McCallum Rules, Exchange Rates, and the Term Structure of Interest Rates

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

March 2010

Abstract

McCallum (1994a) proposes a monetary rule where central banks have some tendency to resist rapid changes in exchange rates to explain the forward premium puzzle. We estimate this monetary policy reaction function within the framework of an affine term structure model to find that, contrary to previous estimates of this rule, the monetary authorities in Canada, Germany and the U.K. respond to nominal exchange rate movements. Our model is also able to replicate the forward premium puzzle.

JEL Classification: E43, F31, G12, G15.
Keywords: Interest Rates, Exchange Rates, Monetary Policy Rules, Uncovered Interest Parity

*I would like to thank Greg Bauer, Scott Hendry, Jose Fernandez-Serrano, Enrique Sentana, Jun Yang, and seminar participants at the Bank of Canada, CEMFI, IE Business School, Universidade Catolica Portuguesa (Lisbon), the 2008 Finance Forum (Barcelona), the 2008 Symposium on Economic Analysis (Zaragoza), and the European Summer Meeting of the Econometric Society (Barcelona, 2009) for useful comments and suggestions. Of course, I remain responsible for any remaining errors. The views expressed in this paper are those of the author and do not necessarily reflect those of the Bank of Canada.

1 Introduction

Over the last twenty-five years the majority of empirical studies of exchange rates have rejected the hypothesis of uncovered interest parity. This hypothesis implies that the (nominal) expected return to speculation in the forward foreign exchange market, conditional on available information, should be zero. Many studies have regressed ex-post rates of depreciation on a constant and the interest rate differential, rejecting the null hypothesis that the slope coefficient is one. In fact, a robust result is that the slope is negative. This phenomenon, known as the "forward premium puzzle", implies that, contrary to the theory, high domestic interest rates relative to those in the foreign country predict a future appreciation of the home currency.

A particularly interesting explanation of this anomaly has been given by McCallum (1994a). In an influential paper, he shows that models which augment the uncovered interest parity hypothesis with a monetary rule where central banks adjust interest rates to keep exchange rates stable are better able to capture the forward premium puzzle. In fact, this policy behavior insight has been widely cited as one of the main explanations for the rejection of uncovered interest parity (see, e.g., Taylor 1995, Engel 1996, Sarno 2005, and Burnside et al. 2006).[1]

Despite its theoretical appeal, the empirical support for this explanation appears tenuous. The estimates of this policy rule in both Mark and Wu (1996) and Christensen (2000)—which we replicate in this paper— imply that short-term interest rates do not react to exchange rate fluctuations. However, both papers employ single-equation approaches to estimate this rule and do not exploit the cross-sectional information contained in the yield curve.

In this paper, we estimate the McCallum (1994a) rule within the framework of an affine term structure model with time-varying risk premia. This approach, introduced by Ang, Dong and Piazzesi (2007) (ADP from now on) in the context of the estimation of a Taylor (1993) rule, has the advantage of exploiting the information contained in the whole yield curve as opposed to the information contained only on short-term interest rates. In particular, long-term interest rates are conditional expected values of future short-rates after adjusting for risk premia, and these risk-adjusted expectations are formed based on a view of how the central bank conducts monetary policy. Thus, the whole curve reflects the monetary actions of the central bank, and the entire term structure of interest rates can be used to estimate a monetary policy rule. In particular, we estimate a two-country affine term structure model using yield curve data over the period January 1979 to December 2005 for Canada, Germany and the U.K, and taking the U.S. as the foreign country in each case. Our estimates indicate that, in contrast to the results in Mark and Wu (1996) and Christensen

[1]Several other explanations for this anomaly are the existence of a rational risk premium in the foreign exchange rate market, "peso problems", and violations of the rational expectations assumption. See Engel (1996) for a review of this literature.

(2000), the monetary authority in these three countries responds to exchange rate movements. The exchange rate stabilization coefficient is significant at the 5% level for Canada and the U.K. and significant at the 10% level for Germany which suggests that the monetary authority interprets a depreciating exchange rate as a signal of higher future inflation and increases the short rate accordingly.[2] Finally, our proposed term structure model with endogenous risk premia, a main difference with respect to the original work of McCallum (1994a), replicates the forward premium puzzle for all three datasets.

Our approach also allows us to study the impact of the U.S. short-term interest rate, the domestic latent factor, and exchange rate on the yield curve. We find that the U.S. short rate tends to be the main driver of the variability of the long-end of the yield curve regardless the country of examination. For example, 95% of the ten-year ahead variance of the Canadian ten-year yield, 65% of the variance of the German ten-year yield and 87% of the variance of the British ten-year yield can be attributed to U.S. shocks. Also, the variability of the short-end of the yield curve is mainly explained by shocks to the exchange rate. Over 56% of the one-month ahead variance of the Canadian one-month yield, 87% of the variance of the German one-month yield, and 90% of the variance of the British one-month yield is due to exchange rate movements. Finally, both bond and foreign exchange risk premia are explained by a combination of domestic and foreign exchange shocks with the U.S. short-rate playing little or no role at all.

The model that we consider in this paper belongs in the literature on international term structure modeling: see e.g. Saa-Requejo (1993), Frachot (1996), Backus et al. (2001), Dewachter and Maes (2001), Ahn (2004), Brennan and Xia (2006), Dong (2006), Leippold and Wu (2007), and Diez de los Rios (2009). These authors exploit the fact that the same factors that determine the risk premium in the term structure of interest rates in each country might also determine the risk premium in exchange rate returns. To do so, one usually starts by specifying the law of motion for the stochastic discount factor in each one of the countries to then use the law of one price to find the process that the exchange rate follows. Using this approach, the exchange rate is an endogenous variable that is fully determined by the state variables of the model. In contrast, under a McCallum (1994a) rule, the monetary authority intervenes in the short-term bond market to respond to exchange rate movements and, therefore, the rate of depreciation in our model has to itself become a state variable. Thus, an important contribution of this paper is to show how to restrict the parameters of the prices of risk to guarantee that the model is consistent: the exchange rate that comes out of the model is the same as the exchange rate we started with as a state variable. In this way, we incorporate a feedback effect from exchange rates to the yield curve, a feature shared with the work of Pericolli and Taboga (2008) who estimate a joint model of bond

[2] Along these lines, Backus et al. (2009) recently point out a close link between both a Taylor (1993) policy rule where the monetary authority respond to inflation and the McCallum (1994a) rule.

yields, macroeconomic variables and the exchange rate.

Finally, we also estimate the McCallum (1994b) yield-curve-smoothing rule, which was proposed to explain the rejection of the expectations-hypothesis of the term structure, to provide a benchmark to compare our results with. To do so, we use the results in Gallmeyer et al. (2005) who show how to rotate the space of state variables in an affine term structure model to relate the short rate to the term premium. Our findings indicate that both McCallum rule models seem to provide a similar fit of the yield curve. If there is any difference, the McCallum (1994a) exchange-rate-stabilization rule seems to do slightly better.

The rest of the paper is organized as follows. In section 2, we briefly review the forward premium puzzle and the McCallum (1994a) exchange-rate-stabilization policy rule. Section 3 describes the affine term structure model and its estimation. Section 4 presents the empirical results. In Section 5 we compare how both McCallum (1994a) exchange-rate-stabilisation and McCallum (1994b) yield-curve-smoothing rules fit the term structure of interest rates. Section 6 concludes.

2 McCallum Rules and The Forward Premium Puzzle

We begin with a review of the forward premium puzzle and the McCallum (1994a) exchange-rate-stabilization policy rule. Denote the price at time t of a domestic default-free pure-discount bond that pays 1 with certainty at date $t+n$ as $P_t^{(n)}$. The continuously compounded yield on this bond, $y_t^{(n)}$, satisfies $P_t^{(n)} \equiv \exp(-ny_t^{(n)})$. Therefore:

$$y_t^{(n)} = -\frac{1}{n}\log P_t^{(n)}.$$

We refer to the short-term interest rate, or short rate, as the yield on the bond with the shortest maturity under consideration, $r_t = y_t^{(1)}$. We also define $P_t^{(n)*}$ and $y_t^{(n)*}$ as the price at time t of a foreign default-free pure-discount bond and its yield, respectively. Similarly, the foreign short-term interest rate is $r_t^* = y_t^{(1)*}$. Finally, S_t is the spot exchange rate expressed as the price in domestic monetary units of a unit of foreign exchange.

Uncovered interest parity relates the expected rate of depreciation of a currency to the interest rate differential between the countries. It recognizes that portfolio investors at any time t have the choice of holding either (i) bonds denominated in domestic currency, or (ii) holding foreign bonds with the same characteristics. Thus, an investor starting with one unit of domestic currency compares two options. One is to invest in a domestic n-period bond to accumulate $1/P_t^{(n)} = \exp(ny_t^{(n)})$ units of domestic currency. Another option is to convert his unit of domestic currency at the spot exchange rate into $1/S_t$ units of foreign currency, invest into foreign bonds to accumulate $1/(S_tP_t^{(n)*}) = \exp(ny_t^{(n)*})/S_t$, and then reconvert these profits into domestic currency at the prevailing spot exchange rate at $t+n$. If agents are risk neutral, we get the condition of uncovered interest parity

$$\exp(ny_t^{(n)}) = E_t\left[\frac{S_{t+n}}{S_t}\exp(ny_t^{(n)*})\right]. \tag{1}$$

Further, if we assume that the spot exchange rate is conditionally log-normal, we can express the uncovered interest parity hypothesis as:

$$E_t\left(s_{t+n} - s_t\right) = -\frac{1}{2}Var_t\left(s_{t+n} - s_t\right) + n(y_t^{(n)} - y_t^{(n)*}), \tag{2}$$

where $-\frac{1}{2}Var_t\left(s_{t+n} - s_t\right)$ is the Jensen's inequality term and s_t denotes the log of the spot exchange rate.

This theory can be validated empirically by regressing the ex-post rate of depreciation on a constant and the interest rate differential to, then, test if the slope coefficient is equal to one. However, such a test reveals that this theory is strongly rejected in the data. In fact, a robust result in many studies is that the estimated slope is negative and statistically different from zero (see Engel, 1996, for a review of the literature). This empirical rejection is known as the forward premium puzzle and it implies that high domestic interest rates relative to those in the foreign country predict a future appreciation of the home currency.

Since this puzzle is usually related to the existence of a rational risk premium in the foreign exchange rate market, the uncovered interest parity is modified as follows:

$$E_t\left(s_{t+n} - s_t\right) = n(y_t^{(n)} - y_t^{(n)*}) + \xi_t^{(n)}, \tag{3}$$

where we have ignored the Jensen's inequality term and included a risk premium, $\xi_t^{(n)}$.

McCallum (1994a) proposes a model which augments uncovered interest parity with a monetary rule where policymakers have some tendency to resist rapid changes in exchange rates. By modeling monetary policy this way, the resulting equilibrium exchange rate process is better able to capture the forward premium puzzle. We refer to this rule as the McCallum exchange-rate-stabilization policy which takes the form:

$$r_t - r_t^* = \psi_1\Delta s_t + \psi_2(r_{t-1} - r_{t-1}^*) + e_t, \tag{4}$$

where e_t is the monetary policy shock that summarizes the other exogenous determinants of monetary policy. This monetary policy rule implies that the central bank intervenes in the short-term bond market to try to achieve two (perhaps conflicting) goals: "exchange rate stabilisation" governed by the parameter $\psi_1 > 0$, and "interest rate differential smoothing" governed by the parameter $|\psi_2| < 1$. Note that in this model a depreciating exchange rate signals higher expected future inflation, and therefore the monetary authority increases the short rate.

Combining equations (3) and (4) for $n = 1$ with a first order autoregressive process for the risk premium such as[3]

$$\xi_t = \rho\xi_{t-1} + e_t^{\xi},$$

[3] McCallum (1994a) also provides a less realistic model for the risk premium where ξ_t is *iid* with zero mean.

where e_t^{ξ} is exogenous white noise, and $|\rho| < 1$, McCallum (1994a) obtains, by using the method of undertermined coefficients, the following reduced form equation for the exchange rates:

$$s_{t+1} - s_t = \frac{\psi_2 - \rho}{\psi_1}(r_t - r_t^*) - \frac{1}{\psi_1}\xi_{t+1} + \frac{1}{\psi_1 + \psi_2 - \rho}e_{t+1}^{\xi}. \tag{5}$$

On this basis McCallum concludes that if ψ_2 is close to 1, ψ_1 is close to 0.2 and $\rho \ll 1$, then a negative slope coefficient on the forward premium regression may be consistent with the uncovered interest parity theory.

Note, however, that a limitation of this analysis is the exogeneity of the risk premium: this theory does not explain how factors driving the risk premium in foreign exchange markets might be related to factors that affect interest rates. For this reason, we now re-interpret McCallum's findings in the context of an affine term structure model.

3 The Model

3.1 General Setup

The McCallum (1994a) exchange-rate-stabilization policy rule captures the notion that central banks tend to resist rapid changes in exchange rates. In particular, this rule states that central banks set short-term interest rates in such a way that the interest rate differential depends on the current rate of depreciation and past values of the interest rate differentials. Yet, long-term interest rates are conditional expected values of future short rates (after adjusting for risk premia) and, therefore, the entire yield curve in such a set-up have to respond to movements in the foreign interest rate and the rate of depreciation. Hence, both the short-term foreign interest rate and the exchange rate have to themselves become state variables in the term structure model.

In particular, we assume that there are three state variables:

$$\mathbf{x}_t = \left[\begin{array}{ccc} r_t^* & f_t & \Delta s_t \end{array}\right]',$$

where r_t^* is the foreign (i.e. U.S.) short-term interest rate which, following ADP, we treat as a latent factor; f_t is a domestic latent term structure factor; and, $\Delta s_t \equiv s_t - s_{t-1}$ is the one-period rate of depreciation. We also assume that these state variables follow a VAR(1) process:

$$\mathbf{x}_{t+1} = \boldsymbol{\theta} + \boldsymbol{\Phi}\mathbf{x}_t + \mathbf{u}_{t+1}, \tag{6}$$

where $\mathbf{u}_t = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\varepsilon}_t \sim iid\ N(0, \mathbf{I})$. Since in our empirical application, we choose the U.S. to be the foreign country, we model the foreign short-rate, r_t^*, as a first-order autoregressive process: $\phi_{12} = \phi_{13} = 0$ in order to guarantee that this variable is not affected by domestic factors. Also, we assume that $\boldsymbol{\Sigma}^{1/2}$ has the following form:

$$\boldsymbol{\Sigma}^{1/2} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix},$$

so that shocks to the foreign short rate and the domestic factor are orthogonal. This assumption guarantees that the model is identified when both r_t^* and f_t are latent factors. Furthermore, notice that the rate of depreciation is affected by both the shocks to the foreign short rate and the domestic factor. In addition, we postulate the existence of a third shock, orthogonal to the previous ones, that only affects the rate of depreciation.

The short rate is related to the set of state variables through an affine relation:

$$r_t = \delta_0 + \boldsymbol{\delta}_1' \mathbf{x}_t, \tag{7}$$

where δ_0 is a scalar and $\boldsymbol{\delta}_1$ is a 3×1 vector.

Finally, the model is completed by specifying the stochastic discount factor (SDF) to take the following form (see Ang and Piazzesi, 2003 and ADP):

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\boldsymbol{\lambda}_t'\boldsymbol{\lambda}_t - \boldsymbol{\lambda}_t'\boldsymbol{\varepsilon}_{t+1}\right), \tag{8}$$

with prices of risk given by:

$$\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_0 + \boldsymbol{\lambda}_1 \mathbf{x}_t, \tag{9}$$

where $\boldsymbol{\lambda}_0$ is 3×1 vector and $\boldsymbol{\lambda}_1$ is a 3×3 matrix.

This (strictly positive) SDF, m_{t+1}, prices any traded asset denominated in domestic currency through the following relationship:

$$P_t = E_t\left[m_{t+1} X_{t+1}\right], \tag{10}$$

where P_t is the value of a claim to a stochastic cash flow of X_{t+1} units of domestic currency one period later. Using this model to price zero coupon bonds, we obtain the following recursive relation:

$$P_t^{(n)} = E_t\left[m_{t+1} P_{t+1}^{(n-1)}\right], \tag{11}$$

where $P_t^{(n)}$ is the price of a zero-coupon bond of maturity n periods at time t.

Similarly, it is possible to show that solving equation (11) is equivalent to solve the following equation to obtain the price of a zero-coupon bond:

$$P_t^{(n)} = E_t^Q\left[\exp\left(-\sum_{i=0}^{n-1} r_{t+i}\right)\right],$$

where E_t^Q denotes the expectation under the risk-neutral probability measure, under which the dynamics of the state vector $\mathbf{x}_t$ are also characterized by a VAR(1):

$$\mathbf{x}_t = \boldsymbol{\theta}^Q + \boldsymbol{\Phi}^Q \mathbf{x}_{t-1} + \mathbf{u}_t, \tag{12}$$

with

$$\begin{aligned}\boldsymbol{\theta}^Q &= \boldsymbol{\theta} - \boldsymbol{\Sigma}^{1/2}\boldsymbol{\lambda}_0, \\ \boldsymbol{\Phi}^Q &= \boldsymbol{\Phi} - \boldsymbol{\Sigma}^{1/2}\boldsymbol{\lambda}_1.\end{aligned}$$

That is, one can price a zero-coupon bond as if agents were risk-neutral by using the (local) expectations hypothesis once the law of motion of the state variables has been modified to account for the fact that agents are not risk neutral.

Yet remember 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. Therefore, uncovered interest parity must be satisfied under the risk-neutral measure. This implies that the parameters under Q must satisfy an equivalent version of equation (2):

$$E_t^Q \Delta s_{t+1} = -\frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3 + (r_t - r_t^*), \tag{13}$$

where $-\frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3$ is the Jensen's inequality term and $\mathbf{e}_i$ is a 3×1 vector of zeros with a one in the ith position. Substituting (7) into (13) and using (12) to compute the expected rate of depreciation under the risk neutral probability measure, we get that

$$\mathbf{e}_3'\left(\boldsymbol{\theta}^Q + \boldsymbol{\Phi}^Q\mathbf{x}_t\right) = -\frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3 + (\delta_0 + \boldsymbol{\delta}_1'\mathbf{x}_t) - \mathbf{e}_1'\mathbf{x}_t,$$

so the following two restrictions apply:

$$\mathbf{e}_3'\boldsymbol{\Phi}^Q = \boldsymbol{\delta}_1' - \mathbf{e}_1', \tag{14}$$

$$\mathbf{e}_3'\boldsymbol{\theta}^Q = -\frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3 + \delta_0. \tag{15}$$

Finally, Ang and Piazzesi (2003) show that the model (6)-(9) implies that the price of a n-period zero coupon bond satisfies:

$$P_t^{(n)} = \exp\left(A_n + \mathbf{B}_n'\mathbf{x}_t\right),$$

where A_n and $\mathbf{B}_n$ satisfy the recursive relations:

$$\begin{aligned}A_{n+1} &= A_n + \mathbf{B}_n'\boldsymbol{\theta}^Q + \frac{1}{2}\mathbf{B}_n'\boldsymbol{\Sigma}\mathbf{B}_n - \delta_0, \\ \mathbf{B}_{n+1}' &= \mathbf{B}_n'\boldsymbol{\Phi}^Q - \boldsymbol{\delta}_1',\end{aligned} \tag{16}$$

with $A_1 = -\delta_0$ and $\mathbf{B}_1 = -\boldsymbol{\delta}_1$. Thus, the continuously compounded yield on an n-period zero coupon bond at time t, $y_t^{(n)}$, is given by

$$y_t^{(n)} = a_n + \mathbf{b}_n'\mathbf{x}_t, \tag{17}$$

where $a_n = -A_n/n$ and $\mathbf{b}_n = -\mathbf{B}_n/n$. Moreover, note that the one-period yield $y_t^{(1)}$ is the same as the short rate r_t in equation (7).

3.2 Stochastic Discount Factors and Exchange Rates

The law of one price tells us that of the three random variables—the domestic SDF, the foreign SDF and the rate of depreciation—one is effectively redundant and can be constructed from the other two. In fact, Backus et al. (2001) show that under complete markets the rate of depreciation and the domestic and foreign stochastic discount factors satisfy the following relation:

$$\Delta s_{t+1} = \log m^*_{t+1} - \log m_{t+1}. \tag{18}$$

In other words, we are implictly assuming a process for the foreign SDF when specifying the domestic SDF and the rate of depreciation. This is clear once we substitute the law of motion for the rate of depreciation in (6) and the domestic SDF in (8) into this last equation and solve for the foreign SDF to obtain:

$$\log m^*_{t+1} = \mathbf{e}_3'(\boldsymbol{\theta} + \boldsymbol{\Phi}\mathbf{x}_t) - r_t - \frac{1}{2}\boldsymbol{\lambda}_t'\boldsymbol{\lambda}_t - \left[(\boldsymbol{\lambda}_t - (\boldsymbol{\Sigma}^{1/2})'\mathbf{e}_3\right]' \boldsymbol{\varepsilon}_{t+1}.$$

If we now define $\boldsymbol{\lambda}^*_t = \boldsymbol{\lambda}_t - (\boldsymbol{\Sigma}^{1/2})'e_3$ and substitute $\boldsymbol{\lambda}_t$ in this equation, we get:

$$\log m^*_{t+1} = \mathbf{e}_3'(\boldsymbol{\theta}^Q + \boldsymbol{\Phi}^Q\mathbf{x}_t) + \frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3 - r_t - \frac{1}{2}(\boldsymbol{\lambda}^*_t)'(\boldsymbol{\lambda}^*_t) - (\boldsymbol{\lambda}^*_t)'\boldsymbol{\varepsilon}_{t+1}.$$

But notice that $E_t^Q\Delta s_{t+1} = \mathbf{e}_3'(\boldsymbol{\theta}^Q + \boldsymbol{\Phi}^Q\mathbf{x}_t) = -\frac{1}{2}\mathbf{e}_3'\boldsymbol{\Sigma}\mathbf{e}_3 + (r_t - r^*_t)$ because uncovered interest parity holds under the risk-neutral measure. Therefore, the foreign SDF has the same form as (8):

$$m^*_{t+1} = \exp\left[-r^*_t - \frac{1}{2}(\boldsymbol{\lambda}^*_t)'(\boldsymbol{\lambda}^*_t) - (\boldsymbol{\lambda}^*_t)'\boldsymbol{\varepsilon}_{t+1.}\right],$$

with a foreign price of risk, $\boldsymbol{\lambda}^*_t$, that is also affine in $\mathbf{x}_t$:

$$\boldsymbol{\lambda}^*_t = \boldsymbol{\lambda}^*_0 + \boldsymbol{\lambda}^*_1\mathbf{x}_t,$$

being $\boldsymbol{\lambda}^*_0 = \boldsymbol{\lambda}_0 - (\boldsymbol{\Sigma}^{1/2})'\mathbf{e}_3$ and $\boldsymbol{\lambda}^*_1 = \boldsymbol{\lambda}_1$.

Thus, it is straightforward to show that under our framework the price of a foreign n-period zero coupon bond is also affine in the set of state variables $\mathbf{x}_t$:

$$P_t^{(n)*} = \exp\left(A^*_n + \mathbf{B}^{*\prime}_n\mathbf{x}_t\right),$$

where the scalar A^*_n and vector $\mathbf{B}^*_n$ satisfy a set of recursive relations similar to those in (16).[4] Furthermore, the continuously compounded yield on a foreign n-period zero coupon bond at time t will be

$$y_t^{(n)} = a^*_n + \mathbf{b}^{*\prime}_n\mathbf{x}_t, \tag{19}$$

where $a^*_n = -A^*_n/n$ and $\mathbf{b}^*_n = -\mathbf{B}^*_n/n$.

[4] Note that, in this case, $r^*_t = e_1'\mathbf{x}_t$. Thus $\delta^*_0 = 0$ and $\boldsymbol{\delta}^* = e_1$.

Finally, we further assume that the foreign (i.e. U.S.) short-rate, r_t^*, is also a first-order autoregressive process under the risk neutral measure: $\phi_{12}^Q = \phi_{13}^Q = 0$. Such an assumption guarantees that the foreign yield curve is not affected by domestic factors, and it follows a one-factor model. This is clearer if we further assume that $\left|\phi_{11}^Q\right| < 1$ (the short rate is stationary under the risk neutral measure) because it is possible to solve for $\mathbf{b}_n^*$ to obtain that:

$$\mathbf{b}_n^* = \left[\frac{1-(\phi_{11}^Q)^n}{n(1-\phi_{11}^Q)}, 0, 0\right]',$$

where both the foreign factor loadings on the domestic latent factor and the rate of depreciation are zero. Such restrictions might seem restrictive at first sight given that it is well known that we need more than one factor to explain the U.S. yield curve. Yet, given these restrictions, our model is still likely to explain well the level of the U.S. curve which, according to the implications of the McCallum (1994a) monetary rule, should be a main driving factor of the domestic term structure of interest rates. In addition, note that under this assumption one avoids the problem of finding potentially different estimates of the parameters governing the U.S. interest rate process depending on the exchange rate under examination. In fact, augmenting the number of factors in our setup would dramatically increase the number of parameters involved in the estimation of the model, rendering the estimation exercise almost impossible.

3.3 Expected Returns

Following ADP, we also analyze expected holding period returns on bonds. Those are defined as:

$$\begin{aligned} rx_{t+1}^{(n)} &\equiv \log\left(\frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}\right) - r_t, \\ &= ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t. \end{aligned}$$

Given that we assume that expectations are rational, the expected value of this variable is the bond risk premium. In particular, ADP show that expected excess holding period returns on bonds are also affine in $\mathbf{x}_t$:

$$E_t rx_{t+1}^{(n)} = A_n^x + \mathbf{B}_n^{x\prime}\mathbf{x}_t,$$

with the scalar $A_n^x = -\frac{1}{2}\mathbf{B}_{n-1}'\mathbf{\Sigma}\mathbf{B}_{n-1} + \mathbf{B}_{n-1}'\mathbf{\Sigma}^{1/2}\boldsymbol{\lambda}_0$ and the 3×1 vector $\mathbf{B}_n^{x\prime} = \mathbf{B}_{n-1}'\mathbf{\Sigma}^{1/2}\boldsymbol{\lambda}_1$. Note that the expected excess return has three terms: (i) a Jensen's inequality term; (ii) a constant risk premium; and, (iii) a time-varying risk premium where time variation is governed by the parameters in matrix $\boldsymbol{\lambda}_1$.

Similarly, we can also compute the foreign exchange risk premium as the expected excess rate of return to a domestic investor on buying a one-period foreign zero-coupon bond:

$$\begin{aligned} sx_{t+1} &\equiv \log\left(\frac{S_{t+1}}{S_t}\right) + y_t^{(1)*} - y_t^{(1)} \\ &= \Delta s_{t+1} + r_t^* - r_t, \end{aligned}$$

and it is possible to show that the value of this expectation is also affine in $\mathbf{x}_t$:

$$E_t sx_{t+1} = A_s + \mathbf{B}_s'\mathbf{x}_t,$$

with the scalar $A_s = -\frac{1}{2}\mathbf{e}_3'\mathbf{\Sigma}\mathbf{e}_3 + \mathbf{e}_3'\mathbf{\Sigma}^{1/2}\boldsymbol{\lambda}_0$ and the 3×1 vector $\mathbf{B}_s' = \mathbf{e}_3'\mathbf{\Sigma}^{1/2}\boldsymbol{\lambda}_1$.[5] As in the case of the bond risk premium expression, this expected excess return has again three terms: (i) a Jensen's inequality term, (ii) a constant risk premium, and (iii) a time-varying risk premium governed by the matrix $\boldsymbol{\lambda}_1$.

3.4 From Affine to McCallum

In this section, we follow the techniques developed in ADP, to modify the short rate equation to take the same form as the McCallum exchange-rate stabilization policy rule. We start by rewriting equation (7) as:

$$r_t = \delta_{11} r_t^* + f_t + \delta_{13}\Delta s_t, \tag{20}$$

where (to ensure that the model is identified) we have set $\delta_0 = 0$ (to free up the mean of the latent factor f_t) and $\delta_{12} = 1$ (to leave the volatility of the unobserved factor unconstrained). Equation (6) implies that

$$f_t = \theta_2 + \phi_{21} r_{t-1}^* + \phi_{22} f_{t-1} + \phi_{23}\Delta s_{t-1} + u_{2t}. \tag{21}$$

Substituting (21) in (20) gives:

$$\begin{aligned} r_t &= \delta_{11} r_t^* + \delta_{13}\Delta s_t + \theta_2 \\ &\quad + \phi_{21} r_{t-1}^* + \phi_{22} f_{t-1} + \phi_{23}\Delta s_{t-1} + u_{2t}, \end{aligned}$$

and substituting again for f_{t-1} in this last expression and rearranging, we obtain:

$$\begin{aligned} r_t &= \theta_2 + \delta_{11} r_t^* + \delta_{13}\Delta s_t \\ &\quad + (\phi_{21} - \phi_{22}\delta_{11}) r_{t-1}^* + (\phi_{23} - \phi_{22}\delta_{13})\Delta s_{t-1} \\ &\quad + \phi_{22} r_{t-1} + u_{2t}. \end{aligned} \tag{22}$$

Under the unrestricted set-up, the short rate depends on (i) current and lagged values of the foreign short rate and the rate of depreciation, (ii) the lagged short rate and (iii)

[5] We have used equation (18) to get that $E_t\Delta s_{t+1} = \frac{1}{2}(\boldsymbol{\lambda}_0'\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0^{*\prime}\boldsymbol{\lambda}_0^*) + (\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0^*)'\boldsymbol{\lambda}_1\mathbf{x}_t$. Substituting $\boldsymbol{\lambda}_0^* = \boldsymbol{\lambda}_0 - (\mathbf{\Sigma}^{1/2})'e_3$ in this expression gives the equation in the text.

a monetary policy shock. Equating the coefficients in equations (4) and (22) allows us to obtain:

$$\delta_{11} = 1, \qquad \delta_{13} = \psi_1, \qquad \phi_{21} = 0, \qquad \phi_{22} = \psi_2, \qquad \phi_{23} = \psi_1 \psi_2 \tag{23}$$

and $\theta_2 = \psi_0$ if a constant in (4) is included, or $\theta_2 = 0$ otherwise; and $u_{2t} = e_t$ is the monetary policy shock. These restrictions imply that a one percent increase in the foreign short-term rate translates one-for-one into the domestic short-rate, and that a one percent increase in the one-period rate of depreciation leads to a ψ_1 percent increase in the short-rate.

Finally note that these restrictions imply that the coefficients in the vector of factor loadings, $\mathbf{b}_n$, in equation (17) are non-linear functions of ψ_1 (and the rest of parameters under the risk-neutral measure). Thus, the yield curve provides additional over-identifying assumptions that can be exploited to obtain more efficient estimates of the reaction of the domestic short-term rate to movements in exchange rates.

3.5 Estimation Method

We estimate our term structure model using the Kalman filter (e.g., de Jong 2000) with both domestic and foreign yield data, and assuming that all (both domestic and foreign) yields are observed with error, so that the equation for each yield is:

$$\widetilde{y}_t^{(n)} = y_t^{(n)} + \eta_t^{(n)}$$

where $y_t^{(n)}$ is the model-implied yield from equations (17) and (19), and $\eta_t^{(n)}$ is a zero-mean observation error that is i.i.d. across time and yields. We specify $\eta_t^{(n)}$ to be normally distributed and denote the standard deviation of the error term as $\sigma_\eta^{(n)}$. However, to reduce the number of parameters to be estimated, we follow Brennan and Xia (1996) to assume the standard deviation of the yield measurement errors to be of the form: $\sigma_\eta^{(n)} = \sigma_\eta$ where σ_η is a single parameter to be estimated.

On the other hand, we could have estimated our model following the usual convention in the literature (Chen and Scott, 1993, Dai and Singleton 2002; Duffee 2002) in assuming that as many yields as unobservable factors are measured without measurement error. In particular, we could have assumed that the domestic and foreign one-month yields were observed without measurement error, while the yields on the remaining maturities were assumed to be measured with serially uncorrelated zero-mean errors. However, such a choice of bonds to use in the estimation would be arbitrary, and do not guarantee that the estimates will be consistent with the yields of other bonds. More importantly, ADP point out that by not assigning several arbitrary yields to have zero measurement error, one does not bias the estimated monetary policy shocks to have undue influence from only those particular yields.

Let $\widetilde{\mathbf{y}}_t$ be the vector of observed variables that collects the yields on the foreign bonds, the yields on the domestic bonds and the rate of depreciation. In particular, $\widetilde{\mathbf{y}}_t = (\mathbf{y}_t^{*\prime}, \mathbf{y}_t^{\prime}, \Delta s_t)^{\prime}$, where $\mathbf{y}_t^* = (y_t^{(1)*}, \ldots, y_t^{(N)*})$ and $\mathbf{y}_t = (y_t^{(1)}, \ldots, y_t^{(N)})$, and N is largest contract maturity under study (which, without loss of generality, we assume to be the same for all countries). Our asset pricing model, joint with our assumption on the measurement errors, implies that the vector $\widetilde{\mathbf{y}}_t$ has the following state-space representation:

$$\begin{gathered}
\widetilde{\mathbf{y}}_t = \mathbf{c} + \mathbf{D}\mathbf{x}_t + \boldsymbol{\eta}_t, \\
\mathbf{x}_t = \boldsymbol{\theta} + \boldsymbol{\Phi}\mathbf{x}_{t-1} + \mathbf{u}_t, \\
\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{u}_t \end{pmatrix} \Bigg| \begin{pmatrix} \mathbf{y}_{t-1} \\ \boldsymbol{\alpha}_{t-1} \end{pmatrix}, \begin{pmatrix} \mathbf{y}_{t-2} \\ \boldsymbol{\alpha}_{t-2} \end{pmatrix}, \ldots \sim N\left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} \end{pmatrix}\right],
\end{gathered} \tag{24}$$

where, again, $\mathbf{x}_t = (r_t^*, f_t, \Delta s_t)'$ and

$$\mathbf{c} = \begin{pmatrix} \delta_0^* \\ \vdots \\ a_N^* \\ \delta_0 \\ \vdots \\ a_N \\ 0 \end{pmatrix} \qquad \mathbf{D} = \begin{pmatrix} \boldsymbol{\delta}^{*\prime} \\ \vdots \\ \mathbf{b}_N^{*\prime} \\ \boldsymbol{\delta}' \\ \vdots \\ \mathbf{b}_N' \\ \mathbf{e}_3' \end{pmatrix} \qquad \boldsymbol{\Omega} = \begin{pmatrix} \sigma_\eta^2 \mathbf{I}_{2N} & 0 \\ 0 & 0 \end{pmatrix},$$

with the coefficients in $\mathbf{c}$ and $\mathbf{D}$, say a_n, a_n^*, $\mathbf{b}_n$ and $\mathbf{b}_n^*$, solving a set of recursive relations as those in (16), and $\mathbf{I}_{2N}$ being the $2N \times 2N$ identity matrix.

Given this state-space formulation, we can evaluate the exact Gaussian likelihood via the usual prediction error decomposition:

$$\ln L(\boldsymbol{\theta}) = \sum_{t=1}^{T} l_t,$$

with

$$l_t = -\frac{(2N+1)}{2} \ln(2\pi) - \frac{1}{2} \ln |\mathbf{F}_t| - \frac{1}{2} \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t, \tag{25}$$

where $\boldsymbol{\theta}$ is the vector of parameters of the continuous-time model, $\mathbf{v}_t$ is the vector of one-step-ahead prediction errors produced by the Kalman filter, and $\mathbf{F}_t$ their conditional variance. The Kalman filter recursions are given by

$$\left.\begin{aligned}
\mathbf{x}_{t|t-1} &= \boldsymbol{\theta} + \boldsymbol{\Phi}\mathbf{x}_{t-1|t-1} \\
\mathbf{P}_{t|t-1} &= \boldsymbol{\Phi}\mathbf{P}_{t-1|t-1}'\boldsymbol{\Phi} + \boldsymbol{\Sigma} \\
\mathbf{v}_t &= \widetilde{\mathbf{y}}_t - \mathbf{c} - \mathbf{D}\mathbf{x}_{t|t-1} \\
\mathbf{F}_t &= \mathbf{D}\mathbf{P}_{t|t-1}\mathbf{D}' + \boldsymbol{\Omega} \\
\mathbf{x}_{t|t} &= \mathbf{x}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{D}'\mathbf{F}_t^{-1}\mathbf{v}_t \\
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{B}'\mathbf{F}_t^{-1}\mathbf{D}\mathbf{P}_{t|t-1}
\end{aligned}\right\} \tag{26}$$

where $\mathbf{x}_{t|t-1} = E_{t-1}(\mathbf{x}_t)$ and $\mathbf{P}_{t|t-1} = E\left[(\mathbf{x}_t - \mathbf{x}_{t|t-1})(\mathbf{x}_t - \mathbf{x}_{t|t-1})'\right]$ are the expectation and covariance matrix of $\mathbf{x}_t$ conditional on information up to time $t-1$, while $\mathbf{x}_{t|t} = E_t(\mathbf{x}_t)$ and $\mathbf{P}_{t|t} = E\left[(\mathbf{x}_t - \mathbf{x}_{t|t})(\mathbf{x}_t - \mathbf{x}_{t|t})'\right]$ are the expectation and covariance matrix of $\mathbf{x}_t$ conditional on information up to time t (see Harvey, 1989). Given that we are assuming that the state variables are covariance stationarity, we initialize the filter using $\mathbf{x}_0 = E(\mathbf{x}_t) = (\mathbf{I} - \mathbf{\Phi})^{-1}\boldsymbol{\theta}$ and $vec(\mathbf{P}_0) = (\mathbf{I} - \mathbf{\Phi} \otimes \mathbf{\Phi})^{-1} vec(\mathbf{\Sigma})$.

The prediction error decomposition in (25) can also be used to obtain first and second derivatives of the log likelihood function (see Harvey, 1989), which we need to estimate the variance of the score and the expected value of the Hessian that appear in the asymptotic distribution of the Gaussian ML estimator of $\boldsymbol{\theta}$. In particular, the score vector takes the following form:

$$\frac{\partial l_t(\boldsymbol{\theta})}{\partial \psi_i} = s_t(\boldsymbol{\theta}) = -\frac{1}{2} tr\left[\left(\mathbf{F}_t^{-1}\frac{\partial \mathbf{F}_t}{\partial \psi_i}\right)\left(\mathbf{I} - \mathbf{F}_t^{-1}\mathbf{v}_t\mathbf{v}_t'\right)\right] - \frac{\partial \mathbf{v}_t'}{\partial \psi_i}\mathbf{F}_t^{-1}\mathbf{v}_t,$$

while the ij-th element of the conditionally expected Hessian matrix satisfies:

$$-E_{t-1}\left(\frac{\partial^2 l_t}{\partial \psi_i \partial \psi_j}\right) = \frac{1}{2} tr\left(\mathbf{F}_t^{-1}\frac{\partial \mathbf{F}_t}{\partial \psi_i}\mathbf{F}_t^{-1}\frac{\partial \mathbf{F}_t}{\partial \psi_{ij}}\right) + \frac{\partial \mathbf{v}_t'}{\partial \psi_i}\mathbf{F}_t^{-1}\frac{\partial \mathbf{v}_t}{\partial \psi_j}.$$

In turn, these two expressions require the first derivatives of $\mathbf{F}_t$ and $\mathbf{v}_t$, which we can evaluate analytically by an extra set of recursions that run in parallel with the Kalman filter. As Harvey (1989, pp 140-3) shows, the extra recursions are obtained by differentiating the Kalman filter prediction and updating equations (26). In our affine term structure model, the analytical derivatives of the Kalman filter equations with respect to the structural parameters require the derivatives of the bond price coefficients $a_n = -A_n/n$ and $\mathbf{b}_n = -\mathbf{B}_n/n$. These are obtained using the following difference equations:

$$\frac{\partial A_{n+1}}{\partial \psi_i} = \frac{\partial A_n}{\partial \psi_i} + \frac{\partial \mathbf{B}_n'}{\partial \psi_i}\boldsymbol{\theta}^Q + \mathbf{B}_n'\frac{\partial \boldsymbol{\theta}^Q}{\partial \psi_i} + \frac{\partial \mathbf{B}_n'}{\partial \psi_i}\mathbf{\Sigma}\mathbf{B}_n + \frac{1}{2}\mathbf{B}_n'\frac{\partial \mathbf{\Sigma}}{\partial \psi_i}\mathbf{B}_n - \frac{\partial \delta_0}{\partial \psi_i},$$

$$\frac{\partial \mathbf{B}_{n+1}'}{\partial \psi_i} = \frac{\partial \mathbf{B}_n'}{\partial \psi_i}\mathbf{\Phi}^Q + \mathbf{B}_n'\frac{\partial \mathbf{\Phi}^Q}{\partial \psi_i} - \frac{\partial \boldsymbol{\delta}_1'}{\partial \psi_i}.$$

with $\partial A_1/\partial \psi_i = -\partial \delta_0/\partial \psi_i$ and $\partial \mathbf{B}_1/\partial \psi_i = -\partial \boldsymbol{\delta}_1/\partial \psi_i$. Finally, it is worth mentioning that we also employ such analytical expressions for the score vector and information matrix in a score algorithm to maximize the exact log-likelihood function of the data.

4 Results

Our data set consists of monthly observations over the period January 1979 to December 2005 of the rates of depreciation of the U.S. dollar bilateral exchange rates against Canadian dollar, the German DM/Euro, and the British pound, along with the appropriate continuously compounded yields of maturities 1, 12, 24, 60 and 120 months for these countries. We use one-month Eurocurrency interest rates as our one-month yields. Data on the rest of the zero-coupon yield curve has been obtained from the Bank of Canada. In our empirical application, we take the U.S. as the foreign country.

Summary statistics for the variables are presented in Table 1. 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. We find that summary statistics of these variables are consistent with those found in previous studies such as, e.g., Backus et al. (2001) and Bekaert and Hodrick (2001). For example, we find that the rates of depreciation have lower means (in absolute value) than the ones corresponding to the interest rates, but, on the contrary, exchange rates are more volatile. In addition, bond yields display a high level of autocorrelation, while the rates of depreciation do not. The rate of depreciation of the U.S. dollar against the Canadian dollar is less volatile than the rates of depreciation of the U.S. dollar against the other two currencies. The United Kingdom ranks first in terms of the highest (average) level of interest rates during the sample period, followed by Canada, the United States, and Germany.

4.1 Parameter Estimates

Tables 2, 3, and 4 present parameter estimates of the affine term structure model for Canada, Germany and the U.K., respectively. These three tables are organized in the same way: Panel a reports the estimates of the McCallum rule; Panel b presents the estimates of the parameters of the model under the physical measure; and Panel c reports the parameters of the model under the risk neutral measure. In Panel d, we test if the coefficients under both the physical and risk neutral measure are the same.

Notice that the estimated coefficients of the exchange-rate stabilisation parameter, ψ_1, in Panel a of Tables 2–4 are positive for all three countries. This indicates that the monetary authority interprets a depreciating exchange rate as a signal of higher expected future inflation and, therefore, it increases the short rate. Also, this coefficient is significant at the 5% level for Canada and the U.K. and significant at the 10% level for Germany. However notice that, while it is positive and significant, the coefficient ψ_1 is well below the hypothesized value of 0.2 in McCallum (1994a). In particular, these estimates imply that a one percent shock to the monthly rate of depreciation leads to an increase of 1.75 basis point (bp) per month in the Canadian short rate, 3.36 bp increase in the German short rate, and 3.13 bp increase in the British short rate. On the other hand, the interest-rate-smoothing parameter, ψ_2, is close to one for Canada, and bigger than one for Germany and the U.K. While this result is counter-appealing (McCallum assumes that $|\psi_2| < 1$), it is reassuring to note that the eigenvalues of the autocorrelation matrix $\mathbf{\Phi}$ in equation (6) are all less than one in absolute value. Therefore, none of the state variables in our model presents an explosive behavior despite having $\psi_2 > 1$ for these two countries.

Comparing coefficients in Panel b of Tables 2–4, we can see that both the U.S. short-term interest rate and the latent factor are very persistent. This is explained by the fact that the estimated U.S. short-term rate is highly correlated with the level of the U.S. yield curve, while the domestic latent factor is higly correlated with the interest rate differential between the two countries. As widely known in the literature, both variables are highly autocorrelated. Also, notice in Panel b of Table 2 that both the U.S. short-rate and the Canadian latent factor significantly Granger-cause the current rate of depreciation. As for the estimates for Germany in Table 3, we find that only the domestic latent factor significantly Granger-causes changes in the exchange rate. We find in Table 4 that both the British domestic latent factor and the past rate of depreciation Granger-cause the current change in the exchange rate. We also find in these three tables that the impact of the domestic latent factor on the rate of depreciation is negative for all three countries. This is consistent with the forward premium puzzle because the latent factor is highly correlated with the interest rate differential. Finally the estimated matrix $\mathbf{\Sigma}^{1/2}$ shows that both shocks to the U.S. short-term rate and the domestic factors are negatively correlated with the rate of depreciation. In addition, shocks to the domestic factor seem to be more volatile than shocks to the U.S. short-rate.

The coefficients of the process that the state variables follow under the risk-neutral measure are reported in Panel c of Tables 2–4. The analysis of these coefficients reveals that the U.S. short-term interest rate and the latent factors are also very persistent under the risk-neutral measure for all three countries. More importantly, we find in Panel d of Tables 2–4 that the parameters under both the physical and risk neutral measure are statistically different. This indicates that there is a significant constant and time-varying price of risk in our model. Hence, the U.S. short rate, the latent factor and the rate of depreciation will play important roles in driving time-varying expected excess returns, as shown below when analyzing the corresponding variance decompositions.

We also formally test the specification of the model by following de Jong (2000) who suggests testing the validity of the constraints imposed by the affine term structure model on the general state-space representation of a model that does not impose the no-arbitrage assumption. In fact, we do not find evidence against the validity of the pricing model using a Lagrange Multipliers (LM) test. In particular, the LM test statistic is 40.686 for Canada, 40.754 for Germany, and 40.689 for the U.K., all smaller than the 5% (and 10%) critical value of a chi-squared distribution with 31 df (the number of constraints imposed by the affine term structure model).

4.2 Back to the Forward Premium Puzzle

While we have found that the monetary authorities in Canada, Germany and the U.K. respond to exchange rate movements, the motivation for a McCallum's (1994a) monetary policy reaction function resides in explaining the forward premium puzzle. Therefore, we now check if, by adding an endogenous time-varying risk premia to the McCallum rule, our model is still able to replicate a negative slope coefficient when regressing the ex post rate of depreciation on a constant and the interest rate differential.

In the spirit of the work by Hodrick (1992) and Bekaert (1995), we obtain an implied slope coefficient (implied beta) from the affine model that is analogous to the OLS regression slope tested in the simple regression approach. This implied beta is simply the ratio of the model implied covariance between the expected future rate of depreciation and the interest rate differential to the model implied variance of the interest rate differential. To compute this statistic, we exploit the state-space representation in (24) to realize that the implied beta from the affine term structure model is given by:

$$\beta^{(n)} = \frac{1}{n} \times \frac{\mathbf{e}_{2N+1}'\mathbf{D\Phi}(\mathbf{I}-\mathbf{\Phi})^{-1}(\mathbf{I}-\mathbf{\Phi}^{n})\mathbf{\Psi D}'(\mathbf{e}_{N+n}-\mathbf{e}_{n})}{(\mathbf{e}_{N+n}-\mathbf{e}_{n})'(\mathbf{D\Psi D}'+\mathbf{\Omega})(\mathbf{e}_{N+n}-\mathbf{e}_{n})}, \tag{27}$$

where, again, $\mathbf{e}_i$ is a conforming vector of zeros with a one in the ith position; and $\mathbf{\Psi}$ is the unconditional covariance matrix of $\mathbf{x}_t$, which again can be obtained from the equation $vec(\mathbf{\Psi}) = (\mathbf{I}-\mathbf{\Phi}\otimes\mathbf{\Phi})^{-1}vec(\mathbf{\Sigma})$. The numerator of equation (27) is just the model implied covariance between the expected future rate of depreciation and the interest rate differential, while the denominator is the model implied variance of the interest rate differential.

Table 5 presents the term structure of uncovered interest parity slopes implied by the affine model. These are computed using equation (27) and taking the parameter estimates in Tables 2-4 as the true values of the model. We find that the estimated implied betas are all negative, as predicted by the forward premium puzzle. Moreover, they become less negative as we increase the maturity of the contracts under consideration. For example, the implied beta for Canada at the one-month horizon is -1.770, while it is -0.104 at the ten-year horizon. Similar patterns can be found for Germany and the U.K.

We also compute sample estimates of these regression slopes using the coefficients of a VAR(1) model on the rate of depreciation and the set of interest rate differentials.[6] This model is akin to the vector-error-correction model in Clarida and Taylor (1997). Moreover, implied uncovered interest parity slope coefficients from a VAR(1) have already been used in Bekaert and Hodrick (2001).[7] When comparing the implied slopes from the affine model and these new estimates, we find that both implied slopes are close. That is, our model is

[6] Similar results are found when choosing a second-order VAR model.

[7] In practice, we would like to compare the implied betas from the affine model to those computed using traditional OLS methods. However, such an approach has the main drawback of largely reducing the number of effective observations when the maturity of the contract under consideration, n, is large. For example, if we were to compute an OLS slope using one-month yields, we would lose one observation while if we were to use ten-year yields, we would then effectively lose 120 observations (which is roughly half of the sample) when computing the ten-year rate of depreciation. Thus a comparison of OLS betas across different maturities would be complicated by the use of different effective samples. Notice also that a similar problem arises when comparing OLS betas and those computed from the affine model because the term structure model parameter estimates are computed using the whole sample. On the other hand, computing implied betas from a VAR do not suffer from this problem given that a VAR model is estimated using the whole sample thus making a fair comparison between those obtained from an affine model and this approach. In any case, it is reassuring to find that OLS and VAR estimates of the slope coefficient are basically the same when the contract period is $n = 1$ (both are computed using the same number of effective observations), and $n = 12$.

able to replicate a negative uncovered interest parity regression slope as predicted by the forward premium puzzle, and it also provides slope estimates close to what we would have found using a more traditional estimation method.

4.3 Latent Factor Dynamics

Figure 1 plots the estimated latent U.S. short-term rate together with the monthly yield on the U.S two-year bond. We plot the time series of the estimate of r_t^* conditional on information up to time t: $r_{t|t}^* = E_t\left(r_t^* | I_t\right)$ where I_t is the information set at time t. These are obtained using the Kalman filter algorithm.[8] This figure highlights the strong relationship between the estimated short-term rate and the level of the yield curve. Notice that, despite the estimated U.S. short rate being slightly above the monthly yield on the U.S two-year bond, both variables follow each other. In effect, we find that the correlation between our estimated factor and the yield curve ranges from 0.941 (one-month bond yield) to 0.977 (two-year bond yield).

Figure 2 plots the estimate of the Canadian latent factor together with the difference between the Canadian and U.S. two-year bond yields, and the rate of depreciation. Figures 3 and 4 plot the same variables for Germany and the U.K., respectively. Again, we plot the time series of the estimate of f_t conditional on information up to time t: $f_{t|t} = E_t\left(f_t | I_t\right)$. Note in these graphs that the domestic latent factor are strongly correlated with the term structure of bond yield differences. For example the correlation with the two-year bond yield difference is 0.903, while it is 0.904 for Germany and 0.863 for the U.K. Moreover, both the German and British factors seem to have inherited some volatility from the exchange rate. In fact, the correlation of the domestic factor with the rate of depreciation is -0.492 for Germany and -0.564 for the U.K., while it is only -0.207 for Canada.

4.4 Variance Decompositions

Tables 6, 7 and 8 present variance decompositions from the model and the data for Canada, Germany and the U.K., respectively. These show the proportion of the forecast variance that is attributed to each factor. Panel a reports variance decompositions of (i) yield levels, $y_t^{(n)}$; (ii) expected bond excess returns, $E_t rx_{t+1}^{(n)}$; and (iii) yield spreads, $y_t^{(n)} - y_t^{(1)}$. Panel b reports variance decompostions of (i) the rate of depreciation, Δs_{t+1}; and (ii) the foreign exchange rate risk premium, $E_t sx_{t+1}^{(n)}$.

Canada. We first focus on the results for Canada in Panel a of Table 6. One interprets the top row of Table 6 as follows: 1.61% of the one-month ahead forecast variance of the

[8]Note that we have three different estimates of r_t^* depending on the country we focus on. Still, these are highly correlated with each other, and the correlation among the three U.S. short rate estimates ranges from 0.999 to 1. Consequently and for simplicity, we plot the estimate obtained from the U.K. model.

one-month yield is explained by the U.S. short-term rate, 41.52% by the domestic latent factor and 56.87% by the rate of depreciation.

Notice that when we look to the one-month ahead variability of bond yields, we find that the proportion of variability accounted by the U.S. short-term yield increases with the maturity of the bond. This ranges from 1.61% for the one-month yield to 67.31% for the ten-year yield. Second, we find that the proportion of forecast variance explained by the domestic factor has a hump-shaped pattern. It explains 41.52% of the one-month ahead forecast variance of the short-rate, the 75.06% of the variability in one-year bond yields, but it explains only 30.88% of the forecast variance of the long-end of the yield curve. Last, shocks to the exchange rate do not explain the one-month ahead variability of the yield curve with the exception of the variance of the one-month yield (56.87%). This picture changes when we increase the forecasting horizon. For example, once we focus on the one-year ahead horizon, we find that shocks to the exchange rate account for almost 45% of the variability of the one-year yield (versus 6.65% when looking to one-month ahead variance decompositions). Yet, this effect decreases as we increase the maturity, and exchange rate shocks only explain around 20% of the variability at the long-end of the yield curve. Finally, the U.S. short-rate has the most explanatory power for ten-year ahead forecast variances at all points of the yield curve.

Turning to the variance decomposition of the bond risk premium, we find that shocks to the exchange rate are by far the main driving force of expected excess bond returns. In effect, the rate of depreciation has more explanatory power than the U.S. short-rate and the domestic factor at all points of the yield curve and for all forecast horizons. Similarly, the last three columns in Panel a of Table 6 document that shocks to the exchange rate tend to be the main driving force of yield spreads. However, we find that the effect of the domestic factor in explaining yield spreads becomes non-negligible and accounts for around 30% of this variability when we increase the maturity of the bond under consideration to one year. If we further increase the forecast horizon to ten years, we notice that shocks to the U.S. short-rate explains around 30% the variability of the ten-year spread,

Panel b of Table 6 presents the variance decomposition for the rate of depreciation and the foreign exchange risk premium, and it is not surprising to find that the main driver of exchange rate variability is the shock to the rate of depreciation. In particular, it explains around 90% of the variability of the depreciation rate for all forecast horizons. Also, we find that both the domestic latent factor and the rate of depreciation have explanatory power over the foreign exchange risk premia. In particular, they account for around 40% and 50% of its variability, respectively. Finally, the U.S. short-rate has little influence on both the exchange rate and its risk premium.

Germany. Focusing on Panel a of Table 7, which presents variance decompositions from the model and German data, we notice that the rate of depreciation has more explanatory power than the U.S. short rate and the domestic factor at all points of the yield curve for the one-month and one-year forecast horizons. Still, the effect of exchange rate shocks decreases with the bond's maturity. It explains the 87.68% of the one-month ahead variability of the short-end of the curve, while it explains 61.08% of the variability of its long-end. Equally important, the effect of the U.S. short-rate grows with the maturity of the bond under consideration for all forecast horizons. In fact, this state variable becomes the main driver of the ten-year ahead forecast variance of the long-end of the German yield curve: Over 65% of the ten-year ahead variability of the ten-year bond yield is due to the U.S. short-rate.

As a difference with the results for Canada, note in columns 4–6 that the domestic latent factor is now the main driving force of expected excess bond returns. It explains over 90% of the variability of bond risk premia at all maturities and for all forecast horizons. The rate of depreciation, which accounts for almost 90% of the variation of Canadian bond risk premia, now explains only 5% of the forecast variance of German excess bond returns. We also find in the last three columns of Panel a, that very little of the forecast variance of bond premia nor yield spreads can be attributed to the U.S. short-term rate. In effect, over 85% of the one-month ahead variability of the one-year spread. Yet, the explanatory power of this variable decreases with bond's maturity, and the domestic latent factor only explains 25% of the ten-year spread. Finally, the effect of the rate of depreciation tends to increase with both the bond's maturity and the forecast horizon.

We also notice another difference with the Canadian dataset when looking to the variance decomposition of the rate of depreciation in Panel b of Table 7: the main driver of exchange rate variability is the domestic latent factor. It explains around 95% of the variability of the depreciation rate for all forecast horizons. When looking to the exchange rate risk premium, we find that its variability at the short horizon can be attributed to both the latent factor and the rate of depreciation. Each of these two variables explains almost a 45% of the one-month ahead forecast variance of the exchange rate risk premia. Besides, the proportion of the risk premium component explained by exchange rate shocks increases to almost 70% and 75% for the one-year and ten-year ahead horizons, respectively. While the influence of the U.S. short-rate on the exchange rate is almost zero, it accounts for almost 10% of the one-month ahead forecast variance of the exchange risk premium and almost 17% of its ten-year ahead variability.

U.K. Last, we focus on the results for the U.K. in Panel a of Table 8. At short maturities, very little of the one-month and one-year ahead forecast variance can be attributed to the U.S. short-term rate. In fact, this variability is mostly explained by shocks to the exchange rate of depreciation. Here, exchange rate movements explain around 95% of the one-year ahead forecast variance of the one-year yield. However, as we increase the maturity of the bond under consideration, the U.S. short-rate becomes the main driver of the long-end of the yield curve, and almost half of the variability of the ten-year bond is due to U.S. shocks.

These results are similar to those for the German variance decomposition.

Also, the domestic latent factor is the main driving force of expected excess bond returns and explains around 87% of the variability of bond risk premia at all maturities and for all forecast horizons. Likewise, the rate of depreciation accounts for 10% of the forecast variance of the U.K. risk premium, and the effect of U.S. shocks are almost negligible. When looking to the variance decomposition of British bond spreads, we find again that very little of the forecast variance of yield spreads can be attributed to U.S. shocks. In fact, the domestic latent factor tends to explain most of the variability of the one-year spread, while the rate of depreciation explains the forecast variance of five and ten-year yields. That is, the effect of the domestic factor tends to decrease and the effect of exchange rates tend to increase with the maturity of the contract under consideration.

Panel b of Table 8 reveals that the the variance decomposition of the rate of depreciation in the U.K. is similar to that of Germany: the main driver of exchange rates is the domestic latent factor which explains around 95% of the variability of the rate of depreciation at all forecast horizons. Turning to the exchange rate risk premium, we find that its variability at the short horizon is explained by both latent factor and exchange rate shocks. For example, the domestic latent factor explains 67.24% of the variance of the foreign exchange risk premia at the one-month horizon. Once we increase the forecast horizon to one year, we find that both the latent factor and the exchange rate have significant explanatory power over the risk premia: 42.74% and 49.49%, respectively. Finally, over 62% of the ten-year ahead forecast variance of the risk premium can be attributed to exchange rate shocks.

Overall comments. There are several messages that emerge from these tables. First, the U.S. short rate tends to be the main driver of the variability of the long-end of the yield curve regardless of the country being examined or the forecast horizon. Second, the forecast variance of the short-end of the yield curve is mainly explained by shocks to the exchange rate. Finally, U.S. shocks do not explain expected excess returns (risk premium). This is true for both bond and foreign exchange risk premia and these are explained by a combination of domestic and foreign exchange shocks.

4.5 Pricing Errors

Table 9 reports mean pricing errors (MPEs) and mean absolute pricing errors (MAPEs) obtained from the affine term structure model. These are computed as $\eta_t^{(n)} = y_t^{(n)} - a_n - \mathbf{b}_n' \mathbf{x}_{t|t}$ where $\mathbf{x}_{t|t}$ is the estimate of the vector of state variables $\mathbf{x}_t$ conditional on information up to time t: $\mathbf{x}_{t|t} = E_t\left(\mathbf{x}_t | I_t\right)$.

Overall, MPEs tend to be small. In fact, they are less than one bp per month (in absolute value) for all countries and maturities with the exception of the one-month and one-year yield in the U.K. These are still close to one bp per month: 1.1 bp and -1.2 bp, respectively. It is also interesting to highlight that MAPEs of bonds at the middle of the yield curve are smaller than those at the long-end of the yield curve. Nonetheless, they tend to be fairly large. For example, the MAPE of the Canadian one-month yield (ten-year yield) is 5.21bp (5.87 bp) per month, it is 2.92 bp (3.94 bp) for Germany, and 4.59 bp (5.79 bp) for the U.K. As in the case of ADP, we do not find these results surprising because our system only has one latent factor. Additionally, we will argue in section 5 that the magnitudes of these pricing errors are similar to those that we would have obtained by estimating a two-factor arbitrage-free Nelson-Siegel model.

Finally, one-month interest rates tend to have larger MAPEs than the rest of the yields. Therefore, constraining these yields to have zero measurement errors in order to recover latent factors from data on selected yields might lead to misspecification issues.

4.6 Comparison with Other Estimation Methods

Finally, we compare our estimates of the McCallum (1994a) exchange-rate-stabilisation rule to those obtained in previous attempts of estimating this rule. Following Christensen (2000), Panel a of Table 10 reports ordinary least squares estimates of this rule, while Panel b reports exponential GARCH estimates of these parameters. Note that, when using these two approaches, the exchange-rate-stabilisation parameter, ψ_1, is small and positive for Canada and negative for Germany and the U.K. However, we argue that these ordinary least squares (and GARCH) estimates can be biased due endogeneity issues. In particular, as noted in equation (22), the monetary policy shock in the McCallum (1994a) rule coincides with u_{2t} which, in turn, is correlated with the expected rate of depreciation (see equation 6).

To overcome such problems, we also follow Mark and Wu (1996)[9] to estimate the policy rule using instrument variables. As in their case, we use the instrument set given by $(1, \Delta s_{t-1}, \Delta s_{t-2}, r_{t-1} - r^*_{t-1}, r_{t-2} - r^*_{t-2})'$. The results can be found in Panel c. We now find that ψ_1 is negative for Canada and Germany, while it is positive for the U.K. Again, it is not possible to reject that this coefficient is equal to zero for any of the three countries that we include in our study. Again, we believe that such an approach potentially delivers biased estimates of the exchange-rate-stabilisation parameter, ψ_1, because this set of instruments seems to have little explanatory power over the rate of depreciation thus casting doubt on the relevance of the instruments. In particular, the sample R^2 from a regression of the rate of depreciation on the set of instruments is 0.032, 0.016 and 0.031 for Canada, Germany and the U.K., respectively[10] Moreover, since exchange rate changes are intrinsically very difficult to predict, it is not clear how to choose a relevant set of intruments to provide reliable

[9] In particular, we refer to the working paper version of the Mark and Wu (1998) paper.

[10] Although we do not formally test for weak intruments in this regression by means, i.e, of a Stock and Yogo (2005) test, the results provided in Panel c of Table 10 do not seem to pass the Staiger and Stock (1997) rule of thumb that, in the case of one endogenous variable, instruments should be considered weak if the first-stage F is less than 10. In particular, the F-statistic of a regression of the rate of depreciation on the set of instruments is 2.65, 1.27 and 2.54 for Canada, Germany and the U.K., respectively.

inferences about the value of the exchange-rate-stabilisation parameter.

Finally and for the sake of comparison, we provide again the estimates of the McCallum rule obtained using an affine term structure model. Here, we find that the exchange rate stabilisation coefficient is positive and significant at the 5% level for Canada and the U.K., and it is positive and significant at the 10% level for Germany. Therefore, by exploiting information from the entire term structure, we are able to estimate the underlying structural parameters in the policy reaction function more efficiently, at the same time that we avoid the issues that might be plaguing previous single-equation approaches to estimate the McCallum (1994a) rule.

5 Which McCallum Rule?

Monetary policy behavior is not only a solution to the forward premium puzzle but also helps solving another major puzzle in financial economics: the drastic inconsistency of data with the expectation hypothesis of the term structure of interest rates highlighted in, e.g., Fama and Bliss (1987). In particular, McCallum (1994b) shows that by augmenting the expectations-hypothesis model with a monetary policy rule that uses a short-term interest rate instrument and that is sensitive to the slope of the yield curve one can reconcile data and theory. We refer to this rule as the McCallum yield-curve-smoothing policy rule and it takes the form:

$$r_t = \varphi_0 + \varphi_1(y_t^{(n)} - r_t) + \varphi_2 r_{t-1} + v_t \tag{28}$$

where v_t is the monetary policy shock. This policy rule is similar in spirit to that in (4) and it implies that the monetary authority intervenes to try to achieve two goals. The first one is "yield-curve smoothing" governed by the parameter $\varphi_1 > 0$. That is, the central bank interprets a widening term spread as a signal of higher future inflation and, therefore, raises the short-rate accordingly. The second objective is "interest-rate smoothing" governed by the parameter $|\varphi_2| < 1$.

Therefore, we have two competing monetary policy rules trying to explain two different puzzles in financial economics. In this section, we compare our results to those that we would have obtained by embedding the McCallum (1994b) yield-curve-smoothing policy rule into an affine term structure. Still, this is a much easier task than the estimation of the exchange-rate-stabilization rule because Gallmeyer et al. (2005) show that one can rotate the space of state variables in an affine term structure model to relate the short rate to the term premium as in equation (28). In particular, they show that a given m factor affine term structure model can be rotated into a new set of state variables that includes the short rate and the yield spread on $m-1$ bonds of longer maturity. This way, one can express the coefficients in McCallum (1994b) rule as non-linear functions of the parameters of the term structure model. Hence, estimating this rule using a no-arbitrage model amounts to (i) estimating a two-factor affine term structure model, (ii) rotating the space of state variables, and (iii) recovering the coefficients φ_0, φ_1 and φ_2 as functions of the parameters of the original term structure model.

5.1 A no-arbitrage discrete-time Nelson-Siegel model

As previously mentioned, the estimation of a McCallum (1994b) rule requires as a first step the estimation of a two-factor affine term structure model. In particular, we choose to estimate a discrete-time version of the arbitrage-free Nelson-Siegel model presented in Christensen et al. (2007) and introduced in Diebold et al. (2005).[11] This model has several advantages. For one, it is parsimonious and provides a good fit of the yield curve with only a few parameters. Second, it is quite easy to estimate. Third, it is constructed under the no-arbitrage hypothesis and thus it imposes the desirable theoretical restrictions that rule out opportunities for riskless arbitrage. Last, the two latent factors in this model can be interpreted as the level and slope of the yield curve.

In this model, the short rate is just the sum of two latent factors:

$$r_t = z_{1t} + z_{2t}, \tag{29}$$

which, under the physical measure, follow independent AR(1) processes with Gaussian errors:

$$\begin{pmatrix} z_{1t+1} \\ z_{2t+1} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix} \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} + \begin{pmatrix} \upsilon_1 & 0 \\ 0 & \upsilon_2 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}, \tag{30}$$

where $|\rho_i| < 1$ for $i = 1, 2$.

The model is completed by specifying the process that $\mathbf{z}_t = (z_{1t}, z_{2t})'$ follows under the risk-neutral measure.[12] Here we assume again that each latent factor follows an independent AR(1) processes with Gaussian errors:

$$\begin{pmatrix} z_{1t+1} \\ z_{2t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \kappa \end{pmatrix} \begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} + \begin{pmatrix} \upsilon_1 & 0 \\ 0 & \upsilon_2 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}. \tag{31}$$

The difference is that z_{1t} has now a unit root under the risk neutral measure, while we assume $|\kappa| < 1$ to guarantee that z_{2t} is stationary.

Notice that this model falls under the general framework of an affine term structure model. In particular, we can use a set of recursions similar to those in (16) to price bonds in this economy and obtain that

$$y_t^{(n)} = a_n^{NS} + (\mathbf{b}_n^{NS})'\mathbf{z}_t,$$

[11] While the original no-arbitrage Nelson-Siegel models were set-up in continuous-time, we have decided to use discrete-time techniques here in order to be consistent with the modeling choice made on the rest of the paper.

[12] One can easily specify the set of restrictions that guarantee that the prices of risk deliver such a process under the risk neutral measure (see Diebold et al. 2005).

with the factor loadings being

$$\mathbf{b}_n^{NS} = \left[1, \frac{1-\kappa^n}{n(1-\kappa)}\right]'.$$

These two coefficients in $\mathbf{b}_n^{NS}$ share the same properties of the first two factor loadings in the Nelson-Siegel model in Diebold and Li (2006). The first loading is unity which implies that the first latent factor, z_{1t}, affects yields of all maturities one-for-one. Thus, it can be viewed as a long-term/level factor. On the other hand, the second factor starts at one for $n = 1$, and goes to zero as the maturity increases ($n \to \infty$). This way, it affects mainly short maturities, and it can be viewed as a short-term/slope factor. The yield-adjustment term, a_n^{NS}, is similar to that in the arbitrage-free Nelson-Siegel model presented in Christensen et al. (2007).

Finally, we rotate the set of latent factors as shown in Gallmeyer et al. (2005) to relate the short rate to the yield spread on the n-period bond as in the McCallum's (1994b) rule.[13] We show in appendix A that the short rate can be expressed as

$$r_t = \varphi_0 + \varphi_1 \left(y_t^{(n)} - r_t\right) + \varphi_2 r_{t-1} + v_t,$$

where the parameters φ_1 and φ_2 satisfy that:

$$\varphi_1 = \frac{n(1-\kappa)}{n(1-\kappa)-(1-\kappa^n)} \times \frac{\rho_1 - \rho_2}{\rho_2}, \tag{32}$$

$$\varphi_2 = \rho_1, \tag{33}$$

and φ_0 is a highly non-linear function of the parameters of the term structure model. This way, we recover the coefficients on the McCallum (1994b) as functions of the estimated underlying parameters of this term structure model and obtain standard errors of these estimates using the delta method.

5.2 Results

We estimate the discrete-time version of the two-factor arbitrage-free Nelson-Siegel model using the Kalman filter. We assume that all yields are observed with measurement error. While not reported for space considerations, we find that our estimated model share many similar features to those in Diebold and Li (2006) and Christensen et al. (2007). For instance, we find that both the level and the slope factor are very persistent and that the slope factor is more volatile than the level factor. Finally, the estimate (standard error) of the parameter κ is 0.961 (0.002) for Canada, 0.974 (0.001) for Germany and 0.915 (0.005) for the U.K. These numbers are similar to the equivalent (discretized) parameter estimates found in Christensen et al. (2007).

[13] We choose $n = 120$ months.

Next, we recover the coefficients of the McCallum (1994b) yield-curve-smoothing policy rule φ_0, φ_1 and φ_2 from the estimated parameters of the Nelson-Siegel model and compute their standard errors using the delta method. These are reported in Panel a of Table 11. Notice that the estimated yield-curve smoothing parameter, φ_1, is positive for all three countries. This suggests that the monetary authority interprets a widening term spread as a signal of higher future inflation and, therefore, intervenes in the short-term debt market raising the short-rate accordingly. This coefficient is significant at the 5% level for Canada and the U.K. and significant at the 10% level for Germany. Yet, this coefficient tends to be small: a one percent change in the spread leads to a 1.68 bp per month increase in the Canadian short rate, 1.01 bp increase in the German short rate, and 2.34 bp increase in the British short rate. On the other hand, the interest rate smoothing parameter, φ_2, is close to one for all three countries under consideration.

To compare how both McCallum rule models fit the yield curve, Panel b of Table 11 reports MPEs and MAPEs obtained from the Nelson-Siegel model. Note that this panel is analogous to Table 8. We find that the MPEs obtained from the Nelson-Siegel model are all larger than those reported for the McCallum (1994a) affine term structure model. They are now larger than one basis point. For example, the MPE of the Canadian one-month yield (ten-year yield) is 3.75bp (2.25 bp) per month, 2.39 bp (0.91 bp) per month for Germany, and 3.61 bp (1.71 bp) per month for the U.K. Looking to MAPEs, we find a similar picture: the McCallum (1994a) affine term structure model still tends to do better. However, we now find that the Nelson-Siegel model provides a better fit for the long-end of yield curve. For example, the MAPE for the Nelson-Siegel model (McCallum exchange-rate-stabilization model) is 4.48 bp (5.87 bp) for Canada, 2.87 bp (3.94 bp) for Germany, and 3.79 bp (5.79 bp).

To conclude, both McCallum rule models seem to provide similar fits of the yield curve. If any, the McCallum (1994a) exchange-rate-stabilisation rule seems to do slightly better.

6 Final Remarks

In this paper we estimate the McCallum (1994a) rule within the framework of an affine term structure model with time varying risk premia. Using yield curve data over the period January 1979 to December 2005 for Canada, Germany and the U.K., we find that the monetary authority in these three countries responded to exchange rate movements. In particular, we find that the exchange rate stabilisation coefficient is significant at the 5% level for Canada and the U.K. and significant at the 10% level for Germany. This indicates that the central bank interprets a depreciating exchange rate as a signal of higher expected future inflation and, therefore, it increases the short rate. More importantly, the proposed affine term structure model replicates the forward premium puzzle, as it is able to replicate a negative slope coefficient on a regression of the ex-post rate of depreciation on a constant and the interest rate differential for all three datasets.

Similarly, we find that the U.S. short-rate tends to be the main driver of the variability of the long-end of the yield curve regardless of the country being examined. For example, 95% of the ten-year ahead variance of the Canadian ten-year yield, 65% of the variance of the German ten-year yield and 87% of the variance of the British ten-year yield can be attributed to movements in the U.S. short-rate. Second, the variability of the short-end of the yield curve is mainly explained by shocks to the exchange rate. Over 56% of the one-month ahead variance of the Canadian one-month yield, 87% of the variance of the German one-month yield, and 90% of the variance of the British one-month yield is due to exchange rate movements. Finally, both bond and foreign exchange risk premia are explained by a combination of domestic and foreign exchange shocks with the U.S. short-rate playing little or no role at all.

While in this paper we only estimate a McCallum (1994a) rule, our modelling framework can be easily extended to estimate other monetary policy reaction functions where the central bank respond to the rate of depreciation (see the open-economy Taylor-rules of Svensson, 2000, and Taylor, 2001). In such cases, the estimation of these rules requires to include the exchange rate into the set of state variables, and, therefore, one has to guarantee again the self-consistency of the model.

We have also found that while the McCallum (1994a) exchange-rate-stabilisation provides a better fit of the curve overall, the McCallum (1994b) yield-curve-smoothing rule provides a better fit of the long-end of the yield curve. Thus, it would be desirable to obtain a rule that combines both aspects of the monetary policy explanation. That is, a rule such that the central bank increases the short-rate in response to a depreciating exchange rate and to a widening term spread. Such a rule was proposed by Kugler (2000) and its estimation using no-arbitrage methods remains an open research question.

Finally, since we do not rely on a microfounded model, our modelling strategy has the main drawback that we are unable to link the prices of risk to individuals' preferences. Constructing an open-economy version of the structural model in Gallmeyer et al. (2005) or Gallmeyer et al. (2008) would allow us to better understand the monetary policy reaction function of such central banks.

References

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

Ang, A., S. Dong and M. Piazzesi (2007): "No-Arbitrage Taylor Rules," Columbia University Mimeo.

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.

Backus D.K., F. Gavazzoni, C.I. Telmer and S.E. Zin (2009): "Monetary Policy and the Uncovered Interest Rate Parity Puzzle," Carnegie Mellon University Mimeo.

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.

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

Brennan, M.J., and Y. Xia (2006): "International Capital Markets and Foreign Exchange Risk," *Review of Financial Studies*, 19, 753-795.

Burnside, C., M. Eichenbaum, I. Kleshchelski, and S. Rebelo (2006). "The Returns to Currency Speculation," NBER Working Paper No. 12489.

Chen, R.R. and L. Scott (1993) "Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14-31

Christensen, M. (2000): "Uncovered Interest Parity and Policy Behavior," *Economics Letters*, 69, 81-87.

Christensen, J.E., F.X. Diebold and G.D. Rudebush (2007): "The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models," University of Pennsylvania Mimeo.

Clarida, Richard H., and Mark 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.

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.

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

Diebold, F.X. and C. Li (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130, 337-364.

Diebold, F.X., M. Piazzesi and G.D. Rudebusch (2005): "Modeling Bond Yields in Finance and Macroeconomics," *American Economic Review*, 95, 415-420.

Diez de los Rios, A. (2009): "Can Affine Term Structure Models Help Us Predict Exchange Rates," *Journal of Money, Credit and Banking*, 41, 755-766.

Dong, S. (2006): "Macro Variables Do Drive Exchange Rate Movements: Evidence from a No-Arbitrage Model," Columbia University Mimeo.

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

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

Fama E.F. and R.R. Bliss (1987): "The Information in Long-Maturity Forward Rates," *American Economic Review*, 77, 680-692.

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

Gallmeyer M.F., B. Hollifield and S.E. Zin (2005): "Taylor Rules, McCallum Rules and the Term Structure of Interest Rates," *Journal of Monetary Economics*, 52, 921-950.

Gallmeyer M.F., B. Hollifield, F. Palomino and S.E. Zin (2008): "Term Premium Dynamics and the Taylor Rule," Carnegie Mellon University Mimeo.

de Jong, F. (2000): "Time Series and Cross-Section Information in Affine Term-Structure Models," J*ournal of Business & Economic Statistics*, 18, 300-314.

Kugler, P. (2000): "The Expectations Hypothesis of the Term Structure of Interest Rates, Open Interest Rate Parity and Central Bank Policy Reaction," *Economics Letters*, 66, 209–214.

Leippold, M. and L. Wu (2007): "Design and Estimation of Multi-Currency Quadratic Models," *Review of Finance*, 11, 167-207.

Mark, N.C. and Y. Wu (1996): "Risk, Policy Rules, and Noise: Rethinking Deviations from Uncovered Interest Parity," Ohio State University Mimeo.

Mark, N.C. and Y. Wu (1998): "Risk, Policy Rules, and Noise: Rethinking Deviations from Uncovered Interest Parity," *Economic Journal* 108, 1686-1706.

McCallum, B.T. (1994a): "A Reconsideration of the Uncovered Interest Rate Parity Relationship," *Journal of Monetary Economics*, 33-105.

McCallum, B.T. (1994b): "Monetary Policy and the Term Structure of Interest Rates," NBER Working Paper No. 4938.

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

Sarno, L. (2005): "Viewpoint: Towards a Solution to the Puzzles in Exchange Rate Economics: Where do We Stand?," *Canadian Journal of Economics*, 38, 673-708.

Staiger, D. and J.H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557-86.

Stock, J.H. and M. Yogo (2005): "Testing for Weak Instruments in Linear IV Regression," in D.W.K Andrews (ed.), *Identification and Inference for Econometric Models*. New York. Cambrigdge University Press

Svensson, L.E.O. (2000): "Open-Economy Inflation Targeting," *Journal of International Economics*, 50, 155–183

Taylor, J.B. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

Taylor, J.B. (2001): "The Role of the Exchange Rate in Monetary-Policy Rules," *American Economic Review*, 91, 263-267.

Taylor, M.P. (1995): "The Economics of Exchange Rates," *Journal of Economic Literature*, 33, 13-47

Appendix

A Latent factor rotation

In this appendix, we use the methodology developped in Gallmeyer et al. (2005) to rotate the space of state variables in our affine term structure model and relate the short rate to the term premium as in equation (28). In particular, a given m factor model can be rotated into a new set of state variable that includes the short rate and the yield spread on $m-1$ bonds of longer maturity. Since in our equation McCallum (1994b) rule we only have the spread on the n-period bond, we focus only on the rotation of models with only two latent factors.

Let $\mathbf{x}_t$ be a 2×1 vector of state variables such that the short rate is:

$$r_t = \delta_0 + \boldsymbol{\delta}'\mathbf{x}_t,$$

and $\mathbf{x}_t$ follows a VAR(1) under both the physical measure:

$$\mathbf{x}_{t+1} = \boldsymbol{\theta} + \boldsymbol{\Phi}\mathbf{x}_t + \boldsymbol{\Sigma}^{1/2}\boldsymbol{\varepsilon}_{t+1}, \tag{34}$$

and the risk-neutral measure:

$$\mathbf{x}_{t+1} = \boldsymbol{\theta}^Q + \boldsymbol{\Phi}^Q\mathbf{x}_t + \boldsymbol{\Sigma}^{1/2}\boldsymbol{\varepsilon}_{t+1}. \tag{35}$$

Under these assumptions, we know that there is an affine term structure for continuously compounded yields:

$$y_t^{(n)} = a_n + \mathbf{b}_n'\mathbf{x}_t,$$

where a_n and $\mathbf{b}_n$ solve some recursive relations. Note that our model in section 4 belongs to this category.

Following Gallmeyer et al. (2005) we define a new 2×1 vector of state variables, $\mathbf{z}_t$, to include the short rate and the yield spread on the n-period bond:

$$\mathbf{z}_t = \left(r_t, s_t^{(n)}\right)',$$

where $s_t^{(n)} = y_t^{(n)} - r_t$. This new vector of state variables is an affine function of the original state variable. That is, $\mathbf{z}_t = \mathbf{d} + \mathbf{H}\mathbf{x}_t$, where

$$\mathbf{d} = \begin{pmatrix} a_1 \\ a_n - a_1 \end{pmatrix}, \qquad \mathbf{H} = \begin{pmatrix} \mathbf{b}_1' \\ \mathbf{b}_n' - \mathbf{b}_1' \end{pmatrix}.$$

Moreover, provided that $\mathbf{H}$ has full rank (as it is in the case of our Nelson-Siegel model), we can rotate the original set of state variable to write:

$$\mathbf{z}_t = \widetilde{\boldsymbol{\theta}} + \widetilde{\boldsymbol{\Phi}}\mathbf{z}_{t-1} + \mathbf{v}_t,$$

where $\widetilde{\boldsymbol{\theta}} = (\mathbf{I} - \mathbf{H}\boldsymbol{\Phi}\mathbf{H}^{-1})\mathbf{d} + \mathbf{H}\boldsymbol{\theta}$, $\widetilde{\boldsymbol{\Phi}} = \mathbf{H}\boldsymbol{\Phi}\mathbf{H}^{-1}$ and $\boldsymbol{\epsilon}_t = \mathbf{H}\boldsymbol{\Sigma}^{1/2}\boldsymbol{\varepsilon}_t$. In particular, this last equation allows us to express the short rate as:

$$r_t = \widetilde{\theta}_1 + \widetilde{\phi}_{11}r_{t-1} + \widetilde{\phi}_{12}s_{t-1}^{(n)} + \epsilon_{1t},$$

as well as the spread, $s_{t-1}^{(n)}$, as

$$s_{t-1}^{(n)} = \frac{1}{\widetilde{\phi}_{22}} s_t^{(n)} - \frac{\widetilde{\theta}_2}{\widetilde{\phi}_{22}} - \frac{\widetilde{\phi}_{21}}{\widetilde{\phi}_{22}} r_{t-1} - \frac{1}{\widetilde{\phi}_{22}} \epsilon_{2t},$$

Substituting $s_{t-1}^{(n)}$ into r_t we get equation a McCallum (1994b) rule:

$$r_t = \varphi_0 + \varphi_1 \left(y_t^{(n)} - r_t \right) + \varphi_2 r_{t-1} + v_t,$$

where φ_0, φ_1, and φ_2 are non-linear functions of the underlying parameters of the term structure model satisfying:

$$\varphi_1 = \frac{\widetilde{\phi}_{12}}{\widetilde{\phi}_{22}},$$

$$\varphi_0 = \widetilde{\theta}_1 - \varphi_1 \widetilde{\theta}_2, \qquad \varphi_2 = \widetilde{\phi}_{11} - \varphi_1 \widetilde{\phi}_{21},$$

and $v_t = \epsilon_{1t} - \varphi_1 \epsilon_{2t}$. Specializing these previous equations to the discrete-time no-arbitrage Nelson-Siegel model in section 4, we obtain equations (32) and (33) in the text.

Table 1
Summary Statistics

Variable	Mean	Std. Dev	Min.	Max.	Autocorrelation 1	2	3
U.S.							
1-month yield	6.844	3.941	1.016	20.250	0.979	0.955	0.932
1-year yield	6.630	3.300	1.060	15.870	0.985	0.966	0.950
2-year yield	6.896	3.161	1.300	15.730	0.987	0.969	0.954
5-year yield	7.394	2.869	2.350	15.310	0.989	0.974	0.962
10-year yield	7.750	2.593	3.580	14.860	0.990	0.978	0.966
Canada							
Rate of Depreciation	-0.093	17.964	-52.402	54.441	0.019	-0.039	0.018
1-month yield	7.667	4.209	2.016	22.313	0.987	0.969	0.947
1-year yield	7.582	3.585	2.020	18.820	0.987	0.971	0.953
2-year yield	7.760	3.341	2.400	18.080	0.986	0.968	0.952
5-year yield	8.160	3.001	3.270	17.420	0.987	0.973	0.960
10-year yield	8.537	2.895	3.830	17.290	0.990	0.979	0.969
Germany							
Rate of Depreciation	-0.452	38.424	-100.314	132.246	0.060	0.054	0.029
1-month yield	5.431	2.618	2.016	15.000	0.985	0.973	0.963
1-year yield	5.534	2.487	1.930	13.170	0.992	0.978	0.962
2-year yield	5.734	2.326	2.040	12.330	0.991	0.977	0.962
5-year yield	6.246	1.985	2.560	11.490	0.990	0.977	0.963
10-year yield	6.648	1.650	3.210	10.240	0.989	0.975	0.963
U.K.							
Rate of Depreciation	0.547	36.552	-163.359	157.402	0.063	0.002	0.010
1-month yield	8.949	3.909	3.375	18.625	0.988	0.976	0.961
1-year yield	8.294	3.186	3.230	14.960	0.988	0.975	0.962
2-year yield	8.352	3.037	3.320	15.120	0.988	0.974	0.960
5-year yield	8.517	2.960	3.770	15.540	0.989	0.976	0.964
10-year yield	8.584	2.983	4.050	15.440	0.993	0.984	0.976

Note: Data are sampled monthly from January 1979 to December 2005. All variables are measured in percentage points per year, and monthly rates of depreciation are annualized by multiplying by 1,200.

Table 2
Estimates of McCallum (1994a) Affine Term Structure Model: Canada

Panel a: McCallum Rule

ψ_0	Δs_t	$r_t - r_t^*$
0.0004	0.0175	0.9934
(0.0022)	(0.0053)	(0.0188)

Panel b: Physical Measure

		$\boldsymbol{\Phi}$			$\boldsymbol{\Sigma}^{1/2}$		
	$\boldsymbol{\theta}$	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
r_t^*	0.0008	0.9983	0	0	0.0129	0	0
	(0.0009)	(0.0017)	-	-	(0.0004)	-	-
f_t	ψ_0	0	ψ_2	$\psi_1\psi_2$	0	0.0266	0
	-	-	-	-	-	(0.0030)	-
Δs_t	-0.0744	0.6571	-3.3871	-0.0537	-0.4786	-0.2086	1.5383
	(0.1927)	(0.3127)	(0.7691)	(0.0645)	(0.0865)	(0.5047)	(0.0840)

Panel c: Risk Neutral Measure

		$\boldsymbol{\Phi}^Q$		
	$\boldsymbol{\theta}^Q$	r_t^*	f_t	Δs_t
r_t^*	0.0166	0.9935	0	0
	(0.0013)	(0.0004)	-	-
f_t	0.0014	0.0006	0.9597	0.0033
	(0.0078)	(0.0008)	(0.0056)	(0.0050)
Δs_t	$-\frac{1}{2}e_3'\boldsymbol{\Sigma}e_3$	0	1	ψ_1
	-	-	-	-

Panel d: Tests

H_0	Wald	d.f	p-value
$\boldsymbol{\Phi} = \boldsymbol{\Phi}^Q$	81.86	7	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^Q$	151.98	3	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^{Q^*}$	160.76	3	<0.0001

Note: This table lists the estimated coefficients for the affine term structure model in equations (6)-(9) subject to the restrictions in equation (22) for Canada. We assume that all (both domestic and foreign) yields are observed with error. Panel a reports the estimates of the McCallum (1994a) rule in equation (4): $r_t - r_t^* = \psi_0 + \psi_1 \Delta s_t + \psi_2(r_{t-1} - r_{t-1}^*) + e_t$. Panel b presents the estimates of the parameters of the model under the physical measure, while panel c reports the parameters of the model under the risk neutral measure. In panel d, we test whether the coefficients under both the physical and risk neutral measure are the same. The estimate (standard error) of the standard deviation of the measurement error is $\sigma_\eta = 0.0634$ (0.0008). Data are sampled monthly from January 1979 to December 2005.

Table 3
Estimates of McCallum (1994a) Affine Term Structure Model: Germany

Panel a: McCallum Rule

ψ_0	Δs_t	$r_t - r_t^*$
0.0063	0.0336	1.0409
(0.0079)	(0.0192)	(0.0410)

Panel b: Physical Measure

		$\mathbf{\Phi}$			$\mathbf{\Sigma}^{1/2}$		
	$\boldsymbol{\theta}$	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
r_t^*	0.0010	0.9980	0	0	0.0130	0	0
	(0.0010)	(0.0019)	-	-	(0.0004)	-	-
f_t	ψ_0	0	ψ_2	$\psi_1\psi_2$	0	0.1049	0
	-	-	-	-	-	(0.0637)	-
Δs_t	-0.2667	0.0211	-1.6227	-0.0674	-0.2886	-3.3188	0.5893
	(0.2093)	(0.1423)	(0.8876)	(0.0464)	(0.1435)	(0.1536)	(0.3709)

Panel c: Risk Neutral Measure

		$\mathbf{\Phi}^Q$		
	$\boldsymbol{\theta}^Q$	r_t^*	f_t	Δs_t
r_t^*	0.0149	0.9920	0	0
	(0.0019)	(0.0003)	-	-
f_t	0.1693	-0.0040	0.9459	0.0285
	(0.1117)	(0.0008)	(0.0194)	(0.0174)
Δs_t	$-\frac{1}{2}e_3'\mathbf{\Sigma}e_3$	0	1	ψ_1
	-	-	-	-

Panel d: Tests

H_0	Wald	d.f	p-value
$\mathbf{\Phi} = \mathbf{\Phi}^Q$	66.69	7	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^Q$	241.80	3	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^{Q^*}$	258.08	3	<0.0001

Note: This table lists the estimated coefficients for the affine term structure model in equations (6)-(9) subject to the restrictions in equation (22) for Germany. We assume that all (both domestic and foreign) yields are observed with error. Panel a reports the estimates of the McCallum (1994a) rule in equation (4): $r_t - r_t^* = \psi_0 + \psi_1\Delta s_t + \psi_2(r_{t-1} - r_{t-1}^*) + e_t$. Panel b presents the estimates of the parameters of the model under the physical measure, while panel c reports the parameters of the model under the risk neutral measure. In panel d, we test whether the coefficients under both the physical and risk neutral measure are the same. The estimate (standard error) of the standard deviation of the measurement error is $\sigma_\eta = 0.0532$ (0.0007). Data are sampled monthly from January 1979 to December 2005.

Table 4
Estimates of McCallum (1994a) Affine Term Structure Model: U.K.

Panel a: McCallum Rule

ψ_0	Δs_t	$r_t - r_t^*$
-0.0194	0.0313	1.0861
(0.0119)	(0.0115)	(0.0520)

Panel b: Physical Measure

		$\mathbf{\Phi}$			$\mathbf{\Sigma}^{1/2}$		
	$\boldsymbol{\theta}$	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
r_t^*	0.0008	0.9985	0	0	0.0130	0	0
	(0.0009)	(0.0016)	-	-	(0.0004)	-	-
f_t	ψ_0	0	ψ_2	$\psi_1\psi_2$	0	0.0931	0
	-	-	-	-	-	(0.0378)	-
Δs_t	0.7826	0.0791	-3.9614	-0.2354	-0.4053	-3.3186	1.0292
	(0.3295)	(0.2044)	(1.1898)	(0.0545)	(0.1061)	(0.1916)	(0.4196)

Panel c: Risk Neutral Measure

		$\mathbf{\Phi}^Q$		
	$\boldsymbol{\theta}^Q$	r_t^*	f_t	Δs_t
r_t^*	0.0159	0.9932	0	0
	(0.0015)	(0.0004)	-	-
f_t	0.1399	0.0048	0.9456	0.0222
	(0.0711)	(0.0008)	(0.0117)	(0.0103)
Δs_t	$-\frac{1}{2}e_3'\mathbf{\Sigma}e_3$	0	1	ψ_1
	-	-	-	-

Panel d: Tests

H_0	Wald	d.f	p-value
$\mathbf{\Phi} = \mathbf{\Phi}^Q$	99.23	7	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^Q$	299.48	3	<0.0001
$\boldsymbol{\theta} = \boldsymbol{\theta}^{Q^*}$	202.81	3	<0.0001

Note: This table lists the estimated coefficients for the affine term structure model in equations (6)-(9) subject to the restrictions in equation (22) for the U.K. We assume that all (both domestic and foreign) yields are observed with error. Panel a reports the estimates of the McCallum (1994a) rule in equation (4): $r_t - r_t^* = \psi_0 + \psi_1 \Delta s_t + \psi_2(r_{t-1} - r_{t-1}^*) + e_t$. Panel b presents the estimates of the parameters of the model under the physical measure, while panel c reports the parameters of the model under the risk neutral measure. In panel d, we test whether the coefficients under both the physical and risk neutral measure are the same. The estimate (standard error) of the standard deviation of the measurement error is $\sigma_\eta = 0.0628$ (0.0008). Data are sampled monthly from January 1979 to December 2005.

Table 5
Implied Betas

Maturity in months (n)	*Canada* Affine	*Canada* Sample	*Germany* Affine	*Germany* Sample	*U.K.* Affine	*U.K.* Sample
1	-1.770	-1.348	-1.261	-1.201	-2.835	-2.556
12	-1.246	-0.699	-1.221	-1.375	-2.283	-2.616
24	-0.872	-0.529	-1.187	-1.294	-2.047	-2.209
60	-0.336	-0.276	-1.006	-1.036	-1.345	-1.191
120	-0.104	-0.082	-0.652	-0.671	-0.655	-0.541

Note: This table presents the term structure of uncovered interest parity regression slopes implied by the affine term structure model in equations (6)-(9) subject to the restrictions in equation (22). These are computed using equation (23) in the main text and by treating the estimates displayed in tables 2-4 as truth. For comparison purposes, we also compute sample estimates of these regression slopes from the coefficientes of a VAR(1) on the rate of depreciation, Δs_t, and the set of interest rate differentials $(y_t^{(1M)} - y_t^{(1M)*}, \ldots, y_t^{(10Y)} - y_t^{(10Y)*})$. Data are sampled monthly from January 1979 to December 2005.

Table 6
Variance Decomposition: Canada

Panel a: Bond Yields

	Yield Levels			Bond Risk Premia			Yield Spreads		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead									
1-month yield	1.61	41.52	56.87	-	-	-	-	-	-
1-year yield	18.29	75.06	6.65	8.45	1.32	90.23	7.76	0.63	96.61
5-year yield	45.69	51.23	3.08	8.35	1.21	90.43	3.89	19.82	76.29
10-year yield	67.31	30.88	1.81	8.20	1.21	90.59	1.78	30.55	67.67
One-year ahead									
1-month yield	8.18	38.49	53.33	-	-	-	-	-	-
1-year yield	12.85	40.40	46.75	8.64	2.08	89.28	6.34	8.76	84.90
5-year yield	39.79	28.00	32.21	8.62	1.82	89.57	1.51	35.86	62.63
10-year yield	64.08	16.70	19.21	8.58	1.80	89.61	0.66	39.49	59.85
Ten-year ahead									
1-month yield	65.76	14.35	19.89	-	-	-	-	-	-
1-year yield	71.19	13.09	15.72	9.08	2.28	88.63	12.50	9.80	77.70
5-year yield	87.74	5.58	6.68	9.38	1.97	88.65	19.63	30.03	50.34
10-year yield	94.18	2.65	3.17	9.87	1.94	88.19	27.62	29.06	43.32

Panel b: Exchange Rates

	Depreciation Rate			FX Risk Premia		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead	8.68	1.65	89.67	7.48	42.89	49.62
One-year ahead	8.68	2.95	88.37	7.61	39.30	53.09
Ten-year ahead	8.69	3.31	88.00	7.42	39.26	53.32

Note: Panel a reports one-month, one-year and ten-year ahead variance decompositions of forecast variance for (i) yield levels, $y_t^{(n)}$, (ii) bond risk premium, $E_t rx_{t+1}^{(n)} = ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t$, and (iii) yield spreads, $y_t^{(n)} - y_t^{(1)}$. Panel b reports forecast variance decompositions of (i) the rate of depreciation, Δs_{t+1}, and (ii) the foreign exchange rate risk premium, $E_t sx_{t+1}^{(n)} = \Delta s_{t+1} + r_t^* - r_t$. We ignore observation errors when computing these variance decompositions. Data are sampled monthly from January 1979 to December 2005.

Table 7
Variance Decomposition: Germany

Panel a: Bond Yields

	Yield Levels			Bond Risk Premia			Yield Spreads		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead									
1-month yield	2.50	9.81	87.68	-	-	-	-	-	-
1-year yield	7.08	3.63	89.29	0.71	95.26	4.03	1.11	85.10	13.79
5-year yield	19.46	4.50	76.04	0.68	95.25	4.07	1.20	42.86	55.94
10-year yield	35.18	3.73	61.08	0.64	95.28	4.08	0.50	25.16	74.34
One-year ahead									
1-month yield	4.36	6.62	89.01	-	-	-	-	-	-
1-year yield	6.33	5.98	87.68	0.74	94.90	4.36	1.03	60.81	38.16
5-year yield	17.79	5.36	76.84	0.73	94.82	4.45	0.62	13.24	86.14
10-year yield	33.02	4.38	62.60	0.74	94.81	4.46	0.13	9.00	90.87
Ten-year ahead									
1-month yield	29.03	4.74	66.23	-	-	-	-	-	-
1-year yield	32.90	4.38	62.72	1.49	93.23	5.28	3.70	34.47	61.83
5-year yield	50.07	3.28	46.65	1.82	92.72	5.46	6.01	8.09	85.90
10-year yield	64.68	2.32	33.00	2.36	92.20	5.45	8.68	6.65	84.67

Panel b: Exchange Rates

	Depreciation Rate			FX Risk Premia		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead	0.73	96.24	3.03	10.78	45.11	44.11
One-year ahead	0.75	96.13	3.12	16.94	13.31	69.76
Ten-year ahead	0.80	95.86	3.34	17.06	7.88	75.06

Note: Panel a reports one-month, one-year and ten-year ahead variance decompositions of forecast variance for (i) yield levels, $y_t^{(n)}$, (ii) bond risk premium, $E_t rx_{t+1}^{(n)} = ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t$, and (iii) yield spreads, $y_t^{(n)} - y_t^{(1)}$. Panel b reports forecast variance decompositions of (i) the rate of depreciation, Δs_{t+1}, and (ii) the foreign exchange rate risk premium, $E_t sx_{t+1}^{(n)} = \Delta s_{t+1} + r_t^* - r_t$. We ignore observation errors when computing these variance decompositions. Data are sampled monthly from January 1979 to December 2005.

Table 8
Variance Decomposition: U.K.

Panel a: Bond Yields

	Yield Levels			Bond Risk Premia			Yield Spreads		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead									
1-month yield	0.01	9.96	90.03	-	-	-	-	-	-
1-year yield	3.34	17.83	78.84	1.57	87.95	10.49	2.89	76.91	20.20
5-year yield	23.49	17.68	58.84	1.52	87.75	10.72	6.14	37.58	56.28
10-year yield	49.69	11.81	38.50	1.43	87.82	10.75	5.56	22.58	71.87
One-year ahead									
1-month yield	1.69	1.46	96.85	-	-	-	-	-	-
1-year yield	4.01	1.78	94.20	1.57	87.96	10.47	4.00	64.66	31.34
5-year yield	26.33	1.74	71.93	1.54	87.74	10.71	7.16	10.65	82.19
10-year yield	53.68	1.11	45.21	1.51	87.76	10.73	4.92	4.55	90.53
Ten-year ahead									
1-month yield	39.17	0.42	60.41	-	-	-	-	-	-
1-year yield	46.75	0.45	52.80	1.62	87.90	10.48	4.47	53.09	42.44
5-year yield	74.28	0.28	25.44	1.78	87.48	10.74	6.45	5.47	88.08
10-year yield	87.44	0.14	12.42	2.21	87.07	10.72	6.65	2.16	91.20

Panel b: Exchange Rates

	Depreciation Rate			FX Risk Premia		
	r_t^*	f_t	Δs_t	r_t^*	f_t	Δs_t
One-month ahead	1.34	90.00	8.66	4.47	67.24	28.29
One-year ahead	1.48	89.02	9.50	7.77	42.74	49.49
Ten-year ahead	1.58	88.08	10.33	8.93	28.69	62.37

Note: Panel a reports one-month, one-year and ten-year ahead variance decompositions of forecast variance for (i) yield levels, $y_t^{(n)}$, (ii) bond risk premium, $E_t rx_{t+1}^{(n)} = ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t$, and (iii) yield spreads, $y_t^{(n)} - y_t^{(1)}$. Panel b reports forecast variance decompositions of (i) the rate of depreciation, Δs_{t+1}, and (ii) the foreign exchange rate risk premium, $E_t sx_{t+1}^{(n)} = \Delta s_{t+1} + r_t^* - r_t$. We ignore observation errors when computing these variance decompositions. Data are sampled monthly from January 1979 to December 2005.

Table 9
Pricing Errors in Basis Points

	1 month	1 year	2 year	5 year	10 year
Canada					
Mean Pricing Error	0.81	-0.77	-0.62	0.11	0.10
Mean Absolute Pricing Error	5.21	2.44	2.71	4.03	5.87
Germany					
Mean Pricing Error	0.55	0.01	-0.60	-0.06	0.23
Mean Absolute Pricing Error	2.92	1.97	2.22	2.96	3.94
U.K.					
Mean Pricing Error	1.09	-1.24	-0.71	0.74	-0.30
Mean Absolute Pricing Error	4.59	3.10	3.19	4.40	5.79

Note: This table reports mean pricing errors and mean absolute pricing errors for the affine term structure model. These are computed as $\eta_t^{(n)} = y_t^{(n)} - a_n - \mathbf{b}_n' \mathbf{x}_{t|t}$ where $\mathbf{x}_{t|t}$ is the estimate of the vector of state variables $\mathbf{x}_t$ conditional on information up to time t: $\mathbf{x}_{t|t} = E_t\left(\mathbf{x}_t \middle| I_t\right)$. Data are sampled monthly from January 1979 to December 2005.

Table 10
Comparison of McCallum Rule Estimates

Panel a: OLS Estimates

	ψ_0	ψ_1	ψ_2
Canada	0.0083	0.0007	0.8735
	(0.0066)	(0.0031)	(0.0585)
Germany	-0.0047	-0.0042	0.9510
	(0.0043)	(0.0012)	(0.0153)
U.K.	0.0144	-0.0040	0.9159
	(0.0057)	(0.0014)	(0.0213)

Panel b: E-GARCH Estimates

	ψ_0	ψ_1	ψ_2
Canada	0.0013	0.0006	0.9458
	(0.0015)	(0.0011)	(0.0105)
Germany	-0.0047	-0.0004	1.0003
	(0.0010)	(0.0004)	(0.0052)
U.K.	0.0023	-0.0020	0.9738
	(0.0018)	(0.0006)	(0.0086)

Panel c: Intrumental Variables Estimates

	ψ_0	ψ_1	ψ_2
Canada	-0.0011	-0.0166	0.9413
	(0.0049)	(0.0336)	(0.0471)
Germany	-0.0047	-0.0032	0.9756
	(0.0037)	(0.0187)	(0.0190)
U.K.	0.0033	0.0221	0.9741
	(0.0327)	(0.0603)	(0.1635)

Panel d: No-Arbitrage Estimates

	ψ_0	ψ_1	ψ_2
Canada	0.0004	0.0175	0.9934
	(0.0022)	(0.0053)	(0.0188)
Germany	0.0063	0.0336	1.0409
	(0.0079)	(0.0192)	(0.0410)
U.K.	-0.0194	0.0313	1.0861
	(0.0119)	(0.0115)	(0.0520)

Note: Panel a reports ordinary least squares of the parameters of the McCallum (1994a) rule in equation (4): $r_t - r_t^* = \psi_0 + \psi_1 \Delta s_t + \psi_2 (r_{t-1} - r_{t-1}^*) + e_t$. Panel b reports exponential GARCH estimates of these parameters. Panel c reports estimates of the McCallum rule when using the instrument set given by $(1, \Delta s_{t-1}, \Delta s_{t-2}, r_{t-1} - r_{t-1}^*, r_{t-2} - r_{t-2}^*)$. Panel d reports again the estimates of the coefficients in the McCallum rule obtained using an affine term structure model in Tables 2-4. Data are sampled monthly from January 1979 to December 2005.

Table 11
Estimates of McCallum (1994b) Rule

Panel a: McCallum Rule

	φ_0	$y_t^{(n)} - r_t$	r_{t-1}
Canada	-0.0011	0.0168	0.9998
	(0.3051)	(0.0071)	(0.0003)
Germany	-0.0013	0.0101	0.9997
	(0.2774)	(0.0060)	(0.0004)
U.K.	0.0008	0.0234	0.9999
	(0.3219)	(0.0105)	(0.0001)

Panel b: Pricing Errors in basis points

	1 month	1 year	2 years	5 years	10 years
Canada					
Mean Pricing Error	3.75	-1.90	-3.38	-2.75	2.25
Mean Absolute Pricing Error	5.38	2.92	4.26	4.43	4.48
Germany					
Mean Pricing Error	2.39	-0.60	-1.82	-1.75	0.91
Mean Absolute Pricing Error	3.62	1.86	2.80	2.96	2.87
U.K.					
Mean Pricing Error	3.61	-3.75	-3.64	-1.64	1.71
Mean Absolute Pricing Error	3.95	4.27	4.70	4.27	3.79

Note: Panel a reports estimates of the McCallum (1994b) yield-curve-smoothing policy rule in equation (24): $r_t = \varphi_0 + \varphi_1 (y_t^{(n)} - r_t) + \varphi_2 r_{t-1} + v_t$. These are functions of the underlying parameters of the no-arbitrage Nelson-Siegel model in equations (25)-(29). Standard errors are computed using the delta method. Panel b reports mean pricing errors and mean absolute pricing errors for the no-arbitrage Nelson-Siegel. Data are sampled monthly from January 1979 to December 2005.

Figure 1: U.S. short-rate latent factor estimate

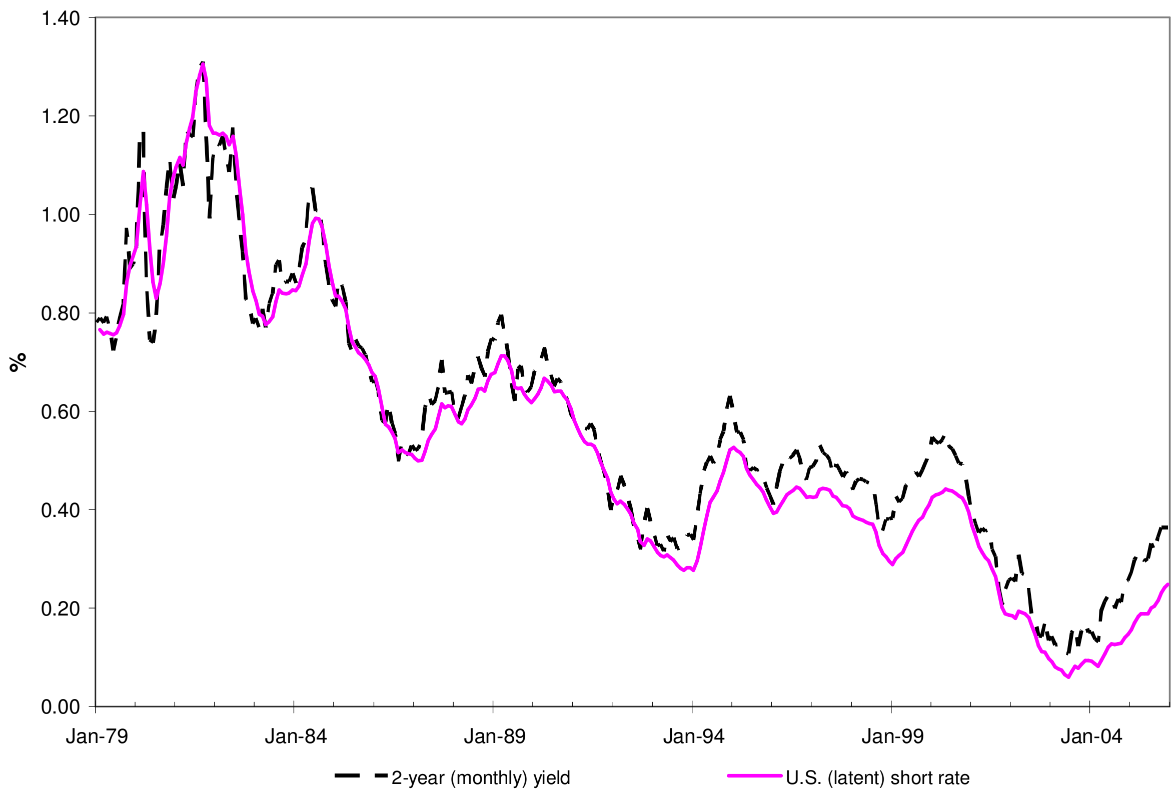

Figure 2: Canadian Latent Factor Estimate

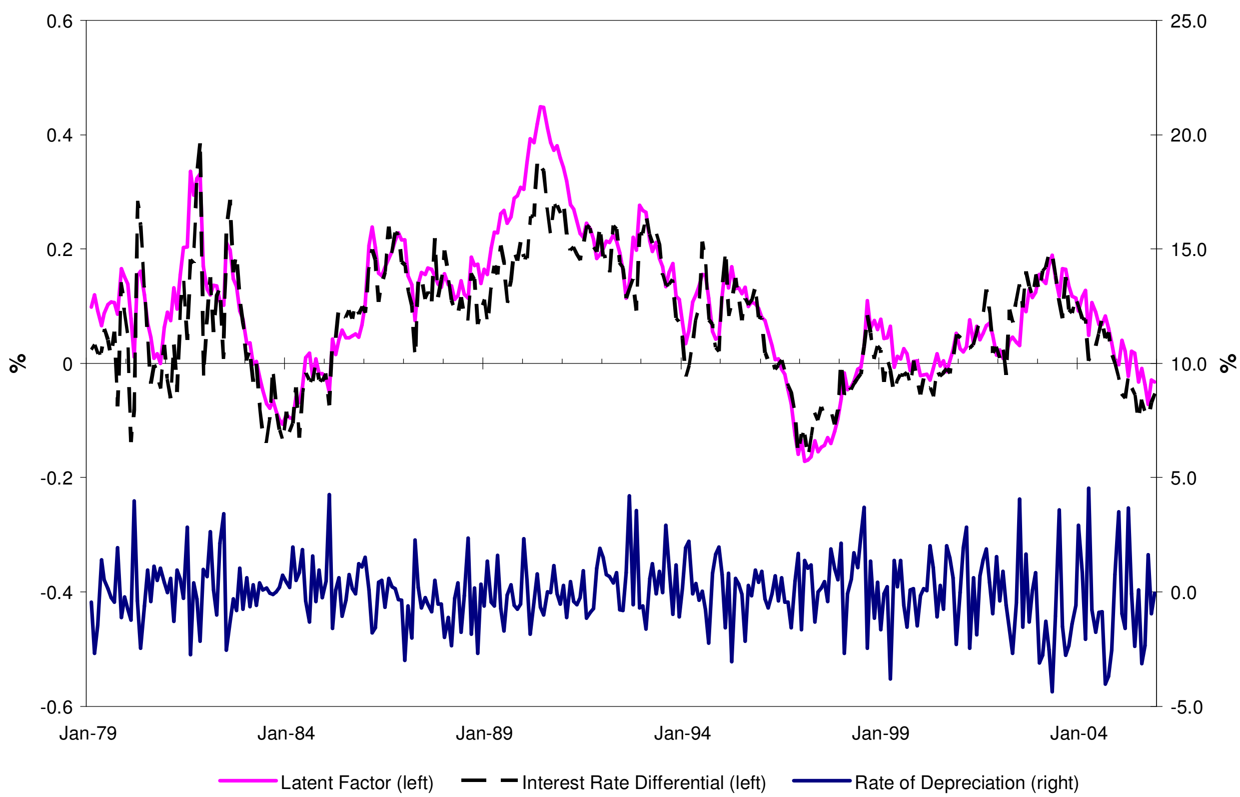

Figure 3: German Latent Factor Estimate

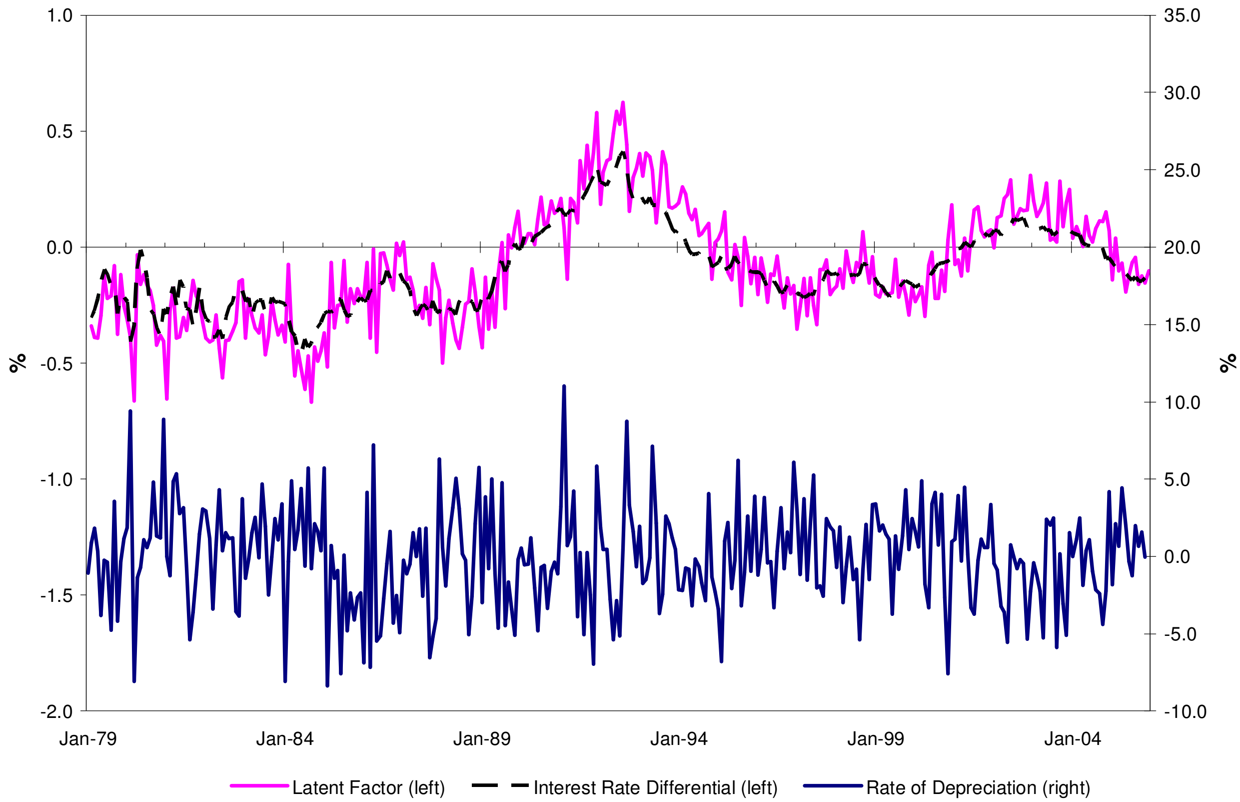

Figure 4: British Latent Factor Estimate