Exchange Rate Dynamics, Learning and Misperception*

Pierre-Olivier Gourinchas
Princeton University

Aaron Tornell
UCLA

April 2000

Abstract. This paper proposes a new explanation for the forward premium puzzle and the delayed overshooting puzzles. We demonstrate empirically and theoretically that both puzzles arise from an under-reaction of short term interest rate forecasts. According to our results, we find that (i) the forward premium is almost always a biased predictor of future depreciation; (ii) the bias can be so severe as to lead to negative coefficients in the "Fama regression"; (iii) delayed overshooting occurs for values of the model's parameters that correspond to empirical estimates (iv) cross country restrictions are not rejected empirically; and (v) the empirical restrictions on the time series behavior of predictable excess returns are strongly validated. Interestingly, the under-reaction of interest rates that we document can also explain part of the empirical puzzles for the Expectation Hypothesis of the term structure of interest rates. Thus, the paper provides a unifying treatment of empirical anomalies on bond and currency markets.

* Comments welcome. This paper is a substantially revised version of "Exchange Rate Dynamics and Learnings", NBER WP 5530. We thank Daron Acemoglu, Olivier Blanchard, Ricardo Caballero, John Campbell, Rudi Dornbusch, Ben Friedman, Graciela Kaminski, John Leahy, Paul O'Connell, Julio Rotemberg, Jiang Wang and seminar participants at Chicago University, Harvard, MIT, Tel Aviv, and the NBER for useful discussions and comments on the previous version. All errors are our own. Pierre-Olivier Gourinchas is affiliated with the NBER and CEPR. Aaron Tornell is affiliated with the NBER. e-mail: `pog@princeton.edu` and `atornell@ucla.edu`.

1. Introduction

This paper proposes a new explanation for the forward premium and delayed overshooting puzzles by demonstrating theoretically and empirically that both puzzles arise from an underreaction of short term interest rate forecasts to innovations in current 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 opposit 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 on 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*, whereby the exchange rate should experience an immediate appreciation, and then depreciate gradually towards its new long run equilibrium value.

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 short rates and exchange rates to a common unanticipated monetary innovation. The later is an *unconditional* statement. Empirically, the forward premium puzzle seem much more prevalent, albeit not always in its most extreme form. 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.

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:

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

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

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 $\mathcal{E}_t e_{t+1}$. According (1), the return on a 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, this arbitrage relationship has no empirical power as it stands 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 very badly (Fama (1984)): regressions of the form $e_{t+1} - e_t = \alpha + \beta(r_t - r_t^*) + u_{t+1}$ typically find a β significantly different from 1, and often negative. 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 (2), predictable excess returns result from the sum of exchange rate expectational errors and the currency risk premium. Accordingly, either a time-varying risk premia or expectational errors can rationalize the forward premium puzzle. 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.[2] 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 Saa-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 without ambiguity that (a) 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

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

appreciation/depreciation. We reproduce their analysis in section 3 and reach substantially similar results. The direct conclusion is that expectational errors must be responsible for the bulk 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.[3] Further, since regime shifts generates forecast errors that die out over time, models based on learning about a *one-time* change in regime cannot be expected to deliver a hump-shaped impulse response of exchange rates to monetary shocks, nor negative coefficients in the Fama regression.

These results suggest that a new theory, based on systematic expectational errors, is needed to understand exchange rate determination and rationalize our two puzzles. This paper presents such a theory. There are two key ingredients: first, agents constantly learn about the duration of monetary policy shocks (transitory versus persistent). The second key feature concerns a specific departure from the assumptions of full rationality: we assume that agents systematically under-react to interest rate innovations. Both assumptions are essential to our results, and we justify them in turn.

Our first assumption, that agents constantly engage in learning about the stance and duration of monetary policy, seems rather uncontroversial. Suppose, for instance, that monetary authorities control target the short term nominal interest rate, as in the US and many other countries. A setup with temporary and persistent interest rate shocks seems appropriate to capture investors' reactions to Federal Open Market Committee (FOMC) meetings.[4] We could think of each meeting as an interest innovation. Since investors only have limited information regarding the

[3]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. This class of models does not deliver a hump-shaped impulse response of the exchange rate to a monetary shock. Moreover, using option prices on Dollar-Deutschemarks between 1984 and 1993 to extract jump-expectations associated with shifts in regimes or bursting bubbles, Baily and Kropywiansky (1994)conclude that although there were significant jump-expectations, enough jumps were present in sample. Therefore peso problems are unlikely to have induced a major bias in exchange rate pricing.

[4]See Batten, Blackwell, Kim, Nocera and Ozeki (1990) for a description of the operating procedures of major Central Banks.

most current monetary decisions, they must conjecture how persistent the last decision of the FOMC will be.[5] This is essential in setting long term rates as well as currency prices. In and by itself, this assumption is not enough to rationalize our puzzles: rational agents cannot be systematically fooled and there are no predictable excess returns, delayed overshooting or forward premium puzzles.

The second assumption is much more controversial: we assume that agents under-react to changes in short term interest rates. An equivalent interpretation is that traders overstate the relative variance of transitory versus persistent innovations in short rates. It is a key assumptions of the model. Ultimately, we argue, its relevance must be judged both by its empirical plausibility and the strength of the results it delivers.

To start with, we document empirically the extent to which interest rate forecasts underreact to innovations in current short rates for all G-7 countries. Using a unique survey data set on interest rate expectations, published by *Currency Forecasters' Digest*, with monthly observations from 1986 to 1995 for all G-7 countries, we find substantial evidence of underreaction to interest rate innovations. In other words, we find (a) no evidence of transitory shocks in forward premium; (b) that market participants implicitly assume a substantial share of the innovation 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).

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- are partly predictable based on publicly available information. Of particular interest to us, numerous studies have documented that asset prices *underreact* 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 underreact to announcements of public information (such as earnings, see Bernard (1992), or Chan, Jegadeesh and Lakonishok (1995)).[6] Thus underreaction to publicly available news seem to be a prevalent fact.

Paralleling resarch in International Finance, the consensus in Finance is that the short term underreaction cannot be explained in terms of risk. Instead, recent

[5]For instance, access to the minutes of a FOMC meeting are only availabe after a six weeks delay, at the following meeting.

[6]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.

models in Behavioral Finance attempt to rationalize asset price puzzles in terms of optimal strategies of boundedly rational agents.[7]

Unlike these behavioral models, we will not take a stand on the origins of the underraction. 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.

What we do, is demonstrate that for typical values of the bias, the equilibrium exchange rates in our model exhibit 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 appreciates, up to the point where the expected depreciation compensates 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, traders believe the domestic interest rate will revert to its equilibrium value fairly rapidly. The exchange rate may initally appreciate moderately. The next period, the interest rate is in fact higher than traders expected, leading to an upward revision in belief. Since they stronger beliefs that the interest rate will remain high, demand for the domestic bond increases. This is a learning effect. However, the domestic interest rate is also reverting to its equilibrium value, lowering the demand for the domestic bond. If the learning effect is strong enough, it dominates the interest rate effect, implying a gradual appreciation of the currency. Eventually, there is not much more to learn and the interest effect dominates as the exchange rate revers 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. Further, 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 arise as soon as there is misperception, albeit not in its most extreme form, delayed overshooting depends upon the parameters of the model.

Intuitively then, hump-shaped dynamics result from the interaction of the *mis-*

[7] 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).

perception about the relative importance of shocks (transitory vs. persistent) and the *gradual response of interest rates*. Misperceptions control the speed at which the agent learns the true nature of past interest rate innovations. By contrast, uncertainty regarding the duration of persistent shocks does not generate delayed overshooting in our model.

To sum up, we propose a *unified* model that rationalizes both the forward discount puzzle and the delayed overshooting puzzles, yields empirically testable predictions regarding the strength of delayed overshooting as well as the behavior of predictable excess returns that are consistent with the data. 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; Interestingly these results open the possibility that exchange rate determination and term structure are consistent with one another; (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 accomodate 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 satsfying since delayed overshooting seem much less prevalent and robust than the forward premium puzzle;

The next section presents a summary of the key results from the model in its simples form. Section 3 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 underreaction of short rate forecasrs. Section 4 presents the full fledged model and our results. Section 5 concludes. Most proofs are included in the appendix.

2. A Simple Case.

Suppose that agents are risk neutral so that there is no risk premium and all predictable excess returns arise from expectational errors. Setting $\zeta_t = 0$, and iterating (2) forward, we can express predictable excess returns as a sequence of forward premium $x_t = r_t - r_t^*$ expectational errors:

$$\xi_t = \sum_{j=1}^{\infty} \mathcal{E}_t^m x_{t+j} - \mathcal{E}_t \mathcal{E}_{t+1}^m x_{t+j}$$

An immediate consequence of the forward premium puzzle is that market expectations of interest rate differentials must *differ* from their statistical expectations, at least at some horizon. To make matters more concrete, assume now that the forward premium follows an AR process with autocorelation λ :

$$x_{t+1} = \lambda x_t + \varepsilon_{t+1}$$

If expectations are rational, it is straightforward to show that predictable excess returns are identically equal to zero and that the exchange rate follows:

$$e_t = \bar{e} - \frac{x_t}{1-\lambda}$$

where $\bar{e}$ is the long run value of the exchange rate. In particular, the response of the nominal exchange rate at time $t+j$ to an innovation ε at time t is:

$$e_{t+j} = \bar{e} - \frac{\lambda^j \varepsilon}{1-\lambda}.$$

The exchange rate overshoots its long run level as in the traditional Dornbusch (1976) model. By constrast, suppose that forward premium forecasts are made according to the following adaptive rule:

$$\mathcal{E}_t^m x_{t+1} = (1-k)\,\lambda\,\mathcal{E}_{t-1}^m x_t + \lambda k\, x_t \tag{3}$$

where k measures the weight given to current observations relative to past expectations. Rational expectations corresponds to the special case $k=1$. This expectation mechanism, while extremely simple will be derived in equilibrium in the following section. Note in particular that it implies underreaction to current innovations when $k<1$. Under these assumptions, predictable excess returns occur as long as statistical and market expectations of short term interest rates do not coincide:

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

and the exchange rate follows:

$$e_t = \bar{e} - \frac{x_t}{1-\lambda} + \frac{1}{1-\lambda}\{\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}\} \tag{4}$$

There are positive predictable excess returns when expected forward premium falls short of its statistical expectation ($\mathcal{E}_t^m x_{t+1} < \mathcal{E}_t x_{t+1}$).

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

$$\begin{aligned} e_{t+j} &= \bar{e} - \frac{\lambda^j \varepsilon}{1-\lambda}\left\{1 - \lambda(1-k)^{j+1}\right\} \\ \xi_{t+j} &= \frac{\varepsilon}{1-\lambda}\lambda^{j+1}(1-k)^{j+1} \geq 0 \end{aligned} \tag{5}$$

The term in λ^j in (5) reflects the interest rate effect: the speed at which the initial disturbance fades away. The term in $(1-k)^{j+1}$ reflects the misperception. A lower k increases the degree of misperception.

From this, we see that (a) there are positive predictable excess returns when $k < 1$, and (b) delayed overshooting at horizon τ -that is $|e_{t+\tau} - \bar{e}| > |e_{t+\tau-1} - \bar{e}|$ – occurs if and only if:

$$\tau < \ln \frac{(1-\lambda)/\lambda}{[1-\lambda(1-k)]} / \ln(1-k). \tag{6}$$

Figure 1 shows the path of the exchange rate in response to an unanticipated decrease in the interest rate at $t = 0$ in both the perfect and imperfect information cases. 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 this result. 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.

- **Learning effect.** When the shock takes place at time 0 agents only observe an increase in x_0 and gradually lower their belief about x_{t+1} using updating equation (3). As $\mathcal{E}_t^m x_{t+1}$ is revised downwards the demand schedule for domestic bonds shifts downwards over time, generating depreciating pressures on the exchange rate.

According to (6), 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. As convergence occurs faster, 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 at the time of the shock beliefs almost converge to the true value of the persistent component of the interest rate. 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 down 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 $\mathcal{E}_t^m x_{t+1}$ increases very little at the time of the shock. Afterwards, although $\mathcal{E}_t^m x_{t+1}$ is updated upwards, the learning effect is too small to dominate the interest rate effect.

Condition (6) defines a 'delayed overshooting' region in the parameter space at horizon τ, D_τ. Figure 2 reports the lower boundary of delayed overshooting regions D_1, D_5 and D_{10} (from the bottom to the top).[8]

We now turn to the issue of the length of time over which the exchange rate moves in the "wrong" direction. 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 on Figure 2.

Thus, our analysis has strong cross-sectional implications. Countries should exhibit conditional delayed overshooting if (a) monetary shocks have high conditional persistence, 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. Whether delayed overshooting occurs at a given horizon depends both upon the learning process (k) and the interest rate process (λ). Further, there are positive predictable excess returns on the domestic currency ($\xi_{t+j} \geq 0$) , even if there is no delayed overshooting.

Lastly, under this simple adaptive expectation scheme, the probability limit of β in the Fama regression mentioned above is always less than 1 and may even be negative:

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

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

To sum up, this very simple 'model' is consistent with the forward discount puzzle -even in its most extreme form with negative Fama coefficient- and with potential delayed overshooting, depending on the parameters that control market expectations (k) and the persistence of interest rate shocks (λ). of the market process and the interest rate.

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

3. An Empirical Exploration of Interest Rate Forecasts

The previous section highlighted the key insights from the model: delayed overshooting and forward premium puzzle can arise in equilibrium from systematic underreaction in interest rates expectations. In this section, we demonstrate that this assumption accords very well with the data.

We start by reviewing the evidence in favor of the expectational error assumption. To do so, we replicate the results from Frankel and Froot (1989). Our results will confirm that the currency risk premium ζ_t is small and relatively constant. In particular, this implies that the forward premium x_t is a good measure of the depreciation expected by the market $\mathcal{E}_t^m e_{t+1} - e_t$, as measured bu survey data. Turning to interest rates, we then show that interest rate survey data may differ substantially from the forward premium implicit in the term structure, especially at short horizon. This validates our use of market based survey data, instead of forward rates, since the latter contain a non negligible time varying risk premium. Once these preliminaries are established, we show that interest rate forecasts systematically underreact to interest rate innovations. Empirically, we demonstrate this result using both a relatively unconstrained ARMA process, as well as a State-Space model that provides the foundation for (3).

3.1. Currency Risk Premium versus Expectational Errors. The standard Fama regression takes the following form:

$$\Delta e_t^k = e_{t+k} - e_t = \alpha + \beta\, x_t^k + \eta_{t+k}^k$$

where $x_t^k = r_t^k - r_t^{*k}$ is the current $k-period$ forward discount, equal to the $k-period$ interest rate differential. Under the null that (a) expectations are rational and (b) the currency risk premium is uncorrelated with x_t^k, $\beta = 1$. We run this regression in table 1. The exchange rate data is from DRI-IFS while the 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.[9] As the horizon increases, the coefficients increase, often becoming not significantly different from 1.

For the sub-period 5/84 to 5/97, we have exchange rate survey forecasts, compiled by the Financial Times Currency Forecaster. Following Frankel and Froot (1989), we can decompose the Fama coefficient for this period between its expectational and risk premium component. Frankel and Froot show that:

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

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

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 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 are very strong: at short horizons (3 to 6 months), for all countries except Italy, b_{re} dominates 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.

3.2. Interest Rate Survey Forecasts and the Term Structure. According to (4), exchange rates depend upon forecasts of future forward premium. Moreover, for this equation to hold, the forward premium forecast $\mathcal{E}_t^m x_{t+1}$ must obey (3). To document this relationship, we need data on expected future short rates. One possibility would be to extract forward rates from the term structure of interest rates. According to the log-linearized version of the theory (see Shiller, Campbell and Schoenholtz (1983)), the forward premium on the $k-period$ ahead one period bond is equal to:[10]

$$fp_t^k = (k+1)\left(r_t^{k+1} - r_t^k\right)$$

and the forward rate is simply:

$$f_t^k = fp_t^k + r_t^k = (k+1)\, r_t^{k+1} - k\, r_t^k$$

Under the expectation hypothesis of the term structure, this forward rate is an unbiased predictor of the future short rate r_{t+k}. As with the uncovered interest rate parity, this assumption may be violated if expectations are systematically biased and/or if there are time varying risk premia on the term structure. While the expectation hypothesis is generally widely rejected, it is possible that forward rates represent market expecations of future short rates. Froot (1989) adopts an approach similar to Frankel and Froot (1989) to investigate this question. Table 6 reports the decomposition of the coefficient from a standard test of the expectation hypothesis: $r_{t+k} - r_t^k = \alpha + \tilde{\beta}\, fp_t^k + \eta_{t+k}^k$ into its expectational error $\tilde{b}_{re}$ and risk premium components $\tilde{b}_{rp}$. To measure the risk premium and the expectational error, we use data from the Financial Times Currency Forecaster on Eurorates forecasts at 3, 6 and 12 month horizon. The data is available monthly from 08/86 to 10/95. Contributors include multinational companies as well as forecasting services from major investment banks, i.e. the most active player on the fixed income and foreign exchange markets.[11] The monthly publication collects interest rates and their forecasts and

[10] Since the Euro bonds have no coupon, the duration is equal to the maturity.

[11] 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,

reports a "market average" weighting individual respondents according to their relative importance. This dataset 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.[12] Unlike the exchange rate decomposition, the results indicate that at short horizons the expectational error term is often dominated by the risk premium term. In fact, the expectational error term is often negative. The direct consequence is that forward rates are not equivalent to survey forecast.

3.3. Modelling the Interest Rate Differential. To characterize the interest rate process, we adopt a state-space representation. The interest rate differential between any two countries, x_t is the sum of a persistent (x_t^p) and a transitory (v_t) components: [13]

$$x_t = \mu + x_t^p + \nu_t \tag{7}$$

μ represents a possible constant. In addition, we assume that the persistent component follows an $AR(q)$ process:

$$\lambda(L)\ x_t^p = \epsilon_t \tag{8}$$

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, define the noise to signal ratio $\eta = \frac{\sigma_\nu^2}{\sigma^2}$.

A more realistic interest rate process would incorporate conditional heteroskedasticity. Conditional heteroscedasticity would 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. In their study of varying risk premia in the term structure, Engle, Lilien and Robins (1987)find strong ARCH effects on the excess holding yield of six-months over three-months T-bills, using quarterly data from 1953:1 to 1971:7. Grier and

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.

[12] 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).

[13] 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.

Perry (1993)look at quarterly interest rate surprises, measured as the difference between one-month T-bill rate and the three month forward for that period. They also find ARCH effects for the sample 1960:III to 1991:IV. This empirical evidence suggests that adding ARCH components would introduce an important element of realism. However, as demonstrated in our previous working paper, adding ARCH elements does not alter the nature of our results. In order to make the presentation of the results more transparent, we decided to report results abstracting from conditional heteroscedasticity.

One possible justification for our interest rate representation lies in its flexibility: depending on the underlying parameters, this representation can accommodate an integrated process -when some of the roots of λ lie on the unit circle- as well as a white noise.[14]

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.[15]

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.[16] 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.

3.3.1. Maximum Likelihood Estimation of the Interest Rate Process. The system (7) -(8) can be estimated by Maximum Likelihood of the associated Kalman Filter. The procedure is standard and summarized in the appendix.[17] Our dataset consists of monthly observations of the 3 months eurorates for Canada,

[14]Indeed, this representation is equivalent to a restricted ARMA process. See our previous working paper version for details.

[15]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.

[16]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. Romer and Romer (1996)argue that the Fed has an informational advantage over the market.

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

France, Germany, Italy, Japan and the U.K. evaluated against the 3 months eurodollar.[18] The sample period is 1974:1 to 1995-12.[19] The results are presented in Table 4, 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. First, the long run autocorrelation ranges from 0.85 for France versus the U.S. to 0.99 for Italy against the U.S. The short run autocorrelation is higher than 1 for the U.K., Germany and Japan against the U.S., indicating further deviations from equilibrium after the initial shock. The table also report the results of a Phillips-Perron Z_t test of a unit root in interest rate differentials. Indeed, we cannot reject the hypothesis of a unit root at conventional levels of significance for Canada, Germany Italy and Japan.[20] Thus, interest rate for those countries and the U.S. rate do not appear to be co-integrated. Second, for all countries and all specifications, the noise to signal ratio is not statistically significant.[21] In only two cases (US-Italy I and US-Canada IV) 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.[22] Further, comparing with table 1, we observe that these correspond to the cases where the Fama coefficient is large and positive.

The use of eurorates may be somewhat problematic, as the forward discount premium may reflect a *country specific* risk premium. Ideally, one would want to estimate (7)-(8) on money market rates. Comparing the money market rate and the eurorate in Figure 6, we see that the two series are virtually identical for most countries in our sample.[23] The main difference, for instance for France between 1981 and 1986, comes from capital controls that decoupled the eurofranc from domestic rates and allowed the French government to ease monetary policy while simultaneously maintaining the Franc within its ERM bands. Clearly, in such cases, the eurorate is indeed the appropriate interest rate one ought to use in order to price exchange rates. However, in order to verify that our results, we estimated (7)-(8) directly on money market rates. The results are reported in Table 5.

It is immediate from this table that the results are virtually unchanged, whether one uses eurorates or money market rates. The long run autocorrelation is very high, and there is no evidence of transitory component, either. The results for France, where the disparities between the two series are strongest, indicate that the long

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

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

[20] It is well know that interest rates may be integrated.

[21] 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.

[22] The long run autocorrelation raises to 0.97 for U.S.-Canada and to 0.99 for U.S.-Italy.

[23] The money market rate is taken from the IFS tape, line l60b.

run autocorrelation is probably larger for the domestic rates than for the offshore rates. Furthermore, we also find that Germany, Italy and Japan rates versus the US federal funds rate might be integrated, as was already the case for eurorates.[24] These results demonstrate that country premium do not affect interest rate differentials sufficiently.

3.3.2. Survey Data. One should be cautious when using survey data. First, there is probably no such thing as a "market expectation". Heterogeneity in forecasts has been well documented in the literature.[25] 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 (Zarnowitz and Braun (1992)). This is not guaranteed, however. 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.[26] The empirical importance and direction of such reputational biases remains an open question.[27]

We assume that agents use linear forecasting formulae, as summarized by the Kalman Filter equations associated with the state-space representation 7-8. It is unlikely that individual forecasters know the exact parameters driving the interest rate differential process. According to one interpretation, these parameters may be time-varying and agents may be in the permanent process of revising their estimates. More generically, we adopt an agnostic view and will estimate the parameters of the filter implicitly used, which we denote the "market filter".[28] Denote $\tilde{\theta} = \left(\left\{ \tilde{\lambda}_i \right\}_{i=1}^{p}, \tilde{\eta}, \tilde{\sigma}^2 \right)$ the parameters of the market filter. For a given market filter we can construct the associated forecast at horizon τ: $x_t^{\tau}\left(\tilde{\theta}\right)$. The forecast con-

[24]The difference for Canada may come from a different sample coverage. The eurorates are only available since 1987.

[25]See Zarnowitz and Braun (1992).

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

[27]Ehrbeck and Waldmann (1996)conclude that models of stretegic bias are rejected by the data, while Lamont (1996), looking at the age dispersion of forecasts, finds that such biases might be important.

[28]The market filter refers to the parameters of the filter used implicitly by market participants, not to the actual data-generating parameters in sample.

structed in such a way uses only information up to time t. We assume that the true forecasts are reported with error: $\hat{x}_t^\tau = x_t^\tau\left(\tilde{\theta}\right) + \upsilon_t^\tau$ where the measurement error is assumed orthogonal, and estimate $\tilde{\theta}$ by Maximum Likelihood.[29] The results for the 3 months eurorates are in Table 6.

Figure 7 shows the actual forecasts and the fitted values for the Canadian US Euro rate differential, together with the forecasts generated according to our estimated data generating process. This figure indicates that our estimated market filter accurately reproduces the dynamics of the market forecasts at all horizons. Actual forecasts exhibit more volatility in the short run, but medium run fluctuations are closely mimicked by our estimates. The comparison between the survey and the fitted forecasts constructed according to the estimated data-generating-process parameters is instructive. At 3 months, there is very little difference between the two series. In other words, the systematic error present in the market filter that will be documented shortly does not affect the forecasts at short horizons in any major way. However, at longer horizons (6 and especially 12 months) the difference between the two series is dramatic: while we do observe significant departures from long run equilibrium in the survey, the data generating process based forecasts indicates that the shocks should have disappeared. Thus, *conditionally on being persistent* (i.e. lasting more than one period) market participants expect the shocks to exhibit *more* persistence than the data suggests. Of course, this *over*-reaction only appears as the horizon lengthens.

Our methodology allows us to distinguish between the *conditional persistence* ($\sum_j \lambda_j$ in terms of our representation), and the *relative importance* of transitory and permanent shocks ($\eta = \frac{\sigma_\nu^2}{\sigma^2},$). As can be seen from the tables, 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 sample, the market estimates a much higher long run autocorrelation sometimes very close to 1.[30] This extra persistence is compensated by a large estimate for the relative variance η. The results in the previous tables indicate that for almost all countries and specifications, the noise-to-signal ratio η is large -and often significant. A substantial share of interest rate changes is allocated to the transitory component.

This stands in sharp contrast with our finding that in the actual data the transitory component is nill: market forecasts initially *under*-react to interest rate changes. Without the excess transitoryness, interest rate shocks would be expected to die out at a much slower rate than observed in the data. However, the presence of the transitory component introduces a dampening effect: a share of the shock disappears extremely rapidly, while the rest decays very slowly.

[29] We use simultaneously the forecasts at all horizons in the estimation process. We report robust standard errors since the horizon is larger than the sampling frequency.

[30] While this indicates that the process might not be perceived as second order stationary, finite sample forecasts are still well defined.

For future reference we will summarize the results of this section.

Stylized Fact In 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 expected 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 actual interest rate differentials is significantly smaller than that of the filter that best replicates interest rate expectations. This is true across countries, and for different interest rates measures. This indicates an 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.

4. A Stylized Model of Exchange Rate Determination

This section presents a model of exchange rate determination similar to Lucas (1982) seminal model, incorporating imperfect information. The world economy consists of two countries and one consumption good per country. Population in each country is constant and equal to 1. This is an endowment economy: households in the domestic (foreign) country receive each period a quantity Y_t (Y_t^*) of the domestic (foreign) good where the driving process for (Y_t, Y_t^*) is exogenous. Agents in each country have identical preferences and maximize at time t :

$$\mathcal{E}_t^m \left\{ \sum_{s=t}^{\infty} \beta^{(s-t)} u\left(c_s, c_s^*\right) \right\}$$

where $0 < \beta < 1$ is an exogenous discount factor and $u\left(c_s, c_s^*\right)$ is the flow utility over the consumption bundle (c_s, c_s^*) of domestic and foreign goods. As in Lucas (1982), we assume markets are complete and all idiosyncratic risks are initally pooled through the appropriate trade in claims on domestic and foreign endowments. The representative agent of the domestic country consumes half the total endowment $\left(\frac{1}{2}Y_t, \frac{1}{2}Y_t^*\right)$ and holds claims to half the world equity.[31] Unlike Lucas, however, we allow for $\mathcal{E}_t^m \left\{x_{t+1}\right\}$, the market expectation, as of time t, of stochastic variable x_{t+1}, to differ from the rational expectation $\mathcal{E}_t \left\{x_{t+1}\right\}$. Both preferences and expectations are common across countries.

[31] We show in the appendix that risks remain fully pooled in equilibrium.

On the monetary side, the market structure is as follows. In each country, the monetary authority issues currency which must be used to buy locally produced goods. This cash-in-advance constraint imposes the following restrictions:

$$p_t Y_t \leq M_t; \quad p_t^* Y_t^* \leq M_t^*$$

where p_t denotes the domestic currency price of domestic output, and M_t the quantity of domestic currency in circulation at time t. Similar notation applies to the foreign country, with p_t^* denoting the foreign currency price of foreign output.

As long as the nominal one-period risk free interest rate is strictly positive, the cash-in-advance constraint will bind. We restrict ourselves to such equilibria. This implies a unitary velocity in the demand for money.

Since this is a well-known model of currency pricing, we relegate the derivation of the results to the appendix and state the main results here.

With complete markets, there is a unique pricing kernel for financial assets between time t and $t+1$, expressed in terms of the domestic good: $\tilde{R}_{t,t+1} = \frac{\beta u_1\left(\frac{1}{2}Y_{t+1}, \frac{1}{2}Y_{t+1}^*\right)}{u_1\left(\frac{1}{2}Y_t, \frac{1}{2}Y_t^*\right)}$ where u_i denotes the partial derivative of u with respect to its ith argument and where we have substituted the equilibrium consumption profiles. $\tilde{R}_{t,t+1}$ satisfies:

$$s_t = \mathcal{E}_t^m \left\{ \tilde{R}_{t,t+1} \left(s_{t+1} + d_{t+1} \right) \right\}$$

where s_t denotes the time-t price of the asset, in terms of the domestic good and d_{t+1} denotes the dividend paid at time $t+1$. The pricing kernel expressed in terms of the domestic currency can be obtained by substituting the nominal value of the payoffs: denoting S_t and D_{t+1} the nominal price and dividend, in domestic currency, a claim to $(S_{t+1} + D_{t+1})$ units of domestic currency next period is equivalent to a claim to $\frac{(S_{t+1}+D_{t+1})}{p_{t+1}}$ units of the domestic good next period. The pricing kernel for nominal claims is therefore: $R_{t,t+1} = \frac{\beta u_1\left(\frac{1}{2}Y_{t+1}, \frac{1}{2}Y_{t+1}^*\right)}{u_1\left(\frac{1}{2}Y_t, \frac{1}{2}Y_t^*\right)} \frac{M_t}{M_{t+1}} \frac{Y_{t+1}}{Y_t}$.

The nominal pricing kernel depends upon the intertemporal marginal rate of substitution for the domestic good, the growth rate of money and of domestic output. Following similar steps, we can define the nominal pricing kernel in terms of the foreign currency: $R_{t,t+1}^* = \frac{\beta u_2\left(\frac{1}{2}Y_{t+1}, \frac{1}{2}Y_{t+1}^*\right)}{u_2\left(\frac{1}{2}Y_t, \frac{1}{2}Y_t^*\right)} \frac{M_t^*}{M_{t+1}^*} \frac{Y_{t+1}^*}{Y_t^*}$.

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

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

To derive the rate of currency depreciation from the model, denote 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)). We can contruct a pricing kernel for nominal claims in the domestic currency as $R^*_{t,t+1}\frac{E_t}{E_{t+1}}$. This implies:

$$R_{t,t+1}\frac{E_{t+1}}{E_t} = R^*_{t,t+1}$$

Note that this relationship holds in every state of nature and not simply in expectations.[32] Expected depreciation is:

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

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

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

The time-varying risk premium ζ_t reflects the difference between conditional means of the pricing kernels.

4.1. An affine interpretation. In this subsection, we restrict further the model to an affine model, similar to (Backus et al. (1998)). This allows us to concentrate directly on the dynamics of the nominal short term interest rates.

Since the focus is on monetary shocks, let's assume that output is constant, equal to its steady state value $\left(\bar{Y}, \bar{Y}^*\right)$.[33] Exchange rate and interest rates are determined by money growth. Assume further that money growth is driven by three factors: a common factor $\bar{z}_t$, as well as the country specific factors z_t and z^*_t. Specifically, we postulate the following money growth process:

$$\ln \frac{M_{t+1}}{M_t} = \ln\beta - \ln R_{t,t+1} = \frac{\bar{\varphi}^2\sigma^2}{2} + \delta\,\bar{z}_t + \frac{\varphi^2\sigma^2}{2} + z_t + \bar{\varphi}\,\bar{\epsilon}_{t+1} + \varphi\epsilon_{t+1} \quad (9)$$

$$\ln \frac{M^*_{t+1}}{M^*_t} = \ln\beta - \ln R^*_{t,t+1} = \frac{\bar{\varphi}^{*2}\sigma^2}{2} + \delta^*\,\bar{z}_t + \frac{\varphi^{*2}\sigma^{*2}}{2} + z^*_t + \bar{\varphi}^*\,\bar{\epsilon}_{t+1} + \varphi^*\epsilon^*_{t+1} \quad (10)$$

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. Further, assume that the state $\mathbf{z}_t = (\bar{z}_t, z_t, z^*_t)'$ obeys the following AR process:

$$\mathbf{z}_{t+1} = \lambda\,\mathbf{z}_t + \boldsymbol{\epsilon}_{t+1} \quad (11)$$

[32] This relation pins down the rate of depreciation, not the level of the exchange rate. If the law of one price holds, the exchange rate level is such that: $E_t = \frac{u_2\left(\frac{1}{2}Y_t, \frac{1}{2}Y^*_t\right)}{u_1\left(\frac{1}{2}Y_t, \frac{1}{2}Y^*_t\right)}\frac{M_t}{Y_t}\frac{Y^*_t}{M^*_t}$. We do not need to make this assumption in what follows.

[33] This assumption is mostly made for simplicity. The model could accomodate output growth as one of the factors.

One possible justification for the money growth process in (9)-(10) is in terms of a monetary policy reaction function: money growth in each country reacts to the common components $\bar{z}_t$ and $\bar{\epsilon}_{t+1}$, as well as the country specific components z_t and ϵ_{t+1}, with loadings reflecting the sensitivity to the various shocks. Together, $\mathbf{z}_t$ and $\boldsymbol{\epsilon}_{t+1}$ span the predictable and unpredictable components in domestic and foreign bond prices and currency movements. As we will see shortly, the assumptions on the constant of the process have been made mostly for simplicity (see Backus et al. (1998)). The coefficients φ determine the correlation between each state variable's innovation and the current money growth rates.

Under (9)-(10), one obtains:

$$\begin{aligned} \mathbf{r}_t &= -\ln\beta + \mathbf{H}'\mathbf{z}_t \qquad (12) \\ \mathcal{E}_t^m\{e_{t+1}\} - e_t &= (\delta - \delta^*)\,\bar{z}_t + z_t - z_t^* + \frac{\bar{\varphi}^2 - \bar{\varphi}^{*2} + \varphi^2 - \varphi^{*2}}{2}\sigma^2 \\ &= r_t - r_t^* + \zeta \qquad (13) \end{aligned}$$

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

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 world factor does not enter expected depreciation since it shifts domestic and foreign interest rates by similar amounts. We make this assumption in what follows.[34] The common shock still affects the determination of interest rates.

The risk premium ζ is constant and depends upon the variance of the innovations and the loadings. 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.

Since the $z's$ are conditionally normally distributed, nothing prevents the nominal interest rates from being negative in this model, which would violate the equilibrium condition on the cash-in-advance. It is possible to fix this problem along the lines of Cox Ingersoll and Ross by assuming a time-varying conditional variance. One problem with this is that the learning problem becomes history dependent and one does not obtain closed-form solutions. For the time-being, we abstract from this problem, but will come back to it later. By ruling out risk premia altogether, this assumption greatly simplifies our analysis.

From the previous assumptions, it follows that $\mathbf{r}_t$ follows an AR(1) process:

$$\mathbf{r}_{t+1} = -(1-\lambda)\ln\beta + \lambda\,\mathbf{r}_t + \mathbf{H}'\boldsymbol{\epsilon}_{t+1} \qquad (14)$$

[34] This assumption simplifies the derivation by ensuring that the gain is the same for each currency.

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 growth rate of money such that the pricing kernel is identical to Vasicek's original pricing kernel. The simplicity of the pricing kernel allows to price simultaneously long bonds as well as currencies.

This literature attempts to replicate the term structure and the forward premium puzzle by postulating appropriate processes for the pricing kernel, in terms of underlying unobservable factors. Since expectations are rational, this is equivalent to deriving properties for the term structure and currency time-varying risk premia that match the data. This paper adopts a different strategy: rather than allowing for complex kernel processes and associated risk-premia, we depart, in a specific sense, from the assumption of full rationality. Hence, we can concentrate on a simpler process that assumes away risk premia, and concentrate on predictable excess returns, whether on long term bonds or currency markets, arising from expectational errors.

4.2. Market Expectations and the Learning Problem. 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. We use a convenient state-space representation to characterize the perceived process for the interest rate process, that is consistent with our empirical specification. In the model presented above, short rates are linear combinations of the underlying persistent unobservable factors.

Assume that households do not perceive (12) but instead perceive:

$$\mathbf{r}_t = -\ln\beta + \mathbf{H}'\mathbf{z}_t + \mathbf{v}_t \tag{15}$$

together with (11), where the elements of $\mathbf{v}_t$ are independent transitory shocks normally distributed with variance σ_v^2, independent from the innovation $\boldsymbol{\epsilon}_t$. Households in this model erroneously perceive a purely transitory component in the determination of future interest rates. In all other respects, expectations are rational. In particular, rational expectations obtain as the special case where $\sigma_v^2 = 0$. $\mathcal{E}_t^m\{.\}$ means precisely that expectations are taken with respect to the household model (15) instead of the true model (12).

As we will see, this assumption implies very naturally that interest rate forecasts under-react to innovations in interest rates, as indicated by the interest rate survey data. The question of interest to us is the extent to which this mispreception affects exchange rate determination in equilibrium.

A case of special interest is when $\delta = 0$. It is immediate that interest rates depend only on the country specific factor and are independent from one another. In the more general case, observations about foreign interest rates convey information about $\bar{z}_t$, the common factor, that needs to be taken into account.

4.3. The Learning Problem. Households form optimal forecasts of future interest rates given their beliefs about the process, summarized in (11)-(15). This is a standard normal-linear filtering problem (references) that can be analyzed with as a Kalman filter (see the appendix). According to (15), households form forecasts of future interest rates according to:

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

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

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:*

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

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:

$$\mathcal{E}_t^m \mathbf{z}_{t+1} = \lambda \mathcal{E}_{t-1}^m \mathbf{z}_t + \lambda \mathbf{P}\mathbf{H}\left(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^2 \mathbf{I}\right)^{-1}\left(\mathbf{r}_t + \ln\beta - \mathbf{H}'\mathcal{E}_{t-1}^m \mathbf{z}_t\right) \qquad (16)$$

In what follows, we assume that the process has been going on for long enough that we have reached the steady state.[35]

[35] By contrast, Albuquerque (1998) focuses on the implications of a similar model in terms of transitional dynamics to its steady state.

- case 1: rational expectations:

  When expectations are rational, it is easy to check that (16) collapses to: $\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda \mathbf{r}_t$, as expected.

- case 2: no common shock:

  When $\delta = 0$, (16) is equivalent to:

$$\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda (1-k) \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda k \, \mathbf{r}_t - (1-\lambda) \ln \beta \tag{17}$$

  where k, the gain of the filter, measures how much weight is given to new observations, relative to past expectations. (17) make 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 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, aka 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 $\eta-$ 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, pre-multiplying (16) by $\mathbf{H}'$, we obtain a similar formula:

$$\mathcal{E}_t^m \mathbf{r}_{t+1} = \lambda (\mathbf{I}-\mathbf{K}) \mathcal{E}_{t-1}^m \mathbf{r}_t + \lambda \mathbf{K} \, \mathbf{r}_t - (1-\lambda) \ln \beta \tag{18}$$

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.

4.4. Equilibrium Exchange Rate and Long Interest Rates. To solve for the equilibrium exchange rate and long interest rates, we conjecture linear solutions of the following form:

$$\begin{aligned} e_t &= \bar{e} + \mathbf{b}'\mathbf{r}_t + \mathbf{c}'\mathcal{E}_t^m \mathbf{z}_{t+1} \\ \mathbf{r}_t^n &= \bar{\mathbf{A}}_n + \mathbf{B}_n' \mathcal{E}_t^m \mathbf{r}_{t+1}; \quad n \geq 2 \end{aligned} \tag{19}$$

where $\mathbf{r}_t^n$ is the vector of domestic and foreign continuously compounded interest rates on n-period zero coupon bonds. $\mathbf{b}$ and $\mathbf{c}$ are 2 x 1 and 3 x 1 vectors respectively, while $\mathbf{B}_n$ is a 3 x 2 matrix, to be determined.

Using the conjectured exchange rate function into (13), we obtain:

$$\begin{aligned} \mathcal{E}_t^m e_{t+1} - e_t &= \left[\mathbf{b}'\mathbf{H}' + \lambda \mathbf{c}' - \mathbf{c}'\right] \mathcal{E}_t^m \mathbf{z}_{t+1} - \mathbf{b}'\mathbf{r}_t \\ &= \mathbf{l}'\mathbf{r}_t \end{aligned}$$

where $\mathbf{l}' = (1, -1)$ is such that $\mathbf{l}'\mathbf{r}_t = r_t - r_t^*$. Identifying the coefficients term by term, we obtain:

$$\begin{aligned} e_t &= \bar{e} - \mathbf{l}'\mathbf{r}_t - \frac{1}{1-\lambda}\mathbf{l}'\mathbf{H}'\mathcal{E}_t^m \mathbf{z}_{t+1} \\ &= \bar{e} - \mathbf{l}'\mathbf{r}_t - \frac{1}{1-\lambda}\mathbf{l}'\mathcal{E}_t^m \mathbf{r}_{t+1} \end{aligned} \tag{20}$$

Thus, exchange rate determination depends only upon the current interest rate differential $\mathbf{l}'\mathbf{r}_t$ and its future expected value $\mathbf{l}'\mathcal{E}_t^m \mathbf{r}_{t+1}$. A high current or expected domestic short rate, relative to the foreign short rate, leads to a current appreciation of the currency. The steady state level of the exchange rate $\bar{e}$ is left undetermined. This is irrelevant for our purpose.[36]

Under rational expectations, the exchange rate follows:

$$e_t^r = \bar{e} - \frac{1}{1-\lambda}\mathbf{l}'\mathbf{r}_t \tag{21}$$

Substracting from (20), we can express the equilibrium exchange rate as the component the rational expectation exchange rate plus a term that reflects misperception of the interest rate process:

$$= e_t^r + \frac{1}{1-\lambda}\mathbf{l}'\left(\mathcal{E}_t \mathbf{r}_{t+1} - \mathcal{E}_t^m \mathbf{r}_{t+1}\right) \tag{22}$$

Since $\mathcal{E}_t^m \mathbf{r}_{t+1}$ under-reacts to $\mathbf{r}_t$, an early result is that exchange rates are *less* volatile, in response to interest rate changes, when there is misperception than under rational expectations.

We solve for bond rates recursively. The arbitrage pricing condition for long bonds is:

$$-(n+1)\,\mathbf{r}_t^{n+1} = \log \mathcal{E}_t^m \left\{\mathbf{R}_{t,t+1} \exp\left(-n\mathbf{r}_{t+1}^n\right)\right\} \tag{23}$$

[36]Note that if the law of one price holds, the exchange rate level is determined by: $\frac{p}{Ep^*} = \frac{M}{Y}\frac{Y^*}{M^*}\frac{1}{E} = \frac{u_1\left(\frac{1}{2}Y, \frac{1}{2}Y^*\right)}{u_2\left(\frac{1}{2}Y, \frac{1}{2}Y^*\right)}$.

Suppose that (19) holds for horizon n. The pricing condition (23) applied to an $n+1$ period zero coupon bond yields:

$$-\left(n+1\right)\,\mathbf{r}_t^{n+1}=\log\mathcal{E}_t^m\left\{\exp\left(\ln\mathbf{R}_{t,t+1}-n\mathbf{r}_{t+1}^n\right)\right\}$$

Substituting using (9), (11) and (15) we obtain:[37]

$$\begin{aligned}\left(n+1\right)\mathbf{B}_{n+1} &= \lambda^{-1}\mathbf{I}+\lambda n\mathbf{B}_n\\ \mathbf{B}_n &= \frac{1}{n}\lambda^{-1}\frac{1-\lambda^n}{1-\lambda}\mathbf{I};\quad \text{for } n\geq 2\end{aligned}$$

and we can write:

$$\mathbf{r}_t^n=\bar{\mathbf{A}}_n+\frac{1}{n}\frac{1-\lambda^n}{1-\lambda}\lambda^{-1}\mathcal{E}_t^m\mathbf{r}_{t+1}$$

Long interest rates depend only upon their *own* expected future short rates. This is so since expected future short rates already incorporates the best forecasts about $\mathbf{z}_{t+1}$.

Under rational expectations, we have:

$$\mathbf{r}_t^{rn}=\bar{\mathbf{A}}_n^r+\frac{1}{n}\frac{1-\lambda^n}{1-\lambda}\mathbf{r}_t$$

where the intercept $\bar{\mathbf{r}}_n^r$ may be different. Combining the two equation, we obtain:

$$\mathbf{r}_t^n=\mathbf{r}_t^{rn}+\frac{1}{n}\frac{1-\lambda^n}{1-\lambda}\lambda^{-1}\left(\mathcal{E}_t^m\mathbf{r}_{t+1}-\mathcal{E}_t\mathbf{r}_{t+1}\right)+\left(\bar{\mathbf{A}}_n-\bar{\mathbf{A}}_n^r\right)$$

Long interest rates differ from their rational expectations counterpart by short rates forecast errors and a term-dependent constant. The latter only matters for the level of the interest rates, not their changes. In particular, taking first differences, since interest rates tend to under-react to innovations in our model, the forecast error error term will on average be negative, so that long rates will fall short of their rational expectation equivalent.

4.5. Predictable Excess Returns Delayed Overshooting and the Fama puzzle.

[37] The solution for $\bar{\mathbf{r}}_{n+1}$ is obtained similarly. Since the expression is ugly and does not really matter, I omit it here. Note simply that $\bar{\mathbf{r}}_{n+1}$ depends upon the conditional variance of the innovations and thus depends upon expectations.

4.5.1. Predictable excess returns. In this subsection, we demonstrate that our simple model is enough to rationalize both Eichenbaum and Evans delayed overshooting puzzle as well as Fama's forward premium puzzle, for some configurations of the underlying parameters of the model. We start by deriving an expression for predictable excess returns. Recall that predictable excess return on the domestic current are defined as:

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

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 (20) and (18):

$$\xi_t = \mathbf{l}' \left(\mathbf{I} + \frac{\lambda \mathbf{K}}{1-\lambda} \right) (\mathcal{E}_t \mathbf{r}_{t+1} - \mathcal{E}_t^m \mathbf{r}_{t+1}) \tag{24}$$

Predictable excess returns depend linearly upon the missperception 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.

We return to our two special cases to gain some intuition on ξ_t.

- case 1: as expected, $\xi_t = 0$ when expectations are rational.
- case 2: In the absence of common factors $(\delta = 0)$, (24) simplifies to:

  $$\xi_t = \left(1 + \frac{\lambda k}{1-\lambda} \right) (\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1}) \tag{25}$$

  Predictable excess return depend upon the forecast error in the forward premium $x_{t+1} = r_t - r_t^*$. When the expected forward premium is lower than predicted according to the true model $(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^m x_{t+1} > 0)$, there are positive excess returns on the domestic currency. The reason is simple: if future interest rate differentials are under-estimated, the currency is artificially depreciated (see (22)) and will subsequently appreciate.

  Using (17), we can write predictable excess returns in a recursive form in this special case:

  $$\xi_t = \lambda (1-k) \xi_{t-1} + \lambda \left(1 + \frac{\lambda k}{1-\lambda} \right) (1-k) \mathbf{l}' \mathbf{H}' \boldsymbol{\epsilon}_t \tag{26}$$

  According to (26) persistence in predictable returns increases with the degree of misperception, measured by k. (26) provides the basis for a simple empirical specification. Since predictable excess returns are not observable empirically, define realized excess returns as $\Upsilon_{t+1} = (r_t - r_t^*) - (e_{t+1} - e_t) =$

$\xi_t - (e_{t+1} - \mathcal{E}_t e_{t+1})$. And according to (20) $e_{t+1} - \mathcal{E}_t e_{t+1} = -\left(1 + \frac{\lambda k}{1-\lambda}\right) \mathbf{l}'\mathbf{H}'\boldsymbol{\epsilon}_{t+1}$, so that:

$$\Upsilon_{t+1} = \lambda (1-k) \Upsilon_t + \left(1 + \frac{\lambda k}{1-\lambda}\right) \mathbf{l}'\mathbf{H}'\boldsymbol{\epsilon}_{t+1}.$$

Realized excess returns follow an AR(1) where the innovation is $\left(1 + \frac{\lambda k}{1-\lambda}\right) \mathbf{l}'\mathbf{H}'\boldsymbol{\epsilon}_t = \left(1 + \frac{\lambda k}{1-\lambda}\right) (\epsilon_t - \epsilon_t^*)$. This derivation is relatively robust in that (1) it does not depend upon survey data and possible measurement error (2) the ARMA(1,1) estimation does not impose the restriction that the elements of $\boldsymbol{\epsilon}_t$ are the innovations of the interest rate process.

4.5.2. Fama regression and forward discount puzzle. The model can also be used to investigate the forward discount puzzle. Recall that the Fama coefficient β_{Fama} converges to:

$$p \lim \beta_{Fama} = \frac{cov\left(e_{t+1} - e_t, r_t - r_t^*\right)}{var\left(r_t - r_t^*\right)}$$

We prove in the appendix the following result:

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

$$p \lim \beta_{Fama} = 1 - \frac{\lambda \mathbf{l}'\left(\mathbf{I} + \frac{\lambda \mathbf{K}}{1-\lambda}\right)(\mathbf{I} - \mathbf{K})\left(\mathbf{I} - \left(\mathbf{I} - \lambda^2 (\mathbf{I} - \mathbf{K})\right)^{-1} \lambda^2 \mathbf{K}\right) var\left(\mathbf{r}_t\right) \mathbf{l}}{\mathbf{l}' var\left(\mathbf{r}_t\right) \mathbf{l}} \tag{27}$$

It is immediate to check that $\beta_{Fama} = 1$ when expectations are rational (since $\mathbf{K} = \mathbf{I}$). In general however, the Fama coefficient is a complex function of the parameters of the model and may be significantly different from 1. We illustrate this result in the simple case where there are no common shocks. (27) simplifies to (27):

$$p \lim \beta_{Fama} = 1 - \frac{\lambda (1 - \lambda (1-k)) (1-k) (1+\lambda)}{1 - \lambda^2 (1-k)} \tag{28}$$

Indeed, it is easy to check that the Fama coefficient is always smaller than 1 and can even be negative. Figure 3 graphs β_{Fama} as a function of λ and k. Figure 5 reports various cross-sections for different values of k. We see from the graph that β_{Fama} fall as the shocks become more persistent (λ increases). The dependence on k is more complex. A low β_{Fama} requires a low k. However, when $k = 0$, which corresponds to an environment where agents believe all shocks are purely transitory, $\beta_{Fama} = 1 - \lambda$ and remains positive. Indeed, it is for small but strictly positive values of k that the minimum of β_{Fama} is attained.

It is easy to understand why β_{Fama} may be smaller than 1 in the limit. By definition, expected depreciation is the difference between the forward premium and predictable excess returns: $\mathcal{E}_t e_{t+1} - e_t = \mathbf{l}'\mathbf{r}_t - \xi_t$. But from (25), we know that ξ_t is a function of $x_t = \mathbf{l}'\mathbf{r}_t$ and $\mathcal{E}^m_{t-1} x_t$ so that:

$$\mathcal{E}_t e_{t+1} - e_t = \left(1 - \lambda \left(1 + \frac{\lambda k}{1-\lambda}\right)(1-k)\right) x_t + \lambda \left(1 + \frac{\lambda k}{1-\lambda}\right)(1-k)\ \mathcal{E}^m_{t-1} x_t \quad (29)$$

where we have substituted (25) as well as (18). Abstract for a moment from the second term on the right.

First, a high forward premium implies positive predictable excess returns (according to (25)). Thus expected depreciation is smaller than the forward premium and β_{Fama} is smaller than 1.

An example might help here. Suppose there is an increase in domestic interest rates, holding foreign interest rates constant. Since households underestimate future domestic interest rates, the currency will initially appreciate less than under rational expectations (see equation (20)). Since the exchange rate moves initially less than warranted by rational expectations, rationally expected depreciation is *smaller* than implied by the forward discount and there are positive predictable excess returns. This is true as soon as $\lambda > 0$ and $k < 1$. If interest rate follow a white noise process, there are no forecast misperceptions (since the agents are biased in favor of a white noise to start with) and no predictable currency movements.

Yet, we also know from figure 3 that β_{Fama} can be negative which is a much stronger result. How can this be? To understand what is going on, return to equation (29). When λ increases, the coefficient on the forward premium can be *negative*, as predictable excess returns become more volatile and dominate expected depreciation. To illustrate what is going on, consider again our increase in domestic interest rates holding foreign interest rates constant . For large enough λ, 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 tightening 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 rates co-exists with an *appreciating* currency. In other words, forward premium and rationally expected depreciation move in opposite directionsand β_{Fama} is negative.

Lastly, our discussion abstracted from the term $\mathcal{E}^m_{t-1} x_t$. But past expected forward premium is correlated with current forward premium. Indeed, one obtains (28) exactly when the correlation is properly taken into account. This term dampens the movements in the predictable excess returns and make it more difficult for β_{Fama} to turn negative (compare the term in x_t in (29) and the Fama coefficient).

4.5.3. 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 the country specific factors z or z^*. For simplicity, asssume also that we start from steady state with $\mathbf{r}_{t-1} = -\ln\beta$. According to (14), short rates follow:

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

Under rational expectation, the exchange rate obeys (21) and follow:

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

The exchange rate overshoots and converges back to its equilibrium value at the same speed as the exchange rate. When agents underreact, the exchange rate follows:

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

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 traders. In fact, since $\mathbf{K}$ is symmetric, one can check that the formula simplifies to:

$$e_{t+j} = \bar{e} - \frac{\lambda^j}{1-\lambda}\left[1 - \lambda\left(k_{11}^{j+1} - k_{21}^{j+1}\right)\right](\epsilon - \epsilon^*) \tag{30}$$

where k_{il}^{j+1} is the (i,l) element of $\mathbf{K}^{j+1}$. (30) makes clear that e_{t+j} is also unaffected by common shocks. In the special case where there are no common shocks, $k_{21} = 0$ and $k_{11}^{j+1} = k^{j+1}$, and the formula simplifies to (5).

Delayed overshooting occurs at horizon τ when $|e_{t+\tau+1} - \bar{e}| > |e_{t+\tau} - \bar{e}|$. Using (30), 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:*

$$\tau < \left[\ln\left(\frac{1-\lambda}{\lambda}\right) - \ln\left(1 - \lambda\left(k_{11} - k_{12}\right)\right)\right] / \ln\left(k_{11} - k_{12}\right)$$

where k_{ij} is the (i,j) element of the matrix $\mathbf{K}$.

This lemma establishes that delayed overshooting depends upon the interaction between the learning process as summarized by $\mathbf{K}$ and λ. The discussion of section 2 applies: delayed overshooting results from the interaction between an interest rate and a learning effect. We obtain (6) as a special case when $\delta = 0$. In that special case, the delayed overshooting region is reported on figure 2.

Given our estimated of λ and η from section 3, we can calculate the implied gain k and the implicit Fama coefficient β for all country pairs. This is only meant as an illustration. The point estimates are presented in table 7, together with their confidence interval, obtained through the delta method. The results indicate that the model can explain a substantial share of the forward discount bias, although our point estimates are almost always positive.

5. Conclusion

We have presented a model of nominal exchange rate determination that solves the forward premium puzzle and exhibits the delayed overshooting pattern of exchange rates found by Eichenbaum and Evans (1995). The key assumption driving our results is the underreaction of bond prices to changes in the short rate. This assumption is strongly supported by the data. Predictable excess returns results from the interaction of learning about the current state of affairs and the intrinsic dynamic response of interest rates to monetary shocks. 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 *less* 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.

Thus we view this paper as serving the useful purpose of uncovering deep rationale for a variety of pathologies on asset markets. In particular, the misperception that we identify holds out the interesting perspective of an integrated understanding of currency, bond and asset markets.

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

References

Albuquerque, Rui, "The Forward Premium Puzzle in a Model of Imperfect Information: Theory and Evidence," 1998. mimeo University of Rochester.

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

Baily, Walter and Leo Kropywiansky, "A Test resolving the 'Peso Problem' in Foreign Exchange Using Option-Implied Jump Expectations," Nov 1994. M.I.T. mimeo.

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.

Engle, Robert, David Lilien, and Russell Robins, "Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model," *Econometrica*, 1987, *55(2)*, 391–407.

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.

Grier, Kevin and Mark Perry, "The Effect of Money Shocks on Interest Rates in the Presence of Conditional Heteroscedasticity," *Journal of Finance*, 1993, *48(4)*, 1445–55.

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.

Lamont, Owen, "Macroeconomic Forecasts and Microeconomic Forecasters," 1996. NBER Working Paper.

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.

Lucas, Robert, "Interest Rates and Currency Prices in a Two-Country World," *Journal of Monetary Economics*, 1982, *10*, 335–359.

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.

Romer, Christina and David Romer, "Federal Reserve Private Information and the Behavior of Interest Rates," Jul 1996. NBER Working Paper 5692.

Saa-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.

Shiller, Robert, John Campbell, and K.L. Schoenholtz, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," *Brookings Papers on Economic Activity*, 1983, pp. 223–242.

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

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

Table 2:
DECOMPOSITION OF THE FAMA COEFFICIENT
Decomposes β according to: $\beta = 1 - b_{re} - b_{rp}$.
with $b_{re} = -\frac{cov\left(u_{t+k}^k, x_t^k\right)}{var\left(x_t^k\right)}$, $b_{rp} = \frac{var\left(\zeta_t^k\right) + cov\left(\mathcal{E}_t^m e_{t+k} - e_t, \zeta_t^k\right)}{var\left(x_t^k\right)}$

Country	Horizon	b_{re}	b_{rp}	β
All	3 months	0.81	0.19	-0.05
Canada		3.03	-0.20	-1.83
France		1.51	-1.87	1.35
Germany		2.30	-1.55	0.25
Italy		-0.50	-1.70	3.20
Japan		4.91	-0.33	-3.58
U.K.		2.82	0.13	-1.95
All	6 months	0.93	0.08	0.15
Canada		2.48	-0.59	-0.89
France		1.03	-1.55	1.51
Germany		2.20	-1.58	0.38
Italy		-0.72	-1.22	2.94
Japan		4.93	-0.70	-3.23
U.K.		2.54	-0.52	-1.02
All	12 months	-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:
EXPECTATION HYPOTHESIS COEFFICIENT
Decomposes $\tilde{\beta}$ according to: $\tilde{\beta} = 1 - \tilde{b}_{re} - \tilde{b}_{rp}$.
with $\tilde{b}_{re} = -\frac{cov\left(u_{t+k}^k, fp_t^k\right)}{var\left(fp_t^k\right)}$, $\tilde{b}_{rp} = \frac{var\left(\tilde{\zeta}_t^k\right) + cov\left(\mathcal{E}_t^m r_{t+k} - r_t^k, \tilde{\zeta}_t^k\right)}{var\left(fp_t^k\right)}$

Country	Horizon	$\tilde{b}_{re}$	$\tilde{b}_{rp}$	β
All	3 months			
Canada		-0.61	0.89	0.71
France		-0.63	0.69	0.93
Germany		-0.21	0.65	0.55
Italy		-0.90	0.78	1.13
Japan		-1.04	0.64	1.4
U.K.		-0.74	0.77	0.97
U.S.		-0.09	0.76	0.34
All	6 months			
Canada		-0.42	0.71	0.71
France		-1.09	0.83	1.26
Germany		-0.35	0.55	0.79
Italy		-1.54	1.24	1.30
Japan		0.61	0.75	-0.37
U.S.		0.47	0.80	-0.27
U.K.		-0.65	0.77	0.87

Table 4:
MAXIMUM LIKELIHOOD ESTIMATION OF THE STATE-SPACE REPRESENTATION OF 3 MONTHS EURORATES
Estimates the following state-space model on monthly data:
$x_t = \mu + x_t^p + \nu_t \quad ; \quad \lambda(L)\, x_t^p = \epsilon_t$.
x_t is the interest rate differential $r_t^* - r_t^{us}$. We report estimates of the autoregressive coefficients (λ) and the noise to signal ratio $\eta = \frac{\sigma_\nu^2}{\sigma_\epsilon^2}$. The sample period is 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. 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**
	(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)			

continued next page

Table 4: 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		
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)			
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)			

continued next page

Table 4: 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		
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 5:
MAXIMUM LIKELIHOOD ESTIMATION OF THE STATE-SPACE REPRESENTATION OF MONEY MARKET RATES
Estimates the following state-space model on monthly data:
$x_t = \mu + x_t^p + \nu_t \quad ; \quad \lambda(L)\, x_t^p = \epsilon_t$.
x_t is the interest rate differential $i_t^* - i_t^{us}$. We report estimates of the autoregressive coefficients (λ) and the noise to signal ratio $\eta = \frac{\sigma_\nu^2}{\sigma_\epsilon^2}$. The sample period is 1973:12 to 1992:12. 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. 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 $60lb$.

	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	208
I	**0.72**				0	**0.72**	-1.055	Z_ρ	**-70.86**
	(0.05)					(0.05)			(0.0)
II	**0.69**	0.04			0	**0.73**	-1.054	Z_t	**-6.34**
	(0.07)	(0.07)				(0.05)			(0.0)
III	**0.69**	0.01	0.04		0	**0.74**	-1.058		
	(0.11)	(0.14)	(0.10)			(0.08)			
IV	**0.68**	0.01	-0.04	0.12	0	**0.77**	-1.054		
	(0.11)	(0.15)	(0.12)	(0.08)		(0.09)			
PANEL B: FRANCE-U.S.								OBS	222
I	**0.93**				0	**0.93**	-0.699	Z_ρ	**-14.74**
	(0.02)					(0.02)			(0.05)
II	**1.17**	**-0.26**			0	**0.91**	-0.659	Z_t	-2.58
	(0.06)	(0.06)				(0.03)			(0.10)
III	**1.19**	**-0.37**	0.09		0	**0.91**	-0.658		
	(0.07)	(0.10)	(0.07)			(0.03)			
IV	**1.20**	**-0.40**	0.20	-0.09	0	**0.91**	-0.650		
	(0.07)	(0.10)	(0.10)	(0.07)		(0.03)			

continued next page

Table 5: MAXIMUM LIKELIHOOD ESTIMATION, MONEY MARKET RATES

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		
PANEL C: GERMANY-U.S.								OBS	221
I	**0.93**				0	**0.93**	-0.861	Z_ρ	-14.11
	(0.03)					(0.03)			(0.06)
II	**1.02**	-0.10			0	**0.92**	-0.857	Z_t	-2.39
	(0.07)	(0.07)				(0.03)			(0.13)
III	**1.02**	-0.14	0.04		0	**0.92**	-0.799		
	(0.07)	(0.10)	(0.07)			(0.03)			
IV	**1.02**	-0.12	-0.04	0.07	0	**0.93**	-0.741		
	(0.07)	(0.10)	(0.10)	(0.07)		(0.03)			
PANEL D: ITALY-U.S.								OBS	221
I	**0.95**				0	**0.95**	-0.815	Z_ρ	**-20.80**
	(0.02)					(0.02)			(0.0)
II	**1.32**	**-0.40**			0	**0.92**	-0.748	Z_t	**-3.49**
	(0.06)	(0.06)				(0.02)			(0.0)
III	**1.38**	**-0.60**	**0.15**		0	**0.93**	-0.740		
	(0.07)	(0.11)	(0.07)			(0.02)			
IV	**1.38**	**-0.61**	0.17	-0.01	0	**0.93**	-0.742		
	(0.07)	(0.11)	(0.11)	(0.07)		(0.02)			
PANEL E: JAPAN-U.S.								OBS	221
I	**0.96**				0	**0.96**	-0.712	Z_ρ	-9.64
	(0.02)					(0.02)			(0.28)
II	**1.39**	**-0.44**			0	**0.95**	-0.614	Z_t	-2.19
	(0.06)	(0.06)				(0.01)			(0.25)
III	**1.44**	**-0.58**	0.09		0	**0.95**	-0.613		
	(0.07)	(0.11)	(0.07)			(0.02)			
IV	**1.43**	**-0.54**	0.01	0.06	0	**0.96**	-0.610		
	(0.07)	(0.12)	(0.12)	(0.07)		(0.02)			

continued next page

Table 5: Maximum Likelihood Estimation, Money Market Rates

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		
Panel F: U.K.-U.S.								Obs	221
I	**0.92**				0.28	**0.92**	-1.301	Z_ρ	**-20.18**
	(0.03)				(0.25)	(0.03)			(0.00)
II	**0.75**	**0.18**			0	**0.93**	-1.281	Z_t	**-3.43**
	(0.07)	(0.07)				(0.03)			(0.01)
III	**0.75**	**0.22**	-0.04		0	**0.93**	-1.282		
	(0.07)	(0.08)	(0.07)			(0.03)			
IV	**0.75**	**0.21**	-0.07	0.04	0	**0.93**	-1.282		
	(0.07)	(0.08)	(0.08)	(0.07)		(0.03)			

Table 6:
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 $\tilde{\theta}$:
$x_t = \mu + x_t^p + \nu_t \quad ; \quad \lambda(L)\, x_t^p = \epsilon_t$.
$\hat{x}_t^\tau = \mathcal{E}_t^s\left\{r_{t+\tau} - r_{t+\tau}^*\right\}$ is the survey based measure of future interest rates differential and differs from market expectations by an iid noise υ_t^τ.

	ESTIMATED COEFFICIENTS					
	λ_1	λ_2	η	$\sum\lambda$		
MODEL	SE	SE	SE	SE	LOG-LIK	OBS
	PANEL A: CANADA-U.S.					
HORIZON:	ALL					
I	**0.972**		**2.99**	**0.972**	0.309	327
	(0.004)		(0.65)	(0.004)		
II	**0.275**	**0.675**	**0.73**	**0.950**	0.323	
	(0.07)	(0.07)	(0.26)	(0.008)		
	3 MONTHS					
I	**0.996**		**2.40**	**0.996**	0.579	109
	(0.009)		(0.788)	(0.009)		
II	**0.155**	**0.828**	0.23	**0.983**	0.684	
	(0.08)	(0.08)	(0.14)	(0.012)		
HORIZON:	6 MONTHS					
I	**0.974**		**3.86**	**0.974**	0.448	109
	(0.008)		(1.335)	(0.008)		
II	**0.385**	**0.573**	**1.31**	**0.958**	0.454	
	(0.034)	(0.031)	(0.586)	(0.013)		
	12 MONTHS					
I	**0.969**		**4.45**	**0.969**	0.072	109
	(0.007)		(2.049)	(0.007)		
II	**0.246**	**0.700**	1.327	**0.946**	0.074	
	(0.047)	(0.043)	(0.774)	(0.012)		

continued next page

Table 6: Pooled Maximum Likelihood Estimation of Market Filter, 3 Months Eurorates

continued from previous page

	Estimated Coefficients					
	λ_1	λ_2	η	$\sum \lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs
	Panel B: France-U.S.					
Horizon:	All					
I	**0.997**		**151.88**	**0.997**	-0.861	291
	(0.001)		(13.43)	(0.001)		
II	0.117	**0.879**	**42.76**	**0.996**	-0.836	
	(0.16)	(0.16)	(8.70)	(0.001)		
	3 Months					
I	**1.027**		**349.75**	**1.027**	-0.723	97
	(0.004)		(80.11)	(0.004)		
II	**-0.111**	**1.150**	**41.68**	**1.039**	-0.639	
	(0.04)	(0.04)	(11.38)	(0.01)		
Horizon:	6 Months					
I	**1.009**		**186.85**	**1.009**	-0.702	97
	(0.004)		(34.05)	(0.004)		
II	**0.477**	**0.536**	**79.67**	**1.013**	-0.702	
	(0.09)	(0.07)	(0.34)	(0.03)		
	12 Months					
I	**1.005**		**190.39**	**1.005**	-0.181	97
	(0.004)		(41.80)	(0.004)		
II	**0.304**	**0.704**	**65.73**	**1.008**	-0.179	
	(0.09)	(0.09)	(16.31)	(0.005)		

continued next page

Table 6: Pooled Maximum Likelihood Estimation of Market Filter, 3 Months Eurorates

continued from previous page

	Estimated Coefficients					
	λ_1	λ_2	η	$\sum\lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs
	Panel C: Germany-U.S.					
Horizon:	All					
I	**0.979**		12.26	**0.979**	0.265	327
	(0.002)		(9.85)	(0.002)		
II	**1.76**	**-0.766**	**14.89**	**0.994**	0.575	
	(0.01)	(0.01)	(2.84)	(0.003)		
	3 Months					
I	**0.993**		**3.853**	**0.993**	0.927	109
	(0.003)		(1.92)	(0.003)		
II	**1.766**	**-0.769**	**14.64**	**0.997**	1.262	
	(0.02)	(0.02)	(4.67)	(0.001)		
Horizon:	6 Months					
I	**0.978**		28.97	**0.978**	0.039	109
	(0.003)		(89.44)	(0.003)		
II	**1.804**	**-0.810**	**21.28**	**0.994**	0.481	
	(0.017)	(0.016)	(6.95)	(0.001)		
	12 Months					
I	**0.978**		22.03	**0.978**	0.206	109
	(0.002)		(52.91)	(0.002)		
II	**1.736**	**-0.742**	**14.69**	**0.994**	0.468	
	(0.01)	(0.02)	(4.63)	(0.001)		

continued next page

Table 6: POOLED MAXIMUM LIKELIHOOD ESTIMATION OF MARKET FILTER, 3 MONTHS EURORATES

continued from previous page

	ESTIMATED COEFFICIENTS					
	λ_1	λ_2	η	$\sum \lambda$		
MODEL	SE	SE	SE	SE	LOG-LIK	OBS
	PANEL D: ITALY-U.S.[38]					
HORIZON:	ALL					
I	**1.060**		NA	**1.060**	-1.569	291
	(0.001)		(NA)	(0.001)		
II	0.008	**1.138**	NA	**1.146**	-1.537	
	(0.01)	(0.01)	(NA)	(0.002)		
	3 MONTHS					
I	**1.048**		NA	**1.048**	-1.342	97
	(0.002)		(NA)	(0.002)		
II	-0.002	**1.160**	NA	**1.157**	-1.081	
	(0.008)	(0.001)	(NA)	(0.003)		
HORIZON:	6 MONTHS					
I	**1.06**		NA	**1.06**	-1.469	97
	(0.001)		(NA)	(0.001)		
II	0.084	1.058	NA	**1.142**	-1.373	
	(6.16)	(6.70)	(NA)	(0.54)		
	12 MONTHS					
I	**1.059**		NA	**1.059**	-1.380	97
	(0.001)		(NA)	(0.001)		
II	**0.119**	**1.015**	NA	**1.134**	-1.267	
	(0.04)	(0.04)	(NA)	(0.004)		

continued next page

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

Table 6: Pooled Maximum Likelihood Estimation of Market Filter, 3 Months Eurorates

continued from previous page

	Estimated Coefficients					
	λ_1	λ_2	η	$\sum \lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs
	Panel E: Japan-U.S.					
Horizon:	All					
I	**0.981**		3.44	**0.981**	0.530	327
	(0.002)		(1.02)	(0.002)		
II	**1.690**	**-0.697**	**8.58**	**0.993**	0.701	
	(0.02)	(0.02)	(1.57)	(0.007)		
	3 Months					
I	**0.996**		**0.36**	**0.996**	1.192	109
	(0.004)		(0.12)	(0.004)		
II	**1.784**	**-0.787**	**14.96**	**0.997**	1.661	
	(0.03)	(0.03)	(5.28)	(0.001)		
Horizon:	6 Months					
I	**0.988**		5.31	**0.988**	0.524	109
	(0.004)		(2.97)	(0.004)		
II	**1.781**	**-0.788**	**14.22**	**0.993**	1.022	
	(0.02)	(0.024)	(3.80)	(0.001)		
	12 Months					
I	**0.978**		0.51	**0.978**	0.229	109
	(0.002)		(0.38)	(0.002)		
II	**1.527**	**-0.537**	**5.28**	**0.990**	0.264	
	(0.06)	(0.06)	(2.52)	(0.001)		

continued next page

Table 6: Pooled Maximum Likelihood Estimation of Market Filter, 3 Months Eurorates

continued from previous page

	Estimated Coefficients					
	λ_1	λ_2	η	$\sum\lambda$		
Model	SE	SE	SE	SE	Log-Lik	Obs
	Panel F: U.K.-U.S.					
Horizon:	All					
I	**0.977**		**1.67**	**0.977**	0.195	327
	(0.002)		(0.32)	(0.002)		
II	**1.766**	**-0.774**	**27.34**	**0.992**	0.414	
	(0.02)	(0.02)	(6.39)	(0.001)		
	3 Months					
I	**0.997**		**1.25**	**0.997**	0.727	109
	(0.005)		(0.33)	(0.005)		
II	**1.690**	**-0.693**	**13.52**	**0.997**	0.852	
	(0.07)	(0.07)	(6.86)	(0.002)		
Horizon:	6 Months					
I	**0.986**		**1.69**	**0.986**	0.093	109
	(0.004)		(0.52)	(0.004)		
II	**1.806**	**-0.812**	**38.72**	**0.994**	0.380	
	(0.03)	(0.03)	(14.74)	(0.001)		
	12 Months					
I	**0.972**		**1.74**	**0.972**	0.079	109
	(0.003)		(0.78)	(0.003)		
II	**1.776**	**-0.784**	**37.43**	**0.992**	0.313	
	(0.02)	(0.02)	(15.09)	(0.001)		

Table 7:
IMPLICIT FAMA COEFFICIENT AND GAIN
This table reports point estimates of the gain of the filter k and the implict Fama coefficient β_{Fama}
Confidence interval are computed using the delta method.

Country	η	λ	$k_{\min}$	k	$k_{\max}$	$\beta_{\min}$	β_{Fama}	$\beta_{\max}$
Canada	2.99	0.90	0.61	0.77	0.93	0.59	0.62	0.65
France	151.8	0.85	-0.11	0.07	0.26	-0.95	0.06	1.08
Germany	12.26	0.96	0.59	0.78	0.97	0.48	0.60	0.72
Japan	3.44	0.94	0.74	0.88	0.94	0.65	0.72	0.78
U.K.	1.67	0.93	0.82	0.88	0.94	0.75	0.79	0.83

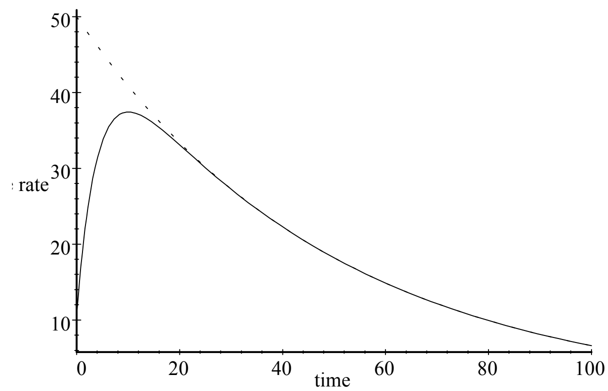


Figure 1: Delayed overshooting (–) and rational expectation (- -)response to monetary innovation. $\lambda = 0.98$, $k = 0.2$.

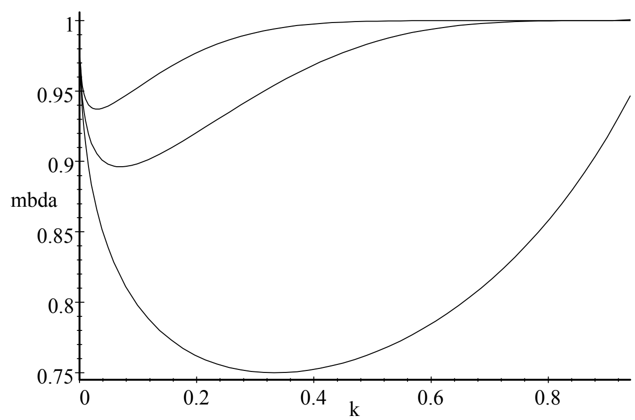


Figure 2: Delayed Overshooting Region as a function of the parameters (λ, k) at horizon $\tau = 1, 5, 10$.

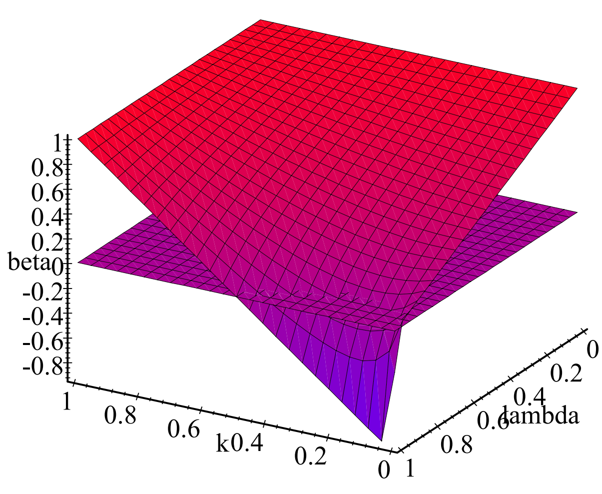


Figure 3: Asymptotic Limit of the Fama coefficient as a function of the parameters: (λ, k)

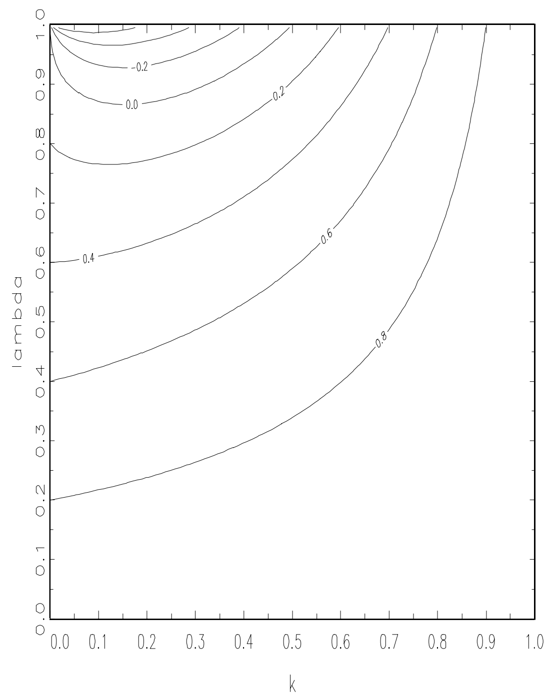


Figure 4: Contour plot of Asymptotic Limit of Fama coefficient as a function of k and λ.

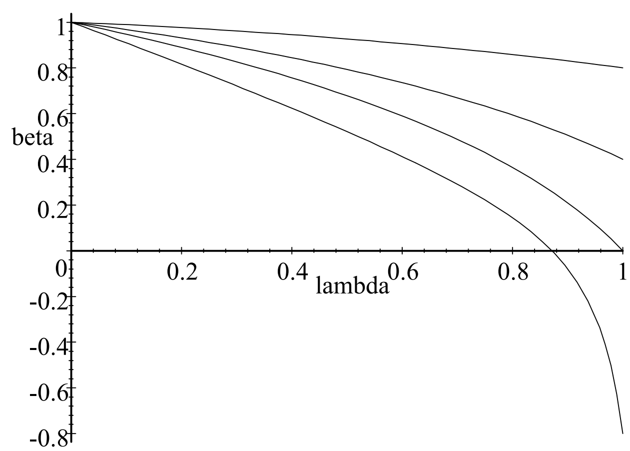


Figure 5: Asymptotic Limit of Fama Coefficient (β) for various values of k (0.9, 0.7, 0.5, 0.1 from top to bottom).

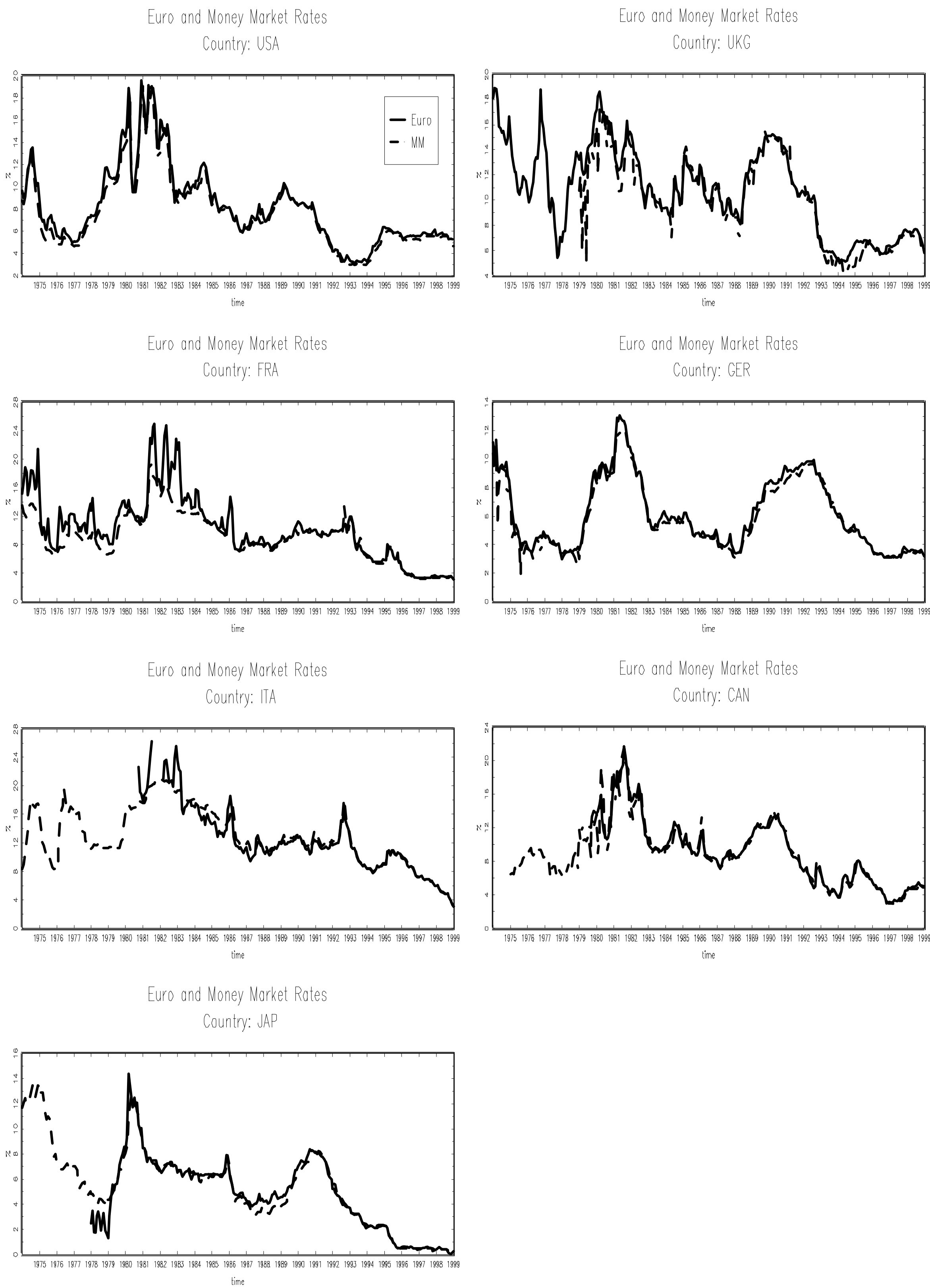


Figure 6: Euro Rates 3 months and Money Market Rates.

Figure 4.A: Euro 3 months and Forecasts
US−CAN

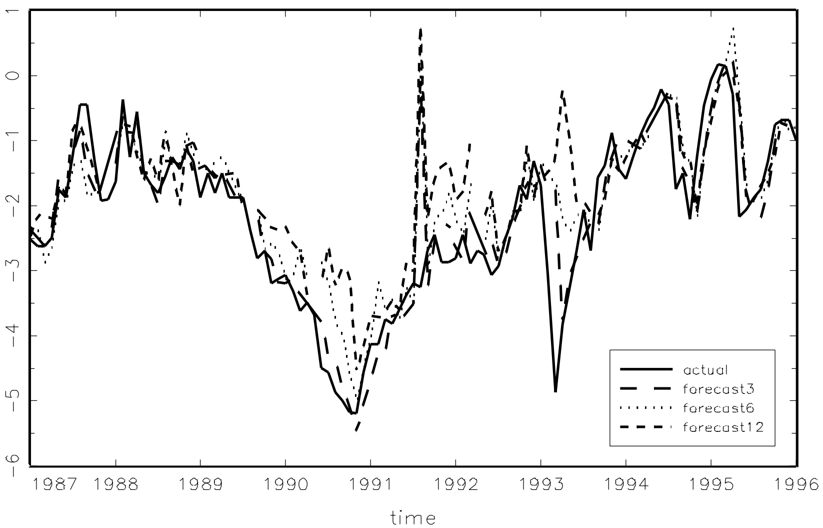


Figure 4.B: Euro 3 months, Forecast 3 Months
US−CAN; AR=3

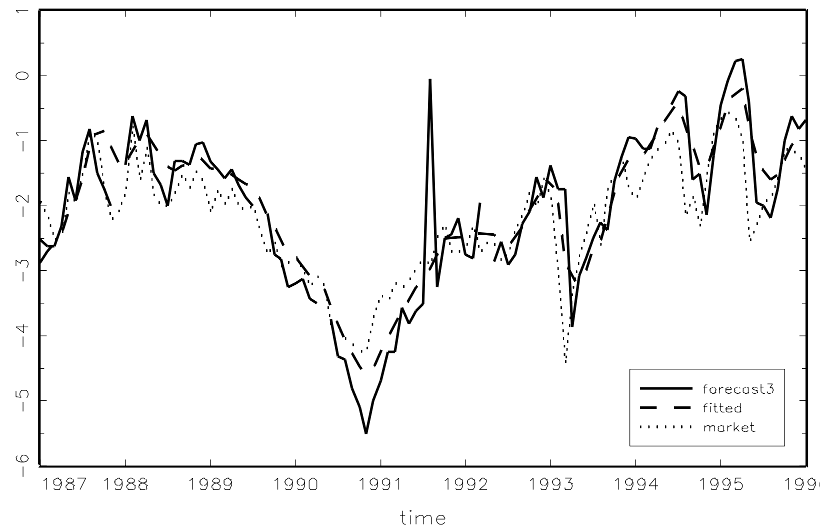


Figure 4.C: Euro 3 months, Forecast 6 Months
US−CAN; AR=3

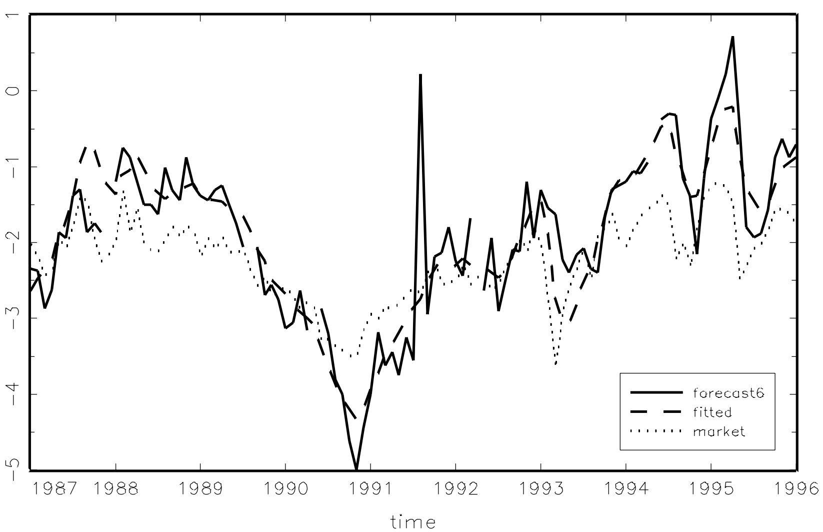
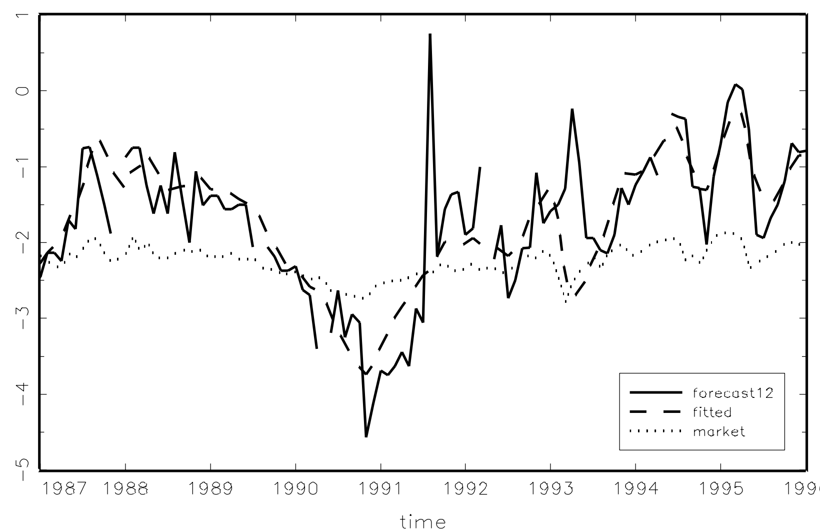


Figure 4.D: Euro 3 months, Forecast 12 Months
US−CAN; AR=3



A. Appendices

A.1. Kalman Filter Estimation. This subsection briefly derives the Kalman Filter equations. We postulate the following process:

$$x_t = \mathbf{H}'\mathbf{z}_t + \nu_t \quad (A.1)$$
$$\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t \quad (A.2)$$

where $\mathbf{z}_t = (z_t, ..., z_{t-p+1})'$, $\mathbf{H}' = (1, 0, ..., 0)'$ is a px1 vector. $\mathbf{z}_t$ is the state vector for the process, (A.1) the measurement equation and (A.2) the space equation. Define the informations set $I_t = \{x_{t-i}, i \geq 0\}$, $\hat{\mathbf{z}}_{t+1|t} = \mathcal{E}\{\mathbf{z}_{t+1}|I_t\}$, 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:

$$\hat{\mathbf{z}}_{t+1|t} = \mathbf{F}\hat{\mathbf{z}}_{t-1|t} + \mathbf{F}\hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \sigma_\nu^2\mathbf{I}\right)^{-1}\left(x_t - \mathbf{H}'\hat{\mathbf{z}}_{t|t-1}\right)$$
$$\hat{\mathbf{P}}_{t+1|t} = \mathbf{F}\left[\hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \sigma_\nu^2\mathbf{I}\right)^{-1}\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\right]\mathbf{F}' + \sigma_\epsilon^2\mathbf{I}$$

The smoother equations are:

$$\hat{\mathbf{z}}_{t|T} = \hat{\mathbf{z}}_{t|t} + \hat{\mathbf{P}}_{t|t}\mathbf{F}'\,\hat{\mathbf{P}}_{t+1|t}^{-1}\left(\hat{\mathbf{z}}_{t+1|T} - \hat{\mathbf{z}}_{t+1|t}\right)$$
$$\hat{\mathbf{P}}_{t|T} = \hat{\mathbf{P}}_{t|t} + \left(\hat{\mathbf{P}}_{t|t}\mathbf{F}'\hat{\mathbf{P}}_{t+1|t}^{-1}\right)\left(\hat{\mathbf{P}}_{t+1|T} - \hat{\mathbf{P}}_{t+1|t}\right)\left(\hat{\mathbf{P}}_{t|t}\mathbf{F}'\hat{\mathbf{P}}_{t+1|t}^{-1}\right)'$$

where

$$\hat{\mathbf{P}}_{t|t} = \hat{\mathbf{P}}_{t|t-1} - \hat{\mathbf{P}}_{t|t-1}\mathbf{H}\left(\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}\mathbf{H} + \sigma_\nu^2\mathbf{I}\right)^{-1}\mathbf{H}'\hat{\mathbf{P}}_{t|t-1}$$

Suppose that the current estimate for the state variable at time t is $\hat{\mathbf{z}}_{t|t}$. According to (**??**)-(**??**), the market forecast for the interest rate differential τ periods hence is:

$$x_t^\tau\left(\tilde{\theta}\right) \equiv \mathcal{E}\{x_{t+\tau}|I_t\} = \mu + \mathbf{H}'\hat{\mathbf{z}}_{t+\tau|t} = \mu + \mathbf{H}'\mathbf{F}^\tau\hat{\mathbf{z}}_{t|t} \quad (A.3)$$

Suppose now that we observe an imprecise measure of this forecast:

$$\hat{x}_t^\tau = x_t^\tau\left(\tilde{\theta}\right) + \upsilon_t^\tau$$

and that the measurement error υ_t^τ is uncorrelated with the true forecast, we can estimate $\tilde{\theta}$ by minimizing:

$$S\left(\tilde{\theta}\right) = \sum_\tau \sum_{t=1}^T \left(\hat{x}_t^\tau - x_t^\tau\left(\tilde{\theta}\right)\right)^2$$

A.2. Empirical Results on Interest Rate Differentials. We develop in this section the estimation procedure for our state-space representation. Assume that the State Space representation (A.1)-(A.2) holds. Under the normality assumption, and assuming additionally that $\hat{\mathbf{z}}_{1|0}$ is normally distributed, $\hat{\mathbf{z}}_{t+1}$ is normally distributed conditionally on I_t, with mean $\hat{\mathbf{z}}_{t+1|t}$ and variance $\hat{\mathbf{P}}_{t+1|t}$. We can then write the conditional likelihood of x_{t+1} as:

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

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 $\hat{z}_0$ 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, $\hat{\mathbf{z}}_0^1 = \mathcal{E}\left\{\mathbf{z}_0|I_T, \hat{\theta}^0\right\}$, 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]

[39] The estimation was modified to take into account a maturity larger than the sampling frequency.

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