2001.

60 “A multi-country trend indicator for euro area inflation: computation and properties” by E. Angelini, J. Henry and R. Mestre, April 2001.

61 “Diffusion index-based inflation forecasts for the euro area” by E. Angelini, J. Henry and R. Mestre, April 2001.

62 “Spectral based methods to identify common trends and common cycles” by G. C. Mendez and G. Kapetanios, April 2001.

63 “Does money lead inflation in the euro area?” by S. N. Altimari, May 2001.

64 “Exchange rate volatility and euro area imports” by R. Anderton and F. Skudelny, May 2001.

65 “A system approach for measuring the euro area NAIRU” by S. Fabiani and R. Mestre, May 2001.

66 “Can short-term foreign exchange volatility be predicted by the Global Hazard Index?” by V. Brousseau and F. Scacciavillani, June 2001.

67 “The daily market for funds in Europe: Has something changed with the EMU?” by G. P. Quiros and H. R. Mendizabal, June 2001.

68 “The performance of forecast-based monetary policy rules under model uncertainty” by A. Levin, V. Wieland and J. C.Williams, July 2001.

69 “The ECB monetary policy strategy and the money market” by V. Gaspar, G. Perez-Quiros and J. Sicilia, July 2001.

70 “Central Bank forecasts of liquidity factors: Quality, publication and the control of the overnight rate” by U. Bindseil, July 2001.

71 “Asset market linkages in crisis periods” by P. Hartmann, S. Straetmans and C. G. de Vries, July 2001.

72 “Bank concentration and retail interest rates” by S. Corvoisier and R. Gropp, July 2001.

73 “Interbank lending and monetary policy transmission – evidence for Germany” by M. Ehrmann and A. Worms, July 2001.

74 “Interbank market integration under asymmetric information” by X. Freixas and C. Holthausen, August 2001.

75 “Value at risk models in finance” by S. Manganelli and R. F. Engle, August 2001.

76 “Rating agency actions and the pricing of debt and equity of European banks: What can we infer about private sector monitoring of bank soundness?” by R. Gropp and A. J. Richards, August 2001.

77 “Cyclically adjusted budget balances: An alternative approach” by C. Bouthevillain, P. Cour-Thimann, G. van den Dool, P. Hernández de Cos, G. Langenus, M. Mohr, S. Momigliano and M. Tujula, September 2001.

78 “Investment and monetary policy in the euro area” by B. Mojon, F. Smets and P. Vermeulen, September 2001.

79 “Does liquidity matter? Properties of a synthetic divisia monetary aggregate in the euro area” by L. Stracca, October 2001.

80 “The microstructure of the euro money market” by P. Hartmann, M. Manna and A. Manzanares, October 2001.

81 “What can changes in structural factors tell us about unemployment in Europe?” by J. Morgan and A. Mourougane, October 2001.

82 “Economic forecasting: some lessons from recent research” by D. Hendry and M. Clements, October 2001.

83 “Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts” by K. F. Wallis, November 2001.

84 “Data uncertainty and the role of money as an information variable for monetary policy” by G. Coenen, A. Levin and V. Wieland, November 2001.

85 “Determinants of the euro real effective exchange rate: a BEER/PEER approach” by F. Maeso-Fernandez, C. Osbat and B. Schnatz, November 2001.

86 “Rational expectations and near rational alternatives: how best to form expecations” by M. Beeby, S. G. Hall and S. B. Henry, November 2001.