Learning, the Forward Premium Puzzle and Market Efficiency

Avik Chakraborty[1]

October 2004

Abstract: The Forward Premium Puzzle is one of the most prominent empirical anomalies in international finance. The forward premium predicts exchange rate depreciation but typically with the opposite sign and smaller magnitude than specified by rational expectations, a result also considered to indicate inefficiency in the foreign exchange market. This paper proposes a resolution of the puzzle based on recursive least squares learning applied to a simple model of exchange rate determination. The key assumption is that risk neutral agents are not blessed with rational expectations and do not have perfect knowledge about the market. Agents learn about the parameters underlying the stochastic process generating the exchange rate using constant gain recursive least squares. When exchange rate data are generated from the model and the empirical tests are performed, for plausible parameter values the results replicate the anomaly along with other observed empirical features of the forward and spot exchange rate data.

Keywords: Spot Exchange Rate, Forward Rate, Constant-gain Recursive Least Squares Learning.

---

[1] Department of Economics, 1285 University of Oregon, OR 97403, (541) 346 4664,
email: achakrab@uoregon.eduhttp://www.uoregon.edu/~achakrab
The paper is adopted from Chapter 4 of doctoral dissertation of Avik Chakraborty.

1. Introduction

The 'Forward Premium Puzzle' is a long-standing empirical paradox in international finance. The puzzle refers to the finding that the forward exchange rate consistently predicts the expected depreciation in the spot exchange rate but with a smaller magnitude and often the opposite sign than specified by rational expectations. A large literature documents and attempts to explain the puzzle, but mostly with very mixed success[2]. This paper proposes a resolution from a new perspective.

The common usage of the assumption of 'Rational Expectations' on the part of market agents seems questionable because it assumes that agents have complete knowledge of the stochastic process that generates the exchange rate. Although, in most cases agents seem thoroughly informed and the market is sufficiently efficient so new information is immediately incorporated into variables, it is unlikely that agents have knowledge about every single parameter value in the stochastic processes that generate the fundamentals and the exchange rate. Moreover, the 'Rational Expectations' model does not seem to provide a satisfactory explanation for the puzzle[3].

Efforts have been made to explain the puzzle dropping the assumption of 'Rational Expectations'. These models introduced different forms of belief distortions on the part of the participating agents. The novelty of this paper is the use of the 'least square learning' rule in expectation formation to explain the puzzle[4]. This assumes that agents do not know the exact value of the parameters and instead attempt to learn them using econometric techniques.

According to theory, if the future rate of depreciation in the exchange rate is regressed on the forward premium (the forward rate less the current spot rate in logarithms), then the slope coefficient on the forward premium should be unity

---

[2] Cumby and Obstfeld (1984), Fama (1984), and Hodrick and Srivastava (1986) are a few of the articles discussing the puzzle. For surveys on the puzzle, see Taylor (1995), Engel (1996) and Lewis (1995).

[3] The approaches used by researchers to try to solve the puzzle are discussed later. Most authors conclude that in order to solve the puzzle rejection of the null hypothesis of rational expectation is helpful.

[4] Using learning to explain exchange rate expectation related anomalies is not completely new. Lewis (1991) used Bayesian Learning in an exchange rate pricing model to study the Peso-Problem. In another paper (1989) she investigated the effect of Bayesian Learning using a structural shift in fundamentals to explain the puzzle. A summary of her methodology is discussed later. In another study Kim (2003) used recursive least squares learning to try to explain the exchange rate determination puzzle.

provided the market is efficient and agents do not make systematic errors in their forecast[5]. More formally, if s_t is the natural log of the current spot exchange rate (defined as the domestic price of foreign exchange), Δs_{t+1} is the depreciation[6] of the natural log of the spot exchange rate from period t to $t+1$, i.e., $\Delta s_{t+1} = s_{t+1} - s_t$, and F_t is the natural log of the one-period forward rate at period t, and e_{t+1} is a white noise error term, then in the following regression equation

$$\Delta s_{t+1} = \alpha + \beta \left(F_t - s_t \right) + e_{t+1}$$

$\hat{\beta}$ should be unity. The hypothesis is based on assumptions of risk-neutrality, efficient markets and rational expectations. If agents are risk neutral then they must set today's forward rate equal to their expectation about the future spot rate. Moreover, if their expectations are rational and the market is efficient, so that new information is readily absorbed, then they should not make systematic forecasting errors. In other words the future spot rate must be equal to the sum of today's forward rate and a random error term with zero expected value, i.e. $s_{t+1} = F_t + u_{t+1}$ where, u_{t+1} is a random error term with $E_t(u_{t+1}) = 0$, denoting the mathematical expectation of u_{t+1} conditioned on information available at time t.

In this model the actual depreciation must equal the expected depreciation on average. In other words, there could be an error in expectation formation in a given period, but over a significant number of observations the average of the expectations about depreciation must equal the average of the actual depreciation.

There are two key assumptions implicit in the above theoretical proposition, viz. risk neutrality and rational expectations by the participating agents[7]. A significant volume of research has empirically tested this hypothesis, and concludes that the OLS estimate of β ($\hat{\beta}$) is significantly less than 1 and in fact in the majority of cases is significantly less than zero. The average $\hat{\beta}$ across 75 early studies was found to be –0.88

---

[5] This regression was originally run to test the efficient market hypothesis.
[6] Given the definition of s_t a positive Δs_{t+1} is a depreciation of the domestic currency and vice-versa.
[7] By rational expectations it is meant that agents do not make systematic forecasting errors.

(Froot, 1990), and only a few of those estimates were positive[8]. The result is paradoxical because the evidence strongly refutes the theoretical prediction that $\beta = 1$, and hence apparently contradicts the efficient market hypothesis. This is the much renowned anomaly.

In this paper 'Constant-gain Recursive Least Square Learning'[9] is attributed to the participating agents in the market. Under this assumption agents form expectations about future values of the spot exchange rate. This assumption is combined with a simple model of exchange rate determination. The spot exchange rate and the forward rate are generated by this model, and the consequent depreciation is regressed on the forward premium. Under certain plausible sets of parameter values, the results are found to be strikingly similar to that obtained from actual exchange rate data, including a negative estimate of β.

The paper is organized as follows. Section 2 presents a summary of the existing literature. Section 3 introduces a simple model of exchange rate determination. Section 4 discusses the learning dynamics used to update the process parameter values and form the expectation. Section 5 presents the simulation results. Section 6 presents empirical results from parameter estimation. Section 7 concludes.

2. The Existing Literature

The enormous research devoted towards finding a solution for the puzzle could be categorized into two different approaches, a risk-premium approach and a distorted belief approach[10].

---

[8] Others in this vast literature have found the average value of the coefficient to diminish in absolute value but remain negative. Interested readers may refer to Froot(1990), Frankel and Froot(1989), Froot and Thaler(1990), Engel(1996), Lewis(1995) and Mark and Wu(1998).

[9] In our set-up decreasing-gain learning converges to rational expectations equilibrium and hence, would not offer a solution. A discussion of the two types of gain in learning behavior is developed in Evans and Honkapohja (2001).

[10] There is a third and less explored approach that addresses an econometric issue related to the puzzle - an omitted variable problem. The idea is that the regression incorporates an omitted variable bias because the error term in the regression equation $\Delta s_{t+1} = \alpha + \beta \left(F_t - s_t \right) + u_{t+1}$ is correlated with s_t under the alternative hypothesis, i.e. $\beta \neq 1$. This was demonstrated by Evans and Lewis (1993). Interested readers may also refer to Oh and Pippenger(1994), Mark and Wu(1998), Mark (2001) and Chakraborty (2004).

Risk-Premium Approach

The risk premium approach assumes that investors in the foreign exchange market are risk-averse. Thus, the forward rate not only incorporates their expectation about the future depreciation but also includes a risk-premium as a hedge against the risk from investing in a more volatile asset characterized by a higher rate of return. Consequently, the interest differential between two countries not only reflects the possibility of future depreciation but also a risk premium. As a result, the expected depreciation erroneously forecasts the actual depreciation, yielding a regression coefficient that is less than one. The situation reaches an extreme when the risk premium so dominates the expectation about depreciation that a negative coefficient is obtained. Fama (1984) demonstrates that, for this to happen, the variance of the risk premium must be greater than the variance of the expected depreciation and their covariance must be negative.

Efforts were made to empirically evaluate the risk premium as a potential explanation for the paradox. This research followed three different paths[11]. Despite its intuitive appeal, these tests demonstrated the difficulty of the risk premium approach to provide a satisfactory explanation of the puzzle. Fama's (1984) decomposition was unable to find clear empirical support. Frankel and Froot (1989) used survey data to test a time varying risk premium hypothesis, with the conclusion that little evidence supported this explanation. Frankel and Chinn (1993) tested the risk-premium hypothesis on a cross section of 17 countries and did not find evidence in support of their hypothesis. Backus, Foresi and Telmer (2001) concluded that affine term structure in the risk-premium model failed to provide a theoretically satisfactory or empirically tractable explanation for the puzzle. Other researchers found similar results[12]. This led to a general rejection of the risk-premium explanation and encouraged researchers to be

---

[11] An account of the research is given in Froot and Thaler(1990) and Backus, Foresi and Telmer (2001).

[12] See Mark(1985), Obstfeld(1990) as mentioned by Froot and Thaler(1990). In an extension of Frankel and Froot (1989), Cavaglia, Verschoor and Wolff (1994) found that a combination of both the risk premium and deviation from rational expectations provides a better explanation of the puzzle. An exception was observed by Bams, Walkowiak and Wolff (2004). Using an econometric panel data model, they found that the dynamics of the forward prediction error was mostly governed by the common factor in the dollar risk premia.

more receptive to the possibility of distorted beliefs or errors in expectations as a potential explanation.

Distorted Belief or Non-rational Expectations Approach

The potential ability of non-rational expectations to explain the results is apparent from some of the other findings related to exchange rate behavior. De Long et. al. (1990) demonstrated that the presence of both rational and non-rational traders in the market tend to distort asset prices significantly away from the fundamental values and therefore has a potential to explain many financial market anomalies. Goldberg and Frydman (1996) showed that, if investors could predict the direction of the change in exchange rate, but did not know the magnitude of the change, monetary models could explain certain idiosyncrasies of the foreign exchange market. Bacchetta and Wincoop (2003) concluded that the comparative variation in exchange rate and its fundamentals, both in short run and long run, could be explained by the existence of informational heterogeneity amongst the investors. Mark and Wu (1998) demonstrated that the behavior of the variance and covariance of the risk premium as suggested by Fama (1984) does not have empirical support. On the other hand the existence of noise traders[13] in the market under certain numerical assumptions yields results that are observed in data. This paper is motivated by this research, which suggests non-rational expectations in foreign exchange markets. In the context of the forward premium puzzle, bounded rationality may justify the systematic under-prediction of future depreciation.

The key idea underlying this paper is that the agents are aware of a stochastic process that generates certain variable values and they also know the form of the process. However, they do not know the actual values of the parameters that govern the stochastic process. Instead, they learn these values over time using econometric regression over past observed data in each period, updating the previously estimated coefficient values as new data become available. They make use of all available information during forecasting. This feature attributes some degree of rationality to the

---

[13] Noise traders deviate from rationality by their inability to distinguish between real and pseudo signals. Therefore, their trading decisions tend to move the exchange rate away from fundamentals. See Mark and Wu (1998), De Long et. al. (1990) for more discussion on noise traders.

agents. However, their lack of knowledge about the parameter values prevents them from being completely rational in the usual sense. Consequently, they are said to have 'bounded rationality', i.e. rationality with some constraints[14].

Lewis (1989) used Bayesian Learning to provide an explanation for the forward premium puzzle. Her model predicted the actual movement in the US dollar during the early 80's in fifty per cent of the cases. In her model the agents were boundedly rational in learning the shift in the money demand structure, but did not know it perfectly. However, the model could not explain the persistence of prediction errors, as the magnitude of the error died down over time. Nonetheless, it provided an indication of the potential of learning dynamics for understanding exchange rate movements.

3. The Model

Small economy case

The model consists of money market equilibrium and the international market parity conditions for a simple economy[15]. The economy is characterized by the following assumptions:

1)   The economy is small and open. Hence, the foreign price and the foreign rate of interest are exogenously given for the economy.

2)   Agents are risk-neutral so that current forward rate is the expected future spot rate.

$$F_t = E^*_t \, s_{t+1} \qquad (1)$$

Together with covered interest parity equation (1) implies uncovered interest parity.

3)   The economy is characterized by flexible prices and flexible nominal exchange rates. Hence, the domestic price level and nominal exchange rate readily adjust in response to a change in the foreign price level to maintain purchasing power parity.

4)   The exchange rate fundamentals follow either a random walk or an AR(1) process with a large auto-regression coefficient.

5)   There are no transaction costs.

---

[14] For more discussion on learning dynamics and corresponding expectation formation, see Evans and Honkapohja (2001).

[15] This version of the model is adopted from Evans (1986).

6)   Agents are attributed 'bounded rationality'. Thus, agents use all available information to form expectations but they have imperfect information about the market. They know the form of the stochastic process that generates the exchange rate but they do not know the underlying parameter values. Each period they update their knowledge about these values using constant gain least squares learning.

The economy is described using three key equations.

$$m_t - p_t = a + by_t - ci_t \qquad (2)$$

$$i_t = i_t^* + E^*_t \, s_{t+1} - s_t \qquad (3)$$

$$p_t = p_t^* + s_t \qquad (4)$$

The first equation represents the money market equilibrium. All lower case letters except the interest rate represent the natural log of the corresponding variables. m_t is the log of money supply, p_t is the log of domestic price level, y_t is log of real output and i_t is the one period domestic interest rate. Parameters b and c are positive constants. The second equation corresponds to the second assumption, i.e. uncovered interest parity. i_t, i_t^*, s_t are the domestic one period interest rate, foreign one period interest rate and log of the spot exchange rate, respectively. The exchange rate is the price of foreign currency in terms of domestic currency. $E^*_t \, s_{t+1}$ is the log of the expected one period future spot exchange rate, i.e. the expectation formed at period 't' about the spot exchange rate in period 't+1'. The third equation represents the third assumption, purchasing power parity. p_t^* is the log of foreign price level. The log of the one period forward rate is given by F_t. Since, all agents are risk neutral by assumption we may rule out the possibility that the forward rate has a risk premium. Therefore, the forward rate is nothing but the expected spot rate, yielding equation (1).

Under rational expectations this would imply $s_t = E^*_t \, s_{t+1} + u_{t+1}$ where $E_t(u_{t+1}) = 0$; i.e.,

$$s_{t+1} - s_t = \alpha + \beta(E^*_t \, s_{t+1} - s_t) + u_{t+1} \quad (5)$$

with $\alpha = 0$ and $\beta = 1$.

However, under the assumption of bounded rationality the value of the two regression coefficients may differ from the rational values. To determine these values the

type of 'bounded rationality' has to be formalized. The first step is to determine how the exchange rate is generated in the economy.

Solving (2)-(4), we obtain

$$s_t = \theta E^*{}_t\, s_{t+1} + v_t \qquad\qquad (6)$$

where, $0 < \theta = c(1+c)^{-1} < 1$ and $v_t = (1+c)^{-1}(m_t - p_t^* - a - by_t + ci_t^*)$.

Large economy case

For the large economy case we retain all the above assumptions, except that the foreign rate of interest i_t^* and foreign price level p_t^* are no longer exogenous. The decisions of the home country impact on the variables of the foreign country and hence, have repercussion effects. The money market equilibrium of the foreign country must also be considered. Thus, the foreign money market equilibrium is

$$m_t^* - p_t^* = a + by_t^* - ci_t^* \qquad\qquad (2a)$$

which is combined with equations (2) – (4)[16]. This yields the same equation (6), with the only difference that in the large economy case we have

$$v_t = (1+c)^{-1}(m_t - m_t^* - b(y_t - y_t^*))$$

Thus, for analytical purposes we will use the general form (6).

In equation (6), v_t is exogenous, representing the exchange rate fundamentals, and is determined from the parameters of the domestic money demand equation and foreign price level and rate of interest in each period. It should be noted that v_t has time varying components and one constant component. Thus, equation (6) could be rewritten separating these two components as

$$s_t = \mu + \theta E^*{}_t\, s_{t+1} + v_t \qquad\qquad (7)$$

where, $\mu = \dfrac{-a}{(1+c)}$ and $v_t = (1+c)^{-1}(m_t - p_t^* - by_t + ci_t^*)$.

Assumption 4 says that the fundamentals follow an AR(1) process with a large auto-regression coefficient. This implies

---

[16] In this model, however, there is an implicit assumption made for convenience. The a, b and c are same for both the countries. The assumption on a and b are not so crucial analytically, but that on c is. However, since, c is the interest semi-elasticity of money demand, does not appear to be greatly different between countries.

$$v_t = \rho v_{t-1} + \varepsilon_t \qquad (8)$$

where, by assumption ρ is large (close to 1). Equations (7) and (8) together govern the exchange rate generating process in this economy and hold even when agents form their expectations with 'bounded rationality'.

4. Learning and Expectation

Suppose that the agents estimate the linear projection of s_t on v_{t-1} and an intercept, i.e., s_t depends on v_{t-1}. Thus, suppose they believe the relationship has the following form:

$$s_t = a + bv_{t-1} + c\varepsilon_t \qquad (9)$$

where, ε_t is the white noise from the v_t process. The assumed form of the relationship is known as the Perceived Law of Motion (PLM).

Under rational expectations, the agents not only know the form of the above equation but they also know the values of a, b and c. In this case their knowledge will produce the following solution[17]:

$$s_t = \overline{a} + \overline{b}v_{t-1} + \overline{c}\varepsilon_t \qquad (10)$$

where, $\overline{a} = (1-\theta)^{-1}\mu$, $\overline{b} = (1-\rho\theta)^{-1}\rho$ and $\overline{c} = (1-\rho\theta)^{-1}$.

Under a learning mechanism the story is different. In this case the agents do not have perfect knowledge about equation (9). They are aware of its functional form but lack knowledge of the values of the key coefficients a and b. At the end of period t-1, they therefore form an estimate a_{t-1} and b_{t-1} of the coefficients by applying econometric techniques using the available information, i.e. the estimated values in the past and the new information in period t-1. In period t they forecast the value of s_{t+1} using a_{t-1}, b_{t-1} and v_t :

$$E^*_t s_{t+1} = a_{t-1} + b_{t-1}v_t \qquad (11)$$

where, $E^*_t s_{t+1}$ denotes the (possibly non-rational) expectation of s_{t+1} formed at time t.

---

[17] The derivation of the rational expectation solution is explained in the Appendix.

At the end of period t, agents calculate the forecast error by measuring the difference between the actual s_t and its predicted value and use this information to update their estimate of the coefficients to a_t and b_t. This process is repeated every period generating a sequence of estimates (a_t, b_t).

Under the E-stability conditions[18], a decreasing gain least squares learning process eventually converges to the rational expectations solution as given by equation (10), and thus would only generate transitional deviations from rational expectations. We therefore instead adopt an alternative form, constant gain least squares learning. A constant gain implies that every period the agents update their estimate of the parameters in a way that remains sensitive, even asymptotically, to forecast errors. Thus, convergence never occurs, and even under E-stability conditions the estimates fluctuate around the rational expectations solution. Since the estimates never converge to the rational value, the expectations formed using these estimates are not perfectly rational even in the long time horizon.

An obvious question concerns the justification for constant gain learning, i.e., what prevents the agents from learning the parameter values perfectly over a long time horizon. The standard rationale provided in the learning literature is that the true parameter values themselves may vary over time due to structural shifts taking an unknown form[19]. This is possible as the parameters in question here, a and b, represent the relationship between the exchange rate and its fundamentals[20].

Under constant gain least squares, parameters are updated using the following algorithms.

If we define matrices $z_t = \begin{pmatrix} 1 \\ v_t \end{pmatrix}$ and $\Phi_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix}$ then the Recursive Least Square estimation is done using the following pair of equations:

---

[18] The meaning of E-stability conditions and their implications for learning are discussed in the Appendix

[19] The rationale for constant gain learning is discussed in Cho, Williams and Sergent (2002), Kasa (2002), Evans and Honkapohja (2001) and Sergent (1999). The explanation is that, when agents know that the structural change may cause the parameters evolving over time, it provides the motivation to use a constant gain sequence that does not diminish over time.

[20] This, of course, is something that one could explore empirically – is there structural instability in the relationship between s_t and fundamentals v_{t-1}?

$$\Phi_t = \Phi_{t-1} + \gamma S_{t-1}^{-1} z_{t-1} (s_t - \Phi_{t-1}' z_{t-1}) \quad (12)$$

$$S_t = S_{t-1} + \gamma (z_t z_t' - S_{t-1}) \quad (13)$$

Equation (12) gives the updating process. At period t the estimates of the parameters Φ_t are obtained by adjusting last period's estimate Φ_{t-1} with a gain factor of the prediction error $(s_t - \Phi_{t-1}' z_{t-1})$. The gain factor is $\gamma S_{t-1}^{-1} z_{t-1}$, where γ is the amount of constant gain each period and S_{t-1}^{-1} comes from equation (13). Equation (13) uses a square matrix S_t to update the second moments' matrix of z_t i.e. of $E(z_t z_t')$.

The above system (consisting of equations (12) and (13)) is a Stochastic Recursive Algorithm[21]. Using this algorithm a pair of values (a_{t-1}, b_{t-1}) are obtained at the end of (t-1). Then agents use these values at t to obtain an expected value for the exchange rate next period, i.e. $E_{t}^{*} s_{t+1}$ according to equation (11). Then substituting this value along with the value of v_t in equation (7) generates an exchange rate for period t.

Thus, by applying constant gain learning, the agents form an expectation about the future exchange rate, which is the same as the forward rate under the assumption of risk neutrality. This expected value helps determine the exchange rate as described above. This process of exchange rate generation and expectation formation takes place under 'bounded rationality', which is a deviation from rational expectations. Thus, if regressions using data generated from this process produce coefficients similar to those from actual data, then the model could provide an explanation of the puzzle. This leads to the next section, the simulation of the model.

5. Simulation

Parameter choices

The following parameter values are to be specified in order to carry out the simulation: μ, θ, ρ and γ. μ was arbitrarily set to 1 (the value of μ does not influence the simulations of the model significantly). $\theta = \dfrac{c}{1+c}$, where c is the interest

---

[21] For a detailed discussion on Stochastic Recursive Algorithms and the intuition behind equations (12) and (13) see, Evans and Honkapohja (2001).

semi-elasticity of money demand. Thus, θ must be less than unity (as c >0). A plausible value of c is 2 for annual data and 8 for quarterly data[22], which implies that θ is 0.67 for annual data and 0.89 for quarterly data. To explore the sensitivity of results to alternative values, θ was assigned three different values (0.1, 0.6 and 0.9) and simulations were run each time to observe the realized $\hat{\beta}$ (the forward premium regression coefficient estimate) values. Similarly, four different values for ρ were chosen, viz. 0.9, 0.95, 0.99 and 1.0 and each time simulations were run[23]. The three γ values chosen were 0.01, 0.05 and 0.1[24].

The initial values of the variables were chosen arbitrarily. s_t was initialized at 0, i.e., the exchange rate was assumed to be 1. a and b were set at their respective rational expectations values in the beginning. v_t was set to zero.

Process and Results

The simulation was run for three different sample sizes, viz. 50, 100 and 400. With each sample size the simulation was run 1000 times. Each simulation generated one $\hat{\beta}$. The averages of the $\hat{\alpha}$ and $\hat{\beta}$ estimates, as well as their average standard errors, were calculated for the 1000 simulations. The averages of the t-statistics for the null hypothesis $\beta = 1$ and the R^2s were also calculated. Each of the parameters θ and γ was assigned three different values and ρ was assigned four different values. A combination for each of those values was chosen for each simulation. Thus, 36 (3 x 4 x 3) possible sets of parameter values were considered and for each combination a simulation was run 1000 times.

---

[22] c is the interest semi-elasticity of money demand. Stock and Watson (1993) estimated c to be .02 for annual data with the interest rate stated in percent per annum.

[23] Many macro variables are known to have a Random Walk component. Thus we focus on cases in which v_t follows a Random Walk or an AR(1) with very high value of coefficient γ. This proposition is tested empirically.

[24] The values of γ for the simulations were chosen following the literature. Orphanides and Williams (2003) suggested that $\gamma = 0.02$ is appropriate. Branch and Evans (2004) stated that for forecasting US inflation and GDP growth γ in the range 0.01 to 0.07 appear to work well for quarterly data.

Table: 1 Simulation Results (No. of simulations = 1000)

				Sample size 50					
θ	ρ	γ	Avg. $\hat{\alpha}$	Avg. SE($\hat{\alpha}$)	Avg. $\hat{\beta}$	Avg. SE($\hat{\beta}$)	S.D.* of $\hat{\beta}$	Avg. t-stat $\beta=1$	Avg. R^2
0.1	0.9	0.01	-0.0063	0.1763	1.1864	0.8621	0.8240	0.2279	0.0547
		0.05	0.0061	0.1624	0.2970	0.6435	0.7164	-1.0344	0.0329
		0.1	-0.0040	0.1608	0.0146	0.4712	0.5041	-2.0802	0.0250
	0.95	0.01	0.0023	0.1797	0.7152	1.3215	1.5167	-0.1897	0.0322
		0.05	-0.0026	0.1636	-0.0221	0.7208	0.7993	-1.4186	0.0251
		0.1	0.0040	0.1600	-0.1231	0.4971	0.5258	-2.2835	0.0235
	0.99	0.01	0.0210	0.1775	-0.3381	1.5247	1.8810	-1.0722	0.0283
		0.05	-0.0028	0.1677	-0.2909	0.7216	0.7786	-1.8624	0.0235
		0.1	0.0088	0.1605	-0.1992	0.4884	0.4954	-2.5258	0.0222
	1.0	0.01	-0.0115	0.1809	-0.5743	1.4679	1.7737	-1.3255	0.0284
		0.05	-0.0155	0.1678	-0.3189	0.7110	0.7673	-1.9508	0.0240
		0.1	-0.0058	0.1613	-0.1935	0.4828	0.4767	-2.5481	0.0216
0.6	0.9	0.01	0.0092	0.3592	1.5586	0.8681	0.7978	0.5856	0.0700
		0.05	-0.0044	0.3275	0.4284	0.8950	1.2819	-0.4754	0.0410
		0.1	-0.0050	0.3155	-0.2412	0.7682	1.0845	-1.4770	0.0327
	0.95	0.01	-0.0402	0.4115	1.6343	1.6418	2.1036	0.4418	0.0475
		0.05	-0.0092	0.3506	-0.5220	1.2689	1.8510	-1.0689	0.0373
		0.1	-0.0054	0.3411	-0.7824	0.8653	1.2718	-1.9869	0.0444
	0.99	0.01	0.0039	0.4151	-1.1064	2.9373	4.2611	-0.8558	0.0408
		0.05	0.0168	0.4098	-1.1940	1.3304	1.7672	-1.7324	0.0417
		0.1	-0.0150	0.3766	-1.0301	0.8651	1.1113	-2.3789	0.0450
	1.0	0.01	-0.0344	0.4400	-2.0072	2.9420	4.0623	-1.3465	0.0435
		0.05	0.0312	0.4271	-1.2488	1.2899	1.2899	-1.8857	0.0407
		0.1	-0.0079	0.3984	-0.9663	0.8340	1.0426	-2.4461	0.0440
0.9	0.9	0.01	-0.0033	0.9009	1.6805	0.8359	0.8358	0.6908	0.0777
		0.05	-0.0132	0.8849	1.2297	0.8739	0.7418	0.2267	0.0506
		0.1	-0.0062	0.8446	0.7001	0.9219	0.9024	-0.2707	0.0334
	0.95	0.01	0.0944	1.3522	2.5745	1.5451	1.6436	0.8894	0.0578
		0.05	-0.0168	1.2439	1.0878	1.7678	2.6799	0.2193	0.0362
		0.1	0.0017	1.1863	-0.9495	1.8568	3.6228	-0.6997	0.0364
	0.99	0.01	0.1485	1.7722	1.0041	6.9409	11.4989	0.0564	0.0441
		0.05	-0.0994	1.6408	-3.9030	4.0591	6.9428	-1.1111	0.0534
		0.1	0.0588	1.6698	-3.8030	2.5986	4.7087	-1.7057	0.0611
	1.0	0.01	-0.0034	1.8561	-7.9905	10.0193	15.2593	-1.1877	0.0538
		0.05	-0.1018	1.9334	-5.0424	4.0030	6.6568	-1.5623	0.0576
		0.1	-0.1680	1.9736	-3.9831	2.4968	4.4488	-1.9368	0.0665

*Standard Deviation across 1000 simulations

$$\theta = \frac{c}{1+c}$$

γ = Size of the constant gain in learning

Table: 2 Simulation Results (No. of simulations = 1000)

θ	ρ	γ	Avg. $\hat{\alpha}$	Avg. SE($\hat{\alpha}$)	Avg. $\hat{\beta}$	Avg. SE($\hat{\beta}$)	S.D.* of $\hat{\beta}$	Avg. t-stat $\beta=1$	Avg. R^2
				Sample size 100					
0.1	0.9	0.01	0.0037	0.1157	0.9851	0.5177	0.3705	-0.0176	0.0441
		0.05	8.3765e-004	0.1128	0.2952	0.4365	0.4468	-1.5524	0.0194
		0.1	-4.5942e-004	0.1116	0.0687	0.3218	0.3651	-2.8547	0.0149
	0.95	0.01	-0.0020	0.1182	0.5653	0.8042	0.8473	-0.4724	0.0189
		0.05	2.6224e-004	0.1137	-0.0710	0.5086	0.6072	-2.0461	0.0140
		0.1	7.3811e-004	0.1117	-0.0813	0.3384	0.3706	-3.1833	0.0122
	0.99	0.01	-0.0021	0.1217	-0.3049	1.0047	1.1306	-1.3442	0.0121
		0.05	0.0031	0.1143	-0.2438	0.5150	0.5427	-2.4275	0.0112
		0.1	0.0013	0.1116	-0.1526	0.3435	0.3379	-3.3823	0.0105
	1.0	0.01	-0.0028	0.1250	-0.5249	0.9383	1.1208	-1.7217	0.0135
		0.05	0.0046	0.1164	-0.2964	0.5060	0.5237	-2.5964	0.0113
		0.1	0.0037	0.1125	-0.1428	0.3310	0.3451	-3.4842	0.0114
0.6	0.9	0.01	0.0013	0.2314	1.1739	0.5302	0.4155	0.3094	0.0529
		0.05	-0.0072	0.2204	0.4987	0.5592	0.6258	-0.8058	0.0245
		0.1	1.7008e-005	0.2169	-0.1868	0.5019	0.7961	-2.2219	0.0187
	0.95	0.01	0.0017	0.2646	1.0654	0.9672	1.1517	0.1472	0.0263
		0.05	0.0072	0.2418	-0.2475	0.8160	1.1334	-1.3891	0.0160
		0.1	0.0088	0.2333	-0.5722	0.5644	0.7716	-2.7294	0.0205
	0.99	0.01	-0.0034	0.2906	-1.0230	1.7868	2.5456	-1.1198	0.0197
		0.05	-0.0075	0.2667	-0.9508	0.8714	1.1746	-2.2303	0.0205
		0.1	-0.0251	0.2490	-0.6926	0.5490	0.7283	-3.1009	0.0233
	1.0	0.01	-0.0227	0.3139	-1.5149	1.7072	2.3245	-1.6140	0.0221
		0.05	-0.0135	0.2896	-0.9656	0.8321	1.1070	-2.4150	0.0209
		0.1	-0.0151	0.2607	-0.7462	0.5349	0.7055	-3.3093	0.0248
0.9	0.9	0.01	0.0261	0.5764	1.2825	0.5154	0.4636	0.4608	0.0597
		0.05	-0.0052	0.5650	0.9718	0.5445	0.3506	-0.1048	0.0366
		0.1	-0.0074	0.5704	0.5120	0.6011	0.6238	-0.8220	0.0186
	0.95	0.01	0.0525	0.8202	1.7076	0.8840	0.8526	0.7046	0.0387
		0.05	1.1558e-004	0.8056	0.6301	1.1023	1.8165	-0.1550	0.0191
		0.1	-0.0441	0.7955	-0.5691	1.0675	2.1758	-1.1919	0.0161
	0.99	0.01	-0.0567	1.2363	-0.6472	4.2435	8.0735	-0.0746	0.0234
		0.05	0.1570	1.1961	-2.9946	2.4004	4.5828	-1.4499	0.0283
		0.1	-0.0067	1.1321	-2.2116	1.4199	2.6652	-2.1200	0.0300
	1.0	0.01	0.1155	1.3792	-5.4923	5.3801	9.3951	-1.2861	0.0295
		0.05	-0.0133	1.4101	-3.2335	2.1617	4.2725	-1.9256	0.0318
		0.1	0.1195	1.3534	-2.3440	1.3238	2.4574	-2.4773	0.0353

*Standard Deviation across 1000 simulations

$\theta = \dfrac{c}{1+c}$

γ = Size of the constant gain in learning

Table: 3 Simulation Results (No. of simulations = 1000)

Sample size 400									
θ	ρ	γ	Avg. $\hat{\alpha}$	Avg. SE($\hat{\alpha}$)	Avg. $\hat{\beta}$	Avg. SE($\hat{\beta}$)	S.D.* of $\hat{\beta}$	Avg. t-stat $\beta=1$	Avg. R^2
0.1	0.9	0.01	8.1798e-005	0.0554	0.9090	0.2317	0.1003	-0.4113	0.0393
		0.05	-7.2450e-005	0.0557	0.4178	0.2046	0.1764	-2.8089	0.0142
		0.1	2.2804e-004	0.0556	0.1110	0.1544	0.1713	-5.7320	0.0050
	0.95	0.01	-9.8485e-004	0.0560	0.7537	0.3481	0.2108	-0.6964	0.0149
		0.05	-2.9238e-004	0.0557	0.0714	0.2527	0.2762	-3.6315	0.0036
		0.1	-3.7457e-004	0.0554	-0.0512	0.1667	0.1822	-6.2948	0.0030
	0.99	0.01	2.2216e-005	0.0572	-0.1816	0.5771	0.6383	-1.9833	0.0028
		0.05	-6.1093e-004	0.0556	-0.1271	0.2663	0.2655	-4.2234	0.0026
		0.1	3.0445e-006	0.0550	-0.0879	0.1661	0.1688	-6.5558	0.0030
	1.0	0.01	0.0027	0.0586	-0.3627	0.5416	0.6109	-2.5059	0.0031
		0.05	-0.0025	0.0560	-0.1227	0.2595	0.2657	-4.3248	0.0027
		0.1	0.0014	0.0551	-0.0740	0.1646	0.1693	-6.5316	0.0029
0.6	0.9	0.01	0.0011	0.1092	0.9958	0.2367	0.1312	-0.0532	0.0442
		0.05	2.2954e-004	0.1085	0.5852	0.2536	0.1683	-1.6174	0.0171
		0.1	-0.0011	0.1072	-6.6718e-005	0.2296	0.3106	-4.2953	0.0043
	0.95	0.01	-3.4222e-004	0.1178	0.9747	0.3768	0.2352	-0.1220	0.0194
		0.05	2.7140e-004	0.1163	0.0199	0.3700	0.4814	-2.5545	0.0043
		0.1	-9.9598e-004	0.1133	-0.4199	0.2609	0.3646	-5.4134	0.0092
	0.99	0.01	2.6424e-004	0.1335	-0.4516	0.8820	1.3262	-1.4910	0.0046
		0.05	-0.0046	0.1230	-0.5369	0.3996	0.5193	-3.8591	0.0060
		0.1	0.0020	0.1179	-0.4367	0.2421	0.3044	-5.9689	0.0095
	1.0	0.01	0.0077	0.1508	-0.8840	0.8391	1.1693	-2.2924	0.0045
		0.05	-0.0044	0.1280	-0.3946	0.3516	0.4579	-4.0130	0.0047
		0.1	0.0026	0.1199	-0.3560	0.2271	0.3098	-6.0069	0.0079
0.9	0.9	0.01	0.0112	0.2669	1.0226	0.2290	0.1549	0.0655	0.0483
		0.05	0.0056	0.2670	0.8312	0.2472	0.1327	-0.7591	0.0293
		0.1	7.6845e-004	0.2769	0.4557	0.2696	0.2158	-2.0864	0.0106
	0.95	0.01	-0.0117	0.3560	1.1089	0.3544	0.2646	0.2406	0.0248
		0.05	-5.2484e-004	0.3571	0.6707	0.4268	0.3570	-0.8236	0.0103
		0.1	-0.0019	0.3756	-0.1811	0.4074	0.7092	-2.8368	0.0042
	0.99	0.01	0.0047	0.5455	-0.0605	1.5703	3.8620	-0.2323	0.0068
		0.05	0.0251	0.5347	-0.9505	0.7978	1.5241	-2.3768	0.0061
		0.1	0.0181	0.5039	-0.7775	0.4434	0.7504	-4.0774	0.0086
	1.0	0.01	0.0604	0.7708	-2.7884	1.9829	4.2431	-1.8977	0.0083
		0.05	-0.0353	0.6291	-1.0649	0.6920	1.3490	-3.1080	0.0067
		0.1	0.0263	0.5452	-0.6876	0.3939	0.6656	-4.4241	0.0084

*Standard Deviation across 1000 simulations

$$\theta = \frac{c}{1+c}$$

γ = Size of the constant gain in learning

Results are presented in Table (1) through Table (3). The most general observation is that all but four of the average $\hat{\beta}$ estimates are numerically less than unity and many of the t statistics for the null hypothesis $H_0 : \beta = 1$ are significantly negative. Furthermore, many of the average $\hat{\beta}$ values are negative. All the average $\hat{\alpha}$ values are insignificantly different from 0. Very low value of R^2 in each of the cases is observed. These patterns seem informally consistent with patterns based on actual data.

The major pattern that is apparent from the simulations is that the average $\hat{\beta}$ value tends to decline with an increase in the size of both the gain in the learning rule γ and the autoregressive parameter ρ, given the values of other parameters. However, there is one peculiarity. When $\rho = 1$ or is very close to 1 (e.g. 0.99), the average $\hat{\beta}$ estimates increase with increasing γ. This non-monotonicity seems puzzling. However, for $\rho = 1$ the average $\hat{\beta}$ is always negative. The implication is that a large value of ρ is crucial in generating a negative regression coefficient, whether or not the fundamentals follow a strict random walk.

The simulations also show that $\hat{\beta}$ values increase with increasing θ. Regarding sample size, the model produces a higher number of negative $\hat{\beta}$ values for smaller samples (50 or 100) compared to larger samples (400). This suggests that, as the sample size increases, the bias declines and $\hat{\beta}$ approaches its RE value. Interestingly, most of the studies on the forward premium puzzle used small samples. Furthermore, it appears that the rational expectations values tend to be approached as γ tends to zero.

Another interesting observation is the similarity between model-generated data on the forward premium, depreciation and forecast errors to those obtained from quarterly exchange rate data[25] on four exchange rates: US dollar price of the French Franc, German Mark, Japanese Yen and UK Pound-sterling for the period of 1973.Q1 through 1994.Q1 (for US dollar price of the French Franc the data spans the period 1973.Q1 through 1992.Q2). The illustrations of movements in the actual data are presented in Fig. (1) – (4), and those obtained from model-generated data are presented

---

[25] The data sources are described in detail in section 6.
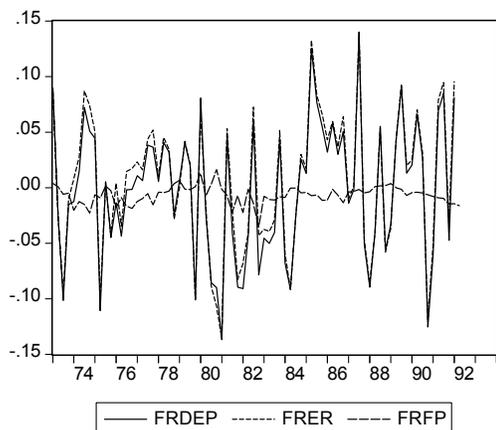
Fig. 1: FRDEP = Depreciation in USD/FRF
FRER = Forecast Error in USD/FRF
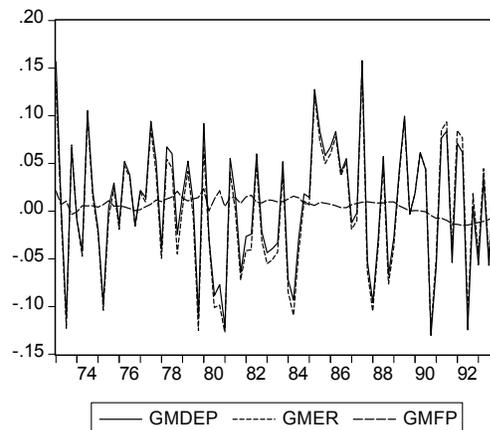FRFP = Forward Premium in USD/FRF

Fig. 2: GMDEP = Depreciation in USD/GMM
GMER = Forecast Error in USD/GMM
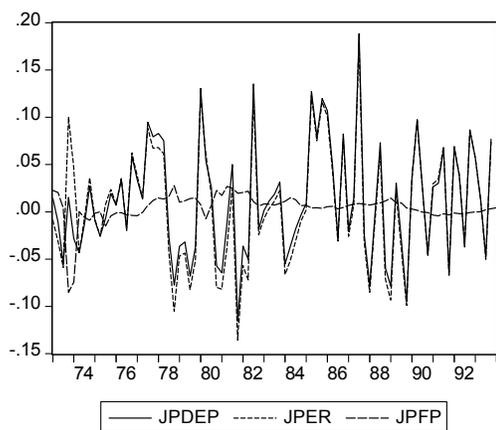GMFP = Forward Premium in USD/GMM

Fig. 3: JPDEP = Depreciation in USD/JPY
JPER = Forecast Error in USD/JPY
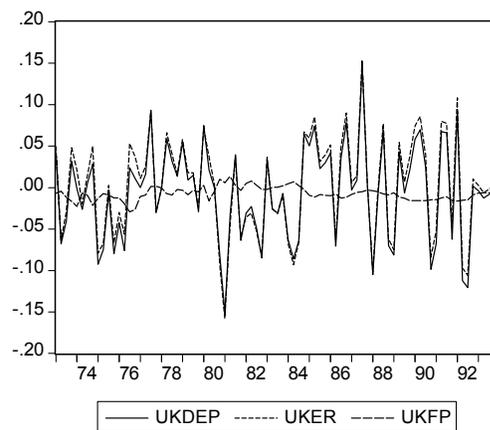JPFP = Forward Premium in USD/JPY

Fig. 4: UKDEP = Depreciation in USD/UKP
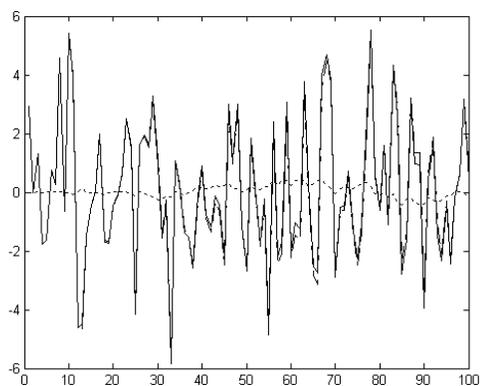UKER = Forecast Error in USD/UKP
UKFP = Forward Premium in USD/UKP

Fig. 5: Model Generated Data: ＿＿ Depreciation, ... Forward Premium, ---- Forecast Error
Sample Size = 100, $\theta = 0.6$, $\rho = 1$, $\gamma = 0.01$ and estimated $\beta = -2.77$
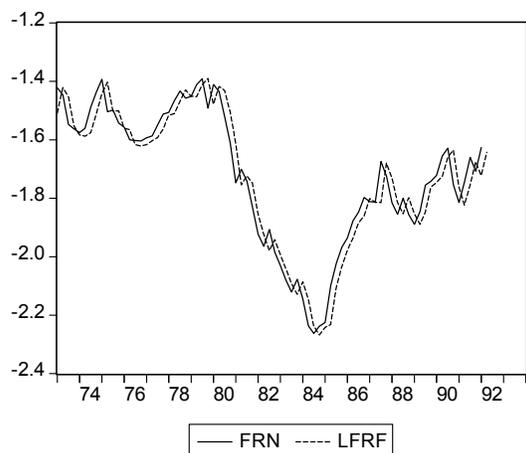
Fig. 6: FRN = Future Spot Rate in USD/FRF
LFRF = Forward Rate in USD/FRF
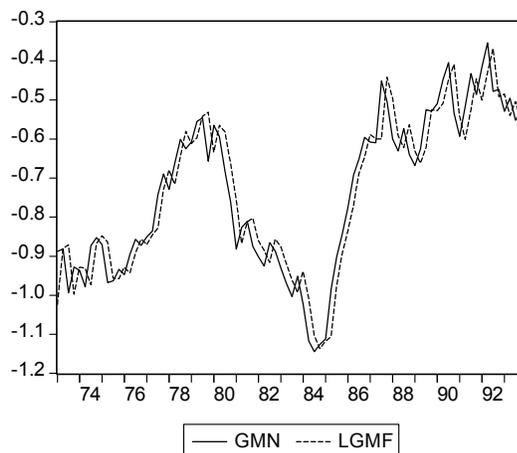
Fig. 7: GMN = Future Spot Rate in USD/GMM
LGMF = Forward in USD/GMM
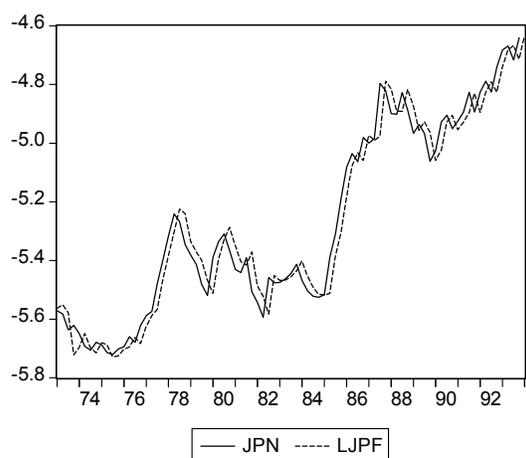
Fig. 8: JPN = Future Spot Rate in USD/JPY
LJPF = = Forward Rate in USD/JPY
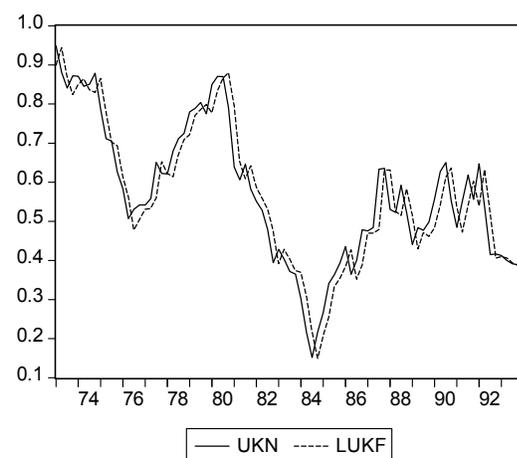
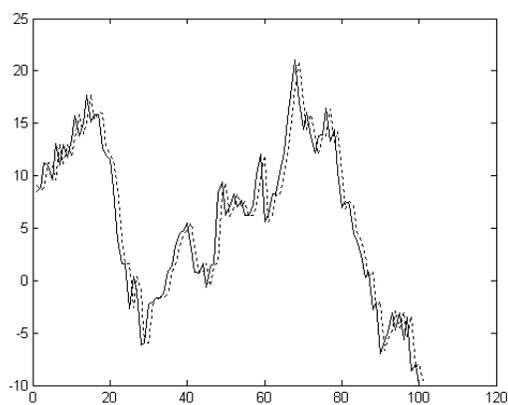Fig. 9: UKN = Future Spot Rate in USD/UKP
LUKF = = Forward Rate in USD/UKP

Fig. 10: Model Generated Data: ____ Future Spot Rate, ---- Forward Rate
Sample Size = 100, $\theta = 0.6$, $\rho = 1$, $\gamma = 0.01$ and estimated $\beta = -0.28$

in Fig. (5). These figures demonstrate, for the forward premium $(F_t - s_t)$, depreciation $(s_{t+1} - s_t)$ and forecast errors $(e_{t+1} = s_{t+1} - F_t = [s_{t+1} - s_t] - [F_t - s_t])$, a strong similarity in the time series pattern in using data generated by the model to analogous patterns using the actual exchange rate data. Fig. (6) – (9) present time series plots in the forward rates and corresponding future spot exchange rates in the observed data on the four exchange rates, and Fig. (10) presents the same two series obtained from the model-generated data. A comparison again reveals a similarity between the actual time series data and model-generated data.

Table 4: Variance and covariance terms for the forecast error, forward rate and forward premium for USD price of French Franc, German Mark, Japanese Yen and UK Pound-sterling (quarterly data) and those generated from the model with $\theta = 0.9$, $\rho = 1$ and $\gamma = 0.1$

Exchange rate (US dollar price of foreign currency)	USD/FRF	USD/GMM	USD/JPY	USD/UKP	Model
Quarterly data					
No. of observations	77	84	84	84	100
$Cov(e_{t+1}, s_{t+1} - s_t)$	0.0688	0.1058	0.0345	0.1069	0.0505
$Cov(e_{t+1}, F_{t+1} - s_t)$	-0.0012	-0.0018	-0.0032	-0.0046	-0.0039
$V(F_{t+1} - s_t)$	0.0012	0.0019	0.0023	0.0019	0.0005
$V(e_{t+1})$	0.0699	0.1076	0.0376	0.1115	0.0509
$V(s_{t+1} - s_t)$	0.0688	0.1059	0.0336	0.1043	0.0502

Note: Variances and covariances are normalized by dividing by variances of the corresponding s_t.
Sources: Quarterly data span the period 1973-quarter 1 to 1992-quarter 2 for French Franc and 1973-quarter 1 to 1994-quarter 1 for German Mark, UK pound-sterling and Japanese Yen. The spot exchange rate and 3 month forward rate data are taken from Harris Bank's Weekly Review. They are drawn from the Fridays occurring nearest to the end of the calendar quarter for quarterly data.

To reinforce the results, when the covariances between forward premium and the forecast errors and the respective variances are compared, the statistics based on the

model-generated data are found to be similar to those based on the actual data[26]. These results are presented in Table (4). Overall, the model seems to do extremely well in explaining the statistics from spot and forward exchange rate data.
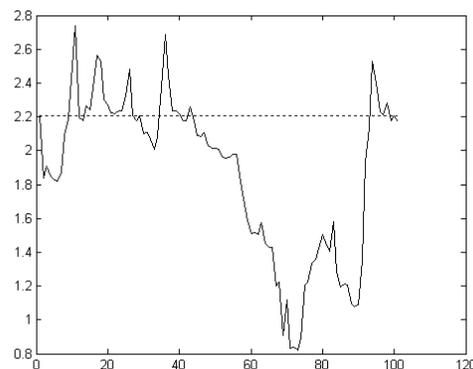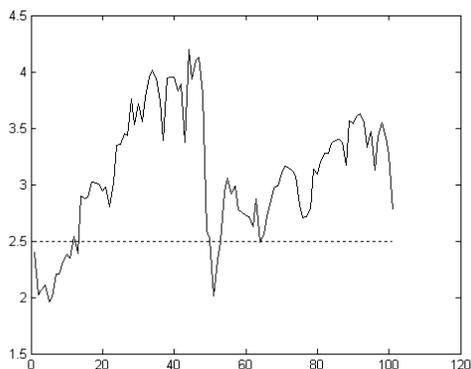

Fig. 11: Time path of parameter estimate for a
under learning
---Time path under rational expectations
Rational Expectations Solution = 2.5
Sample Size = 100, $\theta = 0.6$, $\rho = 1$,
$\gamma = 0.05$ and estimated $\beta = -0.3577$

Fig. 12: Time path of parameter estimate for b
under learning
-- Time path under rational expectations
Rational Expectations Solution = 2.21
Sample Size = 100, $\theta = 0.6$, $\rho = 1$,
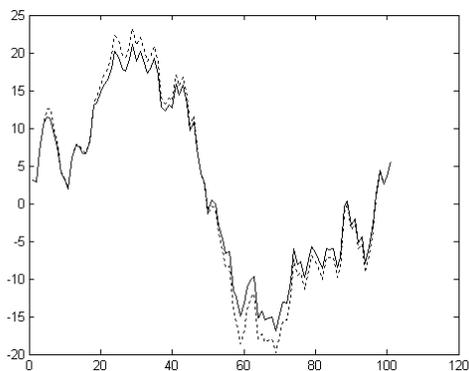$\gamma = 0.05$ and estimated $\beta = -0.3577$

Fig. 13: The time path of log of exchange rate
-----   Time path under constant gain recursive least-square learning
_____   Time path under rational expectations
Sample Size = 100, $\theta = 0.6$, $\rho = 1$, $\gamma = 0.05$ and estimated $\beta = -0.3577$

---

[26] Variances and covariances are normalized by dividing by the variances of the corresponding s_t.

As another illustration, typical movements of the parameters a and b around the rational expectations values are shown in Fig. (11) and Fig. (12), respectively. It shows the fluctuation of the parameters around RE values. In addition, a typical time path of the exchange rate generated from the learning model and that under rational expectations are shown in Fig. (13). The sample size for the simulation was 100 and the parameter values chosen for θ, ρ and γ were 0.6, 0.95 and 0.05, respectively. The corresponding estimates obtained for β and the t-statistic were –0.3577 and –2.6804, respectively. Thus, although a comparison between the two time paths in Fig. (13) suggests small deviations from rational expectations under the learning rule, the estimated values generated from the process indicate that this small deviation is capable of explaining a negative value of $\hat{\beta}$.

Comparing Learning model and Rational Expectations model

The results shown above replicate many observed features of the foreign exchange market using a model in which agents are attributed adaptive learning behavior. A comparison between these results and that obtained from the same model with agents having rational expectations reveals that the learning behavior is the key assumption necessary to generate the negative regression coefficient. This comparison between the regression coefficients and the graphs of the depreciation and the forward premium obtained from the two models are presented in Figures (14) – (15).

For the rational expectations model, two different assumptions are made about the structure of the fundamentals. In one case it is an I (1) variable, i.e. it follows a random walk. In the other case it is an ARIMA (1,1,0) process with very low AR and MA coefficients, i.e., the first difference term of the fundamentals follows an ARMA (1,0) process with low AR and MA coefficients (changing the specification of the fundamentals did not change the key findings). Only when agents exhibit learning behavior does the $\hat{\beta}$ coefficient become negative. The one difference between the pure random walk and ARIMA (1,1,1) cases is that, with the pure random walk, the forward rate exactly equals the current spot rate and hence forward premium is zero. Consequently, in this case there is no movement in the forward premium data over time as evident from Fig. (14), and $\hat{\beta}$ is undefined.
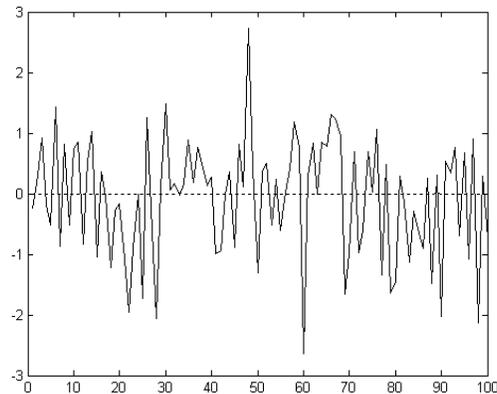
Fig. 14: Model Generated Data: _____ Depreciation, ---- Forward Premium
Sample Size = 100, Rational Expectation with pure random walk in fundamentals.
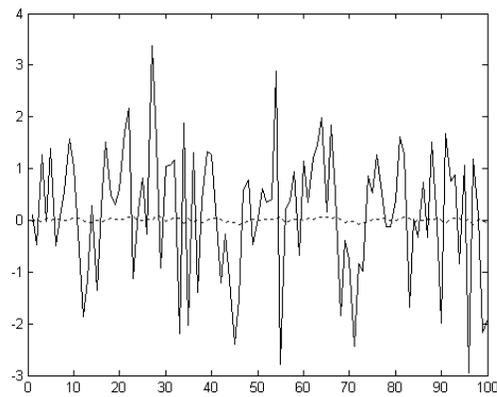$\hat{\beta}$ is undefined.

Fig. 15: Model Generated Data: _____ Depreciation, ---- Forward Premium
Sample Size = 100, Rational Expectation with ARIMA (1,1,1) in fundamentals.
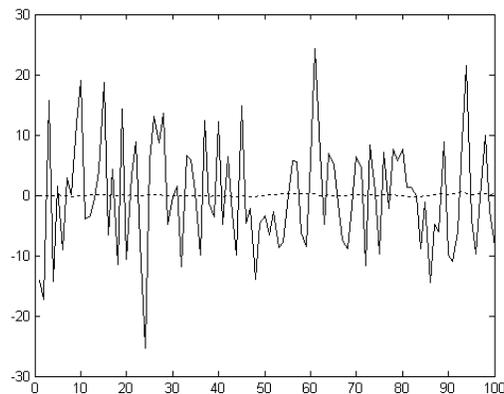$\hat{\beta}$ is 1.1820 (average of 1000 simulations).

Fig. 16: Model Generated Data: _____ Depreciation, ---- Forward Premium
Sample Size = 100, Learning with random walk in fundamentals.
$\gamma = 0.01$ and $\hat{\beta}$ is -5.4904 (average of 1000 simulations).

In contrast, when the fundamentals follow an ARIMA (1,1,0) process there is extra information about the future spot rate compared to current spot rate and there is volatility in the forward premium around zero over time as observed in Fig. (15).

However, this volatility does not generate a negative $\hat{\beta}$ when the agents' expectations are rational. When learning is introduced, however, even for a small gain in learning $\hat{\beta}$ potentially becomes negative. As Fig. (16) shows, $\hat{\beta}$ is in fact negative even when the gain in learning γ is 0.01.

As mentioned before, the autoregressive coefficient ρ in the fundamental process is also important. If $\rho = 1$, the learning model generates negative $\hat{\beta}$ consistently. This result is independent of whether the process is a pure random walk or an ARIMA (1,1,0), as observed in the simulation results. Thus, the general conclusion is that the learning behavior of the agents, when combined with the assumption that the fundamentals follow a process close to a random walk, explains the negative coefficients in the forward premium regressions. Consequently, the nature of the learning mechanism of agents causes the forward premium and the depreciation of the currency to move in opposite directions[27].

Finally, note that the conventional proposition based on the puzzle, that the forward exchange market is not efficient, seems in question. Agents in the forward exchange market are not blessed with rational expectations. However, this does not necessarily contradict the 'efficient market hypothesis' in the sense that new information is immediately absorbed and used efficiently, given the expectations of market participants. They just possess 'bounded rationality'. More discussion on efficiency is presented in the concluding remarks.

6. Estimation Results

The results presented so far are based on certain assumptions about model parameters. As mentioned before, a crucial assumption is that $\rho = 1$ or is near 1. Another assumption was that for quarterly data θ is about 0.9. Although these

---

[27] For a more analytical exposition on how learning dynamics leads to opposite movements in the forward premium and exchange rate changes, see Chakraborty (2004).

assumptions are based on estimates in previous research, direct estimation of these two parameters is warranted.

Data

The data are quarterly series on four exchange rates – US dollar price of the French Franc, German Mark, Japanese Yen and UK pound-sterling. Hence, an increase in the exchange rate means depreciation of the US dollar and vice-versa. Quarterly data are for the period 1973-Q1 to 1992-Q2 for French Franc, and 1973-Q1 to 1994-Q1 for German Mark, Japanese Yen and UK pound-sterling. The spot exchange rate and three month forward rate data are taken from Harris Bank's Weekly Review. They are drawn from the Fridays occurring nearest to the end of the calendar quarter. The future spot rate for a given period is constructed by observing the spot rate on last Friday three months ahead.

However, there are several caveats with this data. There could be significant transaction costs involved in trading currency because of bid-ask spreads and the delivery structure. The rates at which currencies are bought and sold are generally different. Using just one value for each period's spot rate overlooks this difference and hence may potentially introduce bias. Also, entries from Fridays are taken here. When matching the forward rate with the corresponding spot rate in quarterly data, the delivery date for the forward transaction shall be exactly three months from that day. By taking the last Friday of every month this delivery structure could be lost and hence will lead to sampling error. This can introduce bias as well. However, as demonstrated by Bekaert and Hodrick (1993), both of these data problems are unimportant in explaining the forward premium bias, suggesting that this issue can be ignored[28].

Data on fundamentals were collected from International Financial Statistics, published by the International Monetary Fund, for the period of 1971.Q1 through 2000.Q4 (for France and Germany the data are for the period of 1971.Q1 through 1998.Q4). They are quarterly series on the money supply[29] and real GDP for the five countries: US, France, Germany, Japan and the UK. The money supply is adjusted for seasonal variation by averaging it with observations from the last three quarters.

---

[28] The same data were used by Mark and Wu (1998) for the forward premium regression.

[29] For money supply end-of-period data were used as opposed to average during the period.

Logarithmic transformation was made on each series. The fundamentals f_t for each

exchange rate is defined as:

f_t = [log(US Money Supply) – log(Foreign money supply)] – [log(US Real GDP)

– log (Foreign Real GDP)][30]

This measure follows directly from equation (6) and the corresponding large economy

definition of v_t.

Estimation of ρ

Table 5: Augmented Dickey-Fuller test for the fundamentals of four exchange rates.

The null hypothesis in Augmented Dickey-Fuller unit root test is that there exists a unit root in the variables. Dependent variable: Fundamentals (f_t)

Fundamentals for Exchange rates (US dollar price of foreign currency)	USD/FRF	USD/GMM	USD/JPY	USD/UKP
No. of observations (after adjusting endpoints)	109	110	115	114
t-Statistic values				
test-statistic	-2.0168	0.8538	-1.4619	-1.7729
10% Critical value	-2.581	-2.5809	-2.5803	-2.5804
5% Critical value	-2.8882	-2.8879	-2.8867	-2.8870
1% Critical value	-3.4913	-3.4908	-3.4881	-3.4886

Notes: The critical values are the Augmented Dickey-Fuller test critical values, not the conventional t-statistics critical values.
Sources: The fundamentals were constructed based on data on Money Supply and Real GDP for the period of 1971.Q1 through 2000.Q4 for USA, UK and Japan and for the period of 1971.Q1 through 1998.Q4 for Germany and France. The data were collected from International Financial Statistics, published by IMF.

In equation (8) ρ is the autoregressive coefficient of the fundamentals f_t. In the

simulations, ρ was the most important parameter in generating the negative sign of $\hat{\beta}$,

i.e. when $\rho=1$, the magnitudes of most other parameters were minor in generating a

[30] This formula is based on the assumption that the income elasticity of money demand is 1. This assumption appears reasonable since Stock and Watson (1993) demonstrated that this elasticity is not significantly different from 1.

negative $\hat{\beta}$. Thus, an important test of the model is whether ρ is insignificantly different from one.

The fundamental f_t is first tested for existence of unit root. If the test fails to reject the null hypothesis that the fundamental has a unit root, then the conclusion is that $\rho = 1$[31].

The Augmented Dickey-Fuller test for unit roots are presented in Table (5). In none of the four cases can the null hypothesis be rejected even at the 10% level of significance. Thus, fundamentals are found to have unit roots and ρ can be assumed to equal unity.

Table 6: Autoregression results for the fundamentals of four exchange rates.

Dependent variable: Fundamentals (f_t) for four exchange rates				
Fundamentals for Exchange rates (US dollar price of foreign currency)	USD/FRF	USD/GMM	USD/JPY	USD/UKP
No. of observations (after adjusting endpoints)	104	65	109	109
Coefficient on lagged dependent variable (ρ)	0.9739 (0.0152)	0.9320 (0.0412)	0.9726 (0.0306)	0.9834 (0.0148)

Notes: Numbers in parentheses are standard errors.
Sources: The fundamentals were constructed based on data on Money Supply and Real GDP for the period of 1971.Q1 through 2000.Q4 for USA, UK and Japan and for the period of 1971.Q1 through 1998.Q4 for Germany and France. The data were collected from International Financial Statistics, published by IMF.

There is a possibility that the true ρ is close to but less than unity. If so, the ADF test may not be able to reject unit root in f_t because of its low power. In such a situation an autoregression of f_t on its one period lagged value will yield a consistent estimate of ρ.

---

[31] However, there is a possibility that, even though this specification of the fundamentals f_t has a unit root and hence an I (1) variable, it is possible that there exists another linear combination of the money supply and real GDP which is stationary, i.e. that the money supply and real GDP are cointegrated. However, no cointegrating relationship was found between the money supply and real GDP. For more discussion on stationarity and cointegration see Engle and Granger (1987).

The results from auto-regression of fundamentals are summarized in Table (6). Table (5) shows that $H_0 : \rho = 1$ cannot be rejected. Table (6) simply shows that if in fact $\rho < 1$ the estimated value is close to 1. In fact, the presence of a unit root biases the autoregression coefficient downward. This implies that the results in Table (6) are likely to be lower bounds on true estimates of ρ.

In summary, the assumption that the crucial model parameter is close to unity has empirical support.

Estimation of θ

The other parameter to be estimated is θ. Although it is less important than ρ in determining whether $\hat{\beta}$ is negative or not, it does affect the magnitude of $\hat{\beta}$ for a given value of ρ. Also, if ρ is less than unity, the sign of $\hat{\beta}$ may depend on the value of θ. For the simulation, θ was assumed to be 0.9 following Stock and Watson's (1993) estimates of interest semi-elasticity of money demand using quarterly data. Equation (7) gives an alternative way of estimating θ, assuming the validity of the exchange rate model. Thus, regressing the current exchange rate (s_t) on forward rate (F_t) and the fundamentals (f_t) will yield an estimate of θ.

Table (7) presents the results from the estimation. Estimating this simple equation is not straightforward. First there is the stationarity issue. The three variables in this estimation are non-stationary. This raises the question about the consistency of the estimation procedure since it is not wholly clear that the variables are cointegrated[32].

A second complication in estimating θ concerns endogeneity. The forward rate depends on the spot rate. This is addressed by using the lagged value of forward rate as an instrument in place of current value. However, this IV estimation does not necessarily eliminate endogeneity, as there is the possibility that the error reflects a demand shock, which in turn would influence both the fundamentals and the exchange rate. Moreover, the estimation is based on the assumption that equation (7) represents the correct structure.

---

[32] Johansen's cointegration test involving s_t, F_t and f_t together as well as pair wise, produced contradictory results. ADF tests performed on OLS residuals rejected unit roots (except for the case of USD/GMM) indicating in these cases that cointegration is indeed present.

Table 7: Estimation results for θ.

Dependent variable: exchange rate in period t i.e. s_t

Exchange rates (US dollar price of foreign currency)	USD/FRF	USD/GMM	USD/JPY	USD/UKP
No. of observations (after adjusting endpoints)				
OLS	78	68	85	85
IV	74	64	81	81
Coefficient on F_t				
OLS	0.9803	1.0017	0.9698	0.9985
	(0.0054)	(0.0039)	(0.0083)	(0.0043)
IV	0.9781	0.9937	0.9835	1.0012
	(0.0086)	(0.0061)	(0.0098)	(0.0061)
Coefficient on f_t				
OLS	0.0348	0.0304	0.0582	0.0297
	(0.0096)	(0.0102)	(0.0169)	(0.0059)
IV	0.049	0.0263	0.0327	0.0358
	(0.0141)	(0.012)	(0.0179)	(0.0065)
$\bar{R}^2$				
OLS	0.999	0.999	0.998	0.999
IV	0.999	0.999	0.983	0.998

Notes: Numbers in parentheses are standard errors. For the IV estimation the instrument used for F_t is four period lagged values of the same.

Sources: The fundamentals were constructed based on data on Money Supply and Real GDP for the period of 1971.Q1 through 2000.Q4 for USA, UK and Japan and for the period of 1971.Q1 through 1998.Q4 for Germany and France. The data were collected from International Financial Statistics, published by IMF. The spot exchange rate and 3 month forward rate data are taken from Harris Bank's Weekly Review. They are drawn from the Fridays occurring nearest to the end of the calendar quarter.

The estimation results are presented in Table (7). Both OLS and IV estimation generate very high estimates of θ, the coefficient on F_t which is modestly larger than that assumed in the simulations. Moreover, the consistency of the estimates depends

upon strong assumptions. Fortunately, the exact value of θ appears not to be crucial, as the key results appear to hinge primarily upon the value of ρ [33].

7. Conclusion and Comments on Market Efficiency

The objective of the paper was to explore the much renowned 'Forward Premium Puzzle' from a new perspective. Starting from the assumption that bounded rationality is the underlying cause of the anomaly, our model specifies recursive least square learning behavior. Under this learning dynamics, agents form their expectations about future values of the exchange rate, which in turn determines exchange rate movements. Data from the model were generated using computer simulations under various plausible sets of parameter values, and these results were found to be strikingly similar to that observed using actual data. In almost all cases the average $\hat{\beta}$ coefficient in a regression of the change in the spot rate on a forward premium was found to be less than 1, and in the majority of the cases it was negative. Although the results are based on certain assumptions about the model parameters, estimation results for those parameters support the model assumptions. Combining these results it appears that the learning dynamics presented here provides a very satisfactory explanation for the 'forward premium puzzle'.

Often this forward premium regression is considered to be a test for the efficiency of the foreign exchange market. The negative $\hat{\beta}$ is considered to be an indication of systematic errors in expectation and inefficient use of information, and hence a symptom of inefficiency in the market. The model developed here contends that boundedly rational expectations lead to a negative $\hat{\beta}$. This explanation for the paradox contradicts the implication that forward market is inefficient. The 'efficiency market hypothesis' states that a market is efficient if the agents make full use of all available information. This model deviates from rationality in the fact that agents lack complete information about the underlying market parameters. However, they estimate the reduced form parameters over time using a natural econometric procedure, and make

---

[33] Also, the model results seem to be robust for higher values of θ. Preliminary simulations results with $\theta = 0.98$ seem not to alter the model outcomes.

forecasts using their estimated model. The negative $\hat{\beta}$ is a reflection of the learning dynamics, and is not intrinsically an indication of inefficiency in the market. Thus, this finding indicates that the forward premium puzzle is an outcome of 'bounded rationality' in agents' behavior, and is not necessarily inconsistent with the spirit of the 'efficiency market hypothesis'.

There are several natural extensions of the model that could be explored. Two interesting questions are a rationale for 'constant-gain learning' in the empirical data. If there is evidence of periodic structural shifts in the relationship between the fundamentals and the exchange rate, this would justify the rationale for constant gain and the power of the model would be deepened. Also, incorporation of a demand shock in the model dynamics may generate interesting results. Finally, the potential of this model to solve one puzzle indicates the possibility that learning dynamics may have answers for other financial market anomalies.

Appendix

Rational Expectations Solution

Equation (7) is given by

$$s_t = \mu + \theta E^*_t \, s_{t+1} + v_t$$

Equation (10) represents the agents' perceived law of motion (PLM). Taking expectation of equation (10) we get

$$E^*_t \, s_{t+1} = a + bv_t \qquad \text{(10a)}$$

Substituting equation (10a) in equation (7) and rearranging, we get

$$s_t = (\mu + \theta a) + (\theta b + 1)v_t$$
$$\Rightarrow \qquad\qquad\qquad\qquad\qquad \text{(10b)}$$
$$s_t = (\mu + \theta a) + (\theta b + 1)(\rho v_{t-1} + \varepsilon_t)$$

Equation (10b) is the actual law of motion (ALM) as it describes the actual values of the parameters of equation (11). Comparing (10b) with (11) for the Rational Expectations values of a, b and c we get

$$\bar{a} = (\mu + \bar{a}\theta) \text{ or } \bar{a} = (1-\theta)^{-1}\mu$$
$$\bar{b} = (1 + \theta\bar{b})\rho \text{ or } \bar{b} = (1-\theta\rho)^{-1}\rho$$
$$\bar{c} = 1 + \theta\bar{b} \text{ or } \bar{c} = (1-\rho\theta)^{-1}$$

E-stability

E-stability conditions govern the stability conditions under least squares learning behavior around the rational expectations equilibrium. In other words these conditions state when a deviation from rational expectations equilibrium will bring the system back to it.

E-stability condition says that if $\bar{a}$ and $\bar{b}$ are the rational expectations solutions for a and b, respectively, then the solutions are stable if the differential equations

$$\frac{d}{dt}(a) = \tau(a) - a \qquad \text{and}$$

$$\frac{d}{dt}(b) = \tau(b) - b \qquad \text{are converging, i.e. their roots are less than zero.}$$

Where, $\tau(.)$ for a variable corresponds to the respective coefficient in ALM.

So, in the present case we get two equations

$$\frac{da}{dt} = \mu + (\theta - 1)a \qquad (14)$$

$$\frac{db}{dt} = (\theta\rho - 1)b + \rho \qquad (15)$$

Hence, from equations (14) and (15) it is evident that the stability condition for the system is $\theta < 1$ and $\theta\rho < 1$. Since, by assumption $0 < \theta < 1$ and $|\rho| \leq 1$, the solutions are stable in this case.

References:

Bacchetta, P. and Wincoop, E.V. "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?". Working paper (Jan 2003)

Backus, D.K., Foresi, S. and Telmer, C.I. "Affine Term Structure Models and the Forward Premium Anomaly". The Journal of Finance, v51, n1 (Feb 2001): 279-304

Bams, D., Walkowiak, K. and Wolff, C.C.P.I. "More Evidence on the Dollar Risk Premium in the Foreign Exchange Market". Journal of International Money and Finance, v23, n2 (March 2004): 271-282

Bansal, R. "An Exploration of the Forward Premium Puzzle in Currency Markets". The Review of Financial Studies, v10, n2 (Summer 1997): 369-403

Bekaert, G. and Hodrick, R. "On Biases in the Measurement of Foreign Exchange Risk Premiums". Journal of International Money and Finance, v12, n2 (April 2003): 115-138

Branch, W. and Evans, G.W. "A Simple Recursive Forecasting Model". Work in progress (October 2004)

Cavaglia, S.M.F.G., Verschoor, W.F.C. and Wolff, C.C.P. "On the Biasedness of Forward Foreign Exchange rates: Irrationality or Risk Premia?". The Journal of Business, v67, n3 (July 1994): 321-343

Chakraborty, A. "Learning and Forward Premium Puzzle". Doctoral Research Paper, Department of Economics, University of Oregon. (October 2004)

Cho, I.K., Williams, N. and Sargent, T.J. "Escaping Nash Inflation". Review of Economic Studies, v69, (2002) 1-40

Cumby, R.E. and Obstfeld, M. "International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence". Exchange Rate Theory and Practice (1984): 121-51

De Long, J.B., Shleifer, A., Summers, L.H. and Waldmann R.J. "Noise Trader Risk in Financial Markets". The Journal of Political Economy, v98, n4 (Aug 1990): 703-738

Engel, C. "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence". Journal of Empirical Finance 3, (June 1996): 123-192

Engel, C. "A Model of Foreign Exchange Rate Indetermination". National Bureau of Economic Research Working Paper: 5766, (September 1996)

Engle, R.F. and Granger, C.W.J. "Co-integration and Error Correction: Representation, Estimation, and Testing". Econometrica v55, n2 (March 1987): 251-276

Evans, G.W. "A Test for Speculative Bubbles in the Sterling-Dollar Exchange Rate: 1981-84". American Economic Review v76, n4 (September 1986): 621-636

Evans, G.W. and Honkapohja, S. "Learning and Expectations in Macroeconomics", 2001, Princeton University Press

Evans, F. and Lewis, K.K. "Trends in Excess Returns in Currency and Bond Markets". European Economic Review v37, n5 (June 1993): 1005-1019

Fama, E.F. "Forward and Spot Exchange Rates". Journal of Monetary Economics v14, n3 (November 1984): 319-38

Frankel, J.A. and Chinn, M.D. "Exchange Rate Expectations and the Risk Premium: Tests for a Cross Section of 17 Currencies". Review of International Economics, 1(2), (1993): 136-144

Frankel, J.A. and Froot, K.A. "Chartists, Fundamentalists, and Trading in the Foreign Exchange Market". American Economic Review v80, n2 (May 1990): 181-85

Froot, K.A. "Short Rates and Expected Asset Returns". National Bureau of Economic Research Working Paper: 3247, January 1990

Froot, K.A. and Frankel, J.A. "Forward Discount Bias: Is it an Exchange Rate Premium?". The Quarterly Journal of Economics, v104, n1 (Feb 1989): 139-161

Froot, K.A. and Thaler, R.H. "Anomalies: Foreign Exchange". Journal of Economic Perspectives v4, n3 (Summer 1990): 179-92

Goldberg, M.D. and Frydman, R. "Imperfect Knowledge and Behavior in the Foreign Exchange Market". The Economic Journal, v106, n437(Jul 1996): 869-893

Hodrick, R.J. and Srivastava, S. "An Investigation of Risk and Return in Forward Foreign Exchange". Journal of International Money and Finance v3, n1 (April 1984): 5-29

Hodrick, R.J. and Srivastava, S. "The Covariation of Risk Premiums and Expected Future Spot Exchange Rates". Journal of International Money and Finance v5, n0 (Supp. March 1986): S5-21

Kasa, K. "Learning, Large Deviations, and Recurrent Currency Crises". International Economic Review (forthcoming)

Kim, Y.S. "Exchange Rates and Fundamentals Under Adaptive Learning". Doctoral Research Paper, Department of Economics, Ohio State University. (November 2003)

Lewis, K.K. "Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange" American Economic Review v79, n4 (September 1989): 621-36

Lewis, K.K. "Was There a "Peso Problem" in the U.S. Term Structure of Interest Rates: 1979-1982?". International Economic Review v32, n1 (February 1991): 159-73

Lewis, K.K. "Puzzles in International Financial Markets". Handbook of International Economics. Volume 3 (1995): 1913-71

Mark, N.C. "International Macroeconomics and Finance, Theory and Econometric Methods", 2001, Blackwell Publishers

Mark, N.C. "On Time Varying Risk Premia in the Foreign Exchange Market: An Econometric Analysis". Journal of Monetary Economics v16, n1 (July 1985): 3-18

Mark, N.C. and Wu, Y. "Rethinking Deviations from Uncovered Interest parity: The Role of Covariance Risk and Noise". The Economic Journal, v108, (Nov 1998): 1686-1706

Obstfeld, M. "The Effectiveness of Foreign-Exchange Intervention: Recent Experience, 1985-1988". International policy coordination and exchange rate fluctuations (1990): 197-237, A National Bureau of Economic Research Conference Report Chicago and London: University of Chicago Press 1990

Obstfeld, M. and Rogoff, K. "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?". National Bureau of Economic Research Working Paper: 7777, July 2000

Oh, Y.T. and Pippenger, J. "Forward Rates as Predictors of Future Spot Rates: An Econometric Explanation for Sign Reversal". Journal of Economic Development v19, n1 (June 1994): 185-200

Orphanides, A. and Williams, J.C. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy". forthcoming Inflation Targeting, NBER (2003)

Sargent, T.J. "The Conquest of American Inflation". Princeton University Press, Princeton, NJ, (1999)

Stock, J.H. and Watson, M.W. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems". Econometrica v61, n4 (July 1993): 783-820

Taylor, M.P. "The Economics of Exchange Rates". Journal of Economic Literature v33, n1 (March 1995): 13-47