Conditional Correlations Between Cross-country Interest Rates: Evidence from an Affine Term Structure Model with Regime Shifts*

DONG-HYUN AHN†
SIDDHARTHA CHIB‡
KYU HO KANG§

July 2012

Abstract

In this paper we seek to understand the time series dynamics and determinants of the conditional correlations between cross-country interest rates. We develop this understanding within a new and general arbitrage-free multi-factor two-country affine term structure model built upon two novel features: macroeconomic factors and regime shifts in the loadings and market prices of factors. We decompose the macro factors and latent state variables into common and local factors, and allow the term structure dynamics to switch over time among four distinct regimes (two regimes in each country). Our Bayesian MCMC-based estimation results on yield curve data for the U.S. and Canada indicates that the conditional correlation is increasing with the time to maturity, and that its short-end tends to vary with higher volatility, which is driven by the asymmetric business cycles between the countries. (JEL G12, C11, F37)

Keywords: Markov switching process, Bayesian MCMC method, tailored randomized block Metropolis-Hastings, Term structure of conditional correlations

---

*Address for correspondence: School of Economics, Seoul National University, Seoul, South Korea 151-742. E-mail: ahnd@snu.ac.kr.
†Address for correspondence: Olin Business School, Washington University in St. Louis, Campus Box 1133, 1 Bookings Drive, St. Louis, MO 63130. E-mail: chib@wustl.edu.
‡Address for correspondence: Department of Economics and Finance, Hanyang University, Seoul, South Korea 133-791. E-mail: kangkh1121@gmail.com.
1 Introduction

Over the past decade, the two most significant advances in the modeling of term structure dynamics are (i) macro-finance and (ii) regime-shift models. In the macro-finance term structure models, macroeconomic factors are incorporated in the modeling of the stochastic discount factor. These models have shaped our understanding of the fundamental driving forces behind bond yield dynamics and also helped to improve the overall empirical goodness-of-fit of arbitrage-free models of the term structure. In regime-shifting term structure models, on the other hand, which date back to Naik and Lee (1995), the term structure dynamics depend on exogenous regimes, state dependent transition probabilities and priced regime switching risk. These models have been shown to account for the expectation hypothesis puzzle and the predictability puzzle of the excess returns on bonds. Models with both these features have also emerged, as in the work of Ang, Bekaert, and Wei (2008) and Chib and Kang (2012).

The primary goal of this paper is to extend the aforementioned modeling frameworks to an international bond market. We develop an arbitrage-free international affine term structure model of cross-country bond yields in which common and local factors are allowed, macroeconomic variables, in particular, inflation and output growth, along with latent factors, are incorporated as driving factors, and their impact on the joint dynamics of cross-country term structure of interest rates is permitted to structurally shift over time.

In our model, the world economy consists of two countries with different currencies. The common and local factors of macroeconomic and latent variables are assumed to follow a Gaussian vector autoregressive process. The short rate and the market prices

---


3 In the context of two-country models, several papers have treated all underlying factors as latent and shown the importance of using local factors, which influence the short rate of one specific country without affecting the other short rate (for instance, Ahn (2004) and Ahn and Gao (2004)). Meanwhile, common factors have an impact on the short rates of both countries. Recently, a few studies have introduced macroeconomic fundamentals on account of the active interaction between the bond markets and the real economy (for instance, Dong (2006), Chabi-Yo and Yang (2007)). One thing to be noticed is that, unlike the latent factors, observed macro factors such as real GDP growth rate and inflation
of factor risks for each country are affine functions of the factors. In addition, the factor loadings in the short rate specification, and the market price of risk, are subject to Markov switching. In particular, each country switches between two distinct regimes, for a total of four distinct regimes across the two countries. It should be noted that previous work in the international affine term structure context has ignored the possibility of regime changes in the dynamics of the bond prices, although such dynamics have been well documented in closed economy models.

The model we develop is conceptually straightforward and general. It provides rich flexibility with respect to the time-varying conditional variance of interest rates as well as the magnitude and sign of the conditional correlations among the cross-country interest rates. Cappiello, Engle, and Sheppard (2006) document strong evidence on sign-switching correlations among Japan and German bond returns. As discussed by Ahn, Baek, and Gallant (2011), extending this observed behavior to international affine term structure models is extremely important. This is what the model we have developed achieves. Sign-switching correlations and time-varying conditional volatility can be present at the same time. On the other hand, models with correlated square root processes are able to exhibit only non negative conditional correlation between cross-country bond yields and display a certain type of heteroscedastic variation, while models with Gaussian factor processes without Markov switching can generate negative conditional correlations among factors, but not time-varying conditional variance.

We estimate our model by a Bayesian approach with a tuned MCMC (Markov chain Monte Carlo) method based on monthly U.S. and Canadian data over the period from 1986:M1 through 2010:12M. The Bayes approach is particularly relevant because our model has the structure of a high-dimensional non-linear state space model. Non Bayesian state space estimation methodologies could perhaps also be considered but

rate for each country do not directly represent local or common factors because the two economies are likely to be influenced by world-wide shocks (e.g. technological progress and oil prices) and local shocks (e.g. fiscal policy). Thus, in this paper we explicitly decompose the macro variables into common and local factors where those factors are mutually independent.

4In order to accommodate the correlation and volatility structure of cross-country yield curves Mosberger and Schneider (2009) propose an alternative modeling approach by incorporating mixture of Gaussian and square root processes. They find strong empirical evidence for negatively correlated global factors whereas the existence of local factors in the US-UK case is not supported by the data.
only if the maximum number of bond yields is small, say three or four. But in a multi-
country context, the number of yields that must be considered is generally larger. In
our application, for example, we work with six different yields for each country, which
leads, in fact, to a quite high dimensional model that only seems possible to estimate
by the methods we present.

In the empirical analysis, we address the following questions:

• What is the generic shape of the term structure of cross-country conditional cor-
relations? Are the short-rate correlations, on average, higher than the long-rate correlations?

• What are the key determinants of the observed shape of term structure conditional correlations - macroeconomic factors or latent factors?

• Does the term structure of cross-country correlations change over time? Are the short-rate correlations more volatile than the long-rate correlations?

• What are the primary driving forces behind the time-varying term structure of conditional correlations - macroeconomic factors or latent factors?

• Is the model successful for reproducing the sign-switching behavior of correlations? If so, what is the dominant cause of the negative correlations?

Because existing international term structure models have limited ability in generating the sufficiently flexible term structure of correlations, these questions have not been comprehensively addressed before.

The empirical results are striking.

• The overall shape of term structure of correlations is upward sloping and concave. The 3-month cross-country correlation is the lowest, around 0.10. In contrast, the twenty year correlation is as high as 0.77.

• The volatility of correlations is a decreasing function of time-to-maturity. The 3-month correlation ranges from -0.120 to 0.174 whereas the range of 20-year correlation is extremely tight, from 0.767 to 0.782.
• While the latent factors determine the average level of the condition correlations, the time-variation of the conditional correlations is mostly explained by the macroeconomic factors.

• Our model empirically reproduces the sign-switching behavior of correlations among short-term yields ranging between 3 months to 2 years. It is mostly driven by the country-specific regime switching effect of the macroeconomic factors during asymmetric cross-country business cycle periods.

The rest of the paper is organized as follows. In Section 2 we present our two-country affine term structure model and derive the resulting bond prices. We outline the prior-posterior analysis of our model in Section 3. Section 4 deals with the empirical analysis of the real data and Section 5 has concluding comments. Additional details related to the analysis are given in the Appendix.

2 Model

In this section, we propose a cross-country bond pricing model based on the no-arbitrage condition that links the bond prices to stochastic discount factors. The model we propose is parsimonious and flexible enough to capture the distinctive features of cross-country yield curve data.

2.1 Model Specification

In a standard single-country affine term structure model without regime shifts the price of a \( \tau \) period maturity bond at time \( t \) is usually denoted by \( P_t(\tau) \). On letting \( \mathbf{f}_t \) denote the vector of factors, and \( M_{t,t+1} \) the stochastic discount factor (SDF), risk-neutral pricing requires that

\[
P_t(\tau) = \mathbb{E}_t[M_{t,t+1}P_{t+1}(\tau - 1)|\mathbf{f}_t]
\]

In this paper, where we consider two countries, and each country is subject to regime shifts between two states, we need two additional notations, the country indicator and the regime indicator. Assume that the world economy is comprised of two countries, a
domestic country, $d$, and a foreign country $f$, and let $C$ denote the country indicator which takes the values $d$ or $f$. Next, let $q^C_t$ denote the regime indicator, whereby $q^d_t$ is the regime indicator of the domestic country, and $q^f_t$ is the regime indicator of the foreign country. In this context, let $P^C_t(q^C_t, \tau)$ denote the price of a $\tau$ period zero-coupon bond at time $t$ in country $C$ and regime $q^C_t$. Conditioned on the current value of the factors and the regimes, the absence of arbitrage in each country now requires that

\[
P^C_t(q^C_t, \tau) = \mathbb{E}[M^C_{t,t+1} P^C_t(q^C_t, \tau)|q^C_t, f_t] = \mathbb{E}_t[M^C_{t,t+1} P^C_t(q^C_t, \tau)]
\]

where $M^C_{t,t+1}$ is the country-specific nominal pricing kernel and $\mathbb{E}_t$ is the expectation over $(q^C_{t+1}, f_{t+1})$. In pricing bonds with various maturities the economic agents are allowed to observe the current values of the factors and the regimes ($q^C_t$ and $f_t$), as in standard asset pricing models. However, the future factors and regimes are uncertain although their conditional distribution given the most recent values is known. Therefore, in order to solve equation (2.1) for the bond price, we specify the stochastic process of the factors $f_t$ and regimes $q^C_t$, and model the SDF in terms of the risk-free short rate and the market price of risk. With these ingredients, the exchange rate dynamics are endogenously determined.

### 2.1.1 Regime Process

Following Bansal and Zhou (2002) we suppose that the country-specific regime $q^C_t$ is governed independently in each country by a first-order two-state Markov switching process with transition probability matrix

\[
\Pi^C = \begin{bmatrix} q^C_{11} & 1 - q^C_{11} \\ 1 - q^C_{22} & q^C_{22} \end{bmatrix}
\]

where $\Pr[q^C_t = j|q^C_{t-1} = i] = q^C_{ij}$. As a result, the world economy switches among four distinct regimes. The regime indicator appears in the short rate process and the market price of risk for each country. The economic interpretation is that each country can make transitions between high and low spread states. The higher spread can be generated by the more active response of the short rate to the factor shocks or the higher negative market price of risk.
2.1.2 Factor Process

We have two kinds of continuous state variables that govern the stochastic evolution of the domestic and foreign securities: observable macroeconomic variables and unobservable latent variables. As in a standard two-country affine term structure model, we suppose that the latent factors in the global bond markets can be decomposed into one latent common factor \( c_{l,t} \), one domestic local factor \( z_{d,l,t} \), and one foreign local factor \( z_{f,l,t} \).

Our choice of the observable macroeconomic factors are the country-specific real GDP growth rate and inflation rate, as these variables are known to be intimately related to bond markets. Unlike the latent factors, however, these variables do not directly represent local or common factors because the two economies are likely to be influenced by both common and local shocks. For this reason, in each country, we decompose the deviation of the real GDP growth rate at time \( t \) from its unconditional mean, \( g^C_t \), into a common factor \( c_{g,t} \) and idiosyncratic country-specific factors, \( z_{d,g,t} \) and \( z_{f,g,t} \) based on a simple dynamic common factor model specification:

\[
\begin{bmatrix}
g^d_t \\
g^f_t \\
g_t
\end{bmatrix} = \begin{bmatrix}
1 \\
g_c
\end{bmatrix} c_{g,t} + \begin{bmatrix}
z_{d,g,t} \\
z_{f,g,t}
\end{bmatrix}
\] (2.4)

For identification reasons, under this specification, the factors \( (c_{g,t}, z_{d,g,t}, z_{f,g,t}) \) are mutually independent. In the same way, from the demeaned inflation rates \( (\pi^d_t \text{ and } \pi^f_t) \) we can also identify one common inflation factor \( c_{\pi,t} \) and two local inflation factors \( (z_{d,\pi,t} \text{ or } z_{f,\pi,t}) \) as

\[
\begin{bmatrix}
\pi^d_t \\
\pi^f_t
\end{bmatrix} = \begin{bmatrix}
1 \\
\pi_c
\end{bmatrix} c_{\pi,t} + \begin{bmatrix}
z_{d,\pi,t} \\
z_{f,\pi,t}
\end{bmatrix}
\] (2.5)

It is worth noting that the non-zero \( g_c \) and \( \pi_c \) imply the presence of a common factor for each macro variable. Basically, there are six sources of factor risk that are priced in the domestic economy: three common factors and three local factors. It is same for the foreign economy. The description of the factors is summarized in Table 1.

In our formulation, the vector of the factors \( f_t \) includes the nine unobserved components:

\[
f_t = \left( c_{l,t} \ z_{d,l,t} \ z_{f,l,t} \ c_{g,t} \ z_{d,g,t} \ z_{f,g,t} \ c_{\pi,t} \ z_{d,\pi,t} \ z_{f,\pi,t} \right)'
\]

which we assume follow a Gaussian vector autoregressive process with regime switching
common factors in $P_d^t$ (domestic bond prices) and $P_f^t$ (foreign bond prices)
$c^d_t$: common factor in the real GDP growth rate
$c^π_t$: common factor in the inflation rate
$cl,t$: the remaining common factor in $P_d^t$ and $P_f^t$ unexplained by $c^d_t$ and $c^π_t$

domestic local factor in $P_d^t$
$zd,g,t$: local factor in the real GDP growth rate
$zd,π,t$: local factor in the inflation rate
$zd,t$: the remaining local factor in $P_d^t$ unexplained by $zd,g,t$ and $zd,π,t$

foreign local factor in $P_f^t$
$zf,g,t$: local factor in the real GDP growth rate
$zf,π,t$: local factor in the inflation rate
$zf,t$: the remaining local factor in $P_f^t$ unexplained by $zf,g,t$ and $zf,π,t$

| Table 1: Common and local driving factors |

conditional variance

\[
f_t = \mu + G (f_{t-1} - \mu) + \eta_t \tag{2.6}
\]

where

\[
\begin{align*}
\eta_t &\sim iid N(0, \Omega = \Lambda \Gamma \Lambda') \\
\end{align*}
\]

and

\[
\Lambda = diag \left( \sigma_{c,l} \quad \sigma_{l}^d \quad \sigma_{l}^f \quad \sigma_{c,g} \quad \sigma_{g}^d \quad \sigma_{g}^f \quad \sigma_{c,π} \quad \sigma_{π}^d \quad \sigma_{π}^f \right)
\]

\[
\Gamma = \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & \rho_c & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & \rho_d & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \rho_f & 0 \\
0 & 0 & 0 & \rho_c & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & \rho_d & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & \rho_f & 0 & 0 & 1 & 0 & 0 & 0 \\
\end{pmatrix}
\]

In this parameterization, $\Gamma$ is the correlation matrix of the factor shocks, $\eta_t$.

In our setup, the factor volatility $\Lambda$ is assumed to be constant. The zero restrictions in the $\Gamma$ matrix indicate that the six macro factors are independent of the three latent factors. This assumption enables us to decompose the conditional correlation between
cross-country interest rates into the latent factor correlation and the macro factor correlation, so that we can examine their relative importance in accounting for the short rate or long rate conditional correlation dynamics. Our model of $\Gamma$ implies that the common macro factors can be correlated. This is necessary because the common output and inflation factors are likely to be jointly determined by the global aggregate demand and supply innovations such as global financial crisis or oil price shocks. Finally, the country-specific macro factors ($z_{g,t}^d, \pi_t^d$ and $z_{g,t}^f, \pi_t^f$) are also possibly correlated since each can be affected by local aggregate shocks.

2.1.3 Stochastic Discount Factor

We complete our modeling by making assumptions about the SDF. For this, we specify the risk-free short rate process, the market price of the risk and the functional form of the SDF for each country. We first suppose that the short rate in each country is a regime-specific affine function of the vector of nine unobserved continuous state variables

$$r_{q_t^C}^C, t = \delta^C + \beta_{q_t^C}^C (f_t - \mu), \ C \in \{d, f\}$$

where

$$\beta_{q_t^d}^d = \begin{bmatrix} \beta_{1,q_t^d}^d & \beta_{2,q_t^d}^d & 0 & \beta_{4,q_t^d}^d & \beta_{5,q_t^d}^d & 0 & \beta_{7,q_t^d}^d & \beta_{8,q_t^d}^d & 0 \end{bmatrix}'$$

and

$$\beta_{q_t^f}^f = \begin{bmatrix} \beta_{1,q_t^f}^f & 0 & \beta_{3,q_t^f}^f & \beta_{4,q_t^f}^f & 0 & \beta_{6,q_t^f}^f & \beta_{7,q_t^f}^f & \beta_{9,q_t^f}^f \end{bmatrix}'$$

The loadings on the foreign local factors are constrained to be zero, so that the domestic short rate is unaffected by the foreign country-specific factors, and vice versa. That is, the short rate of each country responds to the common factors and the corresponding local factors. We allow for the response of the short rate $\beta_{q_t^C}^C$ to change over time according to the regime process $q_t^C$. This is essential to accommodate the possible changes in the response of the short rate to the underlying factors because the short rate is mostly determined by the monetary authorities in the short run and because the monetary policy has switched between active and less active regimes as is well-known. Hence, each economy at time $t$ is either in a more active or less active regime of the short rate, in essence capturing the time varying conditional correlation of yields and long term risk premium.
Next, we model \( \gamma_{q_{t+1}}^{C} \), a \( 9 \times 1 \) vector, the market prices of factor risk in the country \( C \) associated with the latent and macro factor shocks, as

\[
\gamma_{q_{t+1}}^{C} = \lambda_{q_{t+1}}^{C} + \Phi (f_t - \mu), \quad C \in \{d, f\}
\]

where

\[
\lambda_{q_{t+1}}^{d} = \begin{bmatrix}
\lambda_{1,q_{t+1}}^{d} & \lambda_{2,q_{t+1}}^{d} & 0 & \lambda_{4,q_{t+1}}^{d} & \lambda_{5,q_{t+1}}^{d} & 0 & \lambda_{7,q_{t+1}}^{d} & \lambda_{8,q_{t+1}}^{d} & 0
\end{bmatrix}', \quad (2.9)
\]

\[
\lambda_{q_{t+1}}^{f} = \begin{bmatrix}
\lambda_{1,q_{t+1}}^{f} & 0 & \lambda_{3,q_{t+1}}^{f} & \lambda_{4,q_{t+1}}^{f} & 0 & \lambda_{6,q_{t+1}}^{f} & \lambda_{7,q_{t+1}}^{f} & 0 & \lambda_{9,q_{t+1}}^{f}
\end{bmatrix}', \quad (2.10)
\]

and \( \text{diag}(\Phi) = [\phi_1 \ 0 \ 0 \ \phi_4 \ 0 \ 0 \ \phi_7 \ 0 \ 0] \) \( \quad \) (2.11)

As in Dai et al. (2007), the price of factor risk is assumed to be affine in the factors, which helps detect the time-varying risk premium within regimes. The average market price of risk \( (\lambda_{q_{t+1}}^{C}) \) is also subject to regime shifts.

The impact of the factors on the market price of risk is measured by \( \Phi \). This makes the market price of risk time-varying within regimes while most of previous studies constrain \( \Phi \) to be zero. As Equation (2.11) indicates, we impose that restriction that the common factors have identical effects on both market prices of risk. Under this restriction, the model-implied exchange rate changes, as we discuss shortly, are a linear function of the lagged common factors. Then, we are able to express the resulting econometric model as a linear state space model conditioned on the regimes. Otherwise, the exchange rate changes are a quadratic function of the lagged common and local factors, which makes the calculation of the likelihood very difficult. We refer to \( (\lambda_{q_{t+1}}^{C}, \Phi) \) as the factor-risk parameters. We note that we follow the approach of Ang et al. (2008) in letting the price of risk depend on \( q_{t+1}^{C} \), rather than \( q_{t+1}^{C} \). This implies that in a general equilibrium setting the consumption process depends on the realization of the regimes at time \( t + 1 \). For a thorough discussion, see Dai et al. (2007) and Ang et al. (2008).

Finally, given the short rate process and the market price of risk, we specify the country-specific nominal pricing kernels \( (M_{t,t+1}^{C}) \). We assume complete markets and thus the pricing kernels with minimum variance are uniquely given by

\[
M_{t,t+1}^{C} = \exp\{-r_{q_{t+1}}^{C} - \frac{1}{2} \gamma_{q_{t+1}}^{C} \gamma_{q_{t+1}}^{C} - \gamma_{q_{t+1}}^{C} \epsilon_{t+1}\}, \quad C \in \{d, f\}
\]

(2.12)
where $\tilde{\Gamma}$ is the lower-triangular Cholesky decomposition of $\Gamma$, $L$ is $\Lambda \tilde{\Gamma}$ and $\varepsilon_{t+1} : 9 \times 1$ is equals to $L^{-1} \eta_t$.

### 2.2 Model Solutions for the Bond Prices

Let $y^{C,(\tau)}_{q^C_t,t}$ denote the bond yield of $\tau$ period maturity at time $t$ under regime $q^C_t$ in the country $C$. For the feasibility of the solutions, we assume exponential affine form of the bond prices

$$P^C(q^C_t, \tau) = \exp \left( -\tau y^{C,(\tau)}_{q^C_t,t} \right) \quad (2.13)$$

and

$$y^{C,(\tau)}_{q^C_t,t} = A^C_{q^C_t}(\tau)/\tau + \left( B^C_{q^C_t}(\tau)'/\tau \right) (f_t - \mu) \quad (2.14)$$

where $C \in \{d, f\}$, $a^C_{q^C_t}(\tau) = A^C_{q^C_t}(\tau)/\tau$, and $b^C_{q^C_t}(\tau) = B^C_{q^C_t}(\tau)'/\tau$.

Following Bansal and Zhou (2002) and Chib and Kang (2012), we obtain the solutions by using the law of iterated expectation and the method of undetermined coefficients. This approach gives the following recursive system for the unknown coefficient matrices

$$A^C_{q^C_t}(\tau + 1) = \sum_{j=1}^{2} \Pi^C_{ij} \left( \delta^C_{1,t} + A^C_j(\tau) - B^C_j(\tau)'L\lambda^C_j - \frac{1}{2} B^C_j(\tau)'\Omega B^C_j(\tau) \right) \quad (2.15)$$

and

$$B^C_{q^C_t}(\tau + 1) = \sum_{j=1}^{2} \Pi^C_{ij} \left( \beta^C_{2,t} + (G - L\Phi)'B^C_j(\tau) \right)$$

with $\Pi^C_{ij} = (i,j)$ element of $\Pi^C$

where $c \in \{d, f\}$, and $\tau$ runs over the positive integers. These recursions are initialized by the no-arbitrage condition when $\tau = 0$, i.e., $A^C_{q^C_t}(0) = 0$ and $B^C_{q^C_t}(0) = 0_{3 \times 1}$ for all $q^C_t$. One can see that the resulting intercept and the factor loadings are determined by the weighted average of the two possible regime realizations in the next period where the weights are given by the transition probabilities. This is because agents consider the possibility of the regime shift in the next period.

In summary, a regime exogenously occurs at the beginning of period $t$. This realization is governed by the regime in the previous period and the transition probabilities. Then given the regime at time $t$, the corresponding model parameters are taken from
the full collection of model parameters. These determine the $f_t$ conditioned on $f_{t-1}$ as in (2.6), and the functions $A_{qC}^C(\tau)$ and $B_{qC}^C(\tau)$ according to the recursions in (2.15). Finally, from (2.14), $a_{qC}^C(\tau)$, $b_{qC}^C(\tau)$ and $f_t$ determine the yields of all maturities.

### 2.3 Term Structure of Correlations

Recent studies have found a strong empirical evidence that the conditional correlations among cross-country interest rates are time-varying and switch signs (Ahn et al. (2011) and Cappiello et al. (2006)). This implies that global bond markets are alternating coupling and decoupling over time. One of the goals in this paper is to investigate whether our model with regime shifts and macroeconomic factors is able to generate or capture the sign-switching behavior of cross-country correlations among interest rates.

Since the correlation involves the states in both countries, we begin our discussion by aggregating the regime indicators as follows.

$$s_t = 1 \text{ if } (q^d_t, q^f_t) = (1, 1)$$

$$s_t = 2 \text{ if } (q^d_t, q^f_t) = (2, 1)$$

$$s_t = 3 \text{ if } (q^d_t, q^f_t) = (1, 2)$$

$$s_t = 4 \text{ if } (q^d_t, q^f_t) = (2, 2)$$

The aggregate regime indicator $s_t$ is a four-state Markov process governed by the transition probability

$$\Pi = \Pi^f \otimes \Pi^d$$

Then the one-period ahead conditional correlation between the bond yields with $\tau$ period maturity is given by

$$Cor_{s_t} \left( \frac{y^d_{s_t+1,t+1}, y^f_{s_t+1,t+1}}{y^d_{s_t+1,t+1}, y^f_{s_t+1,t+1}} \right) = \frac{\text{Cov}_{s_t} \left( y^d_{s_t+1,t+1}, y^f_{s_t+1,t+1} \right)}{\text{SD}^d_{s_t} \left( y^d_{s_t+1,t+1} \right) \text{SD}^f_{s_t} \left( y^f_{s_t+1,t+1} \right)}$$

(2.16)

where

$$\text{SD}^d_{s_t=i,t} \left( y^d_{s_t+1,t+1} \right) = \sqrt{\sum_{j=1}^{4} \Pi_{ij} \left( b^d_{s_t+1=j, \tau} \Omega b^d_{s_t+1=j, \tau} \right)}$$

(2.17)

$$\text{SD}^f_{s_t=i,t} \left( y^f_{s_t+1,t+1} \right) = \sqrt{\sum_{j=1}^{4} \Pi_{ij} \left( b^f_{s_t+1=j, \tau} \Omega b^f_{s_t+1=j, \tau} \right)}$$

(2.18)
and \( \text{Cov}_{s_t=i,t} \left( y_{s_{t+1},t+1}^d, y_{s_{t+1},t+1}^f \right) = \sum_{j=1}^{4} \prod_{ij} \left( b_{s_{t+1}=j}^f(\tau) \Omega b_{s_{t+1}=j}^d(\tau) \right) \) (2.19)

Equation (2.16) states that given \( s_t \) and \( f_t \) the conditional correlation at time \( t+1 \) is regime-dependent and so time-varying. The covariance between the cross-country bond yields is computed as in Equation (2.19) as the covariance between \( a_{q_d}^d(\tau) \) and \( a_{q_f}^f(\tau) \) is zero due to the assumption of independence between the country-specific regimes.

A necessary condition to generate the sign-switching correlation is that the product of the factor loadings \( b_{s_{t+1}=1}^d(\tau) \) and \( b_{s_{t+1}=2}^d(\tau) \) for some regimes \( s_{t+1} \) must take a negative value. The cross-country correlation of the short rates is mainly determined by the factor loadings on the short rate. Meanwhile, the relative magnitude of the persistence of the global and local factors plays an important role in determining the correlation of the long term bond yields. If the persistence of the global factors is high, then the factor loadings on the common factors are relatively large and thus the correlation tends to converge to one as the maturity increases. In contrast, if the local factors are more persistent, the correlation shrinks to zero because the local factors of the two countries are independent. It should be noted that \( b_{s_{t+1}=1}^d(\tau) = b_{s_{t+1}=3}^d(\tau) = b_{s_{t+1}=2}^d(\tau) = b_{s_{t+1}=2}^d(\tau) \), \( b_{s_{t+1}=1}^f(\tau) = b_{s_{t+1}=2}^f(\tau) \), and \( b_{s_{t+1}=3}^f(\tau) = b_{s_{t+1}=4}^f(\tau) \).

The principal objective of this paper is identifying the driving nature of generating time-varying conditional correlation. In particular, we are interested in examining the relative importance of the latent and macro factors. For this we decompose the covariance term in Equation (2.16) as the portion due to the common latent factor and the portion due to the common macro factors.

\[
\text{Cov}_{s_t=i,t} \left( y_{s_{t+1},t+1}^d, y_{s_{t+1},t+1}^f \right) = \text{Var}_{s_t=i,t} \left( c_{t+1} \right) + \text{Cov}_{s_t=i,t} \left( b_{4,s_{t+1}=1}^d(\tau)c_{g,t+1} + b_{7,s_{t+1}=1}^d(\tau)c_{\pi,t+1}, b_{4,s_{t+1}=1}^f(\tau)c_{g,t+1} + b_{7,s_{t+1}=1}^f(\tau)c_{\pi,t+1} \right)
\]

where \( b_{i,s_{t+1}}^d(\tau) \) and \( b_{i,s_{t+1}}^f(\tau) \) are the \( i \)th element of \( b_{s_{t+1}}^d(\tau) \) and \( b_{s_{t+1}}^f(\tau) \), respectively.
2.4 Exchange Rate and Exchange Risk Premium

The absence of arbitrage across the two countries uniquely and endogenously determines the exchange rate \( X(t) \), i.e. the number of domestic currency units per one unit of foreign currency.

\[
\frac{X(t + 1)}{X(t)} = \frac{M_{d,t+1}}{M_{f,t+1}} \tag{2.21}
\]

The corresponding exchange rate return under complete markets must equal the difference in the log SDFs:

\[
x_{t+1} = \ln X(t + 1) - \ln X(t) = \left( r_{q,t} - r_{q,t} \right) + \left( \frac{1}{2} \gamma_{q,t+1,t} \gamma_{q,t+1,t} \right) + \left( \gamma_{q,t+1,t} \gamma_{q,t+1,t} \right) \varepsilon_{t+1} \tag{2.22}
\]

Notice that any of the three quantities \( M_{d,t+1}, M_{f,t+1} \) and \( X_t \) can be inferred from the others. There are two important features emerging from (2.22). First, in a risk-averse world the expected exchange rate returns on holding foreign currency depends on a risk premia differential across countries, not just the interest rate differential (i.e. forward premium), \( \left( r_{q,t} - r_{q,t} \right) \). Thus, the uncovered interest rate parity does not hold.\(^5\) More importantly, the exchange rate return or the depreciation rate, \( x_{t+1} \) is subject to regime shifts in both conditional mean and volatility. It should be emphasized that the difference in the market price of risk is responsible for both conditional volatility and the exchange risk premia \( \left( \frac{1}{2} \gamma_{q,t+1,t} \gamma_{q,t+1,t} \right) \), and that the innovations to the both common and local factors cause unexpected changes in the exchange rate.

In our setup we assume a flexible exchange rate system as opposed to a fixed exchange rate system. Therefore the shocks from the foreign country are absorbed into the exchange rate. Consequently, in each period the exchange rate is adjusted to reflect

---

\(^5\)Although the exchange rate is completely determined by the dynamics of the two country-specific pricing kernels, the implied dynamics is different from the observed data. Empirically, most of the exchange rate return is not only unexplained by the interest rate differential between two countries, but also that observed exchange rate volatility is much higher than the model-implied volatility. As in Michael W. Brandt and Santa-Clara (2002), assuming incomplete market is useful to account for the excess volatility of exchange rate. However, it does provide additional information from the term structure of interest rates to explain the exchange rate dynamics. Thus, in this paper we do not consider the exchange rate in our empirical work.
the differentials in the short rate and the market price of risks caused by the asymmetric regime and factor shocks.

Due to the parameterizations of $\lambda_{d_{t}}$, $\lambda_{f_{t}}$, and $\Phi$ the exchange risk premium can be rewritten as

$$0.5 \times \left( \lambda_{f_{t}}^f \lambda_{q_{f_{t}}}^f - \lambda_{d_{t}}^d \lambda_{q_{d_{t}}}^d \right) + \left( \lambda_{f_{t}}^f - \lambda_{d_{t}}^d \right) \Phi \left( f_{t-1} - \mu \right)$$

It is an affine function of one-period lagged global factors and their impact on the exchange risk premium is mainly determined by the differential of the regime-specific parameters in the market prices of factor risks across two countries, $\lambda_{q_{f_{t}}}^f$ and $\lambda_{q_{d_{t}}}^d$. This implies that the risks for holding foreign bonds is compensated through the exchange rate, and the exchange rate compensates not only for the interest rate differential, but also the difference in the market price of risks between the two bond markets as Ahn (2004) points out. One distinguishing feature of our model is that the dynamics of the exchange rate changes is regime-dependent, which implies that the data for the exchange rate changes can possibly help identify the regimes ($q_{d_{t}}^d$, $q_{f_{t}}^f$) as well as the parameters in $\lambda_{q_{d_{t}}}^d$, $\lambda_{q_{f_{t}}}^f$, and $\Phi$ which are usually difficult to estimate.

3 Estimation and Inference

3.1 Data

Our statistical inference is based on the collection of historical yields of treasury bills with six different maturities, real GDP growth and inflation, Canadian dollars to one U.S. dollar exchange rate for the sample period 1986:Q4 to 2010:QII. Inflation is calculated as a quarterly decimal change in the GDP deflator. The data for U.S. zero-coupon bond yields and the exchange rate are available online from the board of governors of the federal reserve system (Gurkaynak, Sack, and Wright (2007)), and the Canadian zero-coupon bond yields are obtained from the bank of Canada. Real GDP growth and inflation are from the Saint Louis Fed for U.S. and from the OECD for Canada. Our estimation of the model over the post great moderation avoids confounding the parameters, factors and regimes with the major oil price shocks during the 1970s, and the Volcker disinflation period in the early 1980s.
3.2 Prior-Posterior Analysis

The set of maturities $\{\tau_1, \tau_2, ..., \tau_6\}$ in quarters are $\{1, 2, 4, 8, 20, 40\}$. The observable quantities are stacked as follows:

$$y_t = \begin{bmatrix} y_t^d(\tau_1) & \cdots & y_t^d(\tau_6) & y_t^f(\tau_1) & \cdots & y_t^f(\tau_6) & x_t & g_t^d & g_t^f & \pi_t^d & \pi_t^f \end{bmatrix}'$$

Let $S_n = \{s_t\}_{t=0,1,...,n}$ denote the sequence of the unobserved regime indicators, $F_n = \{f_t\}_{t=1,...,n}$ the sequence of the factors, $y = \{y_t\}_{t=0,1,...,n}$ the full set of observables (date set), and $\theta$ the collection of the model parameters including the initial factors ($f_0$).

Our econometric inference on $(\theta, S_n, F_n)$ is based on a Bayesian MCMC simulation method. The posterior distribution that we would like to simulate is given by

$$\pi(\theta, S_n, F_n | y) \propto f(y | \theta, S_n, F_n) p(F_n | \theta, S_n) p(S_n | \theta) \pi(\theta) \quad (3.1)$$

where $f(y | \theta, S_n, F_n)$ is the distribution of the data given the regime indicators, the factors and the parameters, $p(F_n | \theta, S_n)$ is the density of the factors given by the Equation (2.6) and $p(S_n | \theta)$ is the density of the Markov switching process conditioned on the transition probabilities. $\pi(\theta)$ is the prior density of $\theta$, which is discussed in Appendix A. Further, the model comparison between the switching and non-switching models is based on the marginal likelihood criterion.

3.2.1 Joint Distribution of the Yields and Macroeconomic Variables

For the complete likelihood ($f(y | \theta, S_n, F_n)$) and the likelihood ($f(y | \theta)$) inference we now express the resulting econometric model in state space form, conditioned on the discrete states and the model parameters. We begin with the transition equation describing the evolution of the common and local factors over time. As in Equation (2.6) these follow a VAR(1) process. The depreciation rate $x_t$ is determined by the lagged factors ($f_{t-1}$) and the current factor shocks ($\eta_t$) as one can see from Equation (2.22). Therefore, $f_{t-1}$
and $\eta_t$ as well as $f_t$ should appear in the measurement equation. On letting

$$
\bar{f}_t = ( f'_t \ f'_{t-1} \ \eta'_t )', \quad \bar{\mu} = ( \mu' \ \mu' \ 0_{1 \times 9} )', \quad \bar{\Omega} = \begin{bmatrix} \Omega & 0_{9 \times 18} \\ 0_{18 \times 9} & 0_{18 \times 18} \end{bmatrix}
$$

$$
\bar{\eta}_t = ( \eta'_t \ 0_{1 \times 9} \ 0_{1 \times 9} )', \quad \bar{\Omega} = \begin{bmatrix} G & 0_{9 \times 9} & 0_{9 \times 9} \\ I_9 & 0_{9 \times 9} & 0_{9 \times 9} \\ 0_{9 \times 9} & 0_{9 \times 9} & 0_{9 \times 9} \end{bmatrix}
$$

the transition equation is specified as

$$
\bar{f}_t = \bar{\mu} + \bar{G} ( \bar{f}_{t-1} - \bar{\mu} ) + \bar{T} \bar{\eta}_t \tag{3.2}
$$

Next, we construct the measurement equation to define the relationship between the observations and the unobserved factors as the model implies. The observations are the country-specific yield curve, depreciation rate and the macroeconomic fundamentals. Their dynamics are driven by the exogenous continuous state variables $\bar{f}_t$ conditioned on the regimes, which can be found in Equations (2.14), (2.22), (2.4) and (2.5). The vector of the observable quantities $y_t$ can be expressed as a linear function of $\bar{f}_t$, which

$$
y_t = \bar{a}_{s_t} + \bar{B}_{s_t} ( \bar{f}_t - \bar{\mu} ) + e_t \tag{3.3}
$$

where the intercept term $\bar{a}_{s_t} : 17 \times 1$ and the factor loadings $\bar{B}_{s_t} : 17 \times 17$ are, respectively

$$
\begin{bmatrix}
a_{q_{1t}}(\tau_1) & \cdots & a_{q_{1t}}(\tau_6) & a_{q_{2t}}(\tau_1) & \cdots & a_{q_{2t}}(\tau_6) & (r'_{f_{t-1}} - r'_{d_{t-1}}) + 0.5 (\lambda'_{q_{1t}} \phi - \lambda'_{q_{2t}} \phi) & 0_{1 \times 4}
\end{bmatrix}'
$$

and

$$
\begin{bmatrix}
b_{q_{1t}}(\tau_1)' \\
\vdots \\
b_{q_{1t}}(\tau_6)'
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
b_{q_{2t}}(\tau_1)' \\
\vdots \\
b_{q_{2t}}(\tau_6)'
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
\begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9}
\end{bmatrix}
$$

and

$$
L^{-1} = \begin{bmatrix}
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & g_c & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \pi_c
\end{bmatrix} \begin{bmatrix}
0_{1 \times 9} \\
\vdots \\
0_{1 \times 9} \\
0_{1 \times 9}
\end{bmatrix}
$$

16
We assume that all yields are priced with errors $e_{it}^C(i = 1, 2, ..., 6$ and $C = d, f$) following Chib and Ergashev (2009) and Dong (2006). Even after taking into account the shocks to the factors, the regime switching conditional mean of the exchange rate changes can explain only a small portion of the variation of $x_t$. For this we introduce an additional measurement error $e_t^x$ to account for the empirical fact that exchange rate volatilities are much higher than interest rate volatilities, as discussed in Anderson, Hammond, and Ramezani (2010) and Michael W. Brandt and Santa-Clara (2002). As a result, the vector of the measurement errors $e_t : 17 \times 1$ is given by

$$
\begin{bmatrix}
  e_{dt}^d & \cdots & e_{dt}^d & e_{ft}^f & \cdots & e_{ft}^f & e_t^x & 0 & 0 & 0 & 0
\end{bmatrix}' \sim iid \mathcal{N}(0, \Sigma)
$$

with

$$
diag(\Sigma) = \begin{bmatrix}
  \sigma_{1}^{2d} & \sigma_{2}^{2d} & \cdots & \sigma_{6}^{2d} & \sigma_{1}^{2f} & \sigma_{2}^{2f} & \cdots & \sigma_{6}^{2f} & \sigma_{x}^{2} & 0 & 0 & 0 & 0
\end{bmatrix}'
$$

### 3.2.2 MCMC sampling

Given the joint dynamics of the observations and the prior density we are able to simulate $(\theta, F_n, S_n)$ sequentially from the posterior distribution. In the first step, we simulate $\theta$ given the data and the most recent value of $S_n$. For this we rely on the TaRB-MH (tailed randomized block Metropolis-Hastings) method. The use of this sampling method is relevant in high dimensional non-linear problems. As is well-known, in affine term structure models the likelihood surface tends to be irregular. The local-modality problem becomes more severe in our cross-country model. The TaRB-MH method relies on randomized grouping and efficient tailoring to form proposal densities. The parameters in $\theta$ are grouped into multiple sub-blocks at the beginning of each MCMC iteration. The number of blocks and its components are randomly determined. Each of these sub-blocks is then sampled in sequence by drawing a value from a tailored proposal density constructed for that particular block. The first and second moments of the proposal density are chosen by a suitably designed version of the simulated annealing algorithm.

---

6In order to reconcile the low volatility of interest rates with the high volatility of exchange rates these papers use an incomplete market approach. In this approach we introduce an extra diffusion process that is orthogonal to the domestic and foreign pricing kernels and assets.
As a result, the MCMC sampling is very efficient in terms of the typical MCMC performance metrics and not sensitive to the starting value. For more technical details, we refer the reader to Chib and Ramamurthy (2010) or Chib and Kang (2012).

In the second step we sample $F_n$ conditioned on $(y, \theta, S_n)$ using the Carter and Kohn (1994) algorithm. Finally, $S_n$ is simulated according to the method of Chib (1996). The technical details can be found in Appendix B.

4 Results

4.1 Model Comparison

In order to evaluate our model, we first consider if our four-regime model ($M_4$) improves on the corresponding single regime model $M_1$). In addition, we compare the proposed model with other candidate models: a model with non-switching average market price of risk ($M_2$) and a model with non-switching short rate factor loadings ($M_3$). This comparison helps us to not only learn which of the multi-regime specification best describes the data, but also if regime shifts occur in either the market price of risk or the short rate equation. Notice that under model $M_3$, the factor loadings $b_{st}^C(\tau)$ are regime-independent and thus the conditional correlation becomes constant over time whereas the intercept term $a_{st}^C(\tau)$ is regime-specific. On the other hand, model $M_2$ is relatively less flexible in estimating the exchange risk premium because of the restriction that $\lambda_{q_{st}^1}^C=1=\lambda_{q_{st}^2}^C$ although the conditional correlation can change over time.

Within the Bayesian context, these models are compared in terms of the marginal likelihood $m(y|M_d)$ and ratios of marginal likelihoods (Bayes factors). Following Chib (1995), an estimate of the log marginal likelihood can be calculated from the following fundamental identity

$$
\ln \hat{m}(y|M_d) = \ln f(y|\theta^*, M_d) + \ln \pi(\theta^*, M_d) - \ln \hat{\pi}(\theta^*|y, M_d)
$$

(4.1)

where $d = 1, 2, 3$ and 4, and $\theta^*$ is a high density point in the support of the parameter space. The first term on the right hand side of this expression is the likelihood ordinate. Notice that the regime and the factor are both unobserved state variables and
so should be integrated out for likelihood inference. Unfortunately, there is no direct way of calculating the likelihood value. We estimate it by simulation using the particle filtering method (Chib, Nardari, and Shephard (2002)). The second term is the prior ordinate, which is readily available. The third term, the posterior ordinate \(\pi(\theta^*|y, M_4)\), is estimated from a marginal-conditional decomposition (Chib (1995)). The specific implementation in this context requires the technique of Chib and Jeliazkov (2001) as modified by Chib and Ramamurthy (2010) for the case of randomized blocks.

<table>
<thead>
<tr>
<th>Model</th>
<th>lnL</th>
<th>lnML</th>
<th>n.s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No switching model ((M_1))</td>
<td>-938.8</td>
<td>-1183.2</td>
<td>0.253</td>
</tr>
<tr>
<td>2-Regime model with (\lambda_{t}^{C} = 1 = \lambda_{t}^{C} = 2 ) ((M_2))</td>
<td>-920.2</td>
<td>-1169.2</td>
<td>0.382</td>
</tr>
<tr>
<td>2-Regime model with (\beta_{t}^{C} = 1 = \beta_{t}^{C} = 2 ) ((M_3))</td>
<td>-942.1</td>
<td>-1176.6</td>
<td>0.327</td>
</tr>
<tr>
<td>4-Regime ((M_4))</td>
<td>-911.9</td>
<td>-1162.1</td>
<td>0.339</td>
</tr>
</tbody>
</table>

Table 2: Log likelihood (lnL), log marginal likelihood (lnML) and numerical standard error (n.s.e)

Table 2 reports the results for the marginal likelihood estimation and confirms that the model with the regime switching in both factor loadings and the average market prices of risks is the most supported by the data. The pairwise comparison of the models supports the importance of incorporating a regime process for the model fitting. Allowing for regime shifts in the factor loading and the price of risk considerably increases the likelihood of the model and improves the marginal likelihood. Thus, in the following subsections, we focus on the estimation results for the four-regime model. In particular, we examine whether our proposed model is capable of detecting the time-varying conditional correlation of the cross-country bond yields with the same maturity, and investigate the driving forces.
Table 3: Estimates of $G$, $\Omega$, $\Pi^C$, and $\Phi$. This table presents the posterior mean and standard deviation based on 1,000 MCMC draws beyond a burn-in of 2,000.

### 4.2 Model Parameters

Figures 3 and 4 display the posterior probability of the regimes over time. This figure indicates that the regime changes have been frequent and drastic over time. Throughout the sample period the three different aggregate regimes ($s_t=1$, 3 and 4) have prevailed.

Table 3 through Table 6 provide insights into the identifying forces behind the estimated two distinct regimes for each country. These tables present the posterior estimates of the model parameters. In particular, Table 4 and 5 show that many of parameters are substantially different across regimes. For both countries, the common latent and macro factors are regime-specific in the short rate equation, and the market price of the common latent factor risk differs across regimes. Therefore, the variation across regimes in the impact of the factors on the short rate and the market price of risk play a critical role in identifying the regimes for each country.

Theoretically, the long-rate conditional correlation is mainly determined by the cor-

---

7Chib et al. (2002) propose a particle filter for the stochastic volatility model that is expressed as a state space model with independent switching. Their particle filtering method has to be modified before it can be applied to our model with first-order Markov switching parameters. For more technical details, refer to Fruhwirth-Schnatter (2006).
| Parameter | $q^d_t = 1$ | | $q^d_t = 2$ | | | | $q^d_t = 1$ | | $q^d_t = 2$ | |
|-----------|-------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \beta^d_{1,q^d_t} \) | 0.143 | 0.006 | 106.517 | 0.256 | 0.011 | 145.546 | \( \beta^d_{1,q^d_t} \) | 0.129 | 0.022 | 0.092 |
| \( \beta^d_{2,q^d_t} \) | 0.617 | 0.079 | 163.358 | 0.967 | 0.047 | 159.534 | \( \beta^d_{2,q^d_t} \) | 1.287 | 0.152 | 1.117 |
| \( \beta^d_{3,q^d_t} \) | 0.142 | 0.163 | 136.497 | -0.549 | 0.079 | 112.474 | \( \beta^d_{3,q^d_t} \) | 0.098 | 0.128 | -0.122 |
| \( \beta^d_{4,q^d_t} \) | -0.455 | 0.116 | 149.385 | 0.161 | 0.094 | 139.235 | \( \beta^d_{4,q^d_t} \) | -0.089 | 0.056 | -0.180 |
| \( \beta^d_{5,q^d_t} \) | 0.477 | 0.387 | 135.719 | -0.592 | 0.382 | 153.618 | \( \beta^d_{5,q^d_t} \) | -1.570 | 0.408 | -2.395 |
| \( \beta^d_{6,q^d_t} \) | -0.280 | 0.135 | 139.091 | -0.086 | 0.145 | 70.250 | \( \beta^d_{6,q^d_t} \) | -0.115 | 0.050 | 91.217 |

(a) U.S.

| Parameter | $q^l_t = 1$ | | $q^l_t = 2$ | | | | $q^l_t = 1$ | | $q^l_t = 2$ | |
|-----------|-------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| \( \beta^l_{1,q^l_t} \) | 0.268 | 0.011 | 138.712 | 0.129 | 0.022 | 0.092 | \( \beta^l_{1,q^l_t} \) | 1.287 | 0.152 | 1.117 |
| \( \beta^l_{3,q^l_t} \) | 0.939 | 0.084 | 177.761 | 1.287 | 0.152 | 1.117 | \( \beta^l_{3,q^l_t} \) | 0.098 | 0.128 | -0.122 |
| \( \beta^l_{4,q^l_t} \) | -0.576 | 0.116 | 116.579 | 0.098 | 0.128 | -0.122 | \( \beta^l_{4,q^l_t} \) | -0.089 | 0.056 | -0.180 |
| \( \beta^l_{6,q^l_t} \) | -0.074 | 0.078 | 132.002 | -0.089 | 0.056 | -0.180 | \( \beta^l_{6,q^l_t} \) | -1.570 | 0.408 | -2.395 |
| \( \beta^l_{7,q^l_t} \) | 3.235 | 0.493 | 124.143 | -1.570 | 0.408 | -2.395 | \( \beta^l_{7,q^l_t} \) | -0.115 | 0.050 | 91.217 |
| \( \beta^l_{9,q^l_t} \) | -0.115 | 0.050 | 91.217 | -0.018 | 0.053 | -0.091 | \( \beta^l_{9,q^l_t} \) | 0.098 | 0.128 | -0.122 |

(b) Canada

Table 4: Estimates of the factor loading in the short rate equation This table presents the posterior mean and standard deviation based on 5,000 MCMC draws beyond a burn-in of 2,000.

responding factor loadings. The size of the factor loadings is an increasing function of the factor persistence. The more persistent factors not only explain a larger variation of the long-rate, but also have an impact on the conditional correlation. Figure 1 plots the dynamics of the common and local factors for the latent and macroeconomic variables. As can be seen in Figure 1, all nine factors are found to display different degrees of persistence, and the latent factors look more persistent. Table 3 confirms that the common latent factor reveals the highest persistence, and thus it has more responsibility in determining the conditional correlation between long term bond yields than the other factors. The common latent factor movements are similar to those of the short rates and less mean-reverting compared to the local latent factors. On the other hand, the local latent factors seem to capture the corresponding country’s term spread dynamics.
\[ q_t^d = 1 \quad q_t^d = 2 \]

<table>
<thead>
<tr>
<th>Parameter ( \lambda_{1,q_t^d}^d )</th>
<th>( q_t^d = 1 )</th>
<th>( q_t^d = 2 )</th>
<th>( \lambda_{1,q_t^d}^d )</th>
<th>( q_t^d = 1 )</th>
<th>( q_t^d = 2 )</th>
<th>( \lambda_{1,q_t^d}^d )</th>
<th>( q_t^d = 1 )</th>
<th>( q_t^d = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{1,q_t^d}^d )</td>
<td>-7.680</td>
<td>0.350</td>
<td>174.493</td>
<td>-9.249</td>
<td>0.301</td>
<td>142.029</td>
<td>-7.680</td>
<td>0.350</td>
</tr>
<tr>
<td>( \lambda_{2,q_t^d}^d )</td>
<td>-0.638</td>
<td>0.653</td>
<td>158.511</td>
<td>-1.891</td>
<td>0.320</td>
<td>145.820</td>
<td>-0.638</td>
<td>0.653</td>
</tr>
<tr>
<td>( \lambda_{3,q_t^d}^d )</td>
<td>-5.309</td>
<td>2.640</td>
<td>161.957</td>
<td>-1.529</td>
<td>1.429</td>
<td>148.267</td>
<td>-5.309</td>
<td>2.640</td>
</tr>
<tr>
<td>( \lambda_{4,q_t^d}^d )</td>
<td>-2.068</td>
<td>0.660</td>
<td>89.447</td>
<td>2.260</td>
<td>1.977</td>
<td>83.620</td>
<td>-2.068</td>
<td>0.660</td>
</tr>
<tr>
<td>( \lambda_{5,q_t^d}^d )</td>
<td>-5.375</td>
<td>1.517</td>
<td>85.524</td>
<td>-6.577</td>
<td>2.168</td>
<td>146.439</td>
<td>-5.375</td>
<td>1.517</td>
</tr>
<tr>
<td>( \lambda_{6,q_t^d}^d )</td>
<td>-4.923</td>
<td>1.966</td>
<td>128.886</td>
<td>-2.102</td>
<td>3.157</td>
<td>148.437</td>
<td>-4.923</td>
<td>1.966</td>
</tr>
</tbody>
</table>

(a) U.S.

\[ q_t^f = 1 \quad q_t^f = 2 \]

<table>
<thead>
<tr>
<th>Parameter ( \lambda_{1,q_t^f}^f )</th>
<th>( q_t^f = 1 )</th>
<th>( q_t^f = 2 )</th>
<th>( \lambda_{1,q_t^f}^f )</th>
<th>( q_t^f = 1 )</th>
<th>( q_t^f = 2 )</th>
<th>( \lambda_{1,q_t^f}^f )</th>
<th>( q_t^f = 1 )</th>
<th>( q_t^f = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_{1,q_t^f}^f )</td>
<td>-10.106</td>
<td>0.488</td>
<td>169.647</td>
<td>-8.469</td>
<td>0.720</td>
<td>159.411</td>
<td>-10.106</td>
<td>0.488</td>
</tr>
<tr>
<td>( \lambda_{2,q_t^f}^f )</td>
<td>-0.546</td>
<td>0.388</td>
<td>165.431</td>
<td>-1.980</td>
<td>0.296</td>
<td>109.626</td>
<td>-0.546</td>
<td>0.388</td>
</tr>
<tr>
<td>( \lambda_{3,q_t^f}^f )</td>
<td>-4.109</td>
<td>1.277</td>
<td>151.041</td>
<td>-3.519</td>
<td>1.465</td>
<td>141.806</td>
<td>-4.109</td>
<td>1.277</td>
</tr>
<tr>
<td>( \lambda_{4,q_t^f}^f )</td>
<td>-3.849</td>
<td>3.230</td>
<td>148.537</td>
<td>-2.860</td>
<td>2.363</td>
<td>126.400</td>
<td>-3.849</td>
<td>3.230</td>
</tr>
<tr>
<td>( \lambda_{5,q_t^f}^f )</td>
<td>-2.645</td>
<td>1.247</td>
<td>173.600</td>
<td>-4.496</td>
<td>1.850</td>
<td>175.960</td>
<td>-2.645</td>
<td>1.247</td>
</tr>
<tr>
<td>( \lambda_{6,q_t^f}^f )</td>
<td>-2.836</td>
<td>3.059</td>
<td>170.202</td>
<td>-5.738</td>
<td>1.295</td>
<td>89.819</td>
<td>-2.836</td>
<td>3.059</td>
</tr>
</tbody>
</table>

(b) Canada

Table 5: Estimates of average market prices of risk This table presents the posterior mean and standard deviation based on 5,000 MCMC draws beyond a burn-in of 2,000.

\[ \Sigma \]

<table>
<thead>
<tr>
<th>Parameter ( \sigma_1^{2,C} )</th>
<th>( \Sigma )</th>
<th>( \sigma_1^{2,C} )</th>
<th>( \Sigma )</th>
<th>( \sigma_1^{2,C} )</th>
<th>( \Sigma )</th>
<th>( \sigma_1^{2,C} )</th>
<th>( \Sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^{2,C} )</td>
<td>4.005</td>
<td>0.463</td>
<td>34.247</td>
<td>3.559</td>
<td>0.592</td>
<td>16.701</td>
<td>4.005</td>
</tr>
<tr>
<td>( \sigma_3^{2,C} )</td>
<td>3.168</td>
<td>1.181</td>
<td>32.614</td>
<td>3.065</td>
<td>0.852</td>
<td>8.492</td>
<td>3.168</td>
</tr>
<tr>
<td>( \sigma_4^{2,C} )</td>
<td>6.980</td>
<td>1.031</td>
<td>30.987</td>
<td>2.567</td>
<td>0.348</td>
<td>18.738</td>
<td>6.980</td>
</tr>
<tr>
<td>( \sigma_5^{2,C} )</td>
<td>4.091</td>
<td>0.462</td>
<td>24.786</td>
<td>2.445</td>
<td>0.370</td>
<td>52.707</td>
<td>4.091</td>
</tr>
<tr>
<td>( \sigma_6^{2,C} )</td>
<td>3.070</td>
<td>0.324</td>
<td>14.867</td>
<td>4.714</td>
<td>0.833</td>
<td>80.906</td>
<td>3.070</td>
</tr>
</tbody>
</table>

Table 6: Estimates of measurement error variances (\( \Sigma \)) This table presents the posterior mean and standard deviation based on 5,000 MCMC draws beyond a burn-in of 2,000.

as indicated by Figure 2. It should also be noted that the common and local macro factors are identified by information contained in the yield curve as well as macroeconomic data. As a result, our estimates for the common and local macroeconomic factors are
Figure 1: Common and Local Factors These graphs plot the estimates of the factors. These graphs are based on 10,000 simulated draws of the posterior simulation.
somewhat different from the estimates that are based on macroeconomic data alone.

Table 7 reports the result for the variance decomposition of the yield curve movement for each country. Table 7 (a) indicates the relative contribution of the macroeconomic factors compared to the latent factors in generating the cross-country yield curve over time. For example, the macro factors account for the variation of the U.S. short-term bond yield by as much as 21.5% in regime 1. Regardless of the country and the regime, the fraction of the macro factors is decreasing with the maturity. The long-term bond yield movement is mostly explained by the latent factors. Meanwhile, Table 7 (b) shows that the contribution of the common factors is increasing with the maturity. 86% of the variation of the U.S. long-term bond yields is explained by the common factors. The remaining 14% is contributed by the local factors. Similarly, 94% of the variation of the Canadian long-term bond yields is attributed to the common factors. Consequently, among the nine driving factors, the common latent factor is the key component in determining the long-term bond yield.

The results thus far reveal that factors with different persistence have different impacts on the bond yields in different regimes through the short rate dynamics and the market price of risk. In what follows, we analyze the term structure of conditional correlations in both cross-section and time series. Further, we investigate the relative importance of the latent and macro factors as the key determinant of the conditional correlations and the driving nature of their time-variation.

4.3 Dynamic Term Structure of Conditional Correlations

Figure 5 plots the time series of the conditional correlation between the cross-country yields with the same maturity. Many interesting features emerge from this figure. We first note that the conditional correlation is monotonically increasing in the maturity, which is the cross-sectional characteristic. This is because, as Table 3 indicates, the common factors are more persistent than the local factors, and consequently the factor loadings on the common factors are bigger than those on the local factors. The second distinctive feature is from the time series perspective. The short-rate conditional correlation is more time-varying than the long-rate conditional correlation. This means that
Figure 2: Common and Country-specific Latent Factors and Term Spread These graphs plot the estimates of the common and local latent factors along with the observed short rates and term spread. These graphs are based on 5,000 draws of the posterior simulation.

the regime changes are associated with the short term bond yields correlation rather than the long term bond yields correlation. In particular, Figure 6 shows that the neg-
### Table 7: Variance decomposition of the yields

These tables present the result for the variance decomposition of the yields based on 5,000 simulated draws of the posterior simulation. Table (a) displays the contribution from the macroeconomic factors, and table (b) displays the contribution from the common factors.

**Table (a) Contribution of the Macro Factors**

<table>
<thead>
<tr>
<th>Maturity</th>
<th>( q_t^d = 1 )</th>
<th>( q_t^d = 2 )</th>
<th>( q_t^f = 1 )</th>
<th>( q_t^f = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q</td>
<td>0.215</td>
<td>0.245</td>
<td>0.738</td>
<td>0.059</td>
</tr>
<tr>
<td>2Q</td>
<td>0.147</td>
<td>0.200</td>
<td>0.632</td>
<td>0.047</td>
</tr>
<tr>
<td>4Q</td>
<td>0.074</td>
<td>0.139</td>
<td>0.460</td>
<td>0.044</td>
</tr>
<tr>
<td>8Q</td>
<td>0.028</td>
<td>0.079</td>
<td>0.276</td>
<td>0.048</td>
</tr>
<tr>
<td>12Q</td>
<td>0.008</td>
<td>0.029</td>
<td>0.106</td>
<td>0.033</td>
</tr>
<tr>
<td>20Q</td>
<td>0.001</td>
<td>0.003</td>
<td>0.010</td>
<td>0.004</td>
</tr>
</tbody>
</table>

**Table (b) Contribution of the Common Factors**

<table>
<thead>
<tr>
<th>Maturity</th>
<th>( q_t^d = 1 )</th>
<th>( q_t^d = 2 )</th>
<th>( q_t^f = 1 )</th>
<th>( q_t^f = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q</td>
<td>0.382</td>
<td>0.378</td>
<td>0.493</td>
<td>0.217</td>
</tr>
<tr>
<td>2Q</td>
<td>0.352</td>
<td>0.361</td>
<td>0.541</td>
<td>0.253</td>
</tr>
<tr>
<td>4Q</td>
<td>0.337</td>
<td>0.354</td>
<td>0.594</td>
<td>0.324</td>
</tr>
<tr>
<td>8Q</td>
<td>0.382</td>
<td>0.391</td>
<td>0.647</td>
<td>0.443</td>
</tr>
<tr>
<td>12Q</td>
<td>0.582</td>
<td>0.575</td>
<td>0.786</td>
<td>0.681</td>
</tr>
<tr>
<td>20Q</td>
<td>0.865</td>
<td>0.862</td>
<td>0.943</td>
<td>0.930</td>
</tr>
</tbody>
</table>

The table 7 shows the variance decomposition of the yields. The contribution from the macroeconomic factors and the common factors are presented in tables (a) and (b), respectively.
ate according to the country-specific regime switching effect of the macro factors on
the bond yields although the size of the conditional correlation between the long term
bond yields explained by the macro factors is relatively small. The time variation of the
short-rate correlation is markedly observed during the early 2000’s recession as its sign
was switching.

One natural question is what drove such sign-switching conditional correlation dy-
namics. It seems to have been caused by the business cycle asymmetry between Canada
and United States. After an unprecedented expansion in the 1990s, the U.S. economy
went into a recession between 2001 and 2003. Because of that, the Federal Reserve Bank
(FRB) lowered the federal fund rate starting 2001:Q3. The target interest rate declined
continuously: 3.07% in 2001:Q3, 1.82% in 2001:Q4 and 1.25% in 2003:Q3. In contrast,
Canada has enjoyed a prolonged expansionary period since the late 1990s (Kose and
Cardarelli (2004)). Bank of Canada increased the key interest rate from 2.00% to 3.20%
during 2001:Q4 through 2003:Q2. Such opposite short-rate movements in the process
of the monetary policy implementation during the asymmetric cross-country business
cycle period seems to cause the occasional negative correlation.

Our estimation results for the model parameters capture such asymmetric reaction
of the cross-country short rates to the macroeconomic fundamentals. As seen in Table
4, during the U.S. recession the economy was occupied by regime \( s_t = 3 \) (i.e. \( q^d_t = 1 \) and
\( q^f_t = 2 \)). The response of the U.S. short rate to the common output factor was relatively
active because \( \beta_{4,q^f_t=1}^{d} \) is much higher than \( \beta_{4,q^f_t=2}^{d} \). Meanwhile,
the short rate in Canada rate was affected little by the the common output factor and
the response to the common inflation factor was even negative (i.e. \( \beta_{4,q^f_t=1}^{f} = -1.570 \)).
In short, these asymmetric responses to the common macro factors during the asym-
metric business cycle period between the countries generated the different conditional
correlations in different periods.

4.4 Exchange Rate Risk

Finally Figure 8 (a) exhibits the sign-switching property of the exchange risk premium.
Combined with the fact that the domestic and foreign average market prices of risks
are subject to change according to the country-specific regime processes, this implies that the difference in the market prices of risks between the countries is substantial at each time point as Table 5 indicates. The figure also displays that the exchange risk premium and forward premium have means that are distinguishable from zero, and are both highly autocorrelated. However, they are much less volatile than the depreciation rates as Figure 8 (b) shows.

5 Conclusion

By proposing and estimating an arbitrage-free cross-country affine term structure model with regime shifts, we analyze the dynamic term structure of conditional correlations. In particular, we identify the driving forces behind the time-varying conditional correlations of the cross-country government bond yields. Our Bayesian analysis based on the U.S. and Canada yield curve data indicates that the conditional correlations, especially the short-end ones, have varied substantially over time with their signs switching. More importantly, their time-variation is mostly driven by regime changes in the effect of the macro factors rather than the latent factors on the yield curve. Those regime changes seem to be associated with the cross-country business cycle asymmetry between the two countries, i.e., the country-specific shifts in macroeconomic fundamentals. The focus of this paper is to investigate the implication of regime changes for time-varying cross-country correlations, i.e., how their dynamics is generated, and also the role of each set of factors, macro and latent, in determining their cross-sectional relationship across time-to-maturity and their time-series dynamics. The time-variation of the cross-country correlations is driven entirely by shifts in regimes in the loadings and the market prices of risk associated with factors in this paper.

As such, their reproduced time-series dynamics are characterized by seemingly discrete moves rather than continuous moves. In contrast, in extant literature, such dynamics are driven either by the stochastic process of factors (international affine term structure model with square-root process) or a nonlinear relationship between Gaussian factors and the short rate (international quadratic term structure model). Whereas the
existing models are designed to produce short-term volatile variation in cross-country conditional correlations, our model draws the long-term variation in correlations coupled with short-term persistency. Thus the suggested mechanisms by which time-varying correlations are generated in the two approaches are complementary, and we expect that merging the two alternatives results in more flexible specification for conditional correlations. We reserve this issue for future research.

A Prior Distribution

The prior distribution on the parameter vector $\theta$ is specified as follows. The transition probabilities in $\Pi^C$ have a beta prior distribution $\text{beta}(\bar{\alpha}, \bar{\beta})$. Next, because some of the volatility parameters in the $\Lambda_{st}$ and $\Sigma$ are liable to be small, we follow Chib and Ergashev (2009) and reparameterize them as

$$
\begin{align*}
\sigma_1^{ds} &= 50 \times \sigma_1^d, & \sigma_2^{ds} &= 300 \times \sigma_2^d, & \sigma_3^{ds} &= 200 \times \sigma_3^d, & \sigma_4^{ds} &= 100 \times \sigma_4^d, & \sigma_5^{ds} &= 25 \times \sigma_5^d, \\
\sigma_6^{ds} &= 10 \times \sigma_6^d, & \sigma_1^{fs} &= 50 \times \sigma_1^f, & \sigma_2^{fs} &= 200 \times \sigma_2^f, & \sigma_3^{fs} &= 125 \times \sigma_3^f, & \sigma_4^{fs} &= 40 \times \sigma_4^f, \\
\sigma_5^{fs} &= 20 \times \sigma_5^f, & \sigma_6^{fs} &= 20 \times \sigma_6^f, & \sigma^{xs} &= 0.3 \times \sigma^x, & \sigma_1^{c,g} &= 5 \times \sigma_{c,g}, & \sigma_2^{ds} &= 2 \times \sigma_g, \\
\sigma_3^{fs} &= 3 \times \sigma_3^f, & \sigma_{c,\pi} &= 10 \times \sigma_{c,\pi}, & \sigma_4^{ds} &= 5 \times \sigma_4^d, & \sigma_5^{fs} &= 1 \times \sigma_5^f, & \sigma\sigma^* &= 0, \\
\sigma_6^{fs} &= 10 \times \sigma_6^f, & \sigma^{c,g} &= 5 \times \sigma_{c,g}, & \sigma\sigma_5^{ds} &= 2 \times \sigma_5^d, & \sigma_6^{fs} &= 2 \times \sigma_6^f, & \sigma_1^{c,g} &= 3 \times \sigma_1^c, \\
\sigma_2^{fs} &= 10 \times \sigma_2^f, & \sigma^{c,g} &= 10 \times \sigma_{c,g}, & \sigma_3^{ds} &= 5 \times \sigma_3^d, & \sigma_4^{fs} &= 1 \times \sigma_4^f, & \sigma_5^{fs} &= 1 \times \sigma_5^f.
\end{align*}
$$

Those rescaled coefficients are assumed to have an inverse gamma prior distribution $IG(\bar{\nu}, \bar{\delta})$. The correlation coefficients in $\Gamma$ and the persistence parameters in $G$ which is assumed to be diagonal for simplicity have a uniform prior $Unif(\bar{\alpha}, \bar{\beta})$. The other parameters have a normal prior distribution, $N(\bar{\mu}, \bar{\sigma})$. To choose the prior parameters we rely on the simulation-based method following Chib and Ergashev (2009). For the factor shock volatility and the market price of risk parameters, the prior distribution is set to generate a positive term premium on average for all countries. We also allow for the parameters to vary considerably. On the other hand, we normalize the labels for the country-specific regimes by restricting that the model-implied average term spread under regime 1 is greater than under regime 2. That is, $(a_{q_f=1}^d(\tau_6) - a_{q_f=1}^d(\tau_1)) > (a_{q_f=2}^d(\tau_6) - a_{q_f=2}^d(\tau_1))$ and $(a_{q_f=1}^f(\tau_6) - a_{q_f=1}^f(\tau_1)) > (a_{q_f=2}^f(\tau_6) - a_{q_f=2}^f(\tau_1))$. It is important to notice that our prior is quite symmetric across regimes in order to avoid the case that the regimes are identified by our prior information. For identifying the unobserved factors we denote the $(i, i)$ element of $G$ by $G_{ii}$ and assume that $|G_{ii}| < 1$, $\beta_{1,q_f=1}^d > 0$, $\mu = 0_{9 \times 1}$.
and $\sigma_{c,t} = \sigma^d_l = \sigma^f_l = 1$. The resulting prior parameters for each model parameter are reported in the table (8).

Through the prior the parameters are constrained to lie in the set $\mathcal{R} = \mathcal{R}_1 \cap \mathcal{R}_2 \cap \mathcal{R}_3$ where

$$
\mathcal{R}_1 = \{ \theta | \beta_{1,q^d_l=1} > 0, |G_{ii}| < 1 \}
$$

$$
\mathcal{R}_2 = \{ \theta | (a_{q^d_l=1}^d(\tau_6) - a_{q^d_l=1}^d(\tau_1)) > (a_{q^d_l=2}^d(\tau_6) - a_{q^d_l=2}^d(\tau_1)) \}
$$

$$
\text{and } \mathcal{R}_3 = \{ \theta | (a_{q^d_l=1}^f(\tau_6) - a_{q^d_l=1}^f(\tau_1)) > (a_{q^d_l=2}^f(\tau_6) - a_{q^d_l=2}^f(\tau_1)) \}
$$

Finally, the initial factors $f_0$ are treated as additional parameters to be sampled and its prior is normally distributed with the unconditional mean $\mu$ and variance $V_0$ implied by the prior distribution of $f_t$ in Equation (2.6).

### B MCMC Sampling

This section discusses the steps involved in sampling the posterior distribution of $(\theta, S_n, F_n)$.

**Algorithm: MCMC sampling**

**Step 1** Initialize $(\theta, S_n, F_n)$ and fix $n_0$ (the burn-in) and $n_1$ (the MCMC sample size)

**Step 2** Sample $\theta$ conditioned on $(y, S_n)$

**Step 3** Sample $F_n$ conditioned on $(y, \theta, S_n)$

**Step 4** Sample $S_n$ conditioned on $(y, \theta, F_n)$

**Step 5** Repeat Steps 2-4, discard the draws from the first $n_0$ iterations and save the subsequent $n_1$ draws.

Full details of each of these steps now follow.
B.1 Sampling $\theta$

Integrating out $F_n$, we sample $\theta$ conditioned on $S_n$ by the TaRB-MH algorithm of (Chib and Ramamurthy, 2010). Specifically speaking, in the $j$th iteration, we have $h_j$ sub-blocks of $\theta$

$$\theta_1, \theta_2, \ldots, \theta_{h_j}$$

The parameters in the standard deviation of the pricing errors $\Sigma$, factor shock volatility $\Lambda$ and the transition probabilities $\Pi$ form three fixed block ($\theta_1$, $\theta_2$ and $\theta_3$), and the others are randomly grouped ($\theta_4, \theta_5, \ldots, \theta_{h_j}$). Then conditioned on the most current value of the remaining blocks $\theta_{-i}$, the proposal density for the $i$th block is constructed by a student $t$ distribution with 15 degrees of freedom, $\pi(\theta_i|\theta_{-i}, y, S_n)$. The mode of this proposal density is obtained by a simulated annealing algorithm. If a proposal value violates any of the constraints in $R$, it is immediately rejected. Otherwise, it is probabilistically taken as the next value in the chain as in a standard M-H (Metropolis–Hastings) algorithm (Chib and Greenberg, 1995). The sampling of $\theta$ is complete when all the sub-blocks

$$\pi(\theta_1|\theta_{-1}, y, S_n), \pi(\theta_2|\theta_{-2}, y, S_n), \ldots, \pi(\theta_{h_j}|\theta_{-h_j}, y, S_n)$$ (B.1)

are sequentially updated. The number of blocks and its components are both randomly chosen within each MCMC cycle.

Now we explain how to calculate the log of $f(y|\theta, S_n)$ integrating out $F_n$.

$$\ln f(y|\theta, S_n) = \sum_{t=1}^{n} \ln f[y_t|I_{t-1}, s_t, s_{t-1}, \theta]$$ (B.2)

where $I_t$ is the history of the observations up to time $t$. First, for given $\bar{f}_{t-1|t-1}$ and $\bar{P}_{t-1|t-1}$, one runs the Kalman filter and obtain the following quantities.

$$\bar{f}_{t|t-1} = \mathbb{E}[f_{t|I_{t-1}, S_n, f_{t+1}, \theta}]$$ (B.3)

$$= \bar{\mu} + \bar{G} (\bar{f}_{t-1|t-1} - \bar{\mu})$$ (B.4)

$$\bar{P}_{t|t-1} = \text{Cov}[f_{t|I_{t-1}, S_n, f_{t+1}, \theta}]$$ (B.5)

$$= \bar{G} \bar{P}_{t-1|t-1} \bar{G} + \bar{T} \bar{\Omega} \bar{T}$$ (B.6)

$$f[y_t|I_{t-1}, s_t, \theta] = \mathcal{N}(y_t|\bar{a}_{s_t} + \bar{b}_{s_t} (\bar{f}_{t|t-1} - \bar{\mu}), \bar{b}_{s_t} \bar{P}_{t|t-1} \bar{b}_{s_t}' + \Sigma)$$
\[ K_t = \bar{P}_{t|t-1} \bar{b}^{'}_s \left( \bar{b}_s \bar{P}_{t|t-1} \bar{b}^{'}_s + \Sigma \right)^{-1} \]

\[
\bar{f}_{t|t} = \mathbb{E} \left[ \bar{f}_t | I_t, S_n, \theta \right] = \bar{f}_{t|t-1} + K_t \left( y_t - \bar{a}_s t - \bar{b}_s (\bar{f}_{t|t-1} - \bar{\mu}) \right) \tag{B.7}
\]

\[
\bar{P}_{t|t} = \text{var} \left[ \bar{f}_t | Y_t, S_n, \theta \right] = (I_k - K_t \bar{b}_s) \bar{P}_{t|t-1} \tag{B.8}
\]

At \( t = 1 \), \( \bar{f}_{0|0} \) and \( \bar{P}_{0|0} \) are initialized as the unconditional mean and variance under regime \( s_0 \). From the outputs of the Kalman filtering, one can calculate the likelihood density for each data point,

\[
f \left[ y_t | I_{t-1}, s_t, \theta \right] = \mathcal{N} \left( y_t | \bar{a}_s t + \bar{b}_s (\bar{f}_{t|t-1} - \bar{\mu}), \bar{b}_s \bar{P}_{t|t-1} \bar{b}^{'}_s + \Sigma \right), \tag{B.9}
\]

which completes the computation of the conditional likelihood given \( S_n \).

### B.2 Simulation of \( F_n \)

We sample \( F_n \) because it is necessary for sampling \( S_n \). To do that, we first run the Kalman filter algorithm to calculate \( \bar{f}_{t|t} \) and \( \bar{P}_{t|t} \) for \( t = 1, 2, \ldots, n \). The last iteration provides us with \( \bar{f}_{n|n} \) and \( \bar{P}_{n|n} \), and these can be used to generate \( f_n \) from \( \mathcal{N} \left( f_{n|n}, P_{n|n} \right) \) where \( P_{n|n} \) is the first \( 9 \times 9 \) subblock of \( \bar{P}_{n|n} \) and \( G^* \) is the first 9 rows of \( \bar{G} \). For \( t = n - 1, n - 2, \ldots, 1 \),

\[
\bar{f}_{t|t, f_{t+1}} = \mathbb{E} \left[ \bar{f}_t | I_t, S_n, f_{t+1}, \theta \right] \tag{B.10}
\]

\[
= \bar{f}_{t|t} + \bar{P}_{t|t} G^{*'} \left( G^{*} \bar{P}_{t|t} G^{*'} + \bar{T} \Omega \bar{T}^{'} \right)^{-1} \left( f_{t+1} - \mu - G^{*} (\bar{f}_{t|t} - \bar{\mu}) \right) \tag{B.11}
\]

\[
\bar{P}_{t|t, f_{t+1}} = \text{cov} \left[ \bar{f}_t | I_t, S_n, f_{t+1}, \theta \right] \tag{B.12}
\]

\[
= \bar{P}_{t|t} - \bar{P}_{t|t} G^{*'} \left( G^{*} \bar{P}_{t|t} G^{*'} + \bar{T} \Omega \bar{T}^{'} \right)^{-1} G^{*} \bar{P}_{t|t} \tag{B.13}
\]

\[
f_t | y, S_n, \theta \sim \mathcal{N}_9 \left( f_{t|t, f_{t+1}}, P_{t|t, f_{t+1}} \right) \tag{B.14}
\]

where \( f_{t|t, f_{t+1}} \) is the first 9 rows of \( \bar{f}_{t|t, f_{t+1}} \) and \( P_{t|t, f_{t+1}} \) the first \( 9 \times 9 \) subblock of \( \bar{P}_{t|t, f_{t+1}} \).

Next we sample the initial factor \( f_0 \). Given the prior \( \mathcal{N}_{9 \times 1} (\mu, V_0) \), \( f_0 \) is updated conditioned on \( \theta \) and \( f_1 \) where \( f_1 \) is obtained from Equation (B.14) for \( t = 1 \). Then

\[
f_0 | f_1, \theta \sim \mathcal{N}_{9 \times 1} \left( \tilde{f}_0, \tilde{V}_0 \right) \tag{B.15}
\]

where

\[
\tilde{V}_0 = (V_0^{-1} + G^{*} \Omega^{-1} G)^{-1} \quad \text{and} \quad \tilde{f}_0 = \mu + \tilde{V}_0 G^{*} \Omega^{-1} (f_1 - \mu)
\]
B.3 Simulation of $S_n$

Let $\tilde{y}_t$ $\tilde{\alpha}_{st}$ and $\tilde{b}_{st}$ the first 17 rows of $\tilde{y}_t$ $\tilde{\alpha}_{st}$ and $\tilde{b}_{st}$, respectively. Then the joint density of $(\tilde{y}_t, f_t)$ is given by

$$f (\tilde{y}_t, f_t | \theta) = \sum_{s_t} p \left[ s_t | I_{t-1}, \theta \right] f \left[ \tilde{y}_t, f_t | I_{t-1}, s_t, \theta \right]$$

(B.16)

where

$$p \left[ s_t, s_{t-1} | I_{t-1}, \theta \right] = p \left[ s_t | s_{t-1}, \theta \right] p \left[ s_{t-1} | I_{t-1}, \theta \right]$$

(B.17)

$$p \left[ s_t | I_{t-1}, \theta \right] = \sum_{s_{t-1}} p \left[ s_t, s_{t-1} | I_{t-1}, \theta \right]$$

(B.18)

$$f \left[ y_t, f_t | I_{t-1}, s_t, \theta \right] = f \left[ y_t | I_{t-1}, s_t, f_t, \theta \right] \times f \left[ f_t | I_{t-1}, s_t, \theta \right]$$

(B.19)

$$\tilde{f}_t = \left( (f_t - \mu) (f_{t-1} - \mu) (f_t - \mu - G (f_{t-1} - \mu)) \right) '$$

(B.20)

$$f \left[ y_t | I_{t-1}, s_t, f_t, \theta \right] = \mathcal{N} \left( y_t | \tilde{\alpha}_{st} + \tilde{b}_{st}, \tilde{f}_t, \Sigma \right)$$

(B.21)

and

$$f \left[ f_t | I_{t-1}, s_t, \theta \right] = \mathcal{N} \left( f_t | \mu + G (f_{t-1} - \mu), \Omega \right)$$

(B.22)

In this step one samples the states from $p[S_n | I_n, \theta]$ where $I_n$ is the history of the outcomes of the observations and the factors up to time $n$. This is done according to the method of Chib (1996) by sampling $S_n$ in a single block from the output of one forward and backward pass through the data.

The forward recursion is initialized at $t = 0$ by setting $Pr[s_0 | I_0, \theta]$ to be the unconditional probability. Then one first obtains $Pr[s_t = j | I_t, \theta]$ for all $j = 1, 2, ..., 4$ and $t = 1, 2, ..., n$ by calculating

$$Pr[s_t = j | I_t, \theta] = \frac{p \left[ \tilde{y}_t, f_t | I_{t-1}, s_t = j, \theta \right] Pr[s_t = j | I_{t-1}, \theta]}{p \left[ \tilde{y}_t, f_t | I_{t-1}, \theta \right]}$$

(B.23)

This can be done by Equation (B.17)-(B.22).

In the backward pass, one simulates $S_n$ by the method of composition. One samples $s_n$ from $Pr[s_n | I_n, \theta]$. In this sampling step, $s_n$ can take any value in $\{1, 2, ..., 4\}$. Then for $t = 1, 2, ..., n-1$ we sequentially calculate

$$Pr[s_t = j | I_t, s_{t+1} = k, S^{t+2}, \theta] = Pr[s_t = j | I_t, s_{t+1} = k, \theta]$$

(B.24)
\[
\Pr[s_{t+1} = k | s_t = j] \Pr[s_t = j | \mathbf{I}_t, \theta] \\
\sum_{j=1}^{4} \Pr[s_{t+1} = k | s_t = j] \Pr[s_t = j | \mathbf{I}_t, \theta]
\]

where \( S_t = \{s_{t+1}, ..., s_n\} \) denotes the set of simulated states from the earlier steps. A value \( s_t \) is drawn from this distribution and it is one of the values \( \{1,2,3,4\} \) conditioned on \( s_{t+1} = k \).

References


Figure 3: Probabilities of Aggregated Regimes ($s_t$) These graphs plot the estimates of the probabilities of regimes. These graphs are based on 5,000 draws of the posterior simulation.
These graphs plot the estimates of the probabilities of regimes. These graphs are based on 5,000 draws of the posterior simulation.

Figure 4: Probabilities of Regimes \((q_t^d, q_t^f)\)
Figure 5: Term Structure of the Conditional Correlations

These graphs plot the estimates of the term structure of the conditional correlations. These graphs are based on 10,000 simulated draws of the posterior simulation. Graph (a) displays the time series of the correlations for the four different maturities, and graphs (b) displays the three-dimensional plot for the term structure of the conditional correlations.
Figure 6: The posterior quantiles of the conditional correlation between the cross-country short rates over time. These graphs plot the estimates of the dynamic conditional correlation of cross-country short rates. These graphs are based on 10,000 simulated draws of the posterior simulation. The dotted red lines are 97.5% and 2.5% quantiles and the solid blue line is the median.
Figure 7: The Decomposition of the Term Structure of the Conditional Correlations

These graphs plot the estimates of the probabilities of regimes. These graphs are based on 10,000 simulated draws of the posterior simulation. Graph (a) displays the contribution from the latent factors, and graphs (b) displays the contribution from the macro factors of the term structure of the conditional correlations.
Figure 8: Exchange Risk Premium This graph plots the estimates of the exchange risk premium. These graphs are based on 10,000 simulated draws of the posterior simulation.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>density</th>
<th>support</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_{11}$</td>
<td>beta</td>
<td>(0,1)</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>$q_{22}$</td>
<td>beta</td>
<td>(0,1)</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>$q_{11}$</td>
<td>beta</td>
<td>(0,1)</td>
<td>40</td>
<td>4</td>
</tr>
<tr>
<td>$q_{22}$</td>
<td>beta</td>
<td>(0,1)</td>
<td>40</td>
<td>4</td>
</tr>
</tbody>
</table>

(a) beta prior

<table>
<thead>
<tr>
<th>Parameters</th>
<th>density</th>
<th>support</th>
<th>$\nu$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{c,g}$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{c,g}^d$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{c,g}^f$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{c,\pi}$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{c,\pi}^d$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{c,\pi}^f$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$\sigma_{i}^{x2,C}$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
<tr>
<td>$0.5 \times \sigma^x$</td>
<td>inverse gamma</td>
<td>(0,\infty)</td>
<td>54</td>
<td>260</td>
</tr>
</tbody>
</table>

(b) inverse gamma prior

<table>
<thead>
<tr>
<th>Parameters</th>
<th>density</th>
<th>support</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{ii}(i = 1,2...,9)$</td>
<td>uniform</td>
<td>(0,1)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\rho_C$</td>
<td>uniform</td>
<td>(-1,1)</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\rho_d$</td>
<td>uniform</td>
<td>(-1,1)</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\rho_f$</td>
<td>uniform</td>
<td>(-1,1)</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

(c) uniform prior

<table>
<thead>
<tr>
<th>Parameters</th>
<th>density</th>
<th>support</th>
<th>mean</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta^d_{i,qt}$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>$\beta^f_{j,qt}$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>$\lambda^d_{i,qt}$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>-1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>$\lambda^f_{j,qt}$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>-1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_4$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\phi_7$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$g_C$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$\pi_C$</td>
<td>normal</td>
<td>$(-\infty, +\infty)$</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

(d) normal prior

Table 8: Prior distributions