Can Affine Term Structure Models Help Us to Predict Exchange Rates?

Antonio Diez de los Rios*
Bank of Canada
diez@bankofcanada.ca

April, 2006

Abstract

This paper proposes an arbitrage-free model of the joint behaviour of interest and exchange rates whose exchange rate forecasts outperform those produced by a random-walk model, a vector autoregression on the forward premiums and the rate of depreciation, or the standard forward premium regression. In addition, the model is able to reproduce the forward premium puzzle.

JEL Classification: E43, F31, G12, G15
Keywords: Interest rates, Exchange rates, Forward Premium Puzzle, Affine term structure, Out-of-sample Predictability

---

*I am very grateful to Nour Meddahi and Enrique Sentana for their useful comments and suggestions during the early stages of this project. This project started while I was a post-doctoral fellow at CIREQ and CIRANO; and their hospitality is greatly appreciated. I also thank Greg Bauer, David Bolder, Qiang Dai, Alfonso Novales, Jun Yang and seminar participants at the Bank of Canada, the 2005 CIRANO-CIREQ Conference on Financial Econometrics and the XIII Foro de Finanzas for their comments. However, I remain solely responsible for any remaining errors. Earlier versions of this paper circulated as "The Term Structure of Uncovered Interest Parity Regression Slopes in an Affine Economy". *Address for correspondence:* Financial Market Department, Bank of Canada, 234 Wellington Street, Ottawa, Ontario, K1A 0G9, Canada, Phone: (+1) 613 782 8965, Fax (+1) 613 782 7136. The views expressed in this paper are those of the author and do not necessarily reflect those of the Bank of Canada

1 Introduction

The objective of this paper is to improve our ability to predict exchange rates. In particular, I present a model where restrictions derived from the assumption of no-arbitrage are imposed on the joint behaviour of interest and exchange rates. In this model, prediction is based on the information embedded in interest rate differentials. An example of such information is the well-established fact that regressing ex-post rates of depreciation of a given currency on a constant and the interest rate differential usually delivers a slope coefficient that is negative (see *i.e.* Hodrick, 1987 and Engel, 1996). This phenomenon is known as the "forward premium puzzle" and it implies that currencies where domestic interest rates are high relative to those in the foreign country tend to appreciate. In fact, Clarida and Taylor (1997), using a linear vector error correction model (VECM) framework for the term structure of forward premiums (interest rate differentials), were able to beat the long-standing and devastating result found by Meese and Rogoff (1983a,b) that standard empirical exchange rate models cannot outperform a simple random walk forecast. Thus, interest rates across countries contain information that is useful to predict exchange rates

However, the VECM framework is based solely on the time series properties of interest and exchange rates; and it does not take into account that, for example, a deposit denominated in foreign (domestic) currency is risky (due to exchange rate variability) for domestic (foreign) investors, and therefore investors will demand compensation for bearing such risk. As a result, movements in interest and exchange rates must be related in such a way that they preclude the existence of arbitrage opportunities.

This paper investigates whether imposing such set of restrictions in the estimation of the joint dynamics of interest and exchange rates helps to improve over the Clarida and Taylor (1997) framework. In principle, imposing cross-equation restrictions will reduce the large number parameters that characterizes traditional time series models and, consequently, it will reduce excessive parameter estimation uncertainty that may adversely affect its out-of-sample forecasting performance. This insight is confirmed by Duffee (2002) and Ang and Piazzesi (2003) who find that imposing no-arbitrage restrictions helps out-of-sample forecasting of yields.

For the sake of tractability, the focus of this paper is on internationally affine term structure models, that is, models where not only interest rates (yields) are affine known functions of a set of state variables, but also the expected rate of depreciation (over any arbitrary period of time) satisfies this property. The main benefit of focusing on this class of models is that we avoid the use of Monte Carlo methods to compute the expected rate of depreciation. Although this is a perfectly valid approach (*cf.* Dong, 2005), the use of Monte Carlo methods for the task of predicting exchange rates can be computationally costly because the model is re-estimated at each point of time (of the out-of-sample period) in order to compute the corresponding dynamic forecasts. A first contribution of this paper is to provide conditions to obtain an expected rate of depreciation that is affine on the set of state variables.

Two families of dynamic term structure models fall within this internationally affine framework. The first subgroup is the so-called completely affine term structure model introduced in Dai and Singleton (2000) and it covers most of the work done on international term structure modeling: *i.e.* Saa-Requejo (1993), Frachot (1996), Backus *et al.* (2001), Dewachter and Maes (2001), Hodrick and Vassalou (2002), and Ahn (2004). However, Backus *et al.* (2001) show that the specification of the prices of risk in these models constrains the relationship between interest rates and the risk premium in such a way that the ability to reproduce the forward premium puzzle, and therefore their ability to capture the properties of interest and exchange rates, is severely limited. Despite this fact, more flexibility in modeling the relationship between interest and exchange rates can be achieved from the second group of models that falls within the internationally affine framework: the quadratic-Gaussian class of term structure models introduced in Ahn, Dittmar and Gallant (2002) and Leippold and Wu (2002). In these models, both the interest rates and the expected rate of depreciation are affine once we augment the set of state variables to include the squares and the cross-products of the original set of factors. The models in Leippold and Wu (2003) and Inci and Lu (2005) belong to this category. It is also worth mentioning at this point that Gaussian essentially affine models can be viewed as a particular case of the quadratic case where interest rates are affine but the expected rate of depreciation is quadratic. Therefore, the models in Brennan and Xia (2004) and Dong (2005) also belong to the internationally affine class. Finally, another model that does not fall between these two families but that generates affine interest rates and an affine expected rate of depreciation is the one in Graveline (2005).

Still, the main disadvantage of an internationally affine model is that tractability comes at the price of imposing more restrictions than the ones that the assumption of no-arbitrage implies. The results obtained in this paper suggest, however, that a two factor Gaussian essentially affine model produces forecasts that are superior, on the basis of Root-mean-square error (RMSE) and Mean-absolute-error (MAE) criteria, to those produced by the random walk model and the Clarida and Taylor (1997) approach. I find that imposing no-arbitrage restrictions reduces the RMSE in forecasting the spot U.S. Dollar – Sterling Pound rate by around 35% at the one-year forecast horizon relative to the VECM approach, and by around 15% for the case of the U.S. Dollar - Canadian Dollar. I also find that the gains (if any) from using a linear VECM model with respect to the use of a random walk model are small. For example, the gain at the one year horizon for the U.S. Dollar – Sterling Pound is only a 2.4% (versus the 40% reported by Clarida and Taylor, 1997). In addition, the model is able to reproduce the forward premium anomaly.

The paper is organized as follows. Section 2 describes the data. Section 3 introduces the concept of internationally affine term structure model and discusses several specifications that fall within this framework. Section 4 presents the empirical exercise. Finally, Section 5 concludes.

2 Data

The data set comprises monthly observations over the period January 1976 to December 2004 of U.S. Dollar – Sterling Pound and U.S. Dollar – Canadian Dollar rates of depreciation, along with the corresponding American, British and Canadian Eurocurrency interest rates of maturities 1, 3, 6 and 12 months. These Eurocurrency deposits are essentially zero coupon bond whose payoffs at maturity are the principal plus the interest payment. Exchange rate (expressed as U.S. dollars per unit of foreign currency) and Eurocurrency interest rate data are obtained from Datastream. However, the estimations are carried out using only data over the period January 1976 to December 1997 in order to reserve the last seven years of data for an out-of-sample forecasting exercise.

Table 1 panel a reports summary statistics for these variables. Following Bekaert and Hodrick (2001), all variables are measured in percentage points per year, and the monthly rates of depreciation are annualized by multiplying by 1,200. Note that the rates of depreciation have lower means (in absolute value) than the one corresponding to the interest rates but, on the contrary, interest rates are less volatile. In addition, interest rates display a high level of autocorrelation while the expected rates of depreciation do not. The rate of depreciation of the U.S. Dollar with respect to the Canadian Dollar is less volatile than the rate of depreciation of the U.S. Dollar with respect to the Sterling Pound. The (average) spread between the one year and the one month interest rate is positive for the case of the U.S. while negative for the case of the U.K. and Canada. Finally, the U.K. ranks first in terms of the highest (average) level of interest rates during our sample period. Canada and the U.S. rank second and third, respectively. These properties are consistent with previous studies such as, *e.g.*, Backus, *et al.* (2001) and Bekaert and Hodrick (2001).

Panel b presents the results of the estimation of the forward premium regressions for the U.S. Dollar – Sterling Pound and the U.S. Dollar – Canadian Dollar for the four different maturities available. These are OLS regressions of the ex-post rate of depreciation on a constant an the forward premium:

$$s_{t+h} - s_t = a + bp_t^{(h)} + u_{t+h} \tag{1}$$

where s_t is the logarithm of the spot exchange rate S_t (*i.e.* dollars per pound), $p_t^{(h)} = f_t^{(h)} - s_t$ is the forward premium and f_t is the logarithm of the forward rate $F_t^{(h)}$ contracted at time t and that matures at $t+h$. The uncovered interest parity (UIP) states that, under risk neutrality, the nominal expected return to speculation in the forward foreign exchange market conditional on the available information must be equal to zero:

$$E_t\left[s_{t+h} - s_t\right] = f_t^{(h)} - s_t \tag{2}$$

and therefore it implies that we should find that the constant term is equal to zero while the slope is equal to one, that is, $a = 0$ and $b = 1$ when running a regression such as the one in equation (1). Moreover, notice that this hypothesis implies that the (log) forward exchange rate is an unbiased predictor of the h-periods ahead (log) spot exchange rate; property that has motivated another name for the uncovered interest parity: the "unbiasedness hypothesis". Most often, the uncovered interest parity is stated in terms of the interest rate differential between two countries. In particular, the covered interest parity states that the forward premium is equal to the interest rate differential between two countries: $f_t^{(h)} - s_t = r_t^{(h)} - r_t^{*(h)}$, where $r_t^{(h)}$and $r_t^{*(h)}$ are the h-period interest rates on a deposit denominated in domestic and foreign currency respectively.

For the case of the Sterling Pound, the dataset implies a slope equal to -1.84 when considering a contract maturity of one month, -1.50 for three-month contracts, -1.36 for six-month contracts and -0.82 for one-year contracts. As for the case of the Canadian

Dollar, the slope is -1.35, -0.83, -0.43 and -0.24 for one, three, six and twelve months contracts, respectively. Moreover, it is possible to reject statistically the equality of these slopes to one on a maturity-by-maturity basis (panel b) and when jointly testing that the four coefficients are equal to one (panel c).

As previously noted, this result is inconsistent with the uncovered interest parity and it has been claimed that the main reason for this rejection lies in the fact that agents are not risk neutral. This idea goes back to the influential work of Fama (1984) who shows that if certain conditions are met then the forward premium puzzle can be explained by the existence of rational (time-varying) risk premia in foreign exchange markets. To illustrate his argument, let's start by the so-called Fama's decomposition of the forward premium into an expected rate of depreciation and a risk premium component:

$$\underbrace{f_t - s_t}_{p_t^{(h)}} = \underbrace{E_t\left[s_{t+h} - s_t\right]}_{q_t^{(h)}} + \underbrace{f_t - E_t s_{t+1}}_{d_t^{(h)}} \tag{3}$$

where $q_t^{(h)}$ is the expected rate of depreciation between time t and time $t-h$, $p_t^{(h)}$ is the forward premium, and $d_t^{(h)}$ is the risk premium in Fama's terminology.

Using the law of iterated expectations and substituting this decomposition of the forward premium into the definition of the uncovered interest parity regression slope in equation (1) we obtain:

$$b(h) = \frac{Cov[q_t^{(h)}, p_t^{(h)}]}{Var[p_t^{(h)}]} = \frac{Var[q_t^{(h)}] + Cov[q_t^{(h)}, d_t^{(h)}]}{Var[p_t^{(h)}]}$$

where I write the slope as a function of the maturity h to emphasize that there is a different slope for each value of h. Then, $b(h)$ can take negative values when the risk premium $d_t^{(h)}$ is time-varying and satisfies the condition $Var[q_t^{(h)}] + Cov[q_t^{(h)}, d_t^{(h)}] < 0$. Fama (1984) translates this inequality into two conditions that have been extensively studied in the literature:

1. Negative covariance between $q_t^{(h)}$ and $d_t^{(h)}$
2. Greater variance of $d_t^{(h)}$ than $q_t^{(h)}$.

Therefore, a model of the joint behaviour of interest and exchange rates needs to satisfy these two conditions in order to be empirical plausible.

3 Internationally Affine Models

The analysis is similar to that in Backus, *et al.* (2001) and Brandt and Santa-Clara (2002)[1]. It is based on a two-country world where assets can be denominated in either domestic currency (*i.e.* "dollars") or foreign currency (*i.e.* "pounds"). As usual, starred $*$ variables are foreign counterparts of domestic variables; and I use $(*)$ to denote domestic and foreign quantities at the same time and without distinction.

Initially, consider by a no arbitrage argument the existence of a (strictly positive) discount factor (SDF) M_t that prices any traded asset denominated in dollars through the following relationship[2]:

$$X_t = E_t \left[\frac{M_{t+h}}{M_t} X_{t+h} \right] \tag{4}$$

where X_t is the value of a claim to a stochastic cash flow of X_{t+h} dollars h periods later. Equivalently, we can also divide both sides of this expression by X_t to reformulate the previous expression as:

$$1 = E_t \left[\frac{M_{t+h}}{M_t} R_{t+h} \right] \tag{5}$$

where $R_{t+h} = X_{t+h}/X_t$ is just the gross h-periods return on the asset.

For example, this relationship can be used to price a zero coupon bond that promises to pay one dollar h-periods ahead ($X_{t+h} = 1$). Let $P_t^{(h)}$ be the price of this bond. In this case, direct application of the pricing relationship in equation (4) gives that $P_t^{(h)}$ must equal the conditional expectation of the ratio of the future and actual value of the SDF:

$$P_t^{(h)} = E_t \left[\frac{M_{t+h}}{M_t} \right]$$

If, instead, we consider a position in a h-period contract in the forward foreign exchange market, which involves no payment at date t while a payoff of $F_t^{(h)} - S_{t+h}$ at time $t+h$, we obtain:

$$0 = E_t \left[\frac{M_{t+h}}{M_t} \left(F_t^{(h)} - S_{t+h} \right) \right]$$

Alternatively, we might also need to price assets denominated in foreign currency such as, for instance, a pound-denominated zero coupon bond. Again, consider a no-arbitrage

[1] See also Guimarães (2006) for a more general setting including jumps.

[2] The SDF is alternatively known as pricing kernel or state price density and it is a concept related to the representative agent's nominal intertemporal marginal rate of substitution of consumption (See Cochrane, 2001 for an extended discussion on the SDF)

approach to postulate the existence of a foreign SDF M_t^* that prices any asset denominated in pounds through the following relationship:

$$1 = E_t\left[\frac{M_{t+h}^*}{M_t^*}R_{t+h}^*\right] \tag{6}$$

where now R_{t+h}^* is the gross h-periods return on an asset denominated in foreign currency.

However, any return denominated in pounds can be expressed in dollars once it is adjusted by the rate of change of the bilateral spot exchange rate S_{t+h}/S_t. Thus, it must be the case that:

$$1 = E_t\left[\frac{M_{t+h}}{M_t}\frac{S_{t+h}}{S_t}R_{t+h}^*\right]$$

In other words, the law of one price implies that any foreign asset must be correctly priced by both the domestic and the foreign SDFs:

$$E_t\left[\frac{M_{t+h}}{M_t}\frac{S_{t+h}}{S_t}R_{t+h}^*\right] = E_t\left[\frac{M_{t+h}^*}{M_t^*}R_{t+h}^*\right] = 1$$

As noted by Backus, *et al.* (2001) and Brandt and Santa-Clara (2002), this equation is trivially satisfied by a foreign SDF such that:

$$M_t^* = M_t S_t \tag{7}$$

and, furthermore, if markets are complete this specification of the foreign SDF is unique. Therefore, the exchange rate S_t is uniquely determined by the ratio of the two pricing kernels; and we can obtain the law of motion of the (log) exchange rate $s_t = \log S_t$ using Itô's lemma on the stochastic processes of M_t and M_t^*.

Hence, assume the following dynamics of the domestic and foreign SDF:

$$\begin{aligned} \frac{dM_t}{M_t} &= -r(\mathbf{x}_t,t)dt - \mathbf{\Lambda}(\mathbf{x}_t,t)'d\mathbf{W}_t \\ \frac{dM_t^*}{M_t^*} &= -r^*(\mathbf{x}_t,t)dt - \mathbf{\Lambda}^*(\mathbf{x}_t,t)'d\mathbf{W}_t \end{aligned} \tag{8}$$

where r_t and r_t^* are the instantaneous domestic and foreign interest rates (also known as short rates); $\mathbf{W}_t$ is a n-dimensional vector of independent Brownian motions that describes the shocks in this economy; and $\mathbf{\Lambda}_t$ and $\mathbf{\Lambda}_t^*$ are two n-vectors that are usually called the market prices of risk because they describe how the domestic and foreign SDFs respond to the shocks given by $\mathbf{W}_t$. In general, the short rates and the prices of risk are functions of time t and a Markovian n-dimensional vector $\mathbf{x}_t$ that describes completely the state of the economy. The law of motion of these state variables $\mathbf{x}_t$ is given by a diffusion such as:

$$d\mathbf{x}_t = \boldsymbol{\mu}_x(\mathbf{x}_t, t)dt + \boldsymbol{\sigma}_x(\mathbf{x}_t, t)d\mathbf{W}_t \tag{9}$$

where $\boldsymbol{\mu}_x$ is a n-dimensional vector of drifts, and $\boldsymbol{\sigma}_x$ is a $n \times n$ state-dependent factor-volatility matrix.

Using Itô's lemma on (8) and substracting we get:

$$ds_t = \left[(r_t - r_t^*) + \frac{1}{2}(\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t - \boldsymbol{\Lambda}_t^{*\prime}\boldsymbol{\Lambda}_t^*)\right]dt + (\boldsymbol{\Lambda}_t - \boldsymbol{\Lambda}_t^*)'d\mathbf{W}_t \tag{10}$$

This equation ties the dynamic properties of the exchange rate to the specific parameterization of the drift (interest rates), the diffusion (price of risk) coefficients in (8), and the dynamic evolution of the set of state variables (because interest rates and the prices of risks are ultimately related to those). In particular, if we focus on an Euler discretization to the process of the exchange rate and take expectations, we find that:

$$q_t = E_t\left[s_{t+1} - s_t\right] \simeq (r_t - r_t^*) + \frac{1}{2}(\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t - \boldsymbol{\Lambda}_t^{*\prime}\boldsymbol{\Lambda}_t^*) \tag{11}$$

Note that since the covered interest parity implies that the first term of q_t is equal to the forward premium $p_t = (r_t - r_t^*)$, this equation states a Fama's decomposition where the risk premium is equal to $d_t = \frac{1}{2}(\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t - \boldsymbol{\Lambda}_t^{*\prime}\boldsymbol{\Lambda}_t^*)$. In other words, Fama's risk premium is proportional to the differential between the instantaneous variances of the SDFs across the two countries; and, more important, it is time-varying because the prices of risks are usually functions of the state-vector $\mathbf{x}_t$. Consequently, the uncovered interest parity does not necessarily hold in this no-arbitrage framework[3].

Despite this result, we should be aware that equation (11) is only an approximation to the true expected rate of depreciation. First, it ignores the internal dynamics of the variables from moment t to $t+1$. That is, if we approximate the expected rate of depreciation over one month using this equation then our approach ignores that exchange rates move, say, "second-by-second" and that shocks accumulate during this time. Second, r_t and r_t^*

[3]One exception where the uncovered interest parity holds is the case when the SDF is conditionally homoscedastic ($\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t = \alpha$ and $\boldsymbol{\Lambda}_t^{*\prime}\boldsymbol{\Lambda}_t^* = \alpha^*$ being α and α^* two positive constants). In this case, we are within the framework of Hansen and Hodrick (1983) who have shown that, with an additional constant term, the uncovered interest parity is consistent with a model of rational maximizing behaviour in which assets are priced by a no arbitrage restriction. The intuition behind why this hypothesis still holds is that agents are risk-averse, but the risk-premia is not time-varying.

are instantaneous interest rates and not the relevant interest rates on an h-period deposit denominated in domestic and foreign currency, $r_t^{(h)}$ and $r_t^{*(h)}$ respectively. Namely, it is hard to think of the instantaneous interest rates as being a good proxy of their one-year counterparts.

Still, we need to compute the expected rate of depreciation for an arbitrary choice of h (say, 1, 3, 6 months or 1 year) in order to predict the future exchange rate. We can potentially resort to Monte Carlo methods and, given a set of parameters and initial conditions, simulate paths of the exchange rate and obtain the expected rate of depreciation over an arbitrary sample period h computing the average change of the exchange rate across the simulated paths. However, this Monte Carlo approach can be computationally costly, especially if the model is re-estimated at each point of time t and then dynamic forecasts of the spot exchange rate are computed. Therefore, this paper follows a different avenue of research that is in the spirit of the literature on exponentially-affine bond pricing: to restrict the specific functional forms of the short rates, prices of risk, and the drift and diffusion terms of the state variables so as to have a closed-form expression for the expected rate of depreciation (h-periods ahead).

In particular, this paper extends the class of affine term structure models to an *internationally affine* setting where not only the interest rate on a h-period deposit denominated in domestic and foreign currency ($r_t^{(h)}$and $r_t^{*(h)}$, respectively) are affine (known) functions of a set of state variables $\mathbf{x}_t$:

$$
\begin{aligned}
r_t^{(h)} &= A(h) + B(h)'\mathbf{x}_t \\
r_t^{*(h)} &= A^*(h) + B^*(h)'\mathbf{x}_t
\end{aligned}
$$

but also the expected rate of depreciation (h-periods ahead) is affine in the set of state variables $\mathbf{x}_t$:

$$
q_t^{(h)} = E_t\left[s_{t+h} - s_t\right] = C(h) + D(h)'\mathbf{x}_t
$$

Nonetheless, the tractability obtained by using an internationally affine model must come at a price. In particular, we need to impose a set of restrictions on the model that will ultimately constrain the specific functional forms of $A(h)$, $B(h)$, $C(h)$, and $D(h)$ and our ability to predict the exchange rates. In particular, we require two set of restrictions. First, those needed to have interest rates in affine form and which can be found in Duffie and Kan (1996) (see the next subsection for some examples). Second, those conditions needed to obtain an affine expected rate of depreciation (h-periods ahead) and which can be found in the next proposition:

Proposition 1 *If the drift of the process that the log exchange rate s_t follows is affine in a set of state variables $\mathbf{x}_t$, that is,*

$$E_t ds_t = (\gamma_0 + \boldsymbol{\gamma}'\mathbf{x}_t)dt \tag{12}$$

with $\gamma_0 \in \mathbb{R}$ and $\boldsymbol{\gamma} \in \mathbb{R}^n$, and $\mathbf{x}_t$ follows an affine diffusion:

$$d\mathbf{x}_t = \boldsymbol{\Phi}(\boldsymbol{\theta} - \mathbf{x}_t)dt + \boldsymbol{\Sigma}^{1/2}V(\mathbf{x}_t)^{1/2}d\mathbf{W}_t \tag{13}$$

where $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ are $n \times n$ matrices, $\boldsymbol{\theta}$ is a n-vector, $V(\mathbf{x}_t)$ is a diagonal $n \times n$ matrix with i-th typical element $v_i(\mathbf{x}_t) = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_t$, and $\mathbf{W}_t$ is a n-dimensional vector of independent Brownian motions; then, the expected rate of depreciation (h-periods ahead) is a (known) affine function of the state vector $\mathbf{x}_t$:

$$q_t^{(h)} = E_t\left[s_{t+h} - s_t\right] = C(h) + D(h)'\mathbf{x}_t \tag{14}$$

where the coefficients $C(h) \in \mathbb{R}$ and $D(h) \in \mathbb{R}^n$ have the following expressions:

$$\begin{aligned} C(h) &= \gamma_0 h + \boldsymbol{\gamma}'\boldsymbol{\theta}h - \boldsymbol{\gamma}'\boldsymbol{\Phi}^{-1}\left[I - e^{-\boldsymbol{\Phi}h}\right]\boldsymbol{\theta} \\ D(h)' &= \boldsymbol{\gamma}'\boldsymbol{\Phi}^{-1}\left[I - e^{-\boldsymbol{\Phi}h}\right] \end{aligned}$$

Proof. See appendix. ■

The result in this proposition is novel because (up to the best of my knowledge) the literature on multi-country affine models has focused almost entirely on Euler approximations to the expected rate of depreciation h-periods ahead and, therefore, their results are subject to the shortcomings mentioned before[4]. In particular, this proposition tells us that an affine expected rate of depreciation requires both the short rates (r_t and r_t^*) and the instantaneous variances of the pricing kernels ($\boldsymbol{\Lambda}_t'\boldsymbol{\Lambda}_t$ and $\boldsymbol{\Lambda}_t^{*\prime}\boldsymbol{\Lambda}_t^*$) to be affine in $\mathbf{x}_t$ (which guarantees that the drift of the log exchange rate s_t is affine); and, at the same time, the process that $\mathbf{x}_t$ follows must be an affine diffusion. If we compare these conditions with

[4]For example, Hodrick and Vassalou (2002) Leippold and Wu (2003) and Ahn (2004) focuses on Euler approximations of the law of motion of the (log) exchange rate, so their *formulae* regarding the expected rate of depreciation is only valid for h arbitrary small. One exception is Dewatcher and Maes (2001) who provide the expressions for the expected rate of depreciation for an arbitrary choice of h. However, their model is just a particular example of the general framework provided here.

those needed to obtain interest rates in affine form we realize that an internationally affine model imposes additional constraints with respect to the class of affine term structure models. For example, it is possible to obtain affine interest rates without having an instantaneous variance of the SDF that is affine in $\mathbf{x}_t$ (see *e.g.* Duffee, 2000 and Cheridito, *et al.* 2005) or without the condition that the state vector must follow an affine diffusion (see *e.g.* Duarte, 2004). The next subsections investigates which models satisfy the internationally affine conditions and to what extent this represents a constraint to predict exchange rates.

3.1 Affine Models of Currency Pricing

Up to this point, we are interested in finding those models that fall within the internationally affine framework, that is, those models where not only interest rates are affine (known) functions of a set of state variables, but also the expected rate of depreciation satisfies this property. Since one of our demanded characteristics is that the interest rates must be affine in a set of factors and it is well known that the standard formulation of the affine term structure models shares this property, I start by establishing the properties of the exchange rate implied by this class of models.

To this end, I focus on a multi-country version of the Dai and Singleton (2000) standard formulation of these affine term structure models that nests most of the work on international term structure modeling[5]. These models can be considered as multivariate extensions of the Cox, Ingersols and Ross (1985) model and they are characterized by an instantaneous interest rate (also known as short rate) that is an affine function of the set of state variables $\mathbf{x}_t$:

$$
\begin{aligned}
r_t &= \delta_0 + \boldsymbol{\delta}_1' \mathbf{x}_t \\
r_t^* &= \delta_0^* + \boldsymbol{\delta}_1^{*\prime} \mathbf{x}_t
\end{aligned}
\tag{15}
$$

where δ_0, δ_0^* are two scalars, and $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_1^*$ are two n-dimensional vectors. The dynamic evolution of these n state variables is given by the following affine diffusion:

$$
d\mathbf{x}_t = \mathbf{\Phi}(\boldsymbol{\theta} - \mathbf{x}_t)dt + \mathbf{\Sigma}^{1/2} V(\mathbf{x}_t)^{1/2} d\mathbf{W}_t \tag{16}
$$

where, again, $\mathbf{\Phi}$ and $\mathbf{\Sigma}$ are $n \times n$ matrices, $\boldsymbol{\theta}$ is a n-vector, and $V(\mathbf{x}_t)$ is a diagonal $n \times n$ matrix with i-th typical element $v_i(\mathbf{x}_t) = \alpha_i + \boldsymbol{\beta}_i' \mathbf{x}_t$. $\mathbf{W}_t$ is a n-dimensional vector of

[5] See *e.g.* Saa-Requejo (1993), Frachot (1996), Backus *et al.* (2001), Dewachter and Maes (2001), Hodrick and Vassalou (2002), and Ahn (2004)

independent Brownian motions. Since it is possible for an arbitrary set of parameters that the state variables $\mathbf{x}_t$ enter in a region where $v_i(\mathbf{x}_t) = \alpha_i + \boldsymbol{\beta}_i'\mathbf{x}_t$ is negative, which would imply that the state vector has a negative conditional variance, Dai and Singleton (2000) provide a set of restrictions on the parameters of the model that guarantees that the dynamics of $\mathbf{x}_t$ are well defined. Finally, the model is completed by an specification of the domestic and foreign prices of risk such that:

$$\begin{aligned} \boldsymbol{\Lambda}_t &= V(\mathbf{x}_t)^{1/2}\boldsymbol{\lambda} \\ \boldsymbol{\Lambda}_t^* &= V(\mathbf{x}_t)^{1/2}\boldsymbol{\lambda}^* \end{aligned} \tag{17}$$

This standard formulation of these affine term structure models is also known as "completely affine" specification (see Duffee, 2002) because it has an instantaneous variance of the SDFs, $\boldsymbol{\Lambda}_t^{(*)\prime}\boldsymbol{\Lambda}_t^{(*)}$, that is affine in the set of factors $\mathbf{x}_t$.

Under this parameterization, interest rates on h-period deposits denominated in domestic and foreign currencies satisfy:

$$r_t^{(h)} = A(h) + B(h)'\mathbf{x}_t$$
$$r_t^{*(h)} = A^*(h) + B^*(h)'\mathbf{x}_t$$

where the coefficients $A^{(*)}(h) \in \mathbb{R}$ and $B^{(*)}(h) \in \mathbb{R}^n$ solve two systems of ordinary differential equations whose details can be found in *e.g.* Duffie and Kan (1996), Dai and Singleton (2003) and Piazzesi (2003).

Substituting the expressions for the short rates and the prices of risk into the law of motion of the (log) exchange rate in equation (10) gives:

$$ds_t = (\gamma_0 + \boldsymbol{\gamma}'\mathbf{x}_t)dt + (\boldsymbol{\lambda} - \boldsymbol{\lambda}^*)'V(\mathbf{x}_t)^{1/2}d\mathbf{W}_t \tag{18}$$

where γ_0 and $\boldsymbol{\gamma}$ are:

$$\begin{aligned} \gamma_0 &= (\delta_0 - \delta_0^*) + \frac{1}{2}\sum_{i=1}^{N}(\lambda_i^2 - \lambda_i^{*2})\alpha_i \\ \boldsymbol{\gamma} &= (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_1^*) + \frac{1}{2}\sum_{i=1}^{N}(\lambda_i^2 - \lambda_i^{*2})\boldsymbol{\beta}_i \end{aligned}$$

Therefore, Proposition 1 holds: the drift of the (log) exchange rate is affine in a set of state variables $\mathbf{x}_t$ and these state variables follow an affine diffusion. This property adds a new meaning to the term "completely affine specification".

However, it has been found that this "completely affine" specification of the prices of risk is empirically restrictive. For example, Duffee (2002) finds that this parameterization produces forecasts of future Treasury yields that are beaten by a random walk specification[6]; and Backus, *et al.* (2001) point out that this model constrains the relationship between interest rates and the risk premium in such a way that the ability of the model to capture the forward premium puzzle, and therefore the ability to predict exchange rates, is severely limited. Therefore, we need to make more flexible assumptions on the form of the prices of risk. However, those models with more flexible specifications of the prices of risk, as the "essentially affine specification" in Duffee (2002) or the "extended affine specification" in Cheridito, *et al.* (2005), do not necessarily have an instantaneous variance of the SDF that is affine in the state variables[7]. In particular, a similar point has been given in Guimarães (2006): "[...] with these [Cheridito *et al.* (2005)] market prices of risk exchange rates will be a nonlinear (not even polynomial) function of latent state variables".

Still, there is hope in reproducing the forward premium puzzle. Dai and Singleton (2002) have been successful in explaining puzzles in a similar conceptual framework to the forward premium anomaly: the rejection of the expectations hypothesis of the term structure of interest rates[8]. In particular, they show that a Gaussian essentially affine model can generate flexible enough time-varying risk premia in holding-period bond returns so as to solve the failure of this traditional "expectation theory". Therefore, the question is whether this model can also reproduce the "forward premium puzzle".

In their model, the specification of the prices of risk is affine in a set of variables[9]:

$$\mathbf{\Lambda}_t^{(*)} = \boldsymbol{\lambda}_0^{(*)} + \boldsymbol{\lambda}_1^{(*)} \mathbf{x}_t$$

where $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_0^*$ are two n-dimensional vectors, and $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_1^*$ are two $n \times n$ matrices; and the

[6] Duffee (2002) claims that this is because *i)* the price of risk variability only comes from $V(\mathbf{x}_t)^{1/2}$ and *ii)* because the sign of $\mathbf{\Lambda}_t^{(*)}$ cannot change as the elements of $V(\mathbf{x}_t)^{1/2}$ are restricted to be nonnegative

[7] One exception is Graveline (2005) whose model is based on the "extended affine specification" in Cheridito, *et al.* (2004). However, he restricts these prices of risk in such a way that the cancelations in equation (10) delivers an affine diffusion for the exchange rate.

[8] See Bekaert and Hodrick (2001) for this relationship between the expectation hypothesis of the term structure of interest rates and the uncovered interest parity (also known as expectation hypothesis of the foreign exchange market).

[9] As claimed by Dai and Singleton (2002): "Since LPY *(the expectations hypothesis of the term structrure)* refers to the properties of the *(conditional)* first moments of yields, and the family of Gaussian models (family $A_0(3)$) gives the most flexibility to the structure of factor correlations and conditional means, one might conjecture a priori that these models would perform at least as well as other affine models". Therefore, since the uncovered interest parity refers also to the properties on the (conditional) first moments, but now of exchange and interest rates, we can apply the same reasoning to our case.

latent state variables follow a multivariate Orstein-Uhlenbeck (Gaussian) process. Again, this model falls within the essentially affine specification in Duffee (2002) and it does not generate an affine expected rate of depreciation because the variance of the SDF $\boldsymbol{\Lambda}_t^{(*)\prime}\boldsymbol{\Lambda}_t^{(*)}$ is quadratic in $\mathbf{x}_t$:

$$\boldsymbol{\Lambda}_t^{(*)\prime}\boldsymbol{\Lambda}_t^{(*)} = \boldsymbol{\lambda}_0^{(*)\prime}\boldsymbol{\lambda}_0^{(*)} + 2\boldsymbol{\lambda}_0^{(*)\prime}\boldsymbol{\lambda}^{(*)}\mathbf{x}_t + \mathbf{x}_t^{\prime}\boldsymbol{\lambda}^{(*)\prime}\boldsymbol{\lambda}^{(*)}\mathbf{x}_t$$

However, quadratic models can be viewed as being "affine" in an augmented set of factors obtained by stacking the original one and their respective squares and cross-products[10]. This idea is exploited in the next subsection.

3.2 Quadratic Models of Currency Pricing

These term structure models are characterized by an instantaneous interest rate that is a quadratic function of the set of state variables $\mathbf{x}_t$:

$$\begin{aligned} r_t &= \delta_0 + \boldsymbol{\delta}_1^{\prime}\mathbf{x}_t + \mathbf{x}_t^{\prime}\boldsymbol{\delta}_2\mathbf{x}_t \\ r_t^* &= \delta_0^* + \boldsymbol{\delta}_1^{*\prime}\mathbf{x}_t + \mathbf{x}_t^{\prime}\boldsymbol{\delta}_2^*\mathbf{x}_t \end{aligned} \tag{19}$$

where δ_0, δ_0^* are two scalars, $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_1^*$ are two n-dimensional vectors, and $\boldsymbol{\delta}_2$, $\boldsymbol{\delta}_2^*$ are two symmetric $n \times n$ matrices. The state variables follow a multivariate Orstein-Uhlenbeck (Gaussian) process:

$$d\mathbf{x}_t = \boldsymbol{\Phi}(\boldsymbol{\theta} - \mathbf{x}_t)dt + \boldsymbol{\Sigma}^{1/2}d\mathbf{W}_t \tag{20}$$

where $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ are $n \times n$ matrices, $\boldsymbol{\theta}$ is a n-vector; and $\mathbf{W}_t$ is a n-dimensional vector of independent Brownian motions. As a difference with the prices of risk in a completely affine framework, the price of risk is a linear function of the state variables $\mathbf{x}_t$:

$$\begin{aligned} \boldsymbol{\Lambda}_t &= \boldsymbol{\lambda}_0 + \boldsymbol{\lambda}_1\mathbf{x}_t \\ \boldsymbol{\Lambda}_t^* &= \boldsymbol{\lambda}_0^* + \boldsymbol{\lambda}_1^*\mathbf{x}_t \end{aligned} \tag{21}$$

where $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_0^*$ are two n-dimensional vectors, and $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_1^*$ are two $n \times n$ matrices. Moreover, note that the Gaussian essentially affine specification in Dai and Singleton (2002), which has both short rates and the prices of risk being affine in a set of Gaussian state

[10] A similar argument has been given in Cheng and Scaillet (2002), Dai and Singleton (2003b), Gouriéroux and Sufana (2003) within the one-country set-up.

variables, is nested by this quadratic formulation when $\boldsymbol{\delta}_2 = \boldsymbol{\delta}_2^* = 0$. I will return to this model shortly.

Ahn, Dittmar and Gallant (2002) and Leippold and Wu (2002) show that in this framework interest rates have a quadratic form:

$$r_t^{(h)} = A(h) + B_1(h)'\mathbf{x}_t + \mathbf{x}_t' B_2(h)\mathbf{x}_t$$
$$r_t^{*(h)} = A^*(h) + B_1^*(h)'\mathbf{x}_t + \mathbf{x}_t' B_2^*(h)\mathbf{x}_t$$

where the coefficients $A^{(*)}(h) \in \mathbb{R}$, $B_1^{(*)}(h) \in \mathbb{R}^n$ and $B_2^{(*)}(h) \in \mathbb{R}^{n\times n}$ solve two systems of ordinary differential equations. Therefore, this model does not fall within the internationally affine framework. However, we can still view any quadratic model as being affine in an augmented set of variables obtained by stacking the original one and their respective squares and cross-products. This point can be illustrated with an example. First, assume that short rates in both countries are quadratic in a global factor x_t:

$$r_t^{(*)} = \delta_0^{(*)} + \delta_1^{(*)} x_t + \delta_2^{(*)} x_t^2$$

and that this state variable follows a Gaussian process such as:

$$dx_t = \phi(\theta - x_t)dt + \sigma dW_t$$

Finally assume that the prices of risk are affine in this global factor:

$$\Lambda_t^{(*)} = \lambda_0^{(*)} + \lambda_1^{(*)} x_t$$

Note that this set of assumptions implies that the h-period domestic and foreign interest rates will satisfy a quadratic relationship:

$$r_t^{(*)(h)} = a^{(*)}(h) + b_1^{(*)}(h)x_t + b_2^{(*)}(h)x_t^2$$

which can be rewritten using a new variable $z_t = x_t^2$ that captures the square of the global factor. In compact notation:

$$r_t^{(*)(h)} = a^{(*)}(h) + B^{(*)}(h)'\widetilde{\mathbf{x}}_t$$

where $B^{(*)}(h) = \left[b_1^{(*)}(h), b_2^{(*)}(h)\right]'$ and $\widetilde{\mathbf{x}}_t = [x_t, z_t]'$. That is, interest rates are affine in the original global factor x_t and its square $z_t = x_t^2$. Therefore, this quadratic model can be viewed as an affine model in the new set of factors given by $\widetilde{\mathbf{x}}_t$.

Similarly, it can be shown that the expected rate of depreciation is also affine in $\widetilde{\mathbf{x}}_t$. Recall that Proposition 1 requires, first, the drift of the (log) exchange rate process to be affine in this new set of state variables and, second, the process that these variables follow must be itself an affine diffusion. The first condition is satisfied because short rates and the instantaneous variance of the SDFs are quadratic in x_t so they can be expressed in terms of $\widetilde{\mathbf{x}}_t = [x_t, z_t]$. Second, it can be shown that if we apply Itô's lemma on $z_t = x_t^2$, then the joint process for x_t and z_t is an affine diffusion. In particular, the law of motion of the augmented set of factors $\widetilde{\mathbf{x}}_t = [x_t, z_t]$ satisfies:

$$d\begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} \phi & 0 \\ -2\phi\theta & 2\phi \end{bmatrix} \left[\begin{pmatrix} \theta \\ \sigma^2/2\phi + \theta^2 \end{pmatrix} - \begin{pmatrix} x_t \\ z_t \end{pmatrix} \right] dt + \begin{bmatrix} \sigma \\ 2\sigma\sqrt{z_t} \end{bmatrix} dW_t$$

Therefore, this model satisfies the two conditions for an expected rate of depreciation.

In brief, the interest rates and the expected rate of depreciation are affine in a new set of state variables. Consequently, this one-factor quadratic model can be interpreted as a two-factor internationally affine model. Furthermore, the generalization of this result to the case where there is more than one factor is straightforward and the details can be found in the appendix. In the general case, $\mathbf{z}_t$ must include the squares of the original set of state variables $\mathbf{x}_t$ and their cross-products. A compact way to do so is to apply the matrix *vech* operator, which stacks the elements on and below the main diagonal of a square matrix, to the matrix given by $\mathbf{x}_t\mathbf{x}_t'$. As a result, quadratic models can be viewed as part of the internationally affine framework because they provide interest rates and an expected rate of depreciation that are affine in the augmented set of state variables given by $\widetilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{z}_t']'$ with $\mathbf{z}_t = vech(\mathbf{x}_t\mathbf{x}_t')$.

3.2.1 Gaussian Essentially Affine Models

As mentioned before, the Gaussian subfamily of the essentially affine models introduced in Duffee (2002) provides a sufficiently flexible risk premia to explain the puzzles of the expectations hypothesis of the term structure of interest rates. Therefore, it can potentially help us to address the forward premium anomaly. These models are characterized by instantaneous interest rates and prices of risk that are affine in a set of state variables $\mathbf{x}_t$:

$$\begin{aligned} r_t^{(*)} &= \delta_0^{(*)} + \boldsymbol{\delta}_1^{(*)\prime}\mathbf{x}_t \\ \boldsymbol{\Lambda}_t^{(*)} &= \boldsymbol{\lambda}_0^{(*)} + \boldsymbol{\lambda}_1^{(*)}\mathbf{x}_t \end{aligned}$$

where the state vector $\mathbf{x}_t$ follow a multivariate Orstein-Uhlenbeck (Gaussian) process:

$$d\mathbf{x}_t = \mathbf{\Phi}(\boldsymbol{\theta} - \mathbf{x}_t)dt + \mathbf{\Sigma}^{1/2}d\mathbf{W}_t$$

Duffee (2002) shows that this model generates interest rates that are affine in the factors $\mathbf{x}_t$. However, notice once again that the instantaneous variance of the SDF $\mathbf{\Lambda}_t^{(*)\prime}\mathbf{\Lambda}_t^{(*)}$ is quadratic in the set of state variables $\mathbf{x}_t$. Still, this model can be viewed as a particular case of the quadratic specification with $\boldsymbol{\delta}_2^{(*)} = \mathbf{0}_{n\times n}$ in equation (19) and therefore it is subject to the results for quadratic models presented before. That is, the expected rate of depreciation is affine in the augmented set of state variables obtained by stacking the original one and their respective squares and cross-products. An example of such Gaussian essentially affine model of currency pricing can be found in Brennan and Xia (2004) and Dong (2005).

To conclude this section, Table 2 summarizes the theoretical findings of this section. First, the completely affine framework implies that interest rates and the expected rate of depreciation are linear (known) functions of a set of state variables. Second, the quadratic framework implies interest rates and an expected rate of depreciation that is linear in an augmented set of factors that includes the original set of factors and their squares and cross-products. Finally, the Gaussian essentially affine model is in middle ground as the interest rates are linear, while the expected rate of depreciation is quadratic (or linear in the augmented set of factors).

4 Empirical Model

4.1 A two factor Gaussian essentially affine model

In this section, I focus on the estimation of a Gaussian essentially affine model. Several reasons justify the choice of this particular model. First, a multivariate Gaussian process gives the most degree of flexibility to the structure of correlations and conditional means of the state vector. That means that this model is going to perform at least as well as the other members of the family of completely affine models whose general structure is restricted by requirement that the conditional variance must always be positive (see *e.g.* Dai and Singleton, 2000). Moreover, since there is no state variable driving the conditional variance in this proposed model, there is no need to worry about these variables entering some non-admissible space where the volatilities are negative. Second, the specification of the prices of risk is the same as in the quadratic models, so similar degree of flexibility is expected in reproducing the forward premium puzzle. Besides and in contrast with these quadratic models, the Gaussian essentially affine model generates a one-to-one mapping from interest rates to the state vector $\mathbf{x}_t$, so the estimation exercise is easier and can be done by quasi maximum likelihood.

Moreover, since the dataset only contains interest rates with maturities up to one year I focus on a two factor model where these two state variables correspond with the instantaneous interest rates in each of the countries. In terms of our general framework this is translated into: $\mathbf{x}_t = [r_t, r_t^*]$, $\delta_0 = \delta_0^* = 0$, $\boldsymbol{\delta} = (1,0)'$ and $\boldsymbol{\delta}^* = (0,1)'$; where the joint process for the short rates is a multivariate Orstein-Uhlenbeck process:

$$d\begin{pmatrix} r_t \\ r_t^* \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}\left[\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} r_t \\ r_t^* \end{pmatrix}\right]dt + \begin{pmatrix} \sigma_{11} & 0 \\ \sigma_{21} & \sigma_{22} \end{pmatrix} d\mathbf{W}_t \tag{22}$$

and the prices of risks are assumed to be affine (known) functions of these state variables

$$\begin{aligned} \mathbf{\Lambda}_t &= \begin{pmatrix} \lambda_{01} \\ \lambda_{02} \end{pmatrix} + \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}\mathbf{x}_t \\ \mathbf{\Lambda}_t^* &= \begin{pmatrix} \lambda_{01}^* \\ \lambda_{02}^* \end{pmatrix} + \begin{pmatrix} \lambda_{11}^* & \lambda_{12}^* \\ \lambda_{21}^* & \lambda_{22}^* \end{pmatrix}\mathbf{x}_t \end{aligned} \tag{23}$$

4.2 Estimation

Under the above assumptions and given the results presented in the previous section, both domestic and foreign interest rates are affine functions of the set of state variables given by $\mathbf{x}_t = [r_t, r_t^*]'$:

$$\begin{aligned} r_t^{(h)} &= A(h) + B(h)'\mathbf{x}_t \\ r_t^{*(h)} &= A^*(h) + B^*(h)'\mathbf{x}_t \end{aligned} \tag{24}$$

and the expected rate of depreciation is quadratic in the same set of state variables or, if preferred, linear in the augmented set of factor given by $\widetilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{z}_t']'$ with $\mathbf{z}_t = vech(\mathbf{x}_t\mathbf{x}_t')$

$$q_t^{(h)} = C(h) + \widetilde{D}(h)'\widetilde{\mathbf{x}}_t \tag{25}$$

In the (hypothetical) absence of exchange rate data, the estimation of this model can be done by maximum likelihood (ML) exploiting that the conditional distribution of the state variables is Gaussian. For example and following the usual convention in the literature (see *e.g.* Dai and Singleton, 2002; and Duffee 2002), I assume that some of the interest rates are observed without measurement error, while the interest rates on the remaining maturities are assumed to be measured with serially uncorrelated, zero-mean errors. In particular, I assume that domestic and foreign one month interest rates do not contain any source of measurement error, which allows me to recover the state variable $\mathbf{x}_t = [r_t, r_t^*]'$ by inversion of the one-to-one mapping given in equation (24) evaluated at (time is measured in months) $h = 1$:

$$\mathbf{x}_t = H_1^{-1}(\mathbf{r}_t^{(1)} - H_0) \tag{26}$$

where $\mathbf{r}_t^{(1)} = [r_t^{(1)}, r_t^{*(1)\prime}]$, $H_0 = [A(1), A^*(1)]'$ and $H_1 = [B(1), B^*(1)]'$.

Given the value of the state vector $\mathbf{x}_t$ obtained in equation (26), the model-implied interest rates for the remaining maturities (three, six and twelve months) can be computed. Denote by $\mathbf{r}_t^{(-1)}$ a vector that contains the observed domestic and foreign interest rates on these remaining maturities, and denote their implied counterparts by $\widehat{\mathbf{r}}_t^{(-1)}$. Then, the measurement error is $\boldsymbol{\epsilon}_t = \mathbf{r}_t^{(-1)} - \widehat{\mathbf{r}}_t^{(-1)}$ and let it be *i.i.d.* zero-mean normally distributed with density given by $f_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t)$.

The loglikelihood has, then, two parts. First, the contribution of the interest rates that are observed without measurement error. This can be computed using that the conditional distribution of the state variables $\mathbf{x}_t$ is given by a first order vector autorregresion with Gaussian innovations:

$$f_{\mathbf{x}}(\mathbf{x}_{t+1} | \, \mathbf{x}_t) \sim N\left[\mu_t, \mathbf{\Omega}\right]$$

where

$$\mu_t = (I - e^{-\mathbf{\Phi}})\boldsymbol{\theta} + e^{-\mathbf{\Phi}}\mathbf{x}_t$$
$$vec\left[\mathbf{\Omega}\right] = \left[\mathbf{\Phi} \otimes I + I \otimes \mathbf{\Phi}\right]^{-1} \left[I \otimes I - e^{-\mathbf{\Phi}} \otimes e^{-\mathbf{\Phi}}\right] vec(\mathbf{\Sigma})$$

Then, the conditional density of $\mathbf{r}_t^{(1)}$ can be obtained by a change of variable:

$$f_{\mathbf{r}^{(1)}}\left(\mathbf{r}_{t+1}^{(1)} \middle| \, \mathbf{r}_t^{(1)}\right) = \frac{1}{|H_1|} f_{\mathbf{x}}\left(\mathbf{x}_{t+1} | \, \mathbf{x}_t\right)$$

Second, we have assumed that $\boldsymbol{\epsilon}_t$ is normal *i.i.d.* Thus, the loglikelihood of an observation at time t is

$$l_t(\Theta) = \log f_{\mathbf{r}^{(1)}}\left(\mathbf{r}_{t+1}^{(1)} \middle| \, \mathbf{r}_t^{(1)}\right) + \log f_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t)$$

where Θ is a vector that contains the parameters of the model. The loglikelihood of the whole sample is constructed as the usual sum of these log densities over the sample

$L_t(\Theta|\,\mathbf{r}_1^{(1)}) = \sum_t l_t(\Theta)$ where, for simplicity, I have conditioned on the first observation of the one-month interest rates.

Still, this approach does not use the information that exchange rates contain on the ratio of the SDFs (see equation 7). Therefore and to exploit such information, notice that the assumption of rational expectations in foreign exchange markets allows me to write:

$$\Delta s_{t+1} = E_t\left[\Delta s_{t+1}\right] + v_{t+1} \tag{27}$$

where v_{t+1} is a rational expectation forecasting error with zero mean and uncorrelated with any variable in the time t information set. Traditionally, the empirical literature in international finance has combined this last equation with the uncovered interest parity to obtain a testable implication of this theory (see Section 2). Instead, note that the assumption on the absence of measurement errors in the one-month interest rates implies that the "backed out" state variables $\mathbf{x}_t$ and any function of those, such as $\mathbf{z}_t = vech(\mathbf{x}_t\mathbf{x}_t')$, belong to the time t information set. Therefore, I can combine equation (27) with (25) to obtain:

$$\Delta s_{t+1} = C(1) + \widetilde{D}(1)'\widetilde{\mathbf{x}}_t + v_{t+1} \tag{28}$$

where again, $\widetilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{z}_t']'$. This equation can form the basis of an estimation by quasi maximum likelihood (QML) if we assume that v_t is $N(0, \sigma_v^2)$ and independent of W_t and the measurement errors ϵ_t. The parameter σ_v^2 can be interpreted as a general characterization of the mean squared error of the (restricted) projection of the rate of depreciation on the factors and their squares. In this case, the loglikelihood has a third component that captures the contribution of the exchange rates:

$$l_t(\Theta) = \log f_{\mathbf{r}^{(1)}}\left(\mathbf{r}_t^{(1)}\middle|\,\mathbf{r}_{t-1}^{(1)}\right) + \log f_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon}_t) + \log f_v(v_t)$$

and the estimated parameter $\widehat{\Theta}$ can be obtained maximizing the loglikelihood of the whole sample $L_t(\Theta|\,\mathbf{r}_1^{(1)}) = \sum_t l_t(\Theta)$.

QML estimation can be viewed as GMM estimation based on the scores of the quasi likelihood function (its first derivatives with respect to the parameter vector); and for the Gaussian model presented in this section the scores can be computed algebraically (see Harvey, 1989 pp 140-2 for a related example). Another advantage of this model is that it is possible to show that both the covariance matrix of the measurement errors and the σ_v^2, that is the mean squared error of the (restricted) projection of the rate of depreciation, can be concentrated out of the likelihood function. In particular, the estimates of the these two objects are $\widehat{\mathbf{\Omega}}_\epsilon = T^{-1}\sum \widehat{\boldsymbol{\epsilon}}_t\widehat{\boldsymbol{\epsilon}}_t'$ and $\widehat{\sigma}_v^2 = T^{-1}\sum \widehat{v}_t^2$ respectively. Moreover, the QML approach allows us to compute the standard errors of the parameters using the standard GMM *formulae* (see Hamilton, 1994 pp 428-9).

Although somewhat extreme, the assumption on the independence of the error term v_t can be justified from either estimation simplicity or from the empirical observation that a predominant portion of the exchange rate movement is independent of the interest rate movements of either country. For example, Lothian and Wu (2003) point out that the forward premium regressions usually show very low R^2 statistics. This feature can be accounted in our model once we allow the pricing kernels to be driven by an additional source of risk that is orthogonal to the forces driving the short rates[11]. Still, the expressions for the expected rate of depreciation in equation (14) are valid because this new source of risk is orthogonal to the interest rates. Therefore, the estimation is based on the right conditional moments so this approach is expected to deliver consistent estimates of the parameters of the model.

4.3 Results

Table 3 presents the QML estimates of the two factor Gaussian essentially affine model, along with robust estimates of the corresponding standard errors. Panel a contains the results of the estimation exercise for the U.S. Dollar – Sterling Pound, while panel b contains those of the U.S. Dollar – Canadian Dollar. Since the objective of this study is to investigate the ability of affine models in the out-of-sample prediction of exchange rates, I follow Dai and Singleton (2002) and Duffee (2002) to re-estimate the model after setting to zero those coefficients with largest relative standard errors (ϕ_{12}, σ_{21}, λ_{11}, λ_{02}^* for the pair U.S. Dollar – Sterling Pound and ϕ_{12}, σ_{21}, λ_{11}, λ_{12}, λ_{01}, λ_{21}^*, λ_{22}^*, λ_{02}^* for the pair U.S. Dollar – Canadian Dollar). This approach will help us in reducing the large number parameters that the model contains and therefore will reduce excessive parameter estimation uncertainty that may adversely affect the out-of-sample forecasts performance of this model. For the sake of saving space, I only report the results for these restricted models.

[11] Brandt and Santa-Clara (2002) attribute this additional source of risk to market incompleteness while Dewachter and Maes (2001) or Leippold and Wu (2003) assume that bond returns do no necessarily span the returns in the foreign exchange market and, therfore, interpret this orthogonal risk factor as related to factors outside the bond market (*e.g.* stocks).

The results for the U.S. Dollar – Sterling Pound (panel a) indicates that the process for the short rates is mean reverting. However, this mean-reversion is slow (both elements in the diagonal of $\mathbf{\Phi}$ are positive but close to zero). The British short rate seems to revert to a higher level than its American counterpart and, in addition, it is more volatile. Since the process that short rates follow is Gaussian, it is possible for them to take on negative values with positive probability. Still, I find that, with estimated parameters, the probability of a negative short rate is small: 2.59% and 0.30% for the case of the U.S. and U.K. respectively. The implied yield curve for the U.S. is upward sloping with an implied long-term yield of 17.75%. On the contrary, the implied yields curve for the U.K. is downward sloping with long-term yields reaching as high as 9.54%. These seem to be reasonable numbers. First, Backus, *et al.* (2001) reports that their estimates implied long-term yields reaching as high as 80%. Second, note that the estimation is done without interest rate data with maturities higher than one year, being those the ones that can potentially help to anchor these long-term yields. In addition, this result is consistent with an (average) spread between the one year and the one month interest rate that is positive for the case of the U.S. while negative for the case of the U.K. (see Table 1).

Focusing on the estimates of the prices of risk, notice that almost all the elements of the matrices $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^*$ in the prices of risk are statistically different from zero. Therefore, and in the line of the work in Dai and Singleton (2002), extending the specification of the prices of risk seems to be an important factor for the estimation of the model. In addition, this model is able to reproduce the forward premium puzzle. Table 4 (panel a) presents the term structure of forward premium regression slopes implied by the model. These are computed using the closed-form *formulae* derived in the appendix and by treating the estimates of the two factor model as truth. The sample OLS estimates of these slopes are reproduced here again for the sake of comparison. The model-implied slopes are negative and reasonably close to their sample counterparts. However, the model tends to generate uncovered interest parity slopes that are more negatives than the ones that we estimated using standard OLS techniques. Therefore, the results produced by this model for the case of the U.S. Dollar –Sterling Pound do not seem to be highly unreasonable.

The results for the U.S. Dollar – Canadian Dollar are presented in Table 3, panel b. Again, the process for the short rates is mean reverting, and the Canadian short term seems to revert to a higher level than its American counterpart. However, and if compared with the estimates obtained for the U.S. Dollar – Sterling Pound, the long run mean for the short rates (given by $\boldsymbol{\theta}$) is unusually high. In fact, both implied yield curves for the U.S. and Canada are downward sloping with implied long-term yields unreasonable low: 1.03% and 3.84% for U.S. and Canada, respectively. On the contrary, the probability of having a negative short rate is almost zero regardless the choice of the country.

Still, the model is able to reproduce the forward premium anomaly. Again, almost all the elements of the matrices $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^*$ are statistically different from zero and the model-implied slopes are negative (Table 4, panel b). The model underestimates (in absolute terms) the slope for the case of the 1-month and 3-month contracts, while the implied slope is more negative than the OLS counterparts for the 6-month and 12-month contracts. Therefore, and contrary to the results for the pair U.S. Dollar – Sterling Pound, here it seems that the good fit of the exchange rate is done at the expense of the interest rates.

4.4 Out-of-Sample Forecasting

In this subsection, an out-of-sample forecasting experiment is conducted over the period January 1998 – December 2004 (7 years) to evaluate the performance of the Gaussian essentially affine term structure model estimated and described in the previous section. These forecasts are computed according to the recursive procedure employed in Clarida and Taylor (1997) and Clarida *et al.* (2003): at each date t, the model is re-estimated using data up to and including time t and then dynamic forecasts of the spot exchange rate up to $t + 12$. These forecasts are computed using equation (14).

The first column in table 5 presents the results of the accuracy of these forecasts using the Root-mean-square error (RMSE) and Mean-absolute-error (MAE) criteria. Panel a contains the results of the forecasting exercise for the U.S. Dollar – Sterling Pound, while panel b contains those of the U.S. Dollar – Canadian Dollar. In addition, this table also presents a comparison of these forecasts with those generated by three alternative benchmarks: a random walk (RW), a vector autoregression on the forward premia and the rate of depreciation (VAR), and the forward premium regression (OLS). Comparing our forecasts with those produced by the random walk model can be motivated by the fact that the random walk model is considered the usual metric in which exchange rate forecasts have been evaluated since the original work of Meese and Rogoff (1983a, b). However, Clarida and Taylor (1997) show that if one uses a linear vector error correction (VECM) model in the spot and forward exchange rates, it is possible to obtain out-of-sample forecasts of spot exchange rates that beat the random-walk model. Therefore, we include as a second benchmark the forecasts obtained by the use of a vector autoregression (VAR) on the forward premia and the rate of depreciation[12] where the number of lags in the VAR is chosen to be equal to $p = 2$ for UK and $p = 1$ for Canada as suggested by the Bayesian Information Criteria (BIC)[13]. Finally, and for completeness, I also include the forecast produced by a standard ordinary least squares regression of the rate of depreciation onto a constant and the lagged forward premium.

Following Clarida and Taylor (1997), I report the level of the RMSE and the MAE for the affine term structure model, while for the alternative forecasts the results are expressed as the ratio of the RMSE or the MAE to that obtained by the alternative method. For example, the level of the RMSE of the affine forecast for the U.S. Dollar-Sterling Pound rate one year ahead is 0.0496, while the ratio of this to the forecast obtained using a random-walk forecast is 0.637. That means a 36.3% reduction in RMSE by using the forecasts produced by the affine term structure model as opposed to the random walk.

The results for the U.S. Dollar – Sterling Pound (panel a) indicate that the affine term structure model produces the best out-of-sample forecasts among the four competing models. The linear vector autoregression and the random walk model rank second and third respectively. On the other side of the spectrum, the forward premium regression forecasts fails to outperform the random walk at all the four horizons. In addition, the improvement of the affine term structure model forecasts with respect to those obtained from alternative models grows with the forecast horizon. For example, the improvement in RMSE with respect to the random-walk at the 1-month horizon is 2.2%, at the 3-month horizon is 9.7%, at the 6-month horizon is 18.2% and, finally, at the 12-month horizon is 36.3%.

That the VAR forecasts are also able to beat the random walk is consistent with the results found by Clarida and Taylor (1997). However, these gains are smaller than those reported by those authors. For example, the improvement of the forecasts in RMSE at the one year horizon produced by the VAR if compared with those of a random walk is only a

[12] This approach is equivalent to the vector error correction model (VECM) in Clarida and Taylor (1997) if we impose that the spot and the forward exchange rates are cointegrated with known cointegration vector (1, -1). See *i.e.* Mark (2001), pp 51.

[13] The results provided below are qualitatively similar to those obtained with other choices of the number of lags.

2.4%.

Similarly to what it is found for the U.S. Dollar - Sterling Pound exchange rate, the affine term structure model produces the best out-of-sample forecasts of the U.S. Dollar – Canadian Dollar among the four competing models (panel b). However, the VAR model fails to outperform the random walk forecasts at all the horizons. Thus, the random walk model ranks second and the VAR and forward premium forecasts rank third and fourth, respectively. Again, the improvement of the affine term structure model forecasts with respect to those obtained from alternative models grows with the forecast horizon. In particular, at the 1-month horizon, the improvement in RMSE with respect to the random-walk is 1.6%, at the 3-month horizon is 4.0%, at the 6-month horizon is 6.2%, and at the 12-month horizon is 9.4%. Being this gain quantitatively smaller than the one reported for the U.S. Dollar - Sterling Pound exchange rate, it must be noted that this model is successful in beating the random walk model while the linear VAR is not. In particular, at the 1-month horizon, the improvement in RMSE with respect to the VAR is 2.5%, at the 3-month horizon is 5.4%, at the 6-month horizon is 8.0%, and at the 12-month horizon is 13.1%.

These results extend those of Clarida and Taylor (1997) who, using a linear VECM model in the spot and forward exchange rates, were able to obtain out-of-sample forecasts of spot exchange rates that beat the random-walk model. In addition, the results presented in this paper also extend those in Ang and Piazzesi (2003) who, imposing the cross-equation restrictions from no-arbitrage, are able to beat the random walk when they forecast bond yields, again, out-of-sample. Therefore, imposing the cross-equation restrictions from no-arbitrage can help us in extracting information contained in the term structure of forward exchange premia that is useful to forecast exchange rates.

5 Final Remarks

This paper provides an arbitrage-free empirical model that produces exchange forecasts that are superior to those produced by purely time-series method such as a random-walk model or a vector autoregression on the forward premiums and the rate of depreciation. The intuition behind this success is that imposing no-arbitrage restrictions in the estimation of the joint dynamics of interest and exchange rates reduces the large number of parameters that characterize traditional time series models. Consequently, it also reduces excessive parameter estimation uncertainty that may adversely affect the out-of-sample forecasting performance of a purely time-series model.

Several questions are left for further research. The first one is the role of nonlinearities. Clarida, *et al.* (2003) shows that there is strong evidence of the presence of nonlinearities in the joint behaviour of interest and exchange rates that can be used to outperform the forecasts obtained by using linear methods. Therefore, imposing no-arbitrage restrictions in non-linear models can potentially improve over the results presented in this paper. A second possible extension is the use of an arbitrage-free joint model of interest, exchange rates and macro variables to extract the information that interest rates and macro variables contain about the future evolution of exchange rates. Along these lines, Dong (2005) presents a structural VAR identified by the assumption of the absence of arbitrage where the macro variables correspond to output gap and inflation; and where the correlation between the model-implied rate of depreciation and the data is over 60%. However, this author does not conduct an out-of-sample prediction exercise. Another possible extension is the use of commodity prices, such as oil prices, as another macro variable that can potentially help us to predict exchange rate movements, especially for a commodity currency such as the Canadian Dollar (see Amano and Van Norden, 1998 and Chen and Rogoff, 2003).

Appendix

A Proof of Proposition 1

In this appendix, I show that if the drift of the log exchange rate s_t is linear in a set of state variables $\mathbf{x}_t$, and $\mathbf{x}_t$ follows an affine diffusion, then the expected rate of depreciation is a (known) linear function of the state vector $\mathbf{x}_t$. However, to show this point I need one previous result.

Lemma 2 *If the process* $\mathbf{x}_t$ *follows the affine diffusion given by (13) then*

$$E_t\left[\int_t^{t+h}\mathbf{x}_\tau d\tau\right] = \boldsymbol{\theta}h + \boldsymbol{\Phi}^{-1}\left[I - e^{-\boldsymbol{\Phi}h}\right]\left[\mathbf{x}_t - \boldsymbol{\theta}\right] \tag{29}$$

Proof. First note that (see *e.g.* Fackler, 2000) when $\mathbf{x}_t$ follows an affine diffusion:

$$E_t\mathbf{x}_{t+h} = \boldsymbol{\theta} + e^{-\boldsymbol{\Phi}h}(\mathbf{x}_t - \boldsymbol{\theta})$$

Second, take expectations respect to the integral form of (13):

$$E_t\left[\int_t^{t+h}d\mathbf{x}_\tau\right] = \boldsymbol{\Phi\theta} + \boldsymbol{\Phi}E_t\left[\int_t^{t+h}\mathbf{x}_\tau d\tau\right]$$

Finally, notice that $E_t\left[\int_t^{t+h}d\mathbf{x}_\tau\right] = E_t\mathbf{x}_{t+h} - \mathbf{x}_t$ and solve this last equation for $E_t\left[\int_t^{t+h}\mathbf{x}_\tau d\tau\right]$ to obtain (29) ■

Note that the variable inside the conditional expectation $\mathbf{y}_t^{(h)} = \int_t^{t+h}\mathbf{x}_\tau d\tau$ is a flow variable. In particular, there has been a lot of attention in obtaining the process that a set of discretely sampled data follows when these observations (whether stock or flow, or a combination of both) have been generated by an underlying continuous time model (see, *e.g.* Bergstrom 1984). However, this literature has relied on the assumption that this underlying continuous time process is a multivariate version of the Orstein-Uhlenbeck process. Here, the distributional assumption is relaxed to allow for an affine diffusion at the cost of restricting the predictions only to the conditional expectation of the flow variable (instead of the whole distribution of $\mathbf{y}_t^{(h)}$). Nonetheless, this result is enough to prove the linearity of the expected rate of depreciation because the expected rate of depreciation satisfies

$$E_t\left[s_{t+h} - s_t\right] = E_t\left[\int_t^{t+h}ds_\tau\right] = \gamma_0 h + \boldsymbol{\gamma}'E_t\left[\int_t^{t+h}\mathbf{x}_\tau d\tau\right]$$

and once we substitute (29) into this last expression we obtain the desired result.

B Closed-form Expressions of the Implied Uncovered Interest Parity Regression Slope

For expositional purposes collect the expected rate of depreciation and the forward premia in a vector $\mathbf{y}_t^{(h)} = [q_t^{(h)}, p_t^{(h)}]$. In addition, denote $\mathbf{\Omega}(h) = Var[\mathbf{y}_t^{(h)}]$ as the unconditional variance of $\mathbf{y}_t^{(h)}$.

The definition of the (population) uncovered interest parity regression slope in equation (1) is

$$b(h) = \frac{Cov[s_{t+h} - s_t, p_t^{(h)}]}{Var[p_t^{(h)}]} = \frac{Cov[q_t^{(h)}, p_t^{(h)}]}{Var[p_t^{(h)}]}$$

where the second equality comes from the law of iterated expectations. For convenience, rewrite this expression as:

$$b(h) = \frac{e_1'\mathbf{\Omega}(h)e_2}{e_2'\mathbf{\Omega}(h)e_2}$$

where $e_1' = (1, 0)$ and $e_2' = (0, 1)$. The numerator of this equation is the covariance between the expected rate of depreciation and the forward premia, while the denominator is the variance of the forward premia.

However, given that $\mathbf{y}_t^{(h)}$ is a linear function of the states $\mathbf{x}_t$:

$$\mathbf{y}_t^{(h)} = \mathbf{\Psi}_0(h) + \mathbf{\Psi}_1(h)'\mathbf{x}_t$$

with

$$\begin{aligned}
\mathbf{\Psi}_0(h) &= \begin{bmatrix} C(h) \\ A(h) - A^*(h) \end{bmatrix} \\
\mathbf{\Psi}_1(h) &= \begin{bmatrix} D(h) \\ B(h) - B^*(h) \end{bmatrix}
\end{aligned}$$

it is straightforward to realize that the unconditional variance of $\mathbf{y}_t^{(h)}$ is related to the unconditional variance of the factors in the following way:

$$\mathbf{\Omega}(h) = Var[\mathbf{z}_t^{(h)}] = \mathbf{\Psi}_1(h)'Var[\mathbf{x}_t]\mathbf{\Psi}_1(h)$$

Therefore, computing the implied uncovered interest parity regression slope amounts to compute the unconditional variance of the factor $\mathbf{x}_t$. If the state-vector $\mathbf{x}_t$ follows an affine diffusion (as in the case of the affine models of currency pricing), we can use the explicit *formulae* for the unconditional variance provided in Fackler (2000). Specializing his results to our case, one obtains:

$$vec\,[Var(\mathbf{x}_t)] =$$
$$[\mathbf{\Phi} \otimes I + I \otimes \mathbf{\Phi}]^{-1} \left[\mathbf{\Sigma}^{1/2} \otimes \mathbf{\Sigma}^{1/2}\right] vec\,[diag\,[\mathbf{a} + \mathbf{B}\boldsymbol{\theta}]]$$

being

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} \beta_1' \\ \beta_2' \\ \vdots \\ \beta_N' \end{bmatrix}$$

For the quadratic models we can exploit that the unconditional distribution of the original state-vector $\mathbf{x}_t$ is Gaussian with mean $\boldsymbol{\theta}$ and variance such that $vec\,[Var(\mathbf{x}_t)] = [\mathbf{\Phi} \otimes I + I \otimes \mathbf{\Phi}]^{-1} vec(\mathbf{\Sigma})$. Still, we need the covariances between $\mathbf{x}_t$ and $\mathbf{z}_t = vech\,[\mathbf{x}_t\mathbf{x}_t']$. The elements of this matrix can be computed using that if three variables x, y and z are jointly normally distributed then:

$$Cov(xy, z) = \mu_x Cov(y, z) + \mu_y Cov(x, z)$$

Finally, the variance-covariance matrix of $\mathbf{z}_t$ can be computed using that if four variables x, y, u and v are jointly normally distributed then:

$$\begin{gathered} Cov(xy, uv) = \mu_x\mu_u Cov(y, v) + \mu_y\mu_u Cov(x, v) + \\ \mu_x\mu_v Cov(y, u) + \mu_y\mu_v Cov(x, u) + \\ Cov(y, u)Cov(x, v) + Cov(x, u)Cov(y, v) \end{gathered}$$

C From Quadratic to Affine in an Augmented Set of State Variables

Quadratic models of the term structure generate interest rates that are quadratic (known) functions of a set of state variables denoted by $\mathbf{x}_t = [x_{1t}, x_{2t}, \ldots, x_{nt}]'$:

$$r_t^{(*)(h)} = A^{(*)}(h) + B^{(*)}(h)'\mathbf{x}_t + \mathbf{x}_t' C^{(*)}(h)\mathbf{x}_t$$

where the coefficients $A^{(*)}(h) \in \mathbb{R}$, $B^{(*)}(h) \in \mathbb{R}^n$ and $C^{(*)}(h) \in \mathbb{R}^{n\times n}$ solve two systems of ordinary differential equations. As claimed in the paper, this expression can be expressed as an affine function of the original set of state variables and a new set of state variables $\mathbf{z}_t = vech\,[\mathbf{x}_t\mathbf{x}_t']$ that includes the squares of the original ones as well as the corresponding cross-products. To realize why this is true, first note that for a given $n \times n$ matrix $\mathbf{\Gamma}$ it can be shown that

$$\mathbf{x}_t'\mathbf{\Gamma}\mathbf{x}_t = tr(\mathbf{x}_t'\mathbf{\Gamma}\mathbf{x}_t) = tr(\mathbf{\Gamma}\mathbf{x}_t\mathbf{x}_t')$$

Then, use that $tr(\mathbf{\Gamma}\mathbf{x}_t\mathbf{x}_t') = vec(\mathbf{\Gamma})'vec(\mathbf{x}_t\mathbf{x}_t')$; and notice that $\mathbf{x}_t\mathbf{x}_t'$ is a $n \times n$ symmetric matrix so that, $vec(\mathbf{x}_t\mathbf{x}_t') = \mathbf{D}_n vech(\mathbf{x}_t\mathbf{x}_t')$, where $\mathbf{D}_n$ is the duplication matrix and whose details can be found in *pp.* 464-5 of Lütkepohl (1993). This makes:

$$\mathbf{x}_t'\mathbf{\Gamma}\mathbf{x}_t = vec(\mathbf{\Gamma})'\mathbf{D}_n vech(\mathbf{x}_t\mathbf{x}_t') = vec(\mathbf{\Gamma})'\mathbf{D}_n\mathbf{z}_t \tag{30}$$

If we specialize this result to the case when $\mathbf{\Gamma} = B_2^{(*)}(h)$, it delivers interest rates that are (known) affine in the augmented set of state variables given by $\widetilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{z}_t']'$ with $\mathbf{z}_t = vech(\mathbf{x}_t\mathbf{x}_t')$:

$$r_t^{(*)(h)} = A^{(*)}(h) + B_1^{(*)}(h)'\mathbf{x}_t + \left[vec\left(B_2^{(*)}(h)\right)\right]' \mathbf{D}_n\mathbf{z}_t$$

By a similar reasoning, it can be shown that the expected rate of depreciation is also affine in this augmented set of factors. This property requires the drift of the (log) exchange rate process to be affine in this new set of state variables and the process that these variables follow must be itself an affine diffusion. First, we can express the short rates and the instantaneous variance of the domestic and foreign log SDF as affine functions of $\mathbf{x}_t$ and $\mathbf{z}_t$ using equation (30) with $\mathbf{\Gamma} = \boldsymbol{\delta}_2^{(*)}$ and $\mathbf{\Gamma} = \boldsymbol{\lambda}^{(*)\prime}\boldsymbol{\lambda}^{(*)}$ respectively. This implies that the drift of the (log) exchange rate process is affine:

$$E_t ds_t = (\gamma_0 + \boldsymbol{\gamma}_{\mathbf{x}}'\mathbf{x}_t + \boldsymbol{\gamma}_{\mathbf{z}}'\mathbf{z}_t)dt \tag{31}$$

with

$$\begin{aligned}
\gamma_0 &= (\delta_0 - \delta_0^*) + \frac{1}{2}(\boldsymbol{\lambda}_0'\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0^{*\prime}\boldsymbol{\lambda}_0^*) \\
\boldsymbol{\gamma}_{\mathbf{x}} &= (\boldsymbol{\delta}_1 - \boldsymbol{\delta}_1^*) + (\boldsymbol{\lambda}_0'\boldsymbol{\lambda} - \boldsymbol{\lambda}_0^{*\prime}\boldsymbol{\lambda}^*) \\
\boldsymbol{\gamma}_{\mathbf{z}} &= \left[vec(\boldsymbol{\delta}_2 - \boldsymbol{\delta}_2^* + \frac{1}{2}(\boldsymbol{\lambda}'\boldsymbol{\lambda} - \boldsymbol{\lambda}^{*\prime}\boldsymbol{\lambda}^*))\right]' \mathbf{D}_n
\end{aligned}$$

or in compact form:

$$E_t ds_t = (\gamma_0 + \widetilde{\boldsymbol{\gamma}}'\widetilde{\mathbf{x}}_t)dt \tag{32}$$

with $\widetilde{\mathbf{x}}_t = [\mathbf{x}_t', \mathbf{z}_t']'$ and $\widetilde{\boldsymbol{\gamma}} = [\boldsymbol{\gamma}_{\mathbf{x}}', \boldsymbol{\gamma}_{\mathbf{z}}']'$.

Second, it can be shown that if we apply Itô's lemma on $\mathbf{z}_t = vech\left[\mathbf{x}_t\mathbf{x}_t'\right]$ then the joint process for $\mathbf{x}_t$ and $\mathbf{z}_t$ is an affine diffusion (see appendix B in Cheng and Scaillet, 2002). In particular, the law of motion of the augmented set of factors $\widetilde{\mathbf{x}}_t$ satisfies:

$$d\begin{bmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi} & 0 \\ \mathbf{\Phi}_{\mathbf{zx}} & \mathbf{\Phi}_{\mathbf{zz}} \end{bmatrix} \left[\begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\theta}_{\mathbf{z}} \end{pmatrix} - \begin{pmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{pmatrix}\right] dt + \begin{bmatrix} \mathbf{\Sigma}^{1/2} \\ \mathbf{\Sigma}_{\mathbf{z}}(\mathbf{x}_t)^{1/2} \end{bmatrix} d\mathbf{W}_t$$

or in compact notation:

$$d\widetilde{\mathbf{x}}_t = \widetilde{\mathbf{\Phi}}(\widetilde{\boldsymbol{\theta}} - \widetilde{\mathbf{x}}_t)dt + \widetilde{\mathbf{\Sigma}}(\mathbf{x}_t)^{1/2} d\mathbf{W}_t$$

and where the drift is linear with:

$$\begin{aligned}
\mathbf{\Phi}_{\mathbf{zz}} &= 2\mathbf{D}_n^+(\mathbf{\Phi} \otimes \mathbf{I}_n)\mathbf{D}_n \\
\mathbf{\Phi}_{\mathbf{zx}} &= -2\mathbf{D}_n^+(\mathbf{\Phi}\boldsymbol{\theta} \otimes \mathbf{I}_n) \\
\boldsymbol{\theta}_{\mathbf{z}} &= \mathbf{\Phi}_{\mathbf{zz}}^{-1}\left(vech(\mathbf{\Sigma}) - \mathbf{\Phi}_{\mathbf{zx}}\boldsymbol{\theta}\right)
\end{aligned}$$

being $\mathbf{D}_n^+$ the Moore-Penrose inverse of matrix $\mathbf{D}_n$: $\mathbf{D}_n^+ = (\mathbf{D}_n'\mathbf{D}_n)^{-1}\mathbf{D}_n'$.

In addition, the diffusion term satisfies:

$$\mathbf{\Sigma}_{\mathbf{z}}(\mathbf{x}_t)^{1/2} = 2\mathbf{D}_n^+(\mathbf{\Sigma}^{1/2} \otimes \mathbf{x}_t)$$

which implies a volatility matrix $\widetilde{\mathbf{\Sigma}}$ whose elements are affine in $\mathbf{x}_t$ and $\mathbf{x}_t\mathbf{x}_t'$ (and therefore, affine in $\mathbf{x}_t$ and $\mathbf{z}_t$):

$$\widetilde{\mathbf{\Sigma}} = \begin{pmatrix} \mathbf{\Sigma} & 2(\mathbf{\Sigma}' \otimes \mathbf{x}_t')\mathbf{D}_n^+ \\ 2\mathbf{D}_n^+(\mathbf{\Sigma} \otimes \mathbf{x}_t) & 4\mathbf{D}_n^+(\mathbf{\Sigma} \otimes \mathbf{x}_t\mathbf{x}_t')\mathbf{D}_n^{+\prime} \end{pmatrix}$$

References

[1] Amano, R. and S. Van-Norden (1998): "Exchange Rates and Oil Prices", *Review of International Economics*, 6, 4, 683-694.

[2] Ang, A, and M. Piazzesi (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables", *Journal of Monetary Economics*, 50, 745-787.

[3] Anh, D.-H. (2004): "Common Factors and Local Factors: Implications for Term Structures and Exchange Rates", *Journal of Financial and Quantitative Analysis*, 39, 69-102.

[4] Anh, D.-H., R. F. Dittmar and A. R. Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence", *Review of Financial Studies*, 15, 242-288.

[5] Backus D. K., S. Foresi and C. I. Telmer (2001): "Affine Term Structure Models and the Forward Premium Anomaly", *Journal of Finance*, 51, 279-304.

[6] Bekaert, G. and R. Hodrick (2001): "Expectations Hypotheses Tests", *Journal of Finance*, 56, 4, 1357-1393.

[7] Bergstrom, A. (1984): "Continuous Time Stochastic Models and Issues of Aggregation Over Time" in *Handbook of Econometric*s, vol 2, edited by Z. Griliches and M. D. Intriligator. Amsterdam: North-Holand, 1984.

[8] Brandt M. W. and P. Santa-Clara (2002): "Simulated Likelihood Estimation of Diffussions with an Application to Exchange Rate Dynamics in Incomplete Markets", *Journal of Financial Economics*, 63, 161-210.

[9] Brennan, M.J., and Y. Xia (2004): "International Capital Markets and Foreign Exchange Risk", Wharton School of University of Pennsylvania Mimeo

[10] Chen, Y. and K. Rogoff: "Commodity Currencies", *Journal of International Economics*, 60, 1, 133-160.

[11] Cheng, P. and O. Scaillet (2002): "Linear-Quadratic Jump-Diffusion Modeling with Application to Stochastic Volatility", FAME Research Paper Series No 67.

[12] Cheridito, P., D. Filipovic and R. L. Kimmel (2005): "Market Price of Risk Specifications for Affine Models: Theory and Evidence", forthcoming in *Journal of Financial Economics*

[13] Clarida, R. H., L. Sarno, M. P. Taylor and G. Valente (2003): "The out-of-sample success of term structure models as exchange rate predictors: a step beyond", *Journal of International Economics,* 60, 1, 61-83.

[14] Clarida, R. H., and M. P. Taylor (1997): "The Term Structure of Forward Exchange Premiums and the Forecastability of Spot Exchange Rates: Correcting the Errors", *Review of Economics and Statistics*, 89, 353-361.

[15] Cochrane, J. (2001): *Asset Pricing*, Princeton, N.J., Princeton University Press.

[16] Cox, J., J. Ingersoll and S. Ross (1985): "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, 385-407.

[17] Dai, Q. and K. J. Singleton (2000): "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, 55, 1943-78.

[18] Dai, Q. and K. J. Singleton (2002): "Expectations Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure", *Journal of Financial Economics*, 63, 415-411.

[19] Dai, Q. and K. J. Singleton (2003): "Term Structure Dynamics in Theory and Reality", *Review of Financial Studies*, 16, 631-78.

[20] Dai, Q. and K. J. Singleton (2003b): "Fixed Income Pricing", in *Handbook of Economics and Finance*, ed. by Constantinides C., M. Harris and R. Stulz: North Holland.

[21] Dewachter, H. and K. Maes (2001): "An Admissible Affine Model for Joint Term Structure Dynamics of Interest Rates", University of Leuven Mimeo.

[22] Dong, S. (2005): "Monetary Policy Rules and Exchange Rates: A Structural VAR Identified by No Arbitrage", Columbia University Mimeo

[23] Duarte, J. (2004): "Evaluating An Alternative Risk Preference in Affine Term Structure Models", *Review of Financial Studies*, 17, 370-404.

[24] Duffee, G. R. (2002): "Term Premia and Interest Rate Forecasts in Affine Models", *Journal of Finance*, 57, 405-443.

[25] Duffie, D. and R. Kan (1996): "A Yield-Factor Model of Interest Rates", Mathematical Finance, 6, 379-406.

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

[27] Fackler, P.L. (2000): "Moments of Affine Diffusions", North Carolina State University Mimeo.

[28] Frachot, A. (1996): "A Reexamination of the Uncovered Interest Rate Parity Hypothesis", *Journal of International Money and Finance*, 15, 419-437.

[29] Fama, E. (1984): "Forward and Spot Exchange Rates", *Journal of Monetary Economics*, 14, 319-338.

[30] Gouriéroux, C. and R. Sufana (2003): "Wishart Quadratic Term Structure Models", Cahiers du CREF 03-10

[31] Graveline, J.J. (2005): "Exchange Rate Volatility and the Forward Premium Puzzle", Stanford University Mimeo

[32] Guimarães, R.P. (2006): "Taking Arbitrage Restrictions on Interest Rates and Exchange Rates Seriously", Princeton University Mimeo

[33] Hansen L.P. and R. Hodrick (1983): "Risk Averse Speculation in the Forward Foreign Exchange Market: An Econometric Analysis of Linear Models" in *Exchange Rates and International Macroeconomics*, ed. by J.A. Frenkel. Chicago: University of Chicago Press for National Bureau of Economic Research.

[34] Hodrick, R. (1987): *The empirical evidence on the efficiency of forward and futures foreign exchange markets*, Harwood Academic Publishers, Chur, Switzerland.

[35] Hodrick, R. and M. Vassalou (2000): "Do We Need Multi-Country Models to Explain Exchange Rate and Interest Rate and Bond Return Dynamics?, *Journal of Economic Dynamics and Control*, 26, 1275-99.

[36] Inci, A. C. and B. Lu (2004): "Exchange Rates and Interest Rates: Can term Structure Models Explain Currency Movements?", *Journal of Economic Dynamics and Control*, 28, 1595-1624

[37] Leippold, M. and L. Wu (2002): "Asset Pricing under the Quadratic Class", *Journal of Financial and Quantitative Analysis*, 37, 271-295.

[38] Leippold, M. and L. Wu (2003): "Design and Estimation of Multi-Currency Quadratic Models", University of Zurich Mimeo.

[39] Meese, R. A. and Rogoff (1983a): "Empirical Exchange Rate Models of the Seventies", *Journal of International Economics* 14, 3-24.

[40] Meese, R. A. and Rogoff (1983b): "The Out-Of-Sample Failure of Empirical Exchange Rate Models: Sampling Error of Misspecification" in *Exchange Rates and International Macroeconomics*, ed. by J.A. Frenkel. Chicago: University of Chicago Press for National Bureau of Economic Research.

[41] Piazzesi, M. (2003): "Affine Term Structure Models", forthcoming in *Handbook of Financial Econometrics*, edited by Y. Ait-Sahalia and L. P. Hansen.

[42] Saá-Requejo, J. (1993): "The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets", INSEAD Mimeo.

Table 1
Summary Statistics for All Variables

Panel a: Summary Statistics

	Mean	Std. Deviation	Autocorr.
I. Depreciation Rate $s_{t+1} - s_t$			
Sterling Pound	-0.940	40.073	0.080
Canadian Dollar	-1.557	15.761	-0.059
II. Interest Rates r_t			
U.S.			
1-month	7.966	3.558	0.969
3-months	8.046	3.488	0.973
6-months	8.103	3.372	0.974
12-months	8.073	3.020	0.977
U.K.			
1-month	10.571	3.437	0.959
3-months	10.526	3.294	0.961
6-months	10.388	3.082	0.963
12-months	10.100	2.715	0.965
Canada			
1-month	9.142	3.628	0.978
3-months	9.199	3.513	0.979
6-months	9.183	3.319	0.978
12-months	9.069	2.987	0.978

Data are monthly and the sample is January 1976 to December 1997 (252 observations). All variables are measured in percentage points per year, and monthly rates of depreciation are annualized by multiplying by 1,200.

Panel b: Forward Premium Regressions

	$a^{(h)}$	$b^{(h)}$	$H_0 : b^{(h)} = 1$
I. Sterling Pound			
1-month	-5.760	-1.840	8.996
	(3.044)	(0.947)	[0.003]
3-months	-4.700	-1.505	7.215
	(2.937)	(0.933)	[0.007]
6-months	-4.218	-1.361	7.573
	(2.647)	(0.858)	[0.006]
12-months	-3.101	-0.817	6.358
	(2.249)	(0.721)	[0.012]
II. Canadian Dollar			
1-month	-3.172	-1.351	32.337
	(0.904)	(0.414)	[0.000]
3-months	-2.526	-0.827	21.941
	(0.818)	(0.390)	[0.000]
6-months	-1.927	-0.425	16.349
	(0.691)	(0.352)	[0.000]
12-months	-1.680	-0.240	8.856
	(0.622)	(0.417)	[0.003]

Panel c: Wald Test of the Joint Equality
of the Four Forward Premium Regression Slopes

$H_0 : b^{(h)} = 1$	$\forall h = 1, 3, 6, 12$
Sterling Pound	12.990
	[0.011]
Canadian Dollar	36.821
	[0.000]

Data are monthly and the sample is January 1976 to December 1997 (252 observations). Forward premium regressions are of the form $s_{t+h} - s_t = a^{(h)} + b^{(h)} p_t^{(h)} + u_{t+h}$ where $p_t^{(h)}$ is the interest rate differential $r_t^{(h)} - r_t^{*(h)}$ (also known as the forward premium). This equation is estimated by GMM and Newey-West standard errors are presented in parenthesis. The last column $H_0 : b^{(h)} = 1$ in panel b presents the value of the Wald test of the null hypothesis that the slope coefficient is equal to one. In large samples, this test is distributed as a χ^2 with one degree of freedom. Panel c presents an equivalent Wald test of the null hypothesis that all the four slope coefficienst are equal to one. In large samples, this test is distributed as a χ^2 with four degrees of freedom. *P-values* are presented in brackets.

Table 2
Summary of the Properties of *Internationally Affine* Models

		Interest Rates $r_t^{(h)}$, $r_t^{*(h)}$	
		Linear	*Quadratic*
Expected Rate of Depreciation $q_t^{(h)}$	*Linear*	"Completely" Affine DTSM	
	Quadratic	Gaussian "Essentially" Affine DTSM	Quadratic DTSM

Table 3
Estimates of the Two Factor Essentially Affine Model:
Panel a: U.S. Dollar - Sterling Pound

	Index Number (i)	
Parameter	1	2
ϕ_{1i}	0.0238	0
	(0.0079)	
ϕ_{2i}	-0.0785	0.0935
	(0.0324)	(0.0266)
θ_i	0.6745	0.9006
	(0.1977)	(0.1832)
σ_{1i}	0.0756	0
	(0.0082)	
σ_{2i}	0	0.0862
		(0.0065)
λ_{1i}	0	0.1261
		(0.0527)
λ_{2i}	-2.0673	-0.7399
	(0.8182)	(0.6636)
λ_{0i}	-0.2412	5.9025
	(0.0637)	(0.9747)
λ^*_{1i}	1.1885	0.6846
	(0.6829)	(0.3947)
λ^*_{2i}	1.3813	-0.4585
	(0.6970)	(0.3224)
λ^*_{0i}	-5.8778	0
	(0.9960)	

This table presents quasi maximum likelihood (QML) estimates of the two factor Gaussian essentially affine model defined in equations (22) and (23). These estimates are based on monthly observations of the rate of depreciation of the U.S. Dollar - Sterling Pound and 1, 3, 6 and 12-month Eurocurrency interest rates in the U.S. and U.K. The sample period is January 1976 to December 1997 (252 observations). American variables correspond with the index number 1 and British ones correspond with the number 2. Robust standard errors are provided in parenthesis.

Table 3
Estimates of the Two Factor Essentially Affine Model:
Panel b: U.S. Dollar - Canadian Dollar

	Index Number (i)	
Parameter	1	2
ϕ_{1i}	0.0458	0
	(0.0037)	
ϕ_{2i}	-0.1995	0.1999
	(0.0283)	(0.0238)
θ_i	1.0570	1.1795
	(0.0589)	(0.0733)
σ_{1i}	0.0639	0
	(0.0063)	
σ_{2i}	0	0.0637
		(0.0043)
λ_{1i}	0	0
λ_{2i}	16.2445	-12.2519
	(3.8478)	(2.8206)
λ_{0i}	0	-1.9329
		(1.1279)
λ^*_{1i}	16.2852	-12.2643
	(3.8170)	(2.8048)
λ^*_{2i}	0	0
λ^*_{0i}	-1.8498	0
	(1.1153)	

This table presents quasi maximum likelihood (QML) estimates of the two factor Gaussian essentially affine model defined in equations (22) and (23). These estimates are based on monthly observations of the rate of depreciation of the U.S. Dollar - Canadian Dollar and 1, 3, 6 and 12-month Eurocurrency interest rates in the U.S. and Canada. The sample period is January 1976 to December 1997 (252 observations). American variables correspond with the index number 1 and Canadian ones correspond with the number 2. Robust standard errors are provided in parenthesis.

Table 4
Implied Forward Premium Regression Slopes

Panel a: US Dollar - Sterling Pound

	1-month	3-month	6-month	12-month
OLS	-1.840	-1.505	-1.361	-0.817
Implied	-2.001	-1.945	-1.878	-1.788

Panel b: US Dollar -Canadian Dollar

	1-month	3-month	6-month	12-month
OLS	-1.351	-0.827	-0.425	-0.240
Implied	-0.578	-0.536	-0.481	-0.411

This table presents the term structure of forward premium regression slopes implied by the two factor Gaussian essentially affine model defined in equations (22) and (23). These are computed using the closed-form formulae derived in the appendix and by treating the estimates displayed in table 3 as truth. The sample OLS estimates of these slopes are reproduced again for the sake of comparison.

Table 5
Comparison of Out-of-sample Forecasting Performance

Panel a: US Dollar - Sterling Pound

	Affine (level)	VAR(2) (ratio)	RW (ratio)	OLS (ratio)
Root Mean Square Error (*RMSE*)				
1-month horizon	0.0205	0.975	0.978	0.980
3-month horizon	0.0300	0.925	0.903	0.909
6-month horizon	0.0403	0.845	0.819	0.820
12-month horizon	0.0496	0.653	0.637	0.584
Mean Absolute Error (*MAE*)				
1-month horizon	0.0167	0.971	0.959	0.975
3-month horizon	0.0255	0.999	0.952	0.988
6-month horizon	0.0312	0.830	0.828	0.818
12-month horizon	0.0414	0.666	0.632	0.601

Panel b: US Dollar - Canadian Dollar

	Affine (level)	VAR(1) (ratio)	RW (ratio)	OLS (ratio)
Root Mean Square Error (*RMSE*)				
1-month horizon	0.0189	0.975	0.984	0.977
3-month horizon	0.0343	0.946	0.960	0.944
6-month horizon	0.0480	0.920	0.938	0.914
12-month horizon	0.0626	0.847	0.907	0.813
Mean Absolute Error (*MAE*)				
1-month horizon	0.0150	0.959	0.967	0.965
3-month horizon	0.0277	0.976	0.995	0.973
6-month horizon	0.0363	0.932	0.933	0.935
12-month horizon	0.0467	0.869	0.887	0.824

This table presents the results of the out-of-sample forecasting exercise during the last seven years of the sample (January 1998 - December 2004). For the two factor Gaussian essentially affine model the RMSE or the MAE is expressed in levels. For the alternative forecasts the RMSE or the MAE is expressed as the inverse of its ratio to the correspoding figure for the affine model. Therefore, a figure less than one indicates superior relative performance by the VECM model.