NBER WORKING PAPER SERIES

EXCHANGE RATE DYNAMICS, LEARNING AND MISPERCEPTION

Pierre-Olivier Gourinchas
Aaron Tornell

Working Paper 9391
http://www.nber.org/papers/w9391

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2002

Comments welcome. This paper supersedes our previous paper, "Exchange Rate Dynamics and Learnings", NBER WP 5530. We thank Daron Acemoglu, Olivier Blanchard, Ricardo Caballero, John Campbell, Rudi Dornbusch, Charles Engel, Jeff Frankel, Ben Friedman, Graciela Kaminski, John Leahy, Paul O'Connell, Julio Rotemberg, Nouriel Roubini and seminar participants at UCLA for useful discussions and comments. All errors are our own. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Exchange Rate Dynamics, Learning and Misperception
Pierre-Olivier Gourinchas and Aaron Tornell
NBER Working Paper No. 9391
December 2002
JEL No. E4, F31, G1

ABSTRACT

We propose a new explanation for the forward-premium and the delayed-overshooting puzzles. Both puzzles arise from a systematic under-reaction of short-term interest rate forecasts to current innovations. Accordingly, the forward premium is always a biased predictor of future depreciation; the bias can be so severe as to lead to negative coeffcients in the “Fama" regression; delayed overshooting may or may not occur depending upon the persistence of interest rate innovations and the degree of under-reaction; lastly, for G-7 countries against the U.S., these puzzles can be rationalized for values of the model's parameters that match empirical estimates.

Pierre-Olivier Gourinchas
Princeton University
311 Fisher Hall
Economics Department
Princeton, NJ 08544
and NBER
pog@princeton.edu

Aaron Tornell
Department of Economics
UCLA
405 Hilgard Avenue, Bunche Hall #8283
Los Angeles, CA 90095-1477
and NBER
atornell@ucla.edu

Exchange Rate Dynamics, Learning and Misperception

Abstract.

We propose a new explanation for the forward-premium and the delayed-overshooting puzzles. Both puzzles arise from a systematic under-reaction of short-term interest rate forecasts to current innovations. Accordingly, the forward premium is always a biased predictor of future depreciation; the bias can be so severe as to lead to negative coefficients in the "Fama" regression; delayed overshooting may or may not occur depending upon the persistence of interest rate innovations and the degree of under-reaction; lastly, for G-7 countries against the U.S., these puzzles can be rationalized for values of the model's parameters that match empirical estimates.

J.E.L. Classification: E4, F31, G1.

This paper proposes a new explanation for the forward premium and delayed overshooting puzzles by demonstrating empirically how both puzzles can arise from a systematic under-reaction of short term interest rate forecasts to innovations in current interest rates.

Over the past twenty years, a large body of empirical literature has documented the existence of large biases in forward premia for predicting future changes in exchange rates.[1] This Forward Premium Puzzle (FPP) implies typically that differentials between the domestic and foreign nominal interest rates bear little predictive power for the future rate of change in spot rates. If anything, forward rates and expected depreciation tend to move in opposite directions: a positive interest rate differential is more often than not associated with a subsequent *appreciation* of the exchange rate, not the *depreciation* that theory predicts. This empirical regularity implies significant predictable excess returns in currency markets.

A lesser known puzzle, the delayed overshooting puzzle, has been uncovered more recently by Eichenbaum and Evans (1995). These authors find that unanticipated contractionary shocks to U.S. monetary policy are followed by (a) a *persistent* increase in U.S. interest rates, and (b) a *gradual appreciation* of the dollar, followed by a *gradual depreciation* several months later. This 'delayed overshooting' pattern is consistent with predictable excess returns and the forward premium puzzle:

for a while, U.S. interest rates are higher than their foreign counterparts, with an associated forward discount, *yet* the dollar appreciates, yielding positive excess returns. This dynamic pattern is also in contradiction with Dornbusch (1976)'s *overshooting* result, whereby the exchange rate should experience an immediate appreciation, and then depreciate gradually towards its new long run equilibrium value.

The existence of predictable excess returns, implied by these two puzzles, reflects time-varying currency risk premia and/or expectational errors. The evidence we describe below indicates that the former channel is not sufficient to explain the highly volatile predictable excess returns implied by the puzzles. In fact Frankel and Froot (1989) find that expectational errors in foreign exchange markets play a key role.

In this paper we investigate whether *interest rate forecast errors* observed in the data can rationalize simultaneously these two *exchange rate puzzles*. Our starting point is a setup where agents constantly learn about the duration of interest rate shocks: transitory versus persistent.[2] In and by itself, the fact that agents constantly learn about the duration of interest rate innovations is not enough to rationalize the two exchange rate puzzles. Since rational agents cannot be systematically fooled, there cannot exist predictable excess returns, which are necessary for the delayed overshooting and the forward premium puzzles.

However, using a unique survey data set on interest rate expectations, published by the *Financial Times Currency Forecaster*, with monthly observations from 1986 to 1995 for all G-7 countries, we find substantial evidence of under-reaction to interest rate innovations. More specifically, we find (a) no evidence of transitory shocks in the forward premium; (b) yet implicitly market participants expect incorrectly a substantial share of interest rate innovations to be purely transitory. This contrast is striking: the relative variance of transitory shocks implicitly assumed by market participants is often significantly larger than one, indicating that most innovations are perceived as transitory. These results mirror those of Campbell and Shiller (1991).

We will not take a stand on the origins of the forecasts' short-run under-reaction. What we do is develop a *unified* model that allows us to link interest rate forecast to the two puzzles alluded to above. The only deviation from full rationality in the model is the misperception about second moments implicit in interest rate forecasts. Given these forecasts, the exchange rate is determined by the standard no arbitrage condition. We demonstrate that for typical values of the misperception, the equilibrium exchange rate in the model exhibits both delayed overshooting and the forward premium puzzle in its most extreme form –i.e., a negative Fama coefficient.

To gain some intuition for this result, consider the following experiment. Suppose that domestic interest rates increase vis-a-vis constant foreign short rates, then return gradually to their equilibrium value. If agents know the exact nature of the shock, the exchange rate immediately appreciates relative to its long run value up to the point where the expected future depreciation compensates for the interest rate differential. It then progressively reverts to equilibrium as the interest rate differential declines. This is the interest rate effect: there is overshooting and uncovered interest parity holds. Suppose now that agents misperceive the shock as transitory. On impact, agents believe the domestic interest rate will revert to its equilibrium value fairly rapidly. The exchange rate may initially appreciate moderately. In the next period, the interest rate is in fact higher than agents expected, leading to an upward revision in beliefs regarding the persistence of the shocks. This revision leads to an appreciation. This is the learning effect. However, the domestic interest rate is also reverting to its equilibrium value, leading to a depreciation. If the learning effect is strong enough so as to dominate the interest rate effect, there is a gradual appreciation of the currency. Eventually, there is not much more to learn and the interest rate effect dominates. Thus, the exchange rate reverts to its equilibrium value. Along this path, there are positive excess returns in the domestic currency, and the forward premium is negatively correlated with expected appreciation.

The previous intuition describing the conditional response to an interest rate innovation carries over to an unconditional statement about the forward premium. While we show that the forward premium bias always arises as soon as there is misperception, the most extreme form of this puzzle (a negative Fama coefficient), as well as delayed overshooting, depend upon the parameters of the model.

Intuitively then, hump-shaped exchange rate dynamics result from the interaction of the *misperception about the relative importance of interest rate shocks* (transitory vs. persistent) and the *gradual mean reversion of interest rates*. Note that whether there exists delayed overshooting and a negative Fama coefficient in our model depends on misperceptions about second moments of the interest rate process. Misperceptions about first moments play no role. Note also that in our model predictable excess returns vanish at long horizons, and the Fama coefficient converges to one. This is in line with the empirical evidence.

To sum up, we propose a model that links under-reaction of interest rate forecasts to the forward discount and the delayed overshooting puzzles. We are successful along a number of dimensions: (i) using a unique survey of interest-rate forecasts for the G-7 countries, we demonstrate empirically a systematic under-reaction of forecasts to short term changes in interest rates;[3] (ii) according to our model, the forward premium is almost always a biased predictor of future changes in the spot rate; (iii) moreover, the model can accommodate the most extreme forms of the bias (negative Fama coefficients) unlike most of the previous literature; (iv) depending upon the parameters of the model, we may or may not obtain a delayed overshooting response of nominal exchange rates to monetary shocks. This is empirically satisfying since delayed overshooting seems much less prevalent and robust than the forward premium puzzle.

The next subsection describes the related literature. Section 1 presents a simple version of the model. Section 2 documents the empirical evidence on the expectational and term structure components of the forward discount, reproducing results from Frankel and Froot (1989). We also present evidence on the systematic under-reaction of short rate forecasts. Section 3 presents the model. Section 4 concludes. Most proofs are included in the appendix.

0.1. Related Literature. Delayed overshooting and the forward premium puzzle are both statements about predictable excess returns. Yet they differ in subtle ways. The former is a statement about the joint *conditional* response of nominal interest rates and exchange rates to a common unanticipated monetary innovation. The later is an *unconditional* statement. Empirically, the forward premium puzzle seems much more prevalent, albeit not always in its most extreme form.[4]

In an accounting sense, there are two possible explanations for the forward premium puzzle: time-varying risk premium and/or expectational errors. To see this, start from the standard log-linearized arbitrage condition:

$$\mathcal{E}_t^m e_{t+1} - e_t = r_t - r_t^* - \zeta_t \tag{1}$$

where e_t is the log of the domestic price of the foreign currency, r_t and r_t^* are respectively the domestic and foreign one-period nominal interest rate, and ζ_t is a domestic currency risk premium. Here, $\mathcal{E}_t^m e_{t+1}$ represents the *market* expectation of next period exchange rate, which may differ from statistical or rational expectations, denoted by $\mathcal{E}_t e_{t+1}$. According to equation (1), the return on the short domestic bond, r_t, is equal to the return on a foreign bond of the same maturity, r_t^*, adjusted for the market's expectations of depreciation $\mathcal{E}_t^m e_{t+1} - e_t$, as well as a risk premium component ζ_t.

Of course, as it stands this arbitrage relationship has no empirical power since both market expectations and the risk premium are unobservable. Empirical tests make two additional assumptions: (a) expectations are rational in the sense that $\mathcal{E}_t^m e_{t+1} = \mathcal{E}_t e_{t+1}$; (b) the risk premium ζ_t is constant or uncorrelated with the forward premium $r_t - r_t^*$. Tests of (1) under these assumptions fare quite badly (Fama (1984)): regressions of the form $e_{t+1} - e_t = \alpha + \beta\,(r_t - r_t^*) + u_{t+1}$ (the 'Fama regression') typically find a β significantly smaller than 1, and often negative. Table 1 reports the results from such a regression.[5] The results are typical of the literature: at short horizons, the Fama coefficient β is often significantly negative. At longer horizons, β is often not statistically different from 1. This 'forward premium puzzle' or 'Fama puzzle' implies time-varying predictable excess returns defined as:

$$\xi_t = (r_t - r_t^*) - (\mathcal{E}_t e_{t+1} - e_t) = (\mathcal{E}_t^m e_{t+1} - \mathcal{E}_t e_{t+1}) + \zeta_t \tag{2}$$

According to equation (2), predictable excess returns result from expectational errors and/or a currency risk premium. The time-varying risk premium school of thought argues that expectations are rational, markets efficient, and fluctuations in the forward discount reflect changes in underlying risks. As Fama (1984) points out, the typical estimated bias in the forward premium implies that the currency risk premium ζ_t must be more volatile than predictable excess returns ξ_t. In equilibrium models, the currency risk premium fluctuates with relative asset supplies, conditional variances, and risk aversion. It is difficult to reconcile the low volatility of the above-mentioned variables with the high volatility of predictable excess returns, unless one invokes unrealistically high risk aversion coefficients.[6]

A more recent line of research, using affine models, characterizes directly restrictions on the time-series properties of pricing kernels and the underlying risk factors consistent with the forward premium anomaly (see Backus, Foresi and Telmer (1998) and Sa'a-Requejo (1994)). While of great interest to the practitioner, it is difficult to establish equilibrium foundations for the implied pricing kernels.

Maybe the best empirical evidence is provided by Frankel and Froot (1989). Using survey data on exchange rate forecasts, Frankel and Froot decomposed predictable excess returns into their currency risk premium and expectational error components. Their results indicate clearly that (a) at short horizons, almost none of the forward premium bias can be attributed to currency risk premium fluctuations and (b) changes in the forward premium reflect almost one for one changes in expected appreciation/depreciation. Table 2 reproduces their analysis over the period 1986-1995, with substantially similar results.[7] The direct conclusion is that expectational errors must be responsible for a large part of the forward premium puzzle.

Learning about a one-time change in regime has been analyzed by Lewis (1989a) and (1989b). In that model, following a change in regime, agents gradually update their beliefs about the current state of the world, generating systematic forecast errors during the transition. These learning models can explain some part of the exchange rate mispredictions implied by the forward premium bias, although not the more extreme form where expected depreciation and forward premium move in opposite directions. In general, models of learning about infrequent regime shifts have a difficult time matching the size of the bias and do not account for the fact that predictable excess returns do not appear to die out over time between infrequent regime switches.[8] Further, since regime shifts generate forecast errors that die out over time, models based on learning about a *one-time* change in regime cannot be expected to deliver neither hump-shaped exchange rate dynamics, nor negative coefficients in the Fama regression.

There is an interesting parallel between our results and well established empirical regularities on asset returns. A large volume of empirical work has documented various ways in which asset returns -especially stock returns- are partly predictable based on publicly available information. Of particular interest to us, numerous studies have documented that asset prices *under-react* to news in the short run. Cutler, Poterba and Summers (1991) show that aggregate indices tend to be positively serially correlated at short horizons, while Jegadeesh and Titman (1993) and Fama and French (1992) present similar evidence on cross-sections of individual stocks. Stock returns also under-react to announcements of public information (such as earnings, see Bernard (1992), or Chan, Jegadeesh and Lakonishok (1995)).[9] Thus under-reaction to publicly available news seem to be a prevalent fact.

Paralleling research in International Finance, the consensus in Finance is that the short term under-reaction cannot be explained in terms of risk. Instead, recent models in Behavioral Finance attempt to rationalize asset price puzzles in terms of optimal strategies of boundedly rational agents.[10] Unlike these behavioral models, we will not take a stand on the origins of the under-reaction. We leave it for future research to investigate the conditions under which such behavior might arise in equilibrium. In particular, we do not rule out that measured expectations reflect statistical expectations conditioned upon a subset of the publicly available information, although we have some doubts as to whether this is a valid research strategy.

1. A Simple Case

In this section we present a simplified version of the model of Section 3. Suppose that agents are risk neutral ($\zeta_t = 0$) so that all predictable excess returns arise from expectational errors. Iterating equation (2) forward, we can express predictable excess returns as function of the sequence of forward premium ($x_t = r_t - r_t^*$) expectational errors:

$$\xi_t = -\sum_{j=1}^{\infty} \left[\mathcal{E}_t^m x_{t+j} - \mathcal{E}_t \mathcal{E}_{t+1}^m x_{t+j}\right] + \left(\bar{e}_t - \mathcal{E}_t \bar{e}_{t+1}\right)$$

where $\bar{e}_t = \lim_{T\to\infty} \mathcal{E}_t^m e_{t+T}$ defines the long run equilibrium value of the exchange rate that satisfies $\mathcal{E}_t^m \bar{e}_{t+1} = \bar{e}_t$. This makes clear that in the absence of currency risk premium $(\zeta_t = 0)$, the forward premium puzzle can be rationalized only if market expectations of interest rate differentials x_t *differ* from their statistical expectations, at least at some horizon, or if forecasts of the long run equilibrium exchange rate are incorrect. In this paper we focus on interest rate forecast errors, as demonstrated by the existence of short term under-reaction of forecasts of interest rate differentials between the US and the other G7 countries. We abstract from long run forecast errors by assuming $\mathcal{E}_t \bar{e}_{t+1} = \bar{e}_t$.[11]

To see whether this forecast under-reaction can explain delayed overshooting and the more extreme form of the forward premium puzzle (a negative coefficient in the Fama regression), let us represent our empirical findings in the following way. Suppose that the forward premium follows an AR process with autocorrelation λ:

$$x_t = z_t = \lambda z_{t-1} + \epsilon_t \tag{3}$$

However, assume that agents perceive instead that the forward premium is given by:

$$x_t = z_t + v_t \tag{4}$$

where the shocks ϵ_t and v_t are i.i.d. normal variables with mean zero and respective variances σ_ϵ^2 and σ_v^2. That is, agents erroneously perceive that, in addition to the persistent component x_t, the forward premium also contains a transitory component v_t.

Given their perceptions, agents form their forecasts and choose their portfolio optimally. Taking the forward premium process (3)-(4) as given, agents solve a standard signal extraction problem using Bayes law. As we show in Section 3, the agents' forecast of the forward premium $\mathcal{E}_t^m x_{t+1}$is given by:

$$\mathcal{E}_t^m x_{t+1} = (1-k)\,\lambda\,\mathcal{E}_{t-1}^m x_t + k\lambda\, x_t \tag{5}$$

The weight given to current observations relative to past expectations is $k = \frac{\lambda^2\sigma^2+\sigma_\epsilon^2}{\lambda^2\sigma^2+\sigma_\epsilon^2+\sigma_v^2}$, where σ^2 is the variance of the estimate $\mathcal{E}_t^m x_{t+1}$.[12] Note that when $\sigma_v^2 = 0$, we have that $k = 1$ and market expectations $\mathcal{E}_t^m x_{t+1}$ correspond to statistical expectations. In contrast, when agents misperceive shocks to be more transitory than what they actually are ($\sigma_v^2 > 0$), we have that $k < 1$.

Given agents' expectations (5), the no-arbitrage condition (1) implies that the exchange rate satisfies

$$e_t = \bar{e}_t - \frac{x_t}{1-\lambda} + \frac{1}{1-\lambda}[\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}] \tag{6}$$

where the stochastic trend $\bar{e}_t$ satisfies $\bar{e}_t = \mathcal{E}_t \bar{e}_{t+1}$. It follows from (6) that there are positive predictable excess returns ($\xi_t > 0$) when the expected forward premium falls short of its statistical expectation ($\mathcal{E}_t^m x_{t+1} < \mathcal{E}_t x_{t+1}$), i.e. when short term interest rate forecasts systematically under-react

$$\xi_t = \left(1 + \frac{\lambda k}{1-\lambda}\right)\left[\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}\right]$$

Notice that under rational expectations agents know that there is no transitory component in the forward premium (i.e., $k = 1$). Thus, the exchange rate follows: $e_t = \bar{e}_t - \frac{x_t}{1-\lambda}$ and there are no predictable excess returns.

The response at time $t + j$ to an interest rate innovation $\varepsilon > 0$ at time t is:

(7)
$$\begin{aligned} e_{t+j} &= \bar{e}_t - \frac{\lambda^j \varepsilon}{1-\lambda}\left[1 - \lambda(1-k)^{j+1}\right] \\ \xi_{t+j} &= \frac{\varepsilon}{1-\lambda}\lambda^{j+1}(1-k)^{j+1} \geq 0 \end{aligned}$$

From this, we see that (a) there are positive predictable excess returns when agents misperceive shocks to be more transitory than what they actually are ($k < 1$); (b) delayed overshooting at horizon τ (i.e., $|e_{t+\tau} - \mathcal{E}_t\bar{e}_{t+\tau}| > |e_{t+\tau-1} - \mathcal{E}_t\bar{e}_{t+\tau-1}|$) occurs if and only if:

(8)
$$\tau < \ln\left(\frac{(1-\lambda)/\lambda}{1-\lambda(1-k)}\right) / \ln(1-k);$$

and (c) the probability limit of β in the Fama regression is always smaller than 1 and may even be negative:

$$p\lim \beta = 1 - \frac{(1+\lambda)(1-\lambda(1-k))\lambda(1-k)}{1-(1-k)\lambda^2} \leq 1$$

Notice that there can be positive predictable excess returns on the domestic currency ($\xi_{t+j} \geq 0$), even if there is no delayed overshooting.

Figure 1 shows the path of the exchange rate in response to an unanticipated decrease in the interest rate at $t = 0$. The exchange rate depreciates for about 10 periods before reverting back to its long run value. If one interprets each period as a week or a month, this graph resembles the impulse response functions estimated by Clarida and Gali (1994), Eichenbaum and Evans (1995) and Grilli and Roubini (1994). The duration of each period should depend on the frequency with which one believes that investors receive "new and relevant" information.

We now describe the intuition behind these results. There are two effects:

- **Interest rate effect**. After an initial downward jump, domestic interest rates follow an increasing path. This induces the exchange rate to experience an immediate depreciation followed by a gradual appreciation to ensure that uncovered interest parity holds. This effect is captured by the term in λ^j in (7): the speed at which the initial disturbance fades away.

- **Learning effect.** When the shock takes place at time t agents only observe a reduction in y_t and gradually lower their belief about x_{t+j} using updating equation (5). A downward revision of $\mathcal{E}^m_{t+j}x_{t+j+1}$ generates depreciating pressures on the exchange rate. This effect is captured by the term in $(1-k)^{j+1}$ in (7). When there is no misperception ($k=1$), this term vanishes.

According to (8), a smaller λ (less persistence) increases the second term proportionally more than the first one, making delayed overshooting less likely. This means that an economy converging more rapidly to its long run equilibrium is less likely to exhibit delayed overshooting. The quicker convergence occurs, the more persistent shocks look like transitory ones. Thus, little weight is given to past observations, weakening the learning effect.

Changes in k (the degree of misperception) have more complex effects. For a sufficiently large k, the learning process works efficiently and beliefs converge to the true value of the persistent component of the interest rate very rapidly. As a consequence the subsequent upward revision of beliefs is very small. Therefore, the learning effect is dominated by the interest rate effect and there is no delayed overshooting. In other words, since beliefs have almost converged at time 0, market participants bid the exchange rate up until it is back on the full information rational expectations path. For sufficiently small k, learning occurs very slowly and interest rates convey little information about their persistent component. Thus, the market forecast $\mathcal{E}^m_t x_{t+1}$ increases very little at the time of the shock. Afterwards, although $\mathcal{E}^m_t x_{t+1}$ is updated upwards, the learning effect is too small to dominate the interest rate effect.

We now turn to the issue of the length of time over which the exchange rate moves in the "wrong" direction. Condition (8) defines a 'delayed overshooting' region at horizon τ, D_τ, in terms of the parameters λ and k.[13] Figure 2 reports the lower boundary of this delayed overshooting region as we vary τ from 1 to 10 periods. As we increase the peak date τ, the conditions on λ and k become more stringent: the frontier of D_τ shifts up, as seen in Figure 2. As expected, delayed overshooting obtains for high persistence and low, but not too low, values of k.

Figure 3 reports the contour plot of the coefficient from the Fama regression as a function of λ and k. The Fama coefficient can be negative for small (but not too small) values of k and large values of λ. The contour plot reveals that the region for negative Fama coefficients coincides roughly with the region for delayed overshooting.

Note that the empirical evidence presented in Table 1 shows that the Fama coefficients are increasing with the horizon. In fact, they are close to one for a number of currencies at horizons of 12 months. This indicates that predictable excess returns disappear as the horizon increases. Figure 4 shows that our model is consistent with these facts. It depicts the limit of the Fama coefficient β_n as a function of the regression horizon n : $e_{t+n} - e_t = \alpha + \beta_n \, n \, r_t^n + \epsilon_{t+n}$ where r_t^n is the continuously compounded yield on a $n - period$ zero coupon bond. This figure makes clear that in our model economy the Fama coefficient can be substantially negative at short horizons, but that it eventually converges to one.

To sum up, our analysis has strong cross-sectional implications. Countries should exhibit unconditional delayed overshooting and the forward discount puzzle in its most extreme form (i.e., a negative Fama coefficient), if (a) monetary shocks have high conditional persistence (λ), resulting, for instance, from a low interest elasticity of money demand, and (b) the degree of misperception $(1 - k)$ is high, but not too high. Further, our analysis indicates that there are always positive predictable excess returns at short horizons ($\xi_t > 0$), even if there is no delayed overshooting.

2. An Empirical Exploration of Interest Rate Forecasts

The previous section highlighted the key insight of the model: delayed overshooting and the forward premium puzzle can arise in equilibrium from a systematic under-reaction of interest rate forecasts. In this section, we demonstrate that this under-reaction is present in the data. More specifically, using a state space model that provides the foundation for equation (5), we establish that interest rate forecasts systematically under-react to interest rate innovations.

2.1. Modelling the Interest Rate Differential. To characterize the interest rate process, we adopt the following state-space representation. The interest rate differential between any two countries (x_t) consists of a persistent (z_t) and a transitory (v_t) components, as well as a constant μ:[14]

$$x_t = \mu + z_t + \nu_t \tag{9}$$

In addition, we assume that the persistent component, z_t, follows an $AR(q)$ process:

$$\lambda(L)\, z_t = \epsilon_t \tag{10}$$

with $\lambda(L) = 1 - \sum_{i=1}^{q} \lambda_i L^i$. The transitory and persistent innovations are independent and normally distributed with mean 0 and variance σ_ν^2 and σ^2 respectively. For future reference, we define the noise to signal ratio $\eta = \frac{\sigma_\nu^2}{\sigma^2}$.[15]

One possible justification for our interest rate representation lies in its flexibility: depending on the underlying parameters, this representation can accommodate an integrated process -where some of the roots of λ lie on the unit circle- as well as a white noise.[16]

A more structural interpretation is also possible. Following Dornbusch (1976), we can interpret the transitory shock ν_t as a relative velocity shock, and the persistent shock ϵ_t as a permanent relative money supply shock. In the presence of sticky prices in the short run, a permanent reduction (increase) in the nominal money stock leads to an increase (reduction) in the domestic interest rate. As prices adjust slowly over time, *real* money supply increases and the interest rate declines gradually until it reaches its steady state value. This interpretation is consistent with the empirical findings of Eichenbaum and Evans (1995): an exogenous shock to the US money supply induces a persistent change in the US interest rate in the opposite direction.[17]

Lastly, we want to emphasize that ν_t and ϵ_t can capture the uncertainty surrounding the conduct of monetary policy. Both the monetary policy target and the information set upon which Central Banks act are imperfectly known to the market.[18] Thus, transitory shocks may arise when the Fed acts on inaccurate forecasts or to reflect balance of power adjustments among the Open Market Committee members. Both elements are not observed by market participants who have then to infer the motivation behind recent policy decisions.

2.1.1. Maximum Likelihood Estimation of the Interest Rate Process. The system (9) -(10) can be estimated by Maximum Likelihood of the associated Kalman Filter. The procedure is standard and summarized in appendix A.3.[19] Our data set consists of monthly observations of the 3 months eurorates for Canada, France, Germany, Italy, Japan and the U.K. evaluated against the 3 months eurodollar.[20] The sample period is 1974:1 to 1995:12.[21] The results are presented in Table 3, for various autoregressive orders.

A quick glance through the table indicates that (a) there is a strong persistent component, already largely documented in the literature, and (b) there is no sign of transitory component, as measured by the noise to signal ratio η. Innovations to the persistent component of interest rate differentials disappear extremely slowly. The long run autocorrelation ranges from 0.85 for France against the U.S., to 0.99 for Italy against the U.S. The short run autocorrelation is higher than one for the U.K., Germany and Japan against the U.S., indicating further deviations from equilibrium after the initial shock. The table also reports the results of two Phillips-Perron Z_t tests of a unit root in interest rate differentials. The results indicate that we cannot reject the hypothesis of a unit root at conventional levels of significance for Canada, Germany Italy and Japan. Thus, the interest rate differential for those countries against the U.S. does not appear co-integrated.

For all countries and all specifications, the noise to signal ratio is not statistically significant.[22] In only two cases (Italy-US AR1 and Canada-US AR4) is the constraint on the noise to signal ratio non-binding. It is interesting to note that in both cases the associated long run autocorrelation is also much higher than for other specifications.[23]

2.1.2. Survey Data. To measure market forecasts, we use consensus data from the *Financial Times Currency Forecaster* on Eurorates forecasts at 3, 6 and 12 months. The data is available monthly from August 1986 to October 1995. Contributors include multinational companies as well as forecasting services from major investment banks, i.e., the most active players on the fixed income and foreign exchange markets.[24] The monthly publication collects interest rates and their forecasts and reports a "market average" weighting individual forecasts according to their relative importance. This data set is unique in its coverage and consistency. We have not found any other source of interest rate forecasts prior to 1986 covering all G-7 countries. While survey data on monthly money market rates are available for the US, and have been used in previous studies, we were unable to find similar survey forecasts for foreign countries.[25] As an alternative to market forecasts, one could use forward rates implicit in the term structure. This would be incorrect however, if the term structure contains a significant time varying component. In the working paper version of this paper, we replicate Froot (1989)'s results and demonstrate that this is indeed the case. Thus, we conclude that it is better to use direct consensus forecasts to measure market expectations.

One should still be cautious when using survey data. First, there is probably substantial heterogeneity in forecasts.[26] Nonetheless, "market expectations" constructed from individual heterogenous forecasts may possess better statistical properties than individual forecasts if the idiosyncratic components "wash out" in the aggregation process. This is not guaranteed. A recent theoretical literature has emphasized that there may be systematic biases in individual forecasts: forecasters who care about their reputation, may have incentives to use forecasts in order to manipulate their clients's belief regarding their ability. Such reputational effects are likely to be stronger for professional forecasters than disinterested parties. The direction of the bias, however, is unclear and depends on the information as well as the payoff structure. Scharfstein and Stein (1990) develop a model where managers have an incentive to mimic the behavior of previous managers, while in Zwiebel (1995)'s model average managers have an incentive to herd while "extreme" types - either good or bad- have incentives to scatter.[27] The empirical importance and direction of such reputational biases remains an open question.[28]

We assume that agents use a linear forecasting formula, as summarized by the Kalman filter equations associated with the state-space representation in equations (9)-(10). We will estimate the parameters of the filter implicitly used by market participants, which we call the "market filter", as opposed to the parameters of the true data generating process. For a given market filter we can construct the associated forecast at horizon τ: $x_t^\tau\left(\tilde{\theta}\right)$, where $\tilde{\theta} = \left(\left\{\tilde{\lambda}_i\right\}_{i=1}^{p}, \tilde{\eta}, \tilde{\sigma}^2\right)$ denotes the parameters of the market filter. The forecasts constructed in such a way use only information up to time t. We assume that the observed forecasts are reported with error: $\hat{x}_t^\tau = x_t^\tau\left(\tilde{\theta}\right) + \upsilon_t^\tau$ where the measurement error is i.i.d. and independent of the true forecast, and estimate $\tilde{\theta}$ by Maximum Likelihood. The results are presented in Table 4 separately for each forecast horizon (3, 6 and 12 months, as well as for all forecast horizon pooled.

Our methodology allows us to distinguish between the *conditional persistence* ($\sum_j \tilde{\lambda}_j$ in terms of our representation), and the *relative importance* of transitory and persistent shocks ($\tilde{\eta} = \frac{\sigma_\mu^2}{\sigma^2}$). As can be seen in Tables 3 and 4, conditional persistence is higher for the market filter than for the data generating process. While the long run autocorrelation is close to 0.9 in the data, the market estimates a much higher long run autocorrelation sometimes very close to 1.[29] This extra persistence is compensated by a large estimate for the relative variance $\tilde{\eta}$. The results in Table 4 indicate that for almost all countries and specifications, the perceived noise-to-signal ratio $\tilde{\eta}$ is large -and often significant.[30]

These results stand in sharp contrast with our finding that in the actual data the transitory component is zero. It implies that market forecasts systematically *under-react* to current interest rate changes. We now summarize the results of this section.

Stylized Fact For the interest rate differentials between the US and the other G7 countries, during the period 1986-1995:

1. Persistent shocks are more frequent in sample than perceived by market participants. In other words, the noise-to-signal ratio η (i.e., the ratio of the variance of transitory shocks to the variance of permanents shocks) of the data generating process is significantly smaller than that implied by the linear filter that best replicates interest rate forecasts. This indicates a systematic initial under-reaction to interest rate innovations.
2. Conditional persistence, as measured by the long run auto correlation of the persistent component, is higher for the market filter than for the data generating process. This indicates an over-reaction to interest rate innovations at longer horizons.

3. An Affine Model of Exchange Rate Determination

This section presents an affine model of exchange rate determination under complete markets, similar to Backus et al. (1998), in which there is imperfect information. Our objective is to connect the forward premium forecast under-reaction to the exchange rate anomalies.

The economy consists of two countries, each with its own currency. In each country, nominal assets are traded. The absence of pure arbitrage opportunities implies the existence of a positive pricing kernel for claims in domestic currency, $R_{t,t+1}$, such that:

$$1 = \mathcal{E}_t^m \{R_{t,t+1} R_{t+1}\} \tag{11}$$

where R_{t+1} is the rate of return -in domestic currency- on any traded asset in the economy between t and $t+1$. If the economy admits a representative agent, the pricing kernel is the nominal intertemporal marginal rate of substitution and (11) corresponds to the first-order condition of the agent's program. Equation (11) allows for market expectations $\mathcal{E}_t^m \{.\}$ to differ from their rational counterpart, denoted $\mathcal{E}_t \{.\}$. Similarly, there exists a nominal pricing kernel for claims expressed in foreign currency, $R^*_{t,t+1}$, that satisfies:

$$1 = \mathcal{E}_t^m \left\{R^*_{t,t+1} R^*_{t+1}\right\}$$

Using the pricing kernels, we can derive the one-period continuously compounded nominal risk free rates as:

$$r_t = -\log \mathcal{E}_t^m \{R_{t,t+1}\}; \quad r_t^* = -\log \mathcal{E}_t^m \left\{R^*_{t,t+1}\right\} \tag{12}$$

To derive the rate of currency depreciation from the model, denote by E_t the domestic price of foreign currency (so that an increase in E_t corresponds to a depreciation of the domestic currency). Under complete markets, and with identical preferences across countries, the pricing kernel is unique (see Duffie (1996)). Since we can construct a pricing kernel for nominal claims in the domestic currency as $R^*_{t,t+1}\frac{E_t}{E_{t+1}}$, unicity implies that $R_{t,t+1}\frac{E_{t+1}}{E_t} = R^*_{t,t+1}$ in all states of nature.[31] Expected depreciation is then:

$$\mathcal{E}_t^m e_{t+1} - e_t = \mathcal{E}_t^m \log R^*_{t,t+1} - \mathcal{E}_t^m \log R_{t,t+1} \tag{13}$$

Substituting for the domestic and foreign nominal interest rates, we obtain:

$$\begin{aligned} \mathcal{E}_t^m e_{t+1} - e_t &= r_t - r_t^* + \zeta_t \\ \text{where } \zeta_t &= \left(\mathcal{E}_t^m \log R^*_{t,t+1} - \log \mathcal{E}_t^m \left\{R^*_{t,t+1}\right\}\right) - \left(\mathcal{E}_t^m \log R_{t,t+1} - \log \mathcal{E}_t^m \{R_{t,t+1}\}\right) \end{aligned}$$

The time-varying risk premium ζ_t reflects the difference between conditional means of the pricing kernels. We now follow the literature on affine-yield models and assume that the pricing kernels follow log-normal processes:

$$\ln R_{t,t+1} = -\ln R - \frac{\bar{\varphi}^2 \sigma^2}{2} - \delta\, \bar{z}_t - \frac{\varphi^2 \sigma^2}{2} - z_t - \bar{\varphi}\, \bar{\epsilon}_{t+1} - \varphi \epsilon_{t+1} \tag{14}$$

$$\ln R^*_{t,t+1} = -\ln R - \frac{\bar{\varphi}^{*2} \sigma^2}{2} - \delta^*\, \bar{z}_t - \frac{\varphi^{*2} \sigma^{*2}}{2} - z^*_t - \bar{\varphi}^*\, \bar{\epsilon}_{t+1} - \varphi^* \epsilon^*_{t+1} \tag{15}$$

where the elements of $\boldsymbol{\epsilon}_{t+1} = \left(\bar{\epsilon}_{t+1}, \epsilon_{t+1}, \epsilon^*_{t+1}\right)'$ are independent and normally distributed with mean 0 and variance σ^2. The positive parameters $\delta, \delta^*, \varphi$, φ^*, $\bar{\varphi}$ and $\bar{\varphi}^*$ represent the loadings on the various shocks while R is a positive scalar. $\bar{z}_t$ and $\bar{\epsilon}_{t+1}$ represent, respectively, the predictable and unpredictable components of a shock common to both countries, while (z_t, ϵ_{t+1}) and $(z^*_t, \epsilon^*_{t+1})$ represent the predictable and unpredictable components of country specific shocks. We also assume that the state $\mathbf{z}_t = (\bar{z}_t, z_t, z^*_t)'$ obeys the following AR process:

$$\mathbf{z}_{t+1} = \lambda\, \mathbf{z}_t + \boldsymbol{\epsilon}_{t+1} \tag{16}$$

The specification in (14)-(16) encompasses a broad class of processes while maintaining analytical tractability. Together, $\mathbf{z}_t$ and $\boldsymbol{\epsilon}_{t+1}$ span the predictable and unpredictable components of domestic and foreign bonds and currency price movements. The φ's determine the correlation between each state variable's innovation and the pricing kernel.

Under (14)-(15), the equilibrium interest rate and depreciation rates satisfy:

$$\mathbf{r}_t = \ln R + \mathbf{H}'\mathbf{z}_t \tag{17}$$

$$\begin{aligned} \mathcal{E}^m_t \left\{e_{t+1}\right\} - e_t &= \left(\delta - \delta^*\right) \bar{z}_t + z_t - z^*_t + \frac{\bar{\varphi}^2 - \bar{\varphi}^{*2} + \varphi^2 - \varphi^{*2}}{2}\sigma^2 \\ &= r_t - r^*_t + \zeta \qquad (18) \end{aligned}$$

where $\mathbf{H}' = \begin{pmatrix} \delta & 1 & 0 \\ \delta^* & 0 & 1 \end{pmatrix}$ and $\mathbf{r}'_t = \left(r_t, r^*_t\right)$.

In each country, the domestic interest rate is a linear combination of the world and country specific factors. The expected depreciation rate depends upon the world factor only insofar as it affects domestic and foreign interest rates differently. When $\delta = \delta^*$, the common factor does not affect expected depreciation because it shifts domestic and foreign interest rates by the same amounts. We make this assumption in what follows.[32]

The risk premium ζ is constant and depends upon the variance of the innovations and the loading factors. Since this risk premium is not time-varying, it is irrelevant for our analysis. For simplicity, we set it to 0 by assuming $\bar{\varphi} = \bar{\varphi}^* = \varphi = \varphi^*$, implying that uncovered interest parity holds exactly, under market expectations. By ruling out risk premia altogether, this assumption greatly simplifies our analysis.[33]

The previous assumptions imply that $\mathbf{r}_t$ follows an AR(1) process:

$$\mathbf{r}_{t+1} = (1-\lambda)\ln R + \lambda\,\mathbf{r}_t + \mathbf{H}'\boldsymbol{\epsilon}_{t+1} \tag{19}$$

This model has the simplicity of affine models of the term structure of interest rates. This is not a coincidence: we have assumed a process for the pricing kernel that is identical to Vasicek's original pricing kernel.

3.1. Market Expectations. As discussed in the previous section, the literature on the rational expectation hypothesis of the term structure of interest rates documents systematic deviations between forward rates and expected future short rates. We take this under-reaction as the starting point of our analysis of market expectations and use a convenient state-space representation to characterize the perceived process for the interest rate process, that is consistent with our empirical specification.

Assume that when making forecasts agent do not use the correct pricing kernels in (14) and (15). Instead, they perceive that the pricing kernels satisfy:

$$\ln \tilde{R}_{t,t+1} = \ln R_{t,t+1} - v_t \tag{20}$$

$$\ln \tilde{R}^*_{t,t+1} = \ln R^*_{t,t+1} - v^*_t \tag{21}$$

where v_t and v^*_t are independent normally distributed shocks with mean 0 and variance σ^2_v. The *perceived* pricing kernels differ from the correct ones by a purely transitory shock.

It is important to observe that interest and exchange rates are still determined by (12) and (13). This implies that agents perceive the following relationship between short rates and fundamentals:

$$\mathbf{r}_t = \ln R + \mathbf{H}'\mathbf{z}_t + \mathbf{v}_t \tag{22}$$

Economically, (22) indicates that agents believe that interest rates contain a persistent component, $\mathbf{H}'\mathbf{z}_t$, and a transitory one, $\mathbf{v}_t$, while in fact interest rates contain only a persistent one –they follow

(17). In all other respects, expectations are rational. In particular, rational expectations obtain as the special case where $\sigma_v^2 = 0$.

Equation (22) together with (16) form a state space representation for the interest rate process perceived by agents. As we shall see, this representation captures very naturally the stylized fact that interest rate forecasts under-react to innovations in interest rates. The question of interest to us is the extent to which this misperception affects exchange rate determination in equilibrium.

3.2. The Learning Problem. Agents form optimal forecasts of future interest rates given their beliefs summarized in (16)-(22). This is a standard normal-linear filtering problem. The solution is given by the Kalman filter (see the appendix for derivations). According to (22), agents form forecasts of future interest rates according to:

$$\mathcal{E}_t^m \mathbf{r}_{t+1} = \ln R + \mathbf{H}' \mathcal{E}_t^m \mathbf{z}_{t+1}$$

Define $\mathbf{P}_{t+1} = \mathcal{E}_t^m \left\{ \left(\mathbf{z}_{t+1} - \mathcal{E}_t^m \mathbf{z}_{t+1} \right)^2 \right\}$, the conditional variance of the market belief. The following lemma is a direct consequence of the properties of the Kalman filter.

Lemma 1. *Assume that beliefs about* $\mathbf{z}_1$ *are initially distributed as* $\mathcal{N}\left(\mathcal{E}_t^m \mathbf{z}_{1|0}, \mathbf{P}_1 \right)$ *where* $\mathcal{E}_t^m \mathbf{z}_{1|0}$ *and* $\mathbf{P}_1$ *are an appropriate vector and matrix respectively. Then:*

1. *Beliefs evolve according to:*

$$\mathcal{E}_t^m \mathbf{z}_{t+1} = \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P}_t \mathbf{H} \left(\mathbf{H}' \mathbf{P}_t \mathbf{H} + \sigma_v^2 \mathbf{I} \right)^{-1} \left(\mathbf{r}_t - \log R - \mathbf{H}' \mathcal{E}_{t-1}^m \mathbf{z}_t \right)$$

2. *The conditional variance* $\mathbf{P}_t$ *evolves according to:*

$$\mathbf{P}_{t+1} = \lambda^2 \left[\mathbf{P}_t - \mathbf{P}_t \mathbf{H} \left(\mathbf{H}' \mathbf{P}_t \mathbf{H} + \sigma_v^2 \mathbf{I} \right)^{-1} \mathbf{H}' \mathbf{P}_t \right] + \sigma^2 \mathbf{I}$$

   in particular, it does not depend upon the actual realizations of the interest rates $\mathbf{r}_t$*.*

3. *In the limit as* $t \to \infty$*, the conditional variance converges to a steady state value* $\mathbf{P}$*, solution of:*

$$\mathbf{P} = \lambda^2 \left[\mathbf{P} - \mathbf{P} \mathbf{H} \left(\mathbf{H}' \mathbf{P} \mathbf{H} + \sigma_v^2 \mathbf{I} \right)^{-1} \mathbf{H}' \mathbf{P} \right] + \sigma^2 \mathbf{I}$$

   and the beliefs evolve according to:

$$\mathcal{E}_t^m \mathbf{z}_{t+1} = \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P} \mathbf{H} \left(\mathbf{H}' \mathbf{P} \mathbf{H} + \sigma_v^2 \mathbf{I} \right)^{-1} \left(\mathbf{r}_t - \log R - \mathbf{H}' \mathcal{E}_{t-1}^m \mathbf{z}_t \right) \tag{23}$$

In what follows, we assume that the process has been going on for long enough so that $\mathbf{P}_{t+1}$ has converged.

Two special cases are of interest. First, when expectations are rational, $\sigma_v^2 = 0$, (23) collapses to: $\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda \mathbf{r}_t$, as expected. Second, when there is no common shock, $\delta = 0$, interest rates depend only upon their country specific factor and are independent from one another. In this case (23) implies:

(24) $$\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda (1-k) \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda k \, \mathbf{r}_t + (1-\lambda) \ln R$$

where k, the gain of the filter, measures how much weight is given to new observations, relative to past expectations. Equation (24) makes clear that expected future short rates under-react to changes in the short rate when $k < 1$. In steady state, the gain of the filter is given by: $k = \frac{1+\Delta-\eta\left(1-\lambda^2\right)}{1+\Delta+\eta\left(1+\lambda^2\right)} \leq 1$, where $\eta = \frac{\sigma_v^2}{\sigma^2}$ is the perceived noise-to-signal ratio and $\Delta^2 = \left[\eta \left(1-\lambda^2\right)+1\right]^2 + 4\,\eta\,\lambda^2$. It follows that the gain depends only on the perceived relative variances of the noise and signal components (η) and the degree of persistence (λ). The gain is zero and no learning occurs when the noise is infinite while learning is immediate when there is no noise –i.e., expectations are rational. The gain decreases monotonically with the noise-to-signal ratio and increases with persistence. Intuitively, with a higher λ, today's interest rates contain more information about the persistent component of interest rates and the current realization of the interest rate gets more weight. Given λ, there is a one-to-one mapping between the agent's misperception -as measured by η– and the weight given to past beliefs. We can thus indifferently analyze the properties of the system in terms of (λ, η) or in terms of (λ, k).

In the general case, $\delta \neq 0$, foreign interest rate observations convey information about $\bar{z}_t$, the common factor. If we pre-multiply (23) by $\mathbf{H}'$, we obtain a formula similar to (24):

(25) $$\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda (\mathbf{I}-\mathbf{K}) \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda \mathbf{K} \, \mathbf{r}_t + (1-\lambda) \ln R$$

where $\mathbf{K} = \mathbf{H}'\mathbf{P}\mathbf{H}\left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2 \mathbf{I}\right)^{-1}$ is the matrix representation of the gain of the filter. The formula for the gain indicates that, in general, forecasts of future short rates depend on realizations of both the domestic and foreign short rates. However, a generalized version of under-reaction obtains as the diagonal elements of $\mathbf{K}$ are strictly smaller than 1 as long as σ_v^2 differs from 0.

3.3. Equilibrium Exchange Rate. The equilibrium exchange rate is obtained by solving forward the uncovered interest parity condition (18):

$$e_t = \sum_{j=0}^{T-1} \mathcal{E}_t^m \left(r_{t+j}^* - r_{t+j} \right) + \mathcal{E}_t^m e_{t+T}$$

If we define the equilibrium long-run value of the exchange rate, $\bar{e}_t$, as $\lim_{T\to\infty} \mathcal{E}_t^m e_{t+T}$, and substitute the market beliefs by (25), we obtain:

$$e_t = \bar{e}_t - \mathbf{l}'\mathbf{r}_t - \frac{1}{1-\lambda}\mathbf{l}'\mathbf{H}'\mathcal{E}_t^m \mathbf{z}_{t+1} = \bar{e}_t - x_t - \frac{1}{1-\lambda}\mathcal{E}_t^m x_{t+1} \tag{26}$$

where $\mathbf{l}' = (1, -1)$, so that $x_t = \mathbf{l}'\mathbf{r}_t = r_t - r_t^*$ is the forward premium. The exchange rate depends only upon the current forward premium x_t, its future expected value $\mathcal{E}_t^m x_{t+1}$ and on $\bar{e}_t$. We do not focus in this paper on the determinants of $\bar{e}_t$. However, we assume that there is no misperception regarding its expectations: $\mathcal{E}_t^m \bar{e}_{t+1} = \mathcal{E}_t \bar{e}_{t+1} = \bar{e}_t$. Hence, long run misperceptions do not contribute to predictable excess returns. This assumption implies that there are no predictable excess returns over long horizons. Indeed the data presented in Table 1 indicates fewer predictable excess returns at longer horizons.

Under rational expectations the exchange rate follows:

$$e_t^r = \bar{e}_t - \frac{x_t}{1-\lambda} \tag{27}$$

Substracting (27) from (26), we can express the equilibrium exchange rate as the rational expectation exchange rate plus a term that reflects misperception of the interest rate process:

$$e_t = e_t^r + \frac{1}{1-\lambda}\left(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}\right) \tag{28}$$

3.4. Foreign Exchange Market Anomalies. In this subsection we demonstrate that our simple model is rich enough to rationalize both Eichenbaum and Evans' delayed overshooting puzzle as well as Fama's forward premium puzzle, for some configurations of the parameters.

3.4.1. Predictable excess returns. Recall that predictable excess returns on the domestic currency are defined as:

$$\xi_t = \left(\mathcal{E}_t^m e_{t+1} - \mathcal{E}_t e_{t+1}\right) - \zeta_t$$

where ζ_t is the risk premium. Since there is no risk premium in our set-up, predictable excess returns originate exclusively from forecast errors. Using (25) and (26):

$$\xi_t = \mathbf{l}'\left(\mathbf{I} + \frac{\lambda\mathbf{K}}{1-\lambda}\right)\left(\mathcal{E}_t\mathbf{r}_{t+1} - \mathcal{E}_t^m\mathbf{r}_{t+1}\right) \tag{29}$$

Predictable excess returns depend linearly upon the misperception in short term interest rates forecasts. The relationship between interest rate forecasts and predictable excess returns is complex since the matrix $\mathbf{K}$ is not diagonal in general. To gain some intuition we return to our two special cases. First, when expectations are rational, $\xi_t = 0$, as expected. Second, in the absence of common factors $(\delta = 0)$, (29) simplifies to:

$$\xi_t = \left(1 + \frac{\lambda k}{1-\lambda}\right)\left(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}\right) \tag{30}$$

When the expected forward premium is lower than predicted according to the true model $\left(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1} > 0\right)$, there are positive excess returns on the domestic currency. The reason is simple: if future forward premia are under-estimated, the currency is artificially depreciated (see (28)) and will subsequently appreciate. Using (24), we can write predictable excess returns in a recursive form in this special case:

$$\xi_t = \lambda(1-k)\xi_{t-1} + \left(1 + \frac{\lambda k}{1-\lambda}\right)\lambda(1-k)\,\mathbf{l}'\mathbf{H}'\boldsymbol{\epsilon}_t \tag{31}$$

According to (31) persistence in predictable excess returns increases with the degree of misperception, measured by $1-k$.

3.4.2. Fama regression and the forward discount puzzle. Here we show how the under-reaction of forward premium forecasts is linked to the forward discount puzzle. Recall that the Fama coefficient β converges to:

$$p\lim\beta = \frac{cov\left(e_{t+1} - e_t, x_t\right)}{var\left(x_t\right)}$$

We prove in the appendix the following result.

Lemma 2. *The coefficient of the regression of realized depreciation rates on the forward premium converges in plim to*

$$p\lim\beta = 1 - \frac{\lambda\mathbf{l}'\left(\mathbf{I} + \frac{\lambda\mathbf{K}}{1-\lambda}\right)(\mathbf{I}-\mathbf{K})\left(\mathbf{I} - \left(\mathbf{I} - \lambda^2(\mathbf{I}-\mathbf{K})\right)^{-1}\lambda^2\mathbf{K}\right)var\left(\mathbf{r}_t\right)\mathbf{l}}{\mathbf{l}'var\left(\mathbf{r}_t\right)\mathbf{l}} \tag{32}$$

It is immediate to check that $\beta = 1$ when expectations are rational (since $\mathbf{K} = \mathbf{I}$). In the case where there are no common shocks ($\delta = 0$), (32) simplifies to

$$p\lim \beta = 1 - \frac{\lambda\left(1-\lambda\left(1-k\right)\right)\left(1-k\right)\left(1+\lambda\right)}{1-\lambda^{2}\left(1-k\right)} \tag{33}$$

Since $k < 1$ and $\lambda < 1$, the Fama coefficient is always smaller than 1 and can even be negative. In the general case the Fama coefficient (32) is a complex function of the parameters of the model and may be significantly different from 1.

Figure 3 reports plots the contour of β as a function of λ and k. We see from the graph that β falls as the shocks become more persistent (λ increases). The dependence on k is more complex. A low β requires a low k. However, when $k = 0$, which corresponds to an environment where agents believe all shocks are purely transitory, $\beta = 1 - \lambda$ and remains strictly positive. Indeed, the minimum of β is attained for small but strictly positive values of k. Figure 3 also reports the point estimates for λ and k for each country pair obtained indirectly from our empirical estimates of ($\tilde{\eta}$ and $\tilde{\lambda}$) obtained in section 2.[34]

We can see that for all currencies against the dollar, the process for interest rate forecasts is consistent with very low Fama coefficients (below 0.2 and even negative for Japan and Germany). We emphasize again that we have not used any exchange rate data in this calculation.

Figure 4 reports the coefficient from a Fama regression at horizon n (β_n). Two results emerge. First, the Fama coefficient can be very low -indeed even negative, at short horizons where the effect of misperception is strongest. At longer horizons, however, the Fama coefficient converges to one. Second, β_n can vary non-monotonically with the horizon, as the learning and interest rate effects interact in complex ways.

It is easy to understand why β must be smaller than 1. By definition, expected depreciation is the difference between the forward premium and predictable excess returns: $\mathcal{E}_t e_{t+1} - e_t = x_t - \xi_t$. But from (30), we know that ξ_t is a function of x_t and $\mathcal{E}^m_{t-1} x_t$ so that:

$$\mathcal{E}_t e_{t+1} - e_t = x_t - \lambda\left(1 + \frac{\lambda k}{1-\lambda}\right)(1-k)\left(x_t - \mathcal{E}^m_{t-1} x_t\right) \tag{34}$$

where we have used (30) as well as (25). Consider now an increase in the forward premium x_t at time t. Since interest rate forecasts under-react, $\mathcal{E}^m_{t-1} x_t$ does not increase as much as x_t and the contribution of the predictable excess return is negative. Hence the expected depreciation adjusts by less than the forward premium, implying a coefficient β smaller than 1.

An example might help here. Suppose there is an increase in domestic interest rates, holding foreign interest rates constant. Since agents underestimate future domestic interest rates, the currency will initially appreciate less than under rational expectations (see equation (26)). Thus, the rationally expected depreciation is *smaller* than that implied by the forward discount, and there are positive predictable excess returns. This is true whenever $\lambda > 0$ and $k < 1$.[35]

Yet, we also know from Figure 3 that β can be negative, which is a much stronger result. How can this be? To understand what is going on, return to equation (34). When λ increases, the coefficient on the forward premium forecast error, $\lambda\left(1+\frac{\lambda k}{1-\lambda}\right)(1-k)$, can *exceed* the coefficient on the forward premium (1), as predictable excess returns become more volatile. To illustrate what is going on, consider again an increase in domestic interest rates holding foreign interest rates constant . For λ large enough, the initial mispricing of the exchange rate is so large that it requires the exchange rate to *appreciate* further in the future. As agents update their beliefs about the domestic interest rate, they realize the change in interest rates is more persistent than initially anticipated. Any upward revision on the persistent component of interest rates has a large effect on exchange rates since agents expect high domestic interest rates to persist in the future. Arbitrage implies that the domestic currency appreciates. This scenario is more extreme: a *high* domestic interest rate co-exists with an *appreciating* currency. In other words, the forward premium and rationally expected depreciation move in opposite directions. Thus, β is negative. Our previous discussion abstracted from the term $\mathcal{E}_{t-1}^{m} x_t$. But past expected forward premia are correlated with the current forward premium. Indeed, one obtains (33) exactly when the correlation is properly taken into account. This term dampens the movements in predictable excess returns and makes it more difficult for β to turn negative (compare the term in x_t in (34) and the Fama coefficient).

3.4.3. Delayed Overshooting. The delayed overshooting path is characterized by Eichenbaum and Evans (1995) as the impulse response of the exchange rate to an unanticipated monetary shock. In the context of our model, we compute the path that the exchange rate would follow if an innovation $\boldsymbol{\epsilon}$ were to take place at time t, followed by no other shocks. This shock may be an innovation to the common factor $\bar{z}$ or to a combination of the country specific factors z and z^*. For simplicity, assume that we start from steady state with $\mathbf{r}_{t-1} = \ln R$. According to (19), short rates follow:

$$\mathbf{r}_{t+j} = \ln R + \lambda^j \mathbf{H}' \boldsymbol{\epsilon}$$

Under rational expectation, the exchange rate obeys (27) and follows:

$$e_{t+j} = \bar{e}_{t+j} - \frac{\lambda^j}{1-\lambda} (\epsilon - \epsilon^*)$$

The exchange rate overshoots and converges back to its equilibrium value at the same speed as the interest rate rate. When forward premium forecasts under-react, the exchange rate follows instead:

$$e_{t+j} = \bar{e}_{t+j} - \frac{\lambda^j \mathbf{l}'}{1-\lambda} \left[\mathbf{I} - \lambda (\mathbf{I} - \mathbf{K})^{j+1} \right] \mathbf{H}' \boldsymbol{\epsilon} \tag{35}$$

The term $\mathbf{H}'\boldsymbol{\epsilon}$ determines how each innovation influences the domestic and foreign interest rates, respectively. Recall that common monetary shocks $\bar{z}$ do not influence the rate of depreciation expected by agents when $\delta = \delta^*$. One can check that (35) simplifies to:

$$e_{t+j} = \bar{e}_{t+j} - \frac{\lambda^j}{1-\lambda} \left[1 - \lambda \left(\pi_{11}^{j+1} - \pi_{21}^{j+1} \right) \right] (\epsilon - \epsilon^*) \tag{36}$$

where π_{il}^{j+1} is the (i, l) element of $(\mathbf{I} - \mathbf{K})^{j+1}$. Equation (36) makes clear that e_{t+j} is also unaffected by common shocks. In the special case where there are no common shocks, the formula simplifies to (7).

Delayed overshooting occurs at horizon τ when $|e_{t+\tau} - \mathcal{E}_t^m \bar{e}_{t+\tau}| > |e_{t+\tau-1} - \mathcal{E}_t^m \bar{e}_{t+\tau-1}|$. Using (36), and the properties of the learning process, we have the following lemma:

Lemma 3. *A necessary and sufficient condition for delayed overshooting after τ periods is:*

$$(1 - \lambda (\pi_{11} - \pi_{21})) (\pi_{11}^{\tau} - \pi_{21}^{\tau}) > \frac{1-\lambda}{\lambda}$$

where π_{ij} and π_{ij}^{τ} are defined respectively the (i, j) element of the matrices $\mathbf{I} - \mathbf{K}$ and $(\mathbf{I} - \mathbf{K})^{\tau}$.

This lemma establishes that delayed overshooting depends upon the parameters driving the learning and interest rate processes as summarized by $\mathbf{K}$ and λ. The discussion of Section 1 applies: delayed overshooting results from the interaction between an interest rate and a learning effect. We obtain (8) as a special case when $\delta = \delta^* = 0$. In this special case the delayed overshooting region is reported in Figure 2.

Figure 2 also reports the implicit delayed overshooting horizon for each country vis a vis the US.[36] The results indicate that, except for the French Franc, all currencies exhibit delayed overshooting between 3 and 8 months. While the point estimates are smaller than the 24-36 months delayed overshooting typically reported by Eichenbaum and Evans (1995), they nonetheless indicate that interest rate forecasts are consistent with delayed overshooting for some currency pairs.

4. Conclusion

In this paper we establish a *link* between currency market and bond market anomalies. We show that the data strongly supports the existence of under-reaction of bond prices to changes in short rates. We then develop a nominal exchange rate determination model in which there is incomplete information about the nature of interest rate shocks (transitory vs. persistent). We demonstrate that for typical empirical estimates of the misperception, the equilibrium exchange rate in the model exhibits both delayed overshooting and the forward premium puzzle in its most extreme form –i.e., a negative Fama coefficient. We see this attempt to investigate whether anomalies in different asset markets are consistent with each other as a novel contribution that might be of interest in areas other than international finance.

The dynamics of the exchange rate are determined by two effects which act in opposite directions: an interest rate effect and a learning effect. The first effect is standard. As in other models, it derives from the mean reverting nature of interest rates differentials. The learning effect is novel, and derives from the fact that agents constantly learn whether interest rate shocks are transitory or persistent. The predictable excess returns and hump-shaped dynamics, implied by the two currency market anomalies, can result if and only if agents misperceive interest rate shocks to be more transitory than what they actually are.

The key to our results is the existence of a misperception about second moments of the interest rate process, not about first moments. Furthermore, currency market anomalies in our model economy only exist in the short run. At long horizons predictable excess returns vanish, and the

Fama coefficient converges to one. This is in line with the empirical evidence.

This interpretation, which is new to our knowledge, has important implications. First, it provides a clear analytical characterization of the factors influencing exchange rate responses to monetary shocks. Countries with rapidly converging interest rates, due to either fast moving prices or a large interest elasticity of money demand, will experience *smaller* predictable excess returns. Countries with either a very small or a very large variance of transitory shocks will also converge without delayed overshooting. In the former case because learning occurs fast, in the latter case because learning does not have a significant effect on the demand for assets.

This paper also serves the useful purpose of uncovering a deeper rationale for a variety of asset market pathologies. In particular, the misperceptions we identify raise the interesting prospect of an integrated understanding of currency and bond markets.

Of course, this paper also raises some intriguing questions: why do agents fail to revise their erroneous beliefs about second moments? Can these misperceptions be arbitraged away or taken advantage of by savvier investors? Ultimately, of course, we will need to reconcile observed behavior with models of optimal behavior. While we may not be there yet, this paper indicates a promising avenue of research.

References

Backus, David, Silverio Foresi, and Chris Telmer, "Affine Models of Currency Pricing: Accounting for the Forward Premium Anomaly," July 1998. mimeo New York University.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, "A Model of Investor Sentiment," *Journal of Financial Economics*, September 1998, *49* (3), 307–43.

Batten, Dallas, Michael Blackwell, In-Su Kim, Simon Nocera, and Yuzuru Ozeki, "The Conduct of Monetary Policy in the Major Industrial Countries," International Monetary Fund, Washington DC 1990. Occasional Paper 70.

Bernard, Victor, "Stock Price Reactions to Earnings Announcements," in Richard Thaler, ed., *Advances in Behavioral Finance*, Russell Sage Foundation, NY 1992.

Bollerslev, Tim, Ray Chou, and Kenneth Kroner, "ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence," *Journal of Econometrics*, 1992, *52*, 5–59.

Campbell, John and Robert Shiller, "Yield Spreads and Interest Rate Movements: a Bird's Eye View," *Review of Economic Studies*, May 1991, *58* (3), 495–514.

Chan, Louis, Narasimhan Jegadeesh, and Josef Lakonishok, "Evaluating the Performance of Value versus Glamour Stocks: the Impact of selection Bias," *Journal of Financial Economics*, Jul 1995, *38* (3), 269–296.

Clarida, Richard and Jordi Gali, "Sources of Real Exchange Rate Fluctuations: How Important Are Nominal Shocks?," 1994. NBER Working Paper 4658.

Cutler, David, James Poterba, and Lawrence Summers, "Speculative Dynamics," *Review of Economic Studies*, May 1991, *58* (3), 529–546.

DeBondt, Werner and Richard Thaler, "Does the Stock Market Overreact?," *Journal of Finance*, 1985, *40*, 793–808.

Dornbusch, Rudiger, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 1976, *84*, 1161–1176.

Duffie, Darrell, *Dynamic Asset Pricing Theory*, Princeton University Press, 1996.

Ehrbeck, Tilman and Robert Waldmann, "Why Are Professional Forecasters Biased? Agency Versus Behavioral Explanations," *Quarterly Journal of Economics*, Feb 1996, *111* (1), 21–40.

Eichenbaum, Martin and Charles Evans, "Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates," *Quarterly Journal of Economics*, Nov 1995, *110*, 975–1009.

Fama, Eugene, "Forward and Spot Exchange Rates," *Journal of Monetary Economics*, 1984, *14* (4), 319–338.

____ **and Kenneth French**, "The Cross-Section of Expected Stock Returns," *Journal of Finance*, Jun 1992, *47* (2), 427–465.

Frankel, Jeffrey and Andy Rose, "A Survey of Empirical Research on Nominal Exchange Rates," 1994. NBER Working Paper 4865.

____ **and Kenneth Froot**, "Forward Discount Bias: Is it an Exchange Risk Premium?," *Quarterly Journal of Economics*, 1989, *104*, 139–161.

Friedman, Benjamin, "Interest Rate Expectations versus Forward Rates: Evidence from an Expectations Survey," *Journal of Finance*, 1980, *34*, 965–973.

Froot, Kenneth, "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates," *Journal of Finance*, 1989, *44* (2), 283–304.

Grilli, Vittorio and Nouriel Roubini, "Liquidity and Exchange Rates: Puzzling Evidence From the G-7 Countries," 1994. mimeo Yale University.

Hamilton, James, *Time Series Analysis*, Princeton University Press, 1994.

Hodrick, Robert, "The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets," in "Fundamentals of Pure and Applied Economics" Harwood Academic Publishers 1988.

Hong, Harrison and Jeremy Stein, "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets," *Journal of Finance*, December 1999, *54* (6), 2143–84.

Jegadeesh, Narasimhan and Sheridan Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 1993, *48*, 65–91.

Kaminski, Graciela, "Is There a Peso Problem? Evidence form the Dollar/Pound Exchange Rate, 1976-1987," *American Economic Review*, 1993, *83*, 450–472.

Lewis, Karen, "Can Learning Affect Exchange Rate Behavior? The Case of the Dollar in the Early 1980's," *Journal of Monetary Economics*, 1989, *23*, 79–100.

____ , "Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange," *American Economic Review*, 1989, *79*, 621–636.

____ , "Puzzles in International Finance," 1994. NBER Working Paper 4951.

Muth, John, "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 1960, *55*, 299–306.

Prendergast, Canice and Lars Stole, "Impetuous Youngsters and Jaded Oldtimers: Acquiring a Reputation for Learning," 1996. forthcoming @JPE.

Sa'a-Requejo, Jesus, "The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets," May 1994. mimeo INSEAD.

Scharfstein, David and Jeremy Stein, "Herd Behavior and Investment," *American Economic Review*, 1990, *80*, 464–79.

Zarnowitz, Victor and Phillip Braun, "Twenty Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," 1992. NBER Working Paper 3965.

Zwiebel, Jeffrey, "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 1995, *103*, 1–25.

Table 1:

TEST OF FORWARD PREMIUM PUZZLE

Estimates the following: $e_{t+n} - e_t = \alpha + \beta_n x_t^n + \eta_{t+n}$.

$e_{t+n} - e_t$ is the n-period realized depreciation rate vis a vis the US dollar. x_t^n is the n-period continuously compounded euro-rate and $x_t^n = r_t^n - r_t^{n,us}$ is the forward premium. The number of observations varies for each country pair between 100 and 299. Coefficients estimated with OLS. Newey-West (1987) Robust standard errors in parentheses. The last column reports the p-value for the test $\beta_n = 1$. Source: DRI. Exchange Rate from IFS; Interest Rates: Eurorates from FACS.

Country	Sample	β_n	SE	p	Sample	β_n	SE	p
	HORIZON: 3 MONTHS				HORIZON: 6 MONTHS			
All	74:1-99:7	-0.72	0.22	0.00	74:1-99:4	-0.64	0.18	0.00
Canada	79:8-99:7	-0.87	0.40	0.00	79:8-99:4	-0.75	0.42	0.00
France	74:1-98:9	-0.16	0.85	0.17	74:1-98:6	-0.11	0.86	0.19
Germany	74:1-98:9	-0.50	0.73	0.04	74:1-98:6	-0.47	0.70	0.04
Italy	80:10-98:9	1.41	1.27	0.75	80:10-98:6	1.59	1.45	0.59
Japan	78:1-99:7	-2.98	0.82	0.00	78:1-99:4	-3.06	0.72	0.00
U.K.	74:1-99:7	-1.25	0.95	0.02	74:1-99:4	-1.00	0.88	0.02
	HORIZON: 12 MONTHS							
All	74:1-97:10	0.93	0.13	0.58				
Canada	78:1-98:10	0.70	0.26	0.25				
France	74:1-97:12	0.49	0.71	0.47				
Germany	80:10-97:12	1.39	0.43	0.36				
Italy	79:8-97:12	0.64	1.20	0.77				
Japan	86:8-95:10	2.58	0.93	1.71				
U.K.	74:1-98:10	0.82	0.37	0.63				

Table 2:

DECOMPOSITION OF THE FAMA COEFFICIENT

Decomposes β_n according to: $\beta_n = 1 - b_{re} - b_{rp}$.

with $b_{re} = -\frac{cov\left(u^n_{t+n}, x^n_t\right)}{var\left(x^n_t\right)}$, $b_{rp} = \frac{var(\zeta^n_t)+cov(\mathcal{E}^m_t e_{t+n} - e_t, \zeta^n_t)}{var\left(x^n_t\right)}$

Country	b_{re}	b_{rp}	β_n	b_{re}	b_{rp}	β_n
	HORIZON: 3 MONTHS			HORIZON: 6 MONTHS		
All	0.81	0.19	-0.05	0.93	0.08	0.15
Canada	3.03	-0.20	-1.83	2.48	-0.59	-0.89
France	1.51	-1.87	1.35	1.03	-1.55	1.51
Germany	2.30	-1.55	0.25	2.20	-1.58	0.38
Italy	-0.50	-1.70	3.20	-0.72	-1.22	2.94
Japan	4.91	-0.33	-3.58	4.93	-0.70	-3.23
U.K.	2.82	0.13	-1.95	2.54	-0.52	-1.02
	HORIZON: 12 MONTHS					
All	-0.98	0.92	1.06			
Canada	-1.83	1.47	1.36			
France	0.03	-0.15	1.11			
Germany	-0.78	0.98	0.80			
Italy	-2.10	2.41	0.7			
Japan	-1.72	1.26	1.45			
U.K.	-0.99	0.90	1.08			

Table 3:

MAXIMUM LIKELIHOOD ESTIMATION OF THE STATE-SPACE REPRESENTATION OF 3 MONTHS EURORATES

Estimates the following state-space model on monthly data: $x_t = \mu + z_t + \nu_t \quad ; \quad \lambda(L)\, z_t = \epsilon_t$. x_t is the 3 months eurorate differential $r_t - r_t^{us}$. We report estimates of the autoregressive coefficients (λ) and the noise to signal ratio $\eta = \frac{\sigma_\nu^2}{\sigma_\epsilon^2}$. Sample period: 1974:1 to 1995:12. The number of observations varies for each country pair. Heteroskedasticity consistent standard errors (SE) reported below the estimate in parenthesis. Coefficients estimated by Iterated Maximum Likelihood of the Kalman Filter. A value of 0 for η indicates that the associated constraint ($\eta \geq 0$) is binding. and the corresponding AR process, estimated directly, maximizes the likelihood. Coefficients significant at the 5% level are reported in bold. Z_ρ and Z_t: Phillips-Perron statistics of the unit root hypothesis H_0 ($\alpha = 0, \beta = 1$) in the regression: $x_t = \alpha + \beta x_{t-1} + u_t$ with 12 lags in the Bartlett window. The associated p-values are reported in parenthesis. Source: IFS line 60ldd and 60ea.

	Estimated Coefficients								
	λ_1	λ_2	λ_3	λ_4	η	$\sum\lambda$			
Model	SE	SE	SE	SE	SE	SE	Log-Lik		
PANEL A: CANADA-U.S.								OBS	100
I	**0.90**				0	**0.90**	-0.229	Z_ρ	-8.31
	(0.04)					(0.04)			(0.15)
II	**0.93**	-0.03			0	**0.90**	-0.226	Z_t	-2.04
	(0.10)	(0.10)				(0.04)			(0.17)
III	**0.92**	0.06	-0.09		0	**0.89**	-0.199		
	(0.10)	(0.14)	(0.10)			(0.05)			
IV	**1.37**	-0.31	**-0.38**	**0.29**	0.42	**0.97**	-0.181		
	(0.25)	(0.26)	(0.07)	(0.09)	(0.51)	(0.02)			
PANEL B: FRANCE-U.S.								OBS	213
I	**0.85**				0	**0.85**	-1.267	Z_ρ	**-30.79**
	(0.03)					(0.03)			(0.0)
II	**0.86**	-0.01			0	**0.85**	-1.265	Z_t	**-4.07**

continued next page

Table 3: Maximum Likelihood Estimation, 3 Months Eurorates

continued from previous page

	Estimated Coefficients								
	λ_1	λ_2	λ_3	λ_4	η	$\sum \lambda$			
Model	SE	SE	SE	SE	SE	SE	Log-Lik		
	(0.07)	(0.07)				(0.04)			(0.0)
III	**0.87**	-0.06	0.05		0	**0.86**	-1.260		
	(0.07)	(0.09)	(0.07)			(0.04)			
IV	**0.87**	-0.06	0.06	-0.01	0	**0.86**	-1.260		
	(0.07)	(0.09)	(0.09)	(0.47)		(0.04)			
Panel C: Germany-U.S.								Obs	213
I	**0.96**				0	**0.96**	-0.717	Z_ρ	-5.25
	(0.02)					(0.02)			(0.5)
II	**1.23**	**-0.28**			0	**0.95**	-0.686	Z_t	-1.67
	(0.07)	(0.07)				(0.02)			(0.45)
III	**1.31**	**-0.63**	**0.29**		0	**0.97**	-0.653		
	(0.07)	(0.10)	(0.07)			(0.02)			
IV	**1.31**	**-0.63**	**0.28**	0.01	0	**0.97**	-0.653		
	(0.07)	(0.11)	(0.11)	(0.07)		(0.02)			
Panel D: Italy-U.S.								Obs	94
I	**0.99**				0.25	**0.99**	-0.610	Z_ρ	-6.58
	(0.02)				(0.29)	(0.02)			(0.5)
II	**0.70**	**0.25**			0	**0.95**	-0.563	Z_t	-1.83
	(0.10)	(0.10)				(0.03)			(0.30)
III	**0.74**	**0.36**	-0.16		0	**0.94**	-0.559		
	(0.10)	(0.12)	(0.10)			(0.03)			

continued next page

Table 3: Maximum Likelihood Estimation, 3 Months Eurorates

continued from previous page

	Estimated Coefficients								
	λ_1	λ_2	λ_3	λ_4	η	$\sum \lambda$			
Model	SE	SE	SE	SE	SE	SE	Log-Lik		
IV	**0.69**	**0.46**	0.06	**-0.28**	0	**0.93**	-0.528		
	(0.10)	(0.12)	(0.12)	(0.10)		(0.03)			
Panel E: Japan-U.S.								Obs	205
I	**0.94**				0	**0.94**	-0.764	Z_ρ	-11.95
	(0.02)					(0.02)			(0.08)
II	**1.15**	**-0.22**			0	**0.93**	-0.725	Z_t	-2.51
	(0.07)	(0.07)				(0.02)			(0.12)
III	**1.14**	-0.16	-0.05		0	**0.93**	-0.723		
	(0.07)	(0.11)	(0.07)			(0.02)			
IV	**1.14**	-0.16	-0.07	0.02	0	**0.93**	-0.722		
	(0.07)	(0.11)	(0.11)	(0.07)		(0.02)			
Panel F: U.K.-U.S.								Obs	213
I	**0.93**				0	**0.93**	-0.776	Z_ρ	**-14.70**
	(0.02)					(0.02)			(0.05)
II	**1.10**	**-0.18**			0	**0.92**	-0.764	Z_t	-2.80
	(0.07)	(0.07)				(0.02)			(0.07)
III	**1.13**	**-0.42**	**0.22**		0	**0.93**	-0.743		
	(0.07)	(0.10)	(0.07)			(0.02)			
IV	**1.14**	**-0.43**	**0.26**	-0.04	0	**0.93**	-0.741		
	(0.07)	(0.10)	(0.10)	(0.07)		(0.02)			

Table 4:

POOLED MAXIMUM LIKELIHOOD ESTIMATION OF THE MARKET FILTER FOR 3 MONTHS EURORATES

Estimates the following state-space model on monthly data: $\hat{x}_t^\tau = x_t^\tau\left(\tilde{\theta}\right) + \upsilon_t^\tau$

where $x_t^\tau\left(\tilde{\theta}\right) = \mathcal{E}_t\left\{x_{t+\tau}|\tilde{\theta}\right\}$ is the forecast of the interest rate differential τ periods hence, according to the Kalman Filter with parameter $\tilde{\theta}$: $x_t = \mu + z_t + \nu_t \quad ; \quad \lambda(L)\, z_t = \epsilon_t$. $x_t = r_t - r_t^{us}$ is the forward premium.

	Estimated Coefficients						Estimated Coefficients					
	λ_1	λ_2	η	$\sum\lambda$			λ_1	λ_2	η	$\sum\lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs	SE	SE	SE	SE	Log-Lik	Obs
	Panel A: Canada-U.S.						Panel B: France-U.S.					
					Horizon:	All						
I	**0.972**		**2.99**	**0.972**	0.309	327	**0.997**		**151.88**	**0.997**	-0.861	291
	(0.004)		(0.65)	(0.004)			(0.001)		(13.43)	(0.001)		
II	**0.275**	**0.675**	**0.73**	**0.950**	0.323		0.117	**0.879**	**42.76**	**0.996**	-0.836	
	(0.07)	(0.07)	(0.26)	(0.008)			(0.16)	(0.16)	(8.70)	(0.001)		
					Horizon:	3 Months						
I	**0.996**		**2.40**	**0.996**	0.579	109	**1.027**		**349.75**	**1.027**	-0.723	97
	(0.009)		(0.788)	(0.009)			(0.004)		(80.11)	(0.004)		
II	**0.155**	**0.828**	0.23	**0.983**	0.684		**-0.111**	**1.150**	**41.68**	**1.039**	-0.639	
	(0.08)	(0.08)	(0.14)	(0.012)			(0.04)	(0.04)	(11.38)	(0.01)		
					Horizon:	6 Months						
I	**0.974**		**3.86**	**0.974**	0.448	109	**1.009**		**186.85**	**1.009**	-0.702	97
	(0.008)		(1.335)	(0.008)			(0.004)		(34.05)	(0.004)		
II	**0.385**	**0.573**	**1.31**	**0.958**	0.454		**0.477**	**0.536**	**79.67**	**1.013**	-0.702	
	(0.034)	(0.031)	(0.586)	(0.013)			(0.09)	(0.07)	(0.34)	(0.03)		
					Horizon:	12 Months						
I	**0.969**		**4.45**	**0.969**	0.072	109	**1.005**		**190.39**	**1.005**	-0.181	97
	(0.007)		(2.049)	(0.007)			(0.004)		(41.80)	(0.004)		
II	**0.246**	**0.700**	1.327	**0.946**	0.074		**0.304**	**0.704**	**65.73**	**1.008**	-0.179	
	(0.047)	(0.043)	(0.774)	(0.012)			(0.09)	(0.09)	(16.31)	(0.005)		

continued next page

Table 4: Pooled Maximum Likelihood Estimation of Market Filter

continued from previous page

	Estimated Coefficients						Estimated Coefficients					
	λ_1	λ_2	η	$\sum\lambda$			λ_1	λ_2	η	$\sum\lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs	SE	SE	SE	SE	Log-Lik	Obs
	Panel C: Germany-U.S.						Panel D: Italy-U.S.[37]					
					Horizon: All							
I	**0.979**		12.26	**0.979**	0.265	327	**1.060**		NA	**1.060**	-1.569	291
	(0.002)		(9.85)	(0.002)			(0.001)		(NA)	(0.001)		
II	**1.76**	**-0.766**	**14.89**	**0.994**	0.575		0.008	**1.138**	NA	**1.146**	-1.537	
	(0.01)	(0.01)	(2.84)	(0.003)			(0.01)	(0.01)	(NA)	(0.002)		
					Horizon: 3 Months							
I	**0.993**		**3.853**	**0.993**	0.927	109	**1.048**		NA	**1.048**	-1.342	97
	(0.003)		(1.92)	(0.003)			(0.002)		(NA)	(0.002)		
II	**1.766**	**-0.769**	**14.64**	**0.997**	1.262		-0.002	**1.160**	NA	**1.157**	-1.081	
	(0.02)	(0.02)	(4.67)	(0.001)			(0.008)	(0.001)	(NA)	(0.003)		
					Horizon: 6 Months							
I	**0.978**		28.97	**0.978**	0.039	109	**1.06**		NA	**1.06**	-1.469	97
	(0.003)		(89.44)	(0.003)			(0.001)		(NA)	(0.001)		
II	**1.804**	**-0.810**	**21.28**	**0.994**	0.481		0.084	1.058	NA	**1.142**	-1.373	
	(0.017)	(0.016)	(6.95)	(0.001)			(6.16)	(6.70)	(NA)	(0.54)		
					Horizon: 12 Months							
I	**0.978**		22.03	**0.978**	0.206	109	**1.059**		NA	**1.059**	-1.380	97
	(0.002)		(52.91)	(0.002)			(0.001)		(NA)	(0.001)		
II	**1.736**	**-0.742**	**14.69**	**0.994**	0.468		**0.119**	**1.015**	NA	**1.134**	-1.267	
	(0.01)	(0.02)	(4.63)	(0.001)			(0.04)	(0.04)	(NA)	(0.004)		

continued next page

Table 4: Pooled Maximum Likelihood Estimation of Market Filter

continued from previous page

	Estimated Coefficients						Estimated Coefficients					
	λ_1	λ_2	η	$\sum\lambda$			λ_1	λ_2	η	$\sum\lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs	SE	SE	SE	SE	Log-Lik	Obs
	Panel E: Japan-U.S.						Panel F: U.K.-U.S.					
					Horizon: All							
I	**0.981**		3.44	**0.981**	0.530	327	**0.977**		**1.67**	**0.977**	0.195	327
	(0.002)		(1.02)	(0.002)			(0.002)		(0.32)	(0.002)		
II	**1.690**	**-0.697**	**8.58**	**0.993**	0.701		**1.766**	**-0.774**	**27.34**	**0.992**	0.414	
	(0.02)	(0.02)	(1.57)	(0.007)			(0.02)	(0.02)	(6.39)	(0.001)		
					Horizon: 3 Months							
I	**0.996**		**0.36**	**0.996**	1.192	109	**0.997**		**1.25**	**0.997**	0.727	109
	(0.004)		(0.12)	(0.004)			(0.005)		(0.33)	(0.005)		
II	**1.784**	**-0.787**	**14.96**	**0.997**	1.661		**1.690**	**-0.693**	**13.52**	**0.997**	0.852	
	(0.03)	(0.03)	(5.28)	(0.001)			(0.07)	(0.07)	(6.86)	(0.002)		
					Horizon: 6 Months							
I	**0.988**		5.31	**0.988**	0.524	109	**0.986**		**1.69**	**0.986**	0.093	109
	(0.004)		(2.97)	(0.004)			(0.004)		(0.52)	(0.004)		
II	**1.781**	**-0.788**	**14.22**	**0.993**	1.022		**1.806**	**-0.812**	**38.72**	**0.994**	0.380	
	(0.02)	(0.024)	(3.80)	(0.001)			(0.03)	(0.03)	(14.74)	(0.001)		
					Horizon: 12 Months							
I	**0.978**		0.51	**0.978**	0.229	109	**0.972**		**1.74**	**0.972**	0.079	109
	(0.002)		(0.38)	(0.002)			(0.003)		(0.78)	(0.003)		
II	**1.527**	**-0.537**	**5.28**	**0.990**	0.264		**1.776**	**-0.784**	**37.43**	**0.992**	0.313	
	(0.06)	(0.06)	(2.52)	(0.001)			(0.02)	(0.02)	(15.09)	(0.001)		

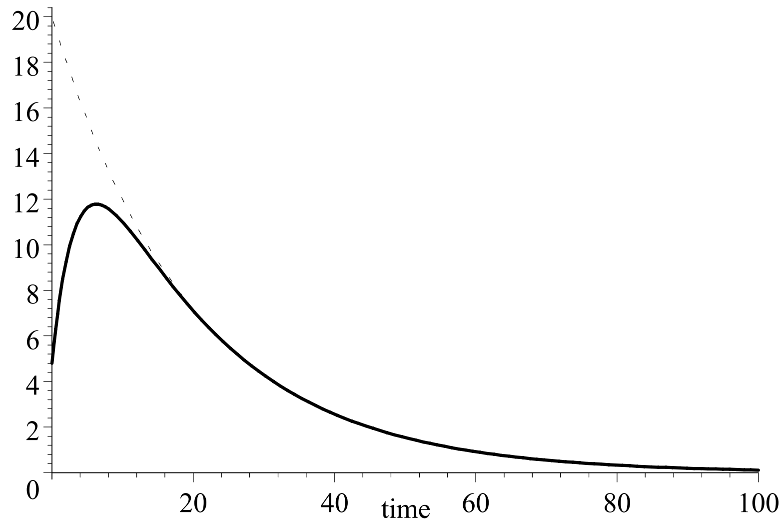

Figure 1: Delayed overshooting (–) and rational expectation (- -)response to monetary innovation.

$\lambda = 0.95,\ k = 0.2.$

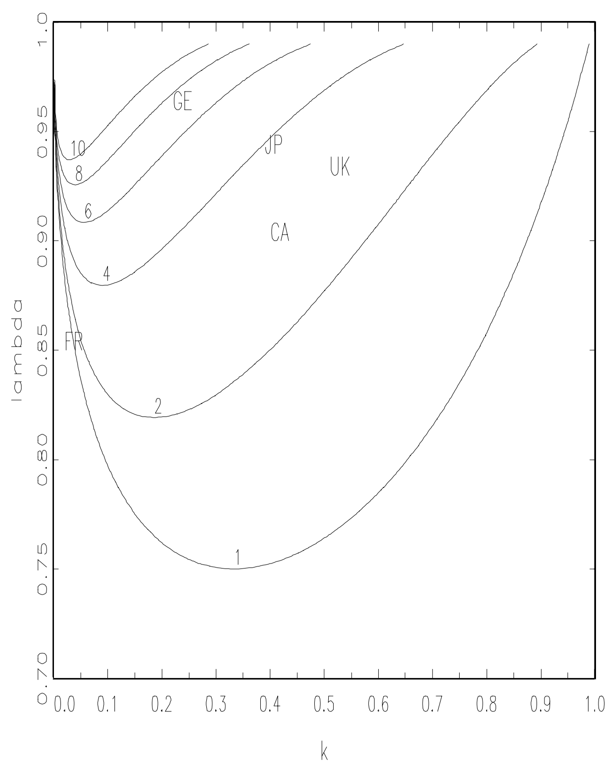

Figure 2: Contour plot of Delayed Overshooting Region as a function of the parameters (λ, k).

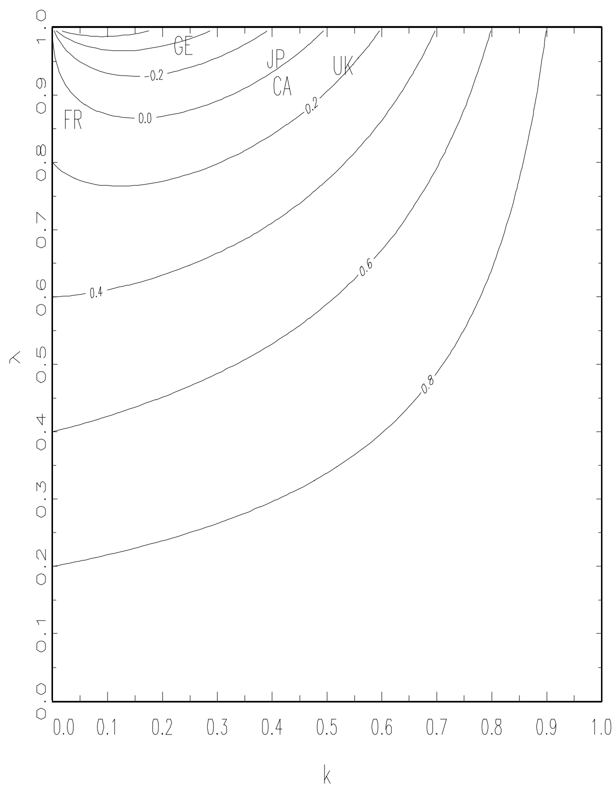

Figure 3: Contour plot of Asymptotic Limit of Fama coefficient as a function of k and λ.

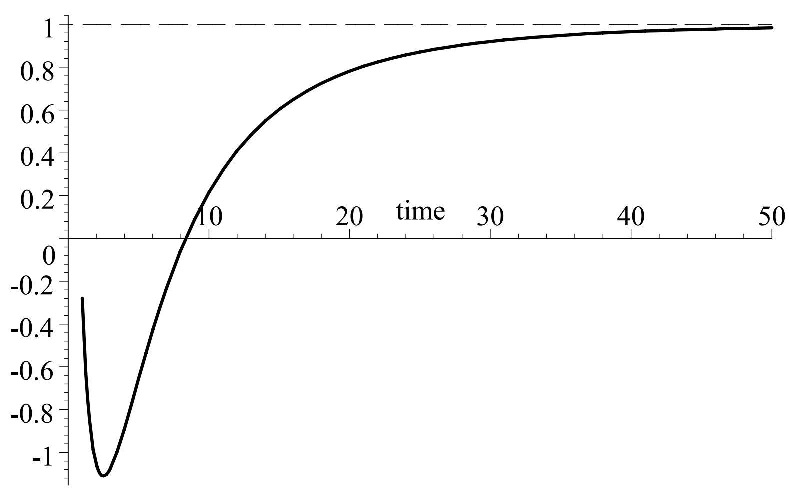

Figure 4: Figure 4: Fama coefficient β_n as a function of the regression horizon. $\lambda = 0.95$ and $k = 0.2$.

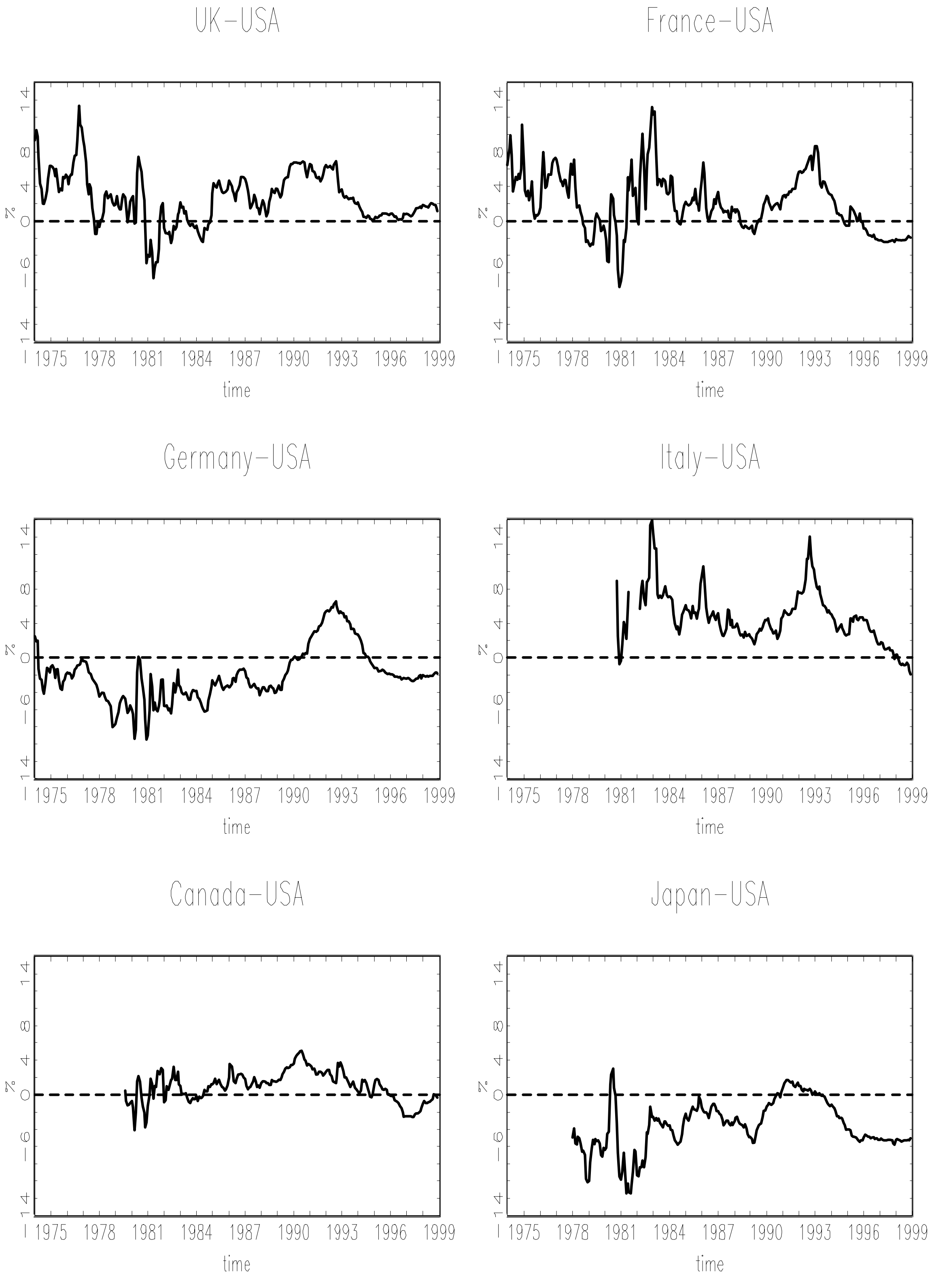

Figure 5: Euro 3 Months Forward Premium against the US dollar.

A. Appendices

A.1. Data sources.

Variable	Definition	Database	Source	Series
Eurorates				
US	3 months US deposit	DRI	IFS	L60LD&D@C111
	London offer rate			
UK	average of the 3 months	DRI	Reuters	UKD03A/B
	eurocurrency deposit			
	bid and ask rate			
France	--	DRI	Reuters	FRD03A/B
Germany	--	DRI	Reuters	WGD03AY/BY
Italy	--	DRI	Reuters	ITD03A/B
Canada	--	DRI	Reuters	CAD03A/B
Japan	--	DRI	Reuters	JAD03A/B.
Consensus Forecasts	3,6 and 12 months consensus	Financial Times		
	forecasts on 3 months Eurorates	Currency Forecaster		
Exchange Rates	spot rates	DRI	IFS	

A.2. Currency Risk Premium versus Expectational Errors.

The standard Fama regression takes the following form:

$$\Delta e_t^n = e_{t+n} - e_t = \alpha + \beta_n x_t^n + \eta_{t+n}^n$$

where $x_t^n = r_t^n - r_t^{*n}$ is the current $n-period$ forward discount, equal to the $n-period$ interest rate differential. Under the null that (a) expectations are rational and (b) the currency risk premium is uncorrelated with x_t^n, $\beta_n = 1$. The estimates of this regression are shown in Table 1. Exchange rates are spot rates from DRI-IFS and interest rates are 3, 6 and 12 months eurorates. The regression uses monthly observations and standard errors correct for the overlap between the sampling frequency and the horizon. The results are typical of this literature. The Fama coefficient is often significantly negative and almost always smaller than one.[38] At longer horizons (12 months), the coefficients are larger, often becoming not significantly different from 1.

For the sub-period 84:5 to 97:5, we also have exchange rate survey forecasts, compiled by the Financial Times Currency Forecaster. Frankel and Froot (1989) show that the Fama coefficient can be decomposed into and expectational component b_{re} and a risk premium component b_{rp} such that:

$$\begin{aligned} \beta &= 1 - b_{re} - b_{rp} \\ b_{re} &= -\frac{cov\left(u_{t+k}^k, x_t^k\right)}{var\left(x_t^k\right)}; \quad b_{rp} = \frac{var\left(\zeta_t^k\right) + cov\left(\mathcal{E}_t^m e_{t+k} - e_t, \zeta_t^k\right)}{var\left(x_t^k\right)} \end{aligned}$$

where ζ_t^k is the risk premium for horizon k and u_{t+k}^k is the market expectational error $e_{t+k} - \mathcal{E}_t^m e_{t+k}$. b_{re} is equal to 0 when this error term is uncorrelated with the forward premium, i.e. expectations are rational. The second term is zero if the risk premium is constant. Table 2 reports the results of this decomposition. The results validate Frankel and Froot (1989)'s earlier analysis: at short horizons (3 to 6 months), for all countries except Italy, the expectational term b_{re} dominates the risk premium term b_{rp} by an order of magnitude. If anything, the risk premium term would tend to *increase* the Fama coefficient. At 12 months, however, the results are reversed with b_{re} contributing positively to β_n.

A.3. Kalman Filter Estimation. In this part we derive the Kalman filter equations by following Hamilton (1994). We postulate the following process for the forward premium:

$$\mathbf{y}_t = \mathbf{H}'\mathbf{z}_t + \boldsymbol{\nu}_t \tag{A.1}$$

$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \tag{A.2}$$

where $\mathbf{y}_t$ is a qx1 vector, $\mathbf{z}_t$ is a px1 state vector, $\mathbf{H}$ and $\mathbf{F}$ are qxp and pxp matrices, respectively. Equation (A.1) is the measurement equation and (A.2) is the space equation with $\mathcal{E}\{\boldsymbol{\nu}_t\} = \mathbf{R}$ and $\mathcal{E}\{\boldsymbol{\epsilon}_t\} = \mathbf{Q}$ where $\mathbf{Q}$ and $\mathbf{R}$ are respectively a pxp and qxq matrices. Define the informations set $I_t = \{\mathbf{y}_{t-i}, i \geq 0\}$, $\hat{\mathbf{z}}_{t+1|t} = \mathcal{E}\{\mathbf{z}_{t+1}|I_t\}$, and $\hat{\mathbf{P}}_{t+1|t} = \mathcal{E}\left\{\left(\mathbf{z}_{t+1} - \hat{\mathbf{z}}_{t+1|t}\right)\left(\mathbf{z}_{t+1} - \hat{\mathbf{z}}_{t+1|t}\right)' |I_t\right\}$. The filtering equations are:

$$\hat{\mathbf{z}}_{t+1|t} = \mathbf{F}\hat{\mathbf{z}}_{t-1|t} + \mathbf{F}\hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \mathbf{R}\right)^{-1}\left(\mathbf{y}_t - \mathbf{H}'\hat{\mathbf{z}}_{t|t-1}\right) \tag{A.3}$$

$$\hat{\mathbf{P}}_{t+1|t} = \mathbf{F}\left[\hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \mathbf{R}\right)^{-1}\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\right]\mathbf{F}' + \mathbf{Q} \tag{A.4}$$

The smoother equations are:

$$\begin{aligned}
\hat{\mathbf{z}}_{t|T} &= \hat{\mathbf{z}}_{t|t} + \hat{\mathbf{P}}_{t|t}\mathbf{F}'\,\hat{\mathbf{P}}_{t+1|t}^{-1}\left(\hat{\mathbf{z}}_{t+1|T} - \hat{\mathbf{z}}_{t+1|t}\right) \\
\hat{\mathbf{P}}_{t|T} &= \hat{\mathbf{P}}_{t|t} + \left(\hat{\mathbf{P}}_{t|t}\mathbf{F}'\hat{\mathbf{P}}_{t+1|t}^{-1}\right)\left(\hat{\mathbf{P}}_{t+1|T} - \hat{\mathbf{P}}_{t+1|t}\right)\left(\hat{\mathbf{P}}_{t|t}\mathbf{F}'\hat{\mathbf{P}}_{t+1|t}^{-1}\right)', \text{ where} \\
\hat{\mathbf{P}}_{t|t} &= \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \mathbf{R}\right)^{-1}\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}
\end{aligned}$$

Suppose that the current estimate for the state variable at time t is $\hat{\mathbf{z}}_{t|t}$. According to (A.1)-(A.2), the market forecast for $\mathbf{y}$ τ-periods hence is:

$$\mathbf{y}_t^\tau\left(\tilde{\theta}\right) \equiv \mathcal{E}\{\mathbf{y}_{t+\tau}|I_t\} = \mathbf{H}'\hat{\mathbf{z}}_{t+\tau|t} = \mathbf{H}'\mathbf{F}^\tau\hat{\mathbf{z}}_{t|t} \tag{A.5}$$

Suppose now that we observe an imprecise measure of this forecast: $\hat{\mathbf{y}}_t^\tau = \mathbf{y}_t^\tau\left(\tilde{\theta}\right) + \boldsymbol{\upsilon}_t^\tau$, where the measurement error υ_t^τ is uncorrelated with the true forecast. We can estimate $\tilde{\theta}$ by minimizing:

$$\text{(A.6)} \qquad S\left(\tilde{\theta}\right) = \sum_{\tau}\sum_{t=1}^{T}\left(\hat{\mathbf{y}}_t^{\tau} - \mathbf{y}_t^{\tau}\left(\tilde{\theta}\right)\right)'\left(\hat{\mathbf{y}}_t^{\tau} - \mathbf{y}_t^{\tau}\left(\tilde{\theta}\right)\right)$$

Applying (A.1)-(A.2) to the interest rate differential with up to three lags in the persistent component, we estimate (9)-(10) as well as the market process, using (A.6).Conversely, applying the Kalman filter to the vector of interest rates $\mathbf{r}_t = (r_t, r_t^*)'$, and the state vector $\mathbf{z}_t = (\bar{z}_t, z_t, z_t^*)$, we obtain lemma 1.

A.4. Empirical Results on Interest Rate Differentials. In this part we develop the estimation procedure for our state-space representation. Assume that the State Space representation (9)-(10) holds. Under the normality assumption, and assuming additionally that $\hat{\mathbf{x}}_{1|0}^p$ is normally distributed, $\hat{\mathbf{x}}_{t+1}^p$ is normally distributed conditionally on I_t, with mean $\hat{\mathbf{x}}_{t+1|t}^p$ and variance $\hat{\mathbf{P}}_{t+1|t}$. We can then write the conditional likelihood of x_{t+1} as:

$$\log f_{x_{t+1}|I_t}\left(x_{t+1}|I_t\right) \propto \log\left|\mathbf{H}'\hat{\mathbf{P}}_{t+1|t}\mathbf{H} + \sigma_\nu^2\mathbf{I}\right| + \left(\frac{\left(x_{t+1} - \mu - \mathbf{H}_{t+1|t}'\hat{\mathbf{x}}_{t+1|t}^p\right)^2}{\mathbf{H}'\hat{\mathbf{P}}_{t+1|t}\mathbf{H} + \sigma_\nu^2\mathbf{I}}\right)$$

We maximize the sample log likelihood $\sum_{t=0}^{T-1}\log f_{x_{t+1}|I_t}\left(x_{t+1}|I_t\right)$ with respect to the vector of parameters $\theta = \left(\{\lambda_i\}_{i=1}^p, \eta, \sigma_\epsilon^2, \mu\right)'$.[39] To initiate the estimation procedure, we need an estimate of the space variable $\hat{\mathbf{x}}_0^p$ and its conditional mean square error. Maximum likelihood estimation over the vector θ is then performed. Once an estimate $\hat{\theta}^0$ is found, we run the smoother in order to revise the initial state vector. That is, the smoother gives us the initial value of the persistent component, conditional on the entire sample information and the filter parameters, $\hat{\mathbf{x}}_0^{p1} = \mathcal{E}\left\{\hat{\mathbf{x}}_0^p|I_T, \hat{\theta}^0\right\}$, and its mean square error. In general, this revised estimate does not correspond to the initial one. We can then iterate the maximum likelihood estimation with this new initial state variable until convergence to $\hat{\theta}^1$. Iterating this procedure will give ultimately a parameter vector consistent with the initial state vector.[40]

A.5. Proofs.

Under rational expectations, interest rates and exchange rates follow:

$$\text{(A.7)} \qquad \mathbf{r}_t = \ln R + \mathbf{H}'\mathbf{z}_t$$

$$\mathcal{E}_t^m\left\{e_{t+1}\right\} - e_t = \left(\delta - \delta^*\right)\bar{z}_t + z_t - z_t^* + \frac{\bar{\varphi}^2 - \bar{\varphi}^{*2} + \varphi^2 - \varphi^{*2}}{2}\sigma^2$$

$$\text{(A.8)} \qquad = r_t - r_t^* + \zeta$$

Proof. Using equations (14) and (15) in the pricing equation (12), we obtain:

$$r_t = -\log\mathcal{E}_t\exp\left(-\ln R - \frac{\bar{\varphi}^2\sigma^2}{2} - \delta\,\bar{z}_t - \frac{\varphi^2\sigma^2}{2} - z_t - \bar{\varphi}\,\bar{\epsilon}_{t+1} - \varphi\epsilon_{t+1}\right)$$

$$
\begin{aligned}
&= \ln R + \frac{\bar{\varphi}^2\sigma^2}{2} + \delta\,\bar{z}_t + \frac{\varphi^2\sigma^2}{2} + z_t - \log \mathcal{E}_t \exp\left(-\bar{\varphi}\,\bar{\epsilon}_{t+1} - \varphi\epsilon_{t+1}\right) \\
&= \ln R + \frac{\bar{\varphi}^2\sigma^2}{2} + \delta\,\bar{z}_t + \frac{\varphi^2\sigma^2}{2} + z_t - \frac{\varphi^2\sigma^2}{2} - \frac{\bar{\varphi}^2\sigma^2}{2} \\
&= \ln R + \delta\,\bar{z}_t + z_t
\end{aligned}
$$

A similar equation obtains for the foreign interest rate. Defining $\mathbf{r}_t$ as the vector $(r_t, r_t^*)'$, we obtain equation (A.7).

From equation (13), we obtain:

$$
\begin{aligned}
\mathcal{E}_t^m e_{t+1} - e_t &= \mathcal{E}_t^m \log R^*_{t,t+1} - \mathcal{E}_t^m \log R_{t,t+1} \\
&= \mathcal{E}_t^m \left(-\ln R - \frac{\bar{\varphi}^{*2}\sigma^2}{2} - \delta^*\,\bar{z}_t - \frac{\varphi^{*2}\sigma^2}{2} - z_t^* - \bar{\varphi}^*\,\bar{\epsilon}_{t+1} - \varphi^*\epsilon^*_{t+1} \right) - \\
&\quad \mathcal{E}_t^m \left(-\ln R - \frac{\bar{\varphi}^2\sigma^2}{2} - \delta\,\bar{z}_t - \frac{\varphi^2\sigma^2}{2} - z_t - \bar{\varphi}\,\bar{\epsilon}_{t+1} - \varphi\epsilon_{t+1} \right) \\
&= \left(\delta - \delta^*\right)\,\bar{z}_t + z_t - z_t^* + \frac{\sigma^2}{2}\left(\bar{\varphi}^2 - \bar{\varphi}^{*2} + \varphi^2 - \varphi^{*2}\right)
\end{aligned}
$$

■

Lemma 4. *Assume that beliefs about* $\mathbf{z}_1$ *are initially distributed as* $\mathcal{N}\left(\mathcal{E}_t^m \mathbf{z}_{1|0}, \mathbf{P}_1\right)$ *where* $\mathcal{E}_t^m \mathbf{z}_{1|0}$ *and* $\mathbf{P}_1$ *are an appropriate vector and matrix respectively. Then:*

1. *Beliefs evolve according to:*

$$
\mathcal{E}_t^m \mathbf{z}_{t+1} = \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P}_t \mathbf{H}\left(\mathbf{H}'\mathbf{P}_t\mathbf{H} + \sigma_v^2 \mathbf{I}\right)^{-1}\left(\mathbf{r}_t - \log R - \mathbf{H}'\mathcal{E}_{t-1}^m \mathbf{z}_t\right)
$$

2. *The conditional variance* $\mathbf{P}_t$ *evolves according to:*

$$
\mathbf{P}_{t+1} = \lambda^2\left[\mathbf{P}_t - \mathbf{P}_t\mathbf{H}\left(\mathbf{H}'\mathbf{P}_t\mathbf{H} + \sigma_v^2\mathbf{I}\right)^{-1}\mathbf{H}'\mathbf{P}_t\right] + \sigma^2\mathbf{I}
$$

in particular, it does not depend upon the actual realizations of the interest rates $\mathbf{r}_t$*.*

3. *In the limit as* $t \to \infty$*, the conditional variance converges to a steady state value* $\mathbf{P}$*, solution of:*

$$
\mathbf{P} = \lambda^2\left[\mathbf{P} - \mathbf{P}\mathbf{H}\left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2\mathbf{I}\right)^{-1}\mathbf{H}'\mathbf{P}\right] + \sigma^2\mathbf{I}
$$

and the beliefs evolve according to:

$$
\mathcal{E}_t^m \mathbf{z}_{t+1} = \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P}\mathbf{H}\left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2\mathbf{I}\right)^{-1}\left(\mathbf{r}_t - \log R - \mathbf{H}'\mathcal{E}_{t-1}^m \mathbf{z}_t\right)
$$

Proof. This is a direct application of the Kalman filter presented in appendix A.3. The Kalman filter consists of the following state and measurement equations:

$$\begin{aligned} \mathbf{r}_t &= -\ln R + \mathbf{H}'\mathbf{z}_t + \mathbf{v}_t \\ \mathbf{z}_t &= \lambda \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t \end{aligned}$$

so we can apply equations (A.3) and (A.4) with $\mathbf{y}_t = \mathbf{r}_t$, $\mathcal{E}_t^m \mathbf{z}_{t+1} = \hat{\mathbf{z}}_{t+1|t}$, $\mathbf{F} = \lambda.\mathbf{I}$, $\mathbf{R} = \sigma_v^2.\mathbf{I}$, and $\mathbf{Q} = \sigma^2.\mathbf{I}$. Define $\mathbf{P}_t = \mathcal{E}_{t-1}^m \left\{ \left(\mathbf{z}_t - \mathcal{E}_{t-1}^m \mathbf{z}_t \right) \left(\mathbf{z}_t - \mathcal{E}_{t-1}^m \mathbf{z}_t \right)' \right\}$. The filtering equation is then

$$\begin{aligned} \mathcal{E}_t^m \mathbf{z}_{t+1} &= \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P}_t \mathbf{H} \left(\mathbf{H}'\mathbf{P}_t \mathbf{H} + \sigma_v^2.\mathbf{I} \right)^{-1} \left(\mathbf{r}_t - \log R - \mathbf{H}' \mathcal{E}_{t-1}^m \mathbf{z}_t \right) \\ \mathbf{P}_{t+1} &= \lambda^2 \left[\mathbf{P}_t - \mathbf{P}_t \mathbf{H} \left(\mathbf{H}'\mathbf{P}_t \mathbf{H} + \sigma_v^2.\mathbf{I} \right)^{-1} \mathbf{H}'\mathbf{P}_t \right] + \sigma^2.\mathbf{I} \end{aligned}$$

The rest of the proof is immediate. ■

Pre-multiplying by $\mathbf{H}'$, and adding $\log R$ on both sides, we obtain:

$$\begin{aligned} \log R + \mathbf{H}'\mathcal{E}_t^m \mathbf{z}_{t+1} &= \lambda \left(\log R + \mathbf{H}'\mathcal{E}_{t-1}^m \mathbf{z}_t \right) + \lambda \mathbf{H}'\mathbf{P}\mathbf{H} \left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2.\mathbf{I} \right)^{-1} \left(\mathbf{r}_t - \log R - \mathbf{H}'\mathcal{E}_{t-1}^m \mathbf{z}_t \right) + (1-\lambda) \log R \\ \mathcal{E}_t^m \mathbf{r}_{t+1} &= \lambda \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda \mathbf{K} \left(\mathbf{r}_t - \mathcal{E}_{t-1}^m \mathbf{r}_t \right) + (1-\lambda) \log R \\ \mathcal{E}_t^m \mathbf{r}_{t+1} &= \lambda \left(\mathbf{I} - \mathbf{K} \right) \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda \mathbf{K} \mathbf{r}_t + (1-\lambda) \log R \end{aligned}$$

where $\mathbf{K} = \mathbf{H}'\mathbf{P}\mathbf{H} \left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2.\mathbf{I} \right)^{-1}$ is the gain of the filter. In the case without common shocks $(\delta = \delta^* = 0)$, this formula simplifies to:

$$\begin{aligned} \mathbf{K} &= k.\mathbf{I} = \frac{\tilde{\sigma}^2}{\tilde{\sigma}^2 + \sigma_v^2}.\mathbf{I} \\ \mathbf{P} &= \tilde{\sigma}^2.\mathbf{I} = \frac{\sigma^2}{\left(1 - \lambda^2 (1-k)\right)}.\mathbf{I} \end{aligned}$$

Substituting for $\tilde{\sigma}^2$ and solving, we obtain:

$$\begin{aligned} k &= \frac{1 + \Delta - \eta \left(1 - \lambda^2\right)}{1 + \Delta + \eta \left(1 + \lambda^2\right)} \leq 1 \\ \tilde{\sigma}^2 &= \frac{1 + \Delta + \eta \left(1 + \lambda^2\right)}{1 + \Delta + \eta \left(1 - \lambda^2\right)} \sigma^2 \end{aligned}$$

where $\Delta = \left(\eta \left(1 - \lambda^2\right) + 1 \right)^2 + 4\eta\lambda^2$ and $\eta = \sigma_v^2 / \sigma^2$ is the noise-to-signal ratio. Under misperception, we postulate that the exchange rate is a function of the current interest rates $\mathbf{r}_t$ and the current forecasts $\mathcal{E}_{t-1}^m \mathbf{z}_t$, as well as a long term equilibrium value $\bar{e}_t$ that satisfies $\bar{e}_t = \lim_{T\to\infty} \mathcal{E}_t^m e_{t+T}$. Iterating on the uncovered interest rate parity condition, we obtain:

$$e_t = -\sum_{j=0}^{\infty} \mathcal{E}_t^m x_{t+j} + \lim_{T\to\infty} \mathcal{E}_t^m e_{t+T} = -\sum_{j=0}^{\infty} \mathcal{E}_t^m x_{t+j} + \bar{e}_t$$

To compute the exchange rate, we solve for $\mathcal{E}_t^m x_{t+j}$ at all $j \geq 2$:

$$
\begin{aligned}
\mathcal{E}_t^m x_{t+j} &= \mathbf{l}'\mathcal{E}_t^m \mathbf{r}_{t+j} \\
&= \mathbf{l}'\mathcal{E}_t^m \left\{\mathcal{E}_{t+j-1}^m \mathbf{r}_{t+j}\right\} \\
&= \mathbf{l}'\mathcal{E}_t^m \left\{\lambda\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_{t+j-2}^m \mathbf{r}_{t+j-1} + \lambda\mathbf{K}\mathbf{r}_{t+j-1} + (1-\lambda)\log R\right\} \\
&= \mathbf{l}'\lambda\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_t^m \mathbf{r}_{t+j-1} + \mathbf{l}'\lambda\mathbf{K}\,\mathcal{E}_t^m \mathbf{r}_{t+j-1} \\
&= \lambda\,\mathbf{l}'\mathcal{E}_t^m \mathbf{r}_{t+j-1} \\
&= \lambda^{j-1}\mathcal{E}_t^m x_{t+1},
\end{aligned}
$$

where the second line uses the law of iterated expectations and the third line uses $\mathbf{l}'(1-\lambda)\log R = 0$. Substituting back into the expression for the exchange rate, we obtain:

$$
\begin{aligned}
e_t &= -x_t - \mathcal{E}_t^m x_{t+1} - \mathcal{E}_t^m x_{t+1}\sum_{j=2}^{\infty}\lambda^{j-1} + \bar{e}_t \\
&= -x_t - \frac{1}{1-\lambda}\mathcal{E}_t^m x_{t+1} + \bar{e}_t \\
&= -\mathbf{l}'\mathbf{r}_t - \frac{1}{1-\lambda}\mathbf{l}'\mathcal{E}_t^m \mathbf{r}_{t+1} + \bar{e}_t \qquad \text{(A.9)}
\end{aligned}
$$

Under rational expectation, $\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda\mathbf{r}_t$ and the exchange rate follows:

$$
e_t^r = -\frac{1}{1-\lambda}\mathbf{l}'\mathbf{r}_t + \bar{e}_t,
$$

where we are assuming implicitly that $\lim_{T\to\infty}\mathcal{E}_t^m e_{t+T} = \lim_{T\to\infty}\mathcal{E}_t e_{t+T} = \bar{e}_t$. In other words, we are assuming that there are no long run prediction errors. To calculate predictable excess returns, recall from equation (2) that $\xi_t = \mathcal{E}_t^m e_{t+1} - \mathcal{E}_t e_{t+1}$. Substituting (A.9), we obtain:

$$
\begin{aligned}
\xi_t &= \mathcal{E}_t^m\left(-\mathbf{l}'\mathbf{r}_{t+1} - \frac{1}{1-\lambda}\mathbf{l}'\mathcal{E}_{t+1}^m \mathbf{r}_{t+2} + \bar{e}_{t+1}\right) \\
&\quad -\mathcal{E}_t\left(-\mathbf{l}'\mathbf{r}_{t+1} - \frac{1}{1-\lambda}\mathbf{l}'\mathcal{E}_{t+1}^m \mathbf{r}_{t+2} + \bar{e}_{t+1}\right) \\
&= -\mathcal{E}_t^m\left(\mathbf{l}'\left(\mathbf{I} + \frac{\lambda}{1-\lambda}\mathbf{K}\right)\mathbf{r}_{t+1} + \frac{\lambda}{1-\lambda}\mathbf{l}'\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_t^m \mathbf{r}_{t+1} - \bar{e}_{t+1}\right) \\
&\quad +\mathcal{E}_t\left(\mathbf{l}'\left(\mathbf{I} + \frac{\lambda}{1-\lambda}\mathbf{K}\right)\mathbf{r}_{t+1} + \frac{\lambda}{1-\lambda}\mathbf{l}'\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_t^m \mathbf{r}_{t+1} - \bar{e}_{t+1}\right) \\
&= \mathbf{l}'\left(\mathbf{I} + \frac{\lambda}{1-\lambda}\mathbf{K}\right)\left(\mathcal{E}_t \mathbf{r}_{t+1} - \mathcal{E}_t^m \mathbf{r}_{t+1}\right),
\end{aligned}
$$

where the second line uses the definition of $\mathcal{E}_{t+1}^m \mathbf{r}_{t+2}$ and the assumption that $\mathcal{E}_t \bar{e}_{t+1} = \mathcal{E}_t^m \bar{e}_{t+1}$.

Lemma 5. *The coefficient from the regression of realized depreciation rates on the forward premium converges in plim to*

$$
p\lim\beta = 1 - \frac{\lambda\mathbf{l}'\left(\mathbf{I} + \frac{\lambda\mathbf{K}}{1-\lambda}\right)\left(\mathbf{I}-\mathbf{K}\right)\left(\mathbf{I} - \left(\mathbf{I} - \lambda^2\left(\mathbf{I}-\mathbf{K}\right)\right)^{-1}\lambda^2\mathbf{K}\right) var\left(\mathbf{r}_t\right)\mathbf{l}}{\mathbf{l}' var\left(\mathbf{r}_t\right)\mathbf{l}} \qquad \text{(A.10)}
$$

Proof. According to its definition,

$$
\begin{aligned}
p\lim\beta &= \frac{cov\left(e_{t+1}-e_t,x_t\right)}{var\left(x_t\right)} \\
&= \frac{cov\left(\mathcal{E}_te_{t+1}-e_t+u_{t+1},x_t\right)}{var\left(x_t\right)} \\
&= \frac{cov\left(\mathcal{E}_te_{t+1}-e_t,x_t\right)}{var\left(x_t\right)} \\
&= \frac{cov\left(\mathcal{E}_te_{t+1}-\mathcal{E}_t^m e_{t+1}+\mathcal{E}_t^m e_{t+1}-e_t,x_t\right)}{var\left(x_t\right)} \\
&= \frac{cov\left(\mathcal{E}_te_{t+1}-\mathcal{E}_t^m e_{t+1}+x_t,x_t\right)}{var\left(x_t\right)} \\
&= 1-\frac{cov\left(\xi_t,x_t\right)}{var\left(x_t\right)}
\end{aligned}
$$

where $u_{t+1}=e_{t+1}-\mathcal{E}_te_{t+1}$ is uncorrelated with x_t and the fifth line uses the uncovered interest parity. Using the definition of ξ_t, we see that

$$
\begin{aligned}
cov\left(\xi_t,x_t\right) &= \mathbf{l}'\left(\mathbf{I}+\frac{\lambda}{1-\lambda}\mathbf{K}\right)cov\left(\mathcal{E}_t\mathbf{r}_{t+1}-\mathcal{E}_t^m\mathbf{r}_{t+1},\mathbf{r}_t\right)\mathbf{l} \\
&= \lambda\mathbf{l}'\left(\mathbf{I}+\frac{\lambda}{1-\lambda}\mathbf{K}\right)\left(\mathbf{I}-\mathbf{K}\right)cov\left(\left(\mathbf{r}_t-\mathcal{E}_{t-1}^m\mathbf{r}_t\right),\mathbf{r}_t\right)\mathbf{l} \\
&= \lambda\mathbf{l}'\left(\mathbf{I}+\frac{\lambda}{1-\lambda}\mathbf{K}\right)\left(\mathbf{I}-\mathbf{K}\right)\left(var\left(\mathbf{r}_t\right)-cov\left(\mathcal{E}_{t-1}^m\mathbf{r}_t,\mathbf{r}_t\right)\right)\mathbf{l}
\end{aligned}
$$

and

$$
\begin{aligned}
cov\left(\mathcal{E}_{t-1}^m\mathbf{r}_t,\mathbf{r}_t\right) &= cov\left(\lambda\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_{t-2}^m\mathbf{r}_{t-1}+\lambda\mathbf{K}\mathbf{r}_{t-1}+(1-\lambda)\log R,\lambda\mathbf{r}_{t-1}+\mathbf{H}'\boldsymbol{\epsilon}_t\right) \\
&= cov\left(\lambda\left(\mathbf{I}-\mathbf{K}\right)\mathcal{E}_{t-2}^m\mathbf{r}_{t-1}+\lambda\mathbf{K}\mathbf{r}_{t-1},\lambda\mathbf{r}_{t-1}\right) \\
&= \lambda^2\left(\mathbf{I}-\mathbf{K}\right)cov\left(\mathcal{E}_{t-2}^m\mathbf{r}_{t-1},\mathbf{r}_{t-1}\right)+\lambda^2\mathbf{K}\,var\left(\mathbf{r}_{t-1}\right)
\end{aligned}
$$

so that

$$
cov\left(\mathcal{E}_{t-1}^m\mathbf{r}_t,\mathbf{r}_t\right)=\left(\mathbf{I}-\lambda^2\left(\mathbf{I}-\mathbf{K}\right)\right)^{-1}\lambda^2\mathbf{K}\,var\left(\mathbf{r}_t\right)
$$

and

$$
cov\left(\xi_t,x_t\right)=\lambda\mathbf{l}'\left(\mathbf{I}+\frac{\lambda}{1-\lambda}\mathbf{K}\right)\left(\mathbf{I}-\mathbf{K}\right)\left(\mathbf{I}-\left(\mathbf{I}-\lambda^2\left(\mathbf{I}-\mathbf{K}\right)\right)^{-1}\lambda^2\mathbf{K}\right)var\left(\mathbf{r}_t\right)\mathbf{l}
$$

Thus, the Fama coefficient equals:

$$
p\lim\beta=1-\frac{\lambda\mathbf{l}'\left(\mathbf{I}+\frac{\lambda}{1-\lambda}\mathbf{K}\right)\left(\mathbf{I}-\mathbf{K}\right)\left(\mathbf{I}-\left(\mathbf{I}-\lambda^2\left(\mathbf{I}-\mathbf{K}\right)\right)^{-1}\lambda^2\mathbf{K}\right)var\left(\mathbf{r}_t\right)\mathbf{l}}{\mathbf{l}'var\left(\mathbf{r}_t\right)\mathbf{l}}
$$

in the case where there are no common shocks, the formula simplifies to

$$
p\lim\beta=1-\frac{\lambda\left(1-\lambda\left(1-k\right)\right)\left(1-k\right)\left(1+\lambda\right)}{1-\lambda^2\left(1-k\right)}
$$

■ In the case without common shocks, one can also derive the Fama coefficient from a regression at horizon n. To do so, we must first derive the yield on a n period bond. Define r_t^n the continuously compounded yield on a n period zero-coupon bond. r_t^n satisfies:

$$r_t^n = -\frac{1}{n}\log \mathcal{E}_t^m R_{t,t+n}$$

Since $R_{t,t+n} = R_{t,t+1}....R_{t+n-1,t+n}$, we observe that r_t^n can be defined recursively as the solution to:

$$-(n+1)\, r_t^{n+1} = \log \mathcal{E}_t^m \left\{ R_{t,t+1} \exp\left(-n r_{t+1}^n\right)\right\} \tag{A.11}$$

where $r_t^1 = r_t$. We conjecture a linear solution of the form:

$$r_t^n = A_n + B_n r_t + C_n \mathcal{E}_t^m r_{t+1} \tag{A.12}$$

with $A_1 = \log R$, $B_1 = 1$ and $C_1 = 0$. Using the fact that $\log R_{t,t+1} = -\ln R - \frac{\varphi^2\sigma^2}{2} - z_t - \varphi\epsilon_{t+1} = -r_t - \frac{\varphi^2\sigma^2}{2} - \varphi\epsilon_{t+1}$, and substituting equation (A.12) into the term structure equation (A.11), we obtain:

$$\begin{aligned}
-(n+1)\; r_t^{n+1} &= \log \mathcal{E}_t^m \left\{ \exp\left(-r_t - \frac{\varphi^2\sigma^2}{2} - \varphi\epsilon_{t+1} - n\left(A_n + B_n r_{t+1} + C_n \mathcal{E}_{t+1}^m r_{t+2}\right)\right)\right\} \\
&= -r_t - \frac{\varphi^2\sigma^2}{2} - nA_n - nB_n \mathcal{E}_t^m r_{t+1} - nC_n\left(\lambda \mathcal{E}_t^m r_{t+1} + (1-\lambda)\log R\right) \\
&\quad + \log \mathcal{E}_t^m \left\{ \exp\left(\varphi\epsilon_{t+1} - n\left(C_n\lambda k + B_n\right)\left(r_{t+1} - \mathcal{E}_t^m r_{t+1}\right)\right)\right\} \\
&= -(n+1)\; \left(A_{n+1} + B_{n+1} r_t + C_{n+1}\mathcal{E}_t^m r_{t+1}\right)
\end{aligned}$$

Identifying coefficients, we obtain:

$$\begin{aligned}
(n+1)\, B_{n+1} &= 1 \\
(n+1)\, C_{n+1} &= \lambda\, n\, C_n + nB_n \\
(n+1)\, A_n &= \frac{\varphi^2\sigma^2}{2} + nA_n - nC_n\,(1-\lambda)\log R \\
&\quad + \log \mathcal{E}_t^m \left\{ \exp\left(\varphi\epsilon_{t+1} - n\left(C_n\lambda k + B_n\right)\left(r_{t+1} - \mathcal{E}_t^m r_{t+1}\right)\right)\right\}
\end{aligned}$$

Solving for B_n and C_n, we obtain

$$r_t^n = A_n + \frac{1}{n} r_t + \frac{1}{n}\frac{1-\lambda^{n-1}}{1-\lambda}\mathcal{E}_t^m r_{t+1}$$

The interpretation is straightforward: except for an intercept term that captures a constant risk premium, the long term yield is a simple average of the short term interest rate and future expected short rates: $\frac{1}{n}\left(r_t + \mathcal{E}_t^m r_{t+1} + ... + \mathcal{E}_t^m r_{t+n-1}\right)$. Similarly, we can calculate predictable excess returns at horizon n, $\xi_t^n = \mathcal{E}_t^m e_{t+n} - \mathcal{E}_t e_{t+n}$. Since we are assuming that there is no misperception about the equilibrium exchange rate

$\bar{e}_t$, we know that $lim_{n\to\infty}\xi_t^n = 0$ by construction. Therefore, the Fama coefficient must converge to one. At shorter horizons, substituting repeatedly from (1), we obtain:

$$\xi_t^n = \sum_{j=0}^{\infty} \mathcal{E}_t \mathcal{E}_{t+n}^m x_{t+n+j} - \sum_{j=0}^{\infty} \mathcal{E}_t^m x_{t+n+j}$$

For $n \geq 1$ and $j \geq 1$ we have:

$$\begin{aligned}
\mathcal{E}_t^m x_{t+n+j} &= \lambda^{n+j-1} \mathcal{E}_t^m x_{t+1} \\
\mathcal{E}_t \mathcal{E}_{t+n}^m x_{t+n+j} &= \lambda^{j-1} \mathcal{E}_t \mathcal{E}_{t+n}^m x_{t+n+1} \\
&= \lambda^{j-1} \mathcal{E}_t \left(\lambda (1-k) \mathcal{E}_{t+n-1}^m r_{t+n} + \lambda k r_{t+n} \right) \\
&= \lambda^{j-1} \mathcal{E}_t \left(\lambda (1-k) \mathcal{E}_{t+n-1}^m r_{t+n} + \lambda k r_{t+n} \right) \\
&= \lambda^{j-1} \left(\mathcal{E}_t r_{t+1} \lambda^n k \left[1 + (1-k) + .. + (1-k)^{n-1} \right] + (\lambda (1-k))^n \mathcal{E}_t^m r_{t+1} \right) \\
&= \lambda^{n+j-1} \left(\mathcal{E}_t r_{t+1} \left(1 - (1-k)^n \right) + (1-k)^n \mathcal{E}_t^m r_{t+1} \right).
\end{aligned}$$

Thus,

$$\begin{aligned}
\xi_t^n &= \lambda^{n-1} \mathcal{E}_t x_{t+1} + \left(\mathcal{E}_t r_{t+1} \left(1 - (1-k)^n \right) + (1-k)^n \mathcal{E}_t^m r_{t+1} \right) \sum_{j=1}^{\infty} \lambda^{n+j-1} \\
&\quad - \mathcal{E}_t^m x_{t+1} \sum_{j=0}^{\infty} \lambda^{n+j-1} \\
&= \frac{\lambda^{n-1}}{1-\lambda} \mathcal{E}_t x_{t+1} + (1-k)^n \left(\mathcal{E}_t^m r_{t+1} - \mathcal{E}_t r_{t+1} \right) \frac{\lambda^n}{1-\lambda} - \frac{\lambda^{n-1}}{1-\lambda} \mathcal{E}_t^m x_{t+1} \\
&= \frac{\lambda^{n-1}}{1-\lambda} \left(1 - \lambda (1-k)^n \right) \left(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1} \right)
\end{aligned}$$

The firs term reflects the interest rate effect, while the second term reflects the learning effect. As the horizon increases, the interest rate effect brings predictable excess returns back to zero. Consider now a Fama regression at horizon n :

$$e_{t+n} - e_t = \alpha_n + \beta_n x_t^n + u_{t+n}^n$$

where $x_t^n = \mathbf{l}'\mathbf{r}_t^n = r_t^n - r_t^{*,n}$ is the n-period forward premium. The Fama coefficient at horizon n converges to:

$$p\lim \beta_n = \frac{cov\left(e_{t+n} - e_t, x_t^n\right)}{var\left(x_t^n\right)} = 1 - \frac{cov\left(\xi_t^n, x_t^n\right)}{var\left(x_t^n\right)}$$

Observe that it is not enough to have $\lim_{n\to\infty} \xi_t^n = 0$ to conclude that $lim_{n\to\infty} p\lim \beta_n = 1$ because the variance of the forward premium also converges to 0. In fact, we can write:

$$\begin{aligned}
cov\left(\xi_t^n, x_t^n\right) &= \frac{\lambda^{n-1}}{1-\lambda} \left(1 - \lambda (1-k)^n \right) \; cov\left(\lambda x_t - \mathcal{E}_t^m x_{t+1}, \frac{1}{n} x_t + \frac{1}{n} \frac{1-\lambda^{n-1}}{1-\lambda} \mathcal{E}_t^m x_{t+1} \right) \\
&= \frac{\lambda^n}{1-\lambda} \left(1 - \lambda (1-k)^n \right) \; (1-k) \; cov\left(x_t - \mathcal{E}_{t-1}^m x_t, \frac{1}{n} x_t + \frac{1}{n} \frac{1-\lambda^{n-1}}{1-\lambda} \left(\lambda (1-k) \mathcal{E}_{t-1}^m x_t + \lambda k x_t \right) \right)
\end{aligned}$$

Solving for each term, we have:

$$\begin{aligned} cov\left(x_t, \mathcal{E}_{t-1}^m x_t\right) &= \frac{\lambda^2 k}{1-\lambda^2(1-k)} var\left(x_t\right) \\ var\left(\mathcal{E}_{t-1}^m x_t\right) &= var\left(\lambda(1-k)\mathcal{E}_{t-2}^m x_{t-1} + \lambda k x_{t-1}\right) \\ &= \frac{\lambda^2 k^2}{1-\lambda^2(1-k)^2}\left(\frac{1+\lambda^2(1-k)}{1-\lambda^2(1-k)}\right) var\left(x_t\right) \end{aligned}$$

Lastly

$$var\left(x_t^n\right) = var\left(\frac{1}{n}x_t + \frac{1}{n}\frac{1-\lambda^{n-1}}{1-\lambda}\mathcal{E}_t^m x_{t+1}\right)$$

Substituting, we obtain the expression for β_n plotted in Figure 4.

Notes

[1]See Hodrick (1988) and Lewis (1994) for surveys.

[2]The view that agents constantly engage in learning about the stance and duration of monetary policy seems rather uncontroversial. Suppose, for instance, that monetary authorities control the short term nominal interest rate, as in the US and many other countries. A setup with transitory and persistent interest rate shocks seems appropriate to capture agents' reactions to Federal Open Market Committee (FOMC) meetings (See Batten, Blackwell, Kim, Nocera and Ozeki (1990) for a description of the operating procedures of major central banks). We could think of each meeting as an interest rate innovation. Since investors only have limited information regarding the most current monetary decisions, they must conjecture how persistent the last decision of the FOMC will be. For instance, access to the minutes of a FOMC meeting are only availabe after a six weeks delay, at the following meeting.

[3]Interestingly these results open the possibility that exchange rate determination and term structure are consistent with one another.

[4]The results of Clarida and Gali (1994), Grilli and Roubini (1994) nuance the original results of Eichenbaum and Evans (1995) and indicate that delayed overshooting may not occur for all country pairs.

[5]See appendix A.2 for details.

[6]See Lewis (1994) and Frankel and Rose (1994) for surveys.

[7]See Appendix A.2 for details.

[8]Another class of models with expectational errors is the so-called "Peso problem," whereby if an expected shift in regime does not materialize in sample, expectations will appear systematically biased to the econometrician. Kaminski (1993) shows that Peso problems can account for part of the forward discount premium in a model in which regime switches follow a Markov process.

[9]This finance literature also emphasizes overreaction at longer horizons, as evidenced by long term negative correlation in returns (DeBondt and Thaler (1985)). The over-reaction is not relevant for our interpretation of the puzzles. However, we note that our results also indicate that conditional persistence, as perceived by market participants, is larger than the data indicate. Thus, at longer horizon, forward premia forecasts overreact to current premium.

[10] Two well known examples are Hong and Stein (1999) and Barberis, Shleifer and Vishny (1998). The former assumes that there are two classes of agents, 'newswatcher' who cannot condition on past prices but observe private signals about the fundamentals, and 'momentum' traders only condition on past prices. The latter presents a model of investor sentiment consistent with experimental psychological evidence. In their model, the representative investor alternates optimally between two mental representations of earnings (a mean reverting and a trend one), neither of which corresponds to the true process (a random walk).

[11] Equivalently, the second source of fluctuations arises if the long run equilibrium exchange rate does not follow a martingale: $\mathcal{E}_t \bar{e}_{t+1} \neq \bar{e}_t$.

[12] Formula (5) is derived under the assumption that initially agents believe that $x_1 \sim N(x_1, \sigma_1)$.

[13] The boundary of D_τ is given by: $\lambda(k, \tau) = \frac{1+(1-k)^\tau - \sqrt{\phi}}{2(1-k)^{\tau+1}}$ where $\phi = [1 + (1-k)^\tau]^2 - 4(1-k)^{\tau+1}$.

[14] Similar processes have often been used in the learning literature, starting with Muth (1960) in his exploration of the link between rational and adaptive expectations.

[15] A more realistic interest rate process would incorporate conditional heteroskedasticity to capture the tendency in financial data for volatility clustering -i.e., the tendency for large (small) price changes to be followed by other large (small) price changes of unpredictable sign. ARCH models and their various extensions have been successfully applied to several financial time series (see Bollerslev, Chou and Kroner (1992) for a survey), including interest rates.

[16] Indeed, this representation is equivalent to a restricted ARMA process.

[17] We do not want to push this interpretation too far. One reason is that the empirical literature on money, output and interest rates tries to separate the exogenous and endogenous components of money shocks. Our univariate representation does not allow for this distinction.

[18] In the U.S., investors have access to the minutes of the Federal Open Market Committee meetings with a six weeks delay. The full transcript is only available after 3 years. See Batten et al. (1990) for a description of of the operating procedures of the major Central Banks.

[19] See Hamilton (1994, chapter 13) for further details.

[20] Our focus on eurorates is driven by the availability of forecast data for all G-7 countries.

[21] The 3 months Eurorates come from the IFS tape (lines 60ldd, and 60ea).

[22]This statement understates our empirical results: in most cases, the constraint that $\eta \geq 0$ is binding. In such cases, we estimated directly an AR process.

[23]The use of eurorates may be somewhat problematic, as the forward discount premium may reflect a *country specific* risk premium. The working paper version of this paper presents the results from a similar estimation on Money market rates. The results are virtually unchanged.

[24]The Forecasting services that contribute to the Currency Forecaster's Digest are: Predex, Merril Lynch, Mellon Bank, Harris Trust, Bank of America, Morgan Grenfell, Chase Manhattan, Royal Bank of Canada, Midland Montagu, Generale de Banque, MMS International, Chemical Bank, Union Bank of Switzerland, Multinational Computer Models, Goldman Sachs International, Business International, M. Murenbeeld, and Westpac Bank. The multinational companies that contribute are: General Electric, Du Pont, WR Grace, Allied Signal, Monsanto, Ingersoll-Rand, General-Motors, Data General, Eli Lilly, Aetna, American Express, Johnson and Johnson, Sterling Drug, Firestone, 3M, Union Carbide, Texaco, United Brands, SmithKline Beckman, American National Can, RJ Reynolds, Colgate-Palmolive, Warner-Lambert, Schering-Plough, Quaker Oats, Beatrice Foods, Hercules, Baxter Travenol, and Interpublic Group.

[25]Froot (1989) uses quarterly data on the three months T-bill from 1969 to 1986 from the Nagan Bond and Money Market Letter. This dataset has also been used in Friedman (1980).

[26]See Zarnowitz and Braun (1992).

[27]Prendergast and Stole (1996) present a model where the incentive to scatter or mimic depends on age, with younger forecasters more likely to "exaggerate".

[28]Ehrbeck and Waldmann (1996) conclude that models of stretegic bias are rejected by the data.

[29]While this indicates that the process might not be perceived as stationary, finite horizon forecasts are still well defined.

[30]Our procedure did not converge for the Italian Lira.

[31]This relation pins down the rate of depreciation, not the level of the exchange rate. It is consistent with environments in which the law of one price or purchasing power parity holds, as well as with environments in which it does not.

[32]The common shock still affects the determination of interest rates.

[33]Since the $z's$ are conditionally normally distributed, nothing prevents the nominal interest rates r_t from

being negative in this model, which is a somewhat unpleasant feature. It is possible to fix this problem along the lines of Cox Ingersoll and Ross by assuming a time-varying conditional variance of ϵ_t. However, the learning problem described in the next section becomes history dependent and one does not obtain closed-form solutions. We assume instead that we remain in the vicinity of $\ln R$ so that this problem is irrelevant.

[34]The gain $\tilde{k}$ is constructed from the value of $\tilde{\eta}$ estimated in table 4 for an AR(1) and forecast data at all horizons, and the value of λ from the true process (table 3, AR(1)). Since the true λ are often smaller than the market ones, we are biasing the results against finding delayed overshooting or the Fama puzzle. We exclude the Italy Lira since the market filter did not converge.

[35]If interest rates follow a white noise process ($\lambda = 0$), there are no forecast misperceptions (since the agents are biased in favor of a white noise to start with) and no predictable currency movements.

[36]With the exception of the Italian Lira. See footnote 34.

[37]The ML estimator of η failed to converge for Italy-US, we report estimates using a value of $\eta = 160000$

[38]The exceptions are Italy at all horizons and Japan at 12 months.

[39]The estimation takes into account the maturity longer than the sampling frequency.

[40]The asymptotic properties are the same whether we iterate or not.