Note

Internationally affine term structure models

Antonio Diez de los Rios

Bank of Canada, Ottawa, Canada

ARTICLE INFO

Article history:
Received 30 July 2009
Accepted 3 May 2010

JEL classification:
E43
F31
G12
G15

Keywords:
Term structure
Interest rates
Exchange rates

1. Introduction

The affine term structure model (ATSM), originally proposed by Duffie and Kan (1996), is widely regarded as the cornerstone of modern fixed income theory thanks to its main advantage: tractability. In particular, an ATSM provides analytical expressions for bond yields that are affine functions of some state vector. As noted by Piazzesi (2009), tractability is important because otherwise one would need to compute yields with Monte Carlo methods or solution methods for partial differential equations, which could be especially costly from a computational point of view when model parameters are estimated using data on bond yields.

This note presents a set of conditions that extends the tractability of the single-country ATSM to the multi-country case in the context of international term structure models as those in Backus et al. (2001), Brandt and Santa-Clara (2002) and Brennan and Xia (2006) among others. In particular, this note focuses on internationally affine term structure models where not only bond yields in each one of the countries are known affine functions of a set of state variables, but also the expected rate of depreciation satisfies this property. The main contribution of the present paper is to provide conditions to obtain an expected rate of depreciation (over any arbitrary period of time) that is affine on the set of state variables (Section 2).

Two main families of ATSMs are shown to satisfy these conditions (Section 3). The first subgroup is the so-called completely affine term structure model introduced in Dai and Singleton (2000). However, such a specification has been found empirically restrictive. We overcome this issue by showing that the more flexible class of quadratic-Gaussian term structure models introduced in Ahn et al. (2002) and Leippold and Wu (2003) can also deliver an affine expected rate of depreciation when interpreted as being affine in the original set of variables and their respective squares and cross-products. As shown in Diez de los Rios (2009), these results can be used to estimate ATSMs in a multi-country setting, and to study the exchange rate forecasting ability of such models.

2. An affine expected rate of depreciation

The analysis is similar to that in Backus et al. (2001) and Brandt and Santa-Clara (2002). It is based on a two-country world where assets can be denominated in either domestic currency $j = 1$ (i.e., “dollars”) or foreign currency $j = 2$ (i.e., “pounds”). In particular, consider, based on a no-arbitrage argument, the existence of a (strictly positive) stochastic discount factor (SDF), $M_t^{(j)}$, for each country. This SDF prices any traded asset denominated in currency $j$ through the following relationship:

$$1 = E_t \left[ \frac{M_{t+h}^{(j)}}{M_t^{(j)}} R_{t+h}^{(j)} \right] \quad j = 1, 2; \quad (1)$$

where $R_{t+h}^{(j)}$ is just the gross $h$-period return on the asset.
In this set-up, the law of one price implies that any foreign asset must be correctly priced by both the domestic and the foreign SDFs, which, under complete markets, translates into the fact that the exchange rate \( S_t \) ("currency 1" per unit of "currency 2") is uniquely determined by the ratio of the two pricing kernels:

\[
S_t = \frac{M_t^{(2)}}{M_t^{(1)}},
\]

(2)

Therefore, one can obtain the law of motion of the (log) exchange rate, \( S_t = \log S_t \), using Itô’s lemma on the stochastic processes of \( M_t^{(j)} \). To this end, assume the following dynamics of the domestic and foreign SDF:

\[
\frac{dM_t^{(j)}}{M_t^{(j)}} = -r_t^{(j)}(x_t, t)dt + \Lambda^{(j)}(x_t, t) dW_t, \quad j = 1, 2;
\]

(3)

where \( r_t^{(j)}(x_t, t) \) is the instantaneous interest rate (also known as short rate) in country \( j \); \( W_t \) is an \( n \)-dimensional vector of independent Brownian motions that describes the shocks in this economy; and \( \Lambda^{(j)}(x_t, t) \) is an \( n \)-dimensional vector that is usually called the market price of risk because it describes how the SDF responds to the shocks given by \( W_t \). In general, the short rates and the prices of risks are functions of time, \( t \), and a Markovian \( n \)-dimensional vector, \( x_t \), that describes completely the state of the global economy. The law of motion of these state variables, \( x_t \), is given by a diffusion such as:

\[
dx_t = \mu_t(x_t, t)dt + \sigma_t(x_t, t)dW_t,
\]

(4)

where \( \mu_t \) is an \( n \)-dimensional vector of drifts, and \( \sigma_t \) is an \( n \times n \) state-dependent factor-volatility matrix.

Using Itô’s lemma on (3) and subtracting, one gets:

\[
dS_t = \left[ (r_t^{(1)} - r_t^{(2)}) + \frac{1}{2} \sum \left( \Lambda_t^{(1)} - \Lambda_t^{(2)} \right) \right] dt + \left( \Lambda_t^{(1)} - \Lambda_t^{(2)} \right) dW_t. 
\]

(5)

This equation ties the dynamic properties of the exchange rate to the specific parameterization of the drift (interest rates), the diffusion (price of risk) coefficients in (3), and the dynamic evolution of the set of state variables (because interest rates and the prices of risk are ultimately related to those).

While the conditions needed to have bond yields in affine form can be found in Duffie and Kan (1996), the following proposition summarizes the conditions needed to get an expected rate of depreciation that is affine in the set of state variables given by \( x_t \).

**Proposition 1.** If the drift of the process that the log exchange rate \( s_t \) follows is affine in a set of state variables \( x_t \), then,

\[
E_t s_{t+h} = \gamma h + y'x_t dt ,
\]

with \( \gamma \in \mathbb{R} \) and \( y \in \mathbb{R}^n \), and \( x_t \) follows an affine diffusion under the physical measure:

\[
dx_t = \Phi(\theta - x_t)dt + \Sigma^{1/2} V(x_t)^{1/2} dW_t,
\]

where \( \Phi \) and \( \Sigma \) are \( n \times n \) matrices, \( \theta \) is an \( n \)-vector, \( V(x_t) \) is a diagonal \( n \times n \) matrix with \( i \)-th typical element \( v_i(x_t) = \alpha_i + \beta_i x_t \); \( W_t \) is an \( n \)-dimensional vector of independent Brownian motions, and all the eigenvalues of \( \Phi \) are positive to guarantee the stationarity of the process; then, the expected rate of depreciation \( h \)-periods ahead is a (known) affine function of the state vector \( x_t \):

\[
q_t^{(h)} = E_t [s_{t+h} - s_t] = C(h) + D(h)'x_t,
\]

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

\[
C(h) = \gamma h + y'\theta - y'\Phi^{-1}(I - e^{-\Phi h})\theta, \\
D(h) = y'\Phi^{-1}(I - e^{-\Phi h}). 
\]

**Proof.** First note that the expected rate of depreciation satisfies

\[
E_t [s_{t+h} - s_t] = E_t \left[ \int_t^{t+h} ds_t \right] = \gamma h + y'E_t \left[ \int_t^{t+h} x_t d\tau \right],
\]

then take expectations with respect to the integral form of (7):

\[
E_t \left[ \int_t^{t+h} dx_t \right] = \Phi \theta - E_t \left[ \int_t^{t+h} x_t d\tau \right],
\]

and use that \( E_t \left[ \int_t^{t+h} x_t d\tau \right] = E_t x_{t+h} - x_t \) along with the fact that \( E_t x_{t+h} = \theta + e^{-\Phi h} (x_t - \theta) \) and that \( \Phi \) is invertible in order to obtain the desired result. \( \Box \)

The result in this proposition is novel because (to the best of our knowledge) the literature on continuous-time multi-country affine models has focused almost entirely on Euler approximations to the expected rate of depreciation \( h \)-periods ahead. For example, Hodrick and Vassalou (2002), Leippold and Wu (2007) and Ahn (2004) use an Euler approximation of the law of motion of the (log) exchange rate to obtain a formula for the expected rate of depreciation that is valid only for an arbitrary small period \( h \). Yet Eq. (8) has the advantage of being exact and, hence, any model parameter estimates based on this result will not be subject to discretization biases. Similarly, Backus et al. (2001) only provides an expression for the one-period ahead expected rate of depreciation \( (h = 1) \) and, thus, this proposition generalizes their results to the case of an arbitrary choice of \( h \). For example, Diez de los Rios (2009) exploits Eq. (8) to estimate a two-country ATSM and analyze its forecasting ability when predicting exchange rates up to one year ahead.

Also notice that this proposition states that one can obtain an affine expected rate of depreciation when both the short rates, \( r_t^{(j)} \), and the instantaneous variances of the pricing kernels, \( \Lambda_t^{(j)} \), are affine in \( x_t \) (which guarantees that the drift of the log exchange rate, \( s_t \), is affine); and, at the same time, the process that \( x_t \) follows must be an affine diffusion under the physical measure. Note, however, that these conditions are restrictive with respect to the general class of ATSMs. For example, it is possible to obtain affine bond yields without assuming a model where the instantaneous variance of the SDF is affine in \( x_t \) (see Duffee, 2002; Cheridito et al., 2007) or without the condition that the state vector must follow an affine diffusion under the physical measure (see Duarte, 2004).

### 3. Examples

This section presents additional details on the two main families of ATSMs that belong to the internationally affine class.

#### 3.1. Affine models of currency pricing

In this subsection, we focus on a multi-country version of the Dai and Singleton (2000) standard formulation of the ATSMs that nests most of the work on international term structure modelling.\(^1\) These models can be considered as multivariate extensions of the

---

Cox et al. (1985) model, and they are characterized by the following set of assumptions:

1. $r_t^{(j)} = \delta_0^{(j)} + \delta_1^{(j)} x_t + \zeta_0^{(j)} x_t$, where $\delta_0^{(j)}$ is a scalar, and $\delta_1^{(j)}$ is an $n$-dimensional vector.
2. $dx_t = \Phi(\theta - x_t)dt + \Sigma^{1/2} V(x_t)^{1/2} dW_t$, where $\Phi$ and $\Sigma$ are $n \times n$ matrices, $\theta$ is an $n$-vector, $V(x_t)$ is a diagonal $n \times n$ matrix with $i$-th typical element $v_i(x_t) = \epsilon_t + \beta_i x_t$, and $W_t$ is an $n$-dimensional vector of independent Brownian motions.\(^2\)
3. $A^{(j)} = V(x_t)^{1/2} \lambda^{(j)}$ where $\lambda^{(j)}$ is an $n$-dimensional vector.\(^3\)

Under these assumptions, one can show that bond yields satisfy:

$$y_t^{(j,h)} = A^{(j)}(h) + B_t^{(j)} h x_t$$

where $y_t^{(j,h)}$ is the yield on an $h$-period zero-coupon bond in country $j$, and the coefficients $A^{(j)}(h) \in \mathbb{R}$ and $B_t^{(j)}(h) \in \mathbb{R}^n$ solve a system of ordinary differential equations whose details can be found in Duffie and Kan (1996) or Piazzi (2009).

Notice that this model satisfies the conditions in Proposition 1, and therefore the expected rate of depreciation $h$-periods ahead is also an affine function of the state vector $x_t$. Such a formulation is also known as a “completely affine” specification (Duffee, 2002), because it has an instantaneous variance of the SDFs, $A_t^{(j)}$, that is affine in the set of factors $x_t$. The fact that $E_t [S_{th} - s_t]$ is also affine adds a new meaning to the term “completely affine specification.” The problem is that such a specification has been found to be empirically restrictive. For example, Duffee (2002) finds that this parameterization produces forecasts of future Treasury yields that are beaten by a random walk specification\(^4\); and Backus et al. (2001) point out that this model constrains the relationship between interest rates and the risk premium in such a way that the ability of the model to capture the forward premium puzzle is severely limited. In the next section, we analyze a more flexible family of dynamic term structure models that has been found empirically more plausibly.

### 3.2. Quadratic models of currency pricing

The quadratic term structure model was introduced by Ahn et al. (2002) and Leippold and Wu (2003) in order to accommodate rich nonlinear and time-varying dynamics in bond yields. In particular, these models are characterized by the following set of assumptions:\(^5\)

1. $r_t^{(j)} = \delta_0^{(j)} + \delta_1^{(j)} x_t + \zeta_0^{(j)} x_t$, where $\delta_0^{(j)}$ is a scalar, $\delta_1^{(j)}$ is an $n$-dimensional vector, and $\zeta_0^{(j)}$ is a symmetric $n \times n$ matrix.
2. $dx_t = \Phi(\theta - x_t)dt + \Sigma^{1/2} V_t^{1/2} dW_t$, where $\Phi$ and $\Sigma$ are $n \times n$ matrices, $\theta$ is an $n$-vector, and $W_t$ is an $n$-dimensional vector of independent Brownian motions.
3. $A_t^{(j)} = V(x_t)^{1/2} \lambda_t^{(j)}$, where $\lambda_0^{(j)}$ is a $n$-dimensional vector, and $\lambda_1^{(j)}$ is an $n \times n$ matrix.

It can be shown that in this framework bond yields have a quadratic form:

$$y_t^{(j,h)} = A^{(j)}(h) + B_t^{(j)} h x_t$$

where the coefficients for each country $j$, $A^{(j)}(h) \in \mathbb{R}$ and $B_t^{(j)}(h) \in \mathbb{R}^n$, solve a system of ordinary differential equations. And, it is possible to view any quadratic model as being affine in the original set of variables and their respective squares and cross-products. To do so, just express (9) as:

$$y_t^{(j,h)} = A^{(j)}(h) + B_t^{(j)} h x_t$$

where $x_t = (x_t, z_t)'$, $z_t = vec(x_t)$, $B_t^{(j)}(h) = \{B_t^{(j)}, \Sigma_z^{1/2} \Sigma_z^{1/2} D_n \}$, and $D_n$ is the duplication matrix.\(^6\)

Similarly, it can be shown that the expected rate of depreciation is also affine in this augmented set of factors. To do so, first note that the drift of the (log) exchange rate process can be expressed as:

$$E_t ds_t = (\gamma_0 + y_1 x_t + x_2 y_2 x_t)dt$$

for some $\gamma_0$, $y_1$, and $y_2$. Second, we can use the same tools as before to show that the drift of the exchange rate is affine in the augmented set of state variables:

$$E_t ds_t = (\gamma_0 + y_1 x_t')dt$$

with $x_t = (x_t, z_t')$, and $y_1 = vec(y_2)'. Finally, it can be shown that if one applies Ito’s lemma on $z_t = vec(x_t)$ then the joint process for $x_t$ and $z_t$ is an affine diffusion (see Appendix B in Cheng and Scaillet, 2002). In particular, the law of motion of the augmented set of factors $x_t$ satisfies:

$$d\left(\begin{array}{c} x_t \\ z_t \end{array} \right) = \left( \begin{array}{cc} \Phi & \Omega \\ \Phi \Sigma & \Phi \Sigma \end{array} \right) \left( \begin{array}{c} \theta \\ \theta_t \end{array} \right) - \left( \begin{array}{c} x_t \\ z_t \end{array} \right) dt + \left( \begin{array}{c} \Sigma_x^{1/2} \\ \Sigma_z x_t^{1/2} \end{array} \right) dW_t$$

where the drift is linear with $\Phi x_t = 2D_n (\Phi \otimes I_n) D_n$, $\Phi z_t = -2D_n (\Phi \otimes I_n)$, and $\theta = \Phi z_t vec(\Sigma) - \Phi \theta_t$ and $D_n$ being the Moore-Penrose inverse of matrix $D_n$. In addition, the diffusion term satisfies $\Sigma_x x_t^{1/2} = 2D_n^{1/2} \Sigma_x^{1/2} \otimes x_t$, which implies a volatility matrix $\Sigma$ whose elements are affine in $x_t$ and $x_t' x_t$ (and, therefore, affine in $x_t$ and $z_t$). Therefore, the quadratic model also satisfies the conditions given in Proposition 1 if one interprets this model as being affine in an augmented set of state variables.

It is also interesting to note that this quadratic framework also nests the Gaussian essentially affine specification used in Dai and Singleton (2002) and Duffie (2002) when $\delta_1^{(j)} = 0$ for $j = 1, 2$. Such a model combines Gaussian state variables and an affine market price of risk, and it has been shown to be flexible enough to explain the rejection of the expectations hypothesis of the term structure of interest rates as well as to produce forecasts of future Treasury yields that beat the random walk specification. In this case, bond yields are affine in the set of state variables, while the expected rate of depreciation is quadratic.\(^7\)

---

\(^2\) Dai and Singleton (2000) provide a set of restrictions on the parameters of the model that guarantees that $v_t(x_t)$ cannot take on negative values.

\(^3\) As noted earlier in the main text, this formulation rules out specifications of the price of risk such as those in Duffie (2002) and Cheridito et al. (2007).

\(^4\) Duffee (2002) claims that this is because (i) the price of risk variability only comes from $V(x_t)^{1/2}$ and (ii) because the sign of $A_t^{(j)}$ cannot change as the elements of $V(x_t)^{1/2}$ are restricted to be nonnegative.


\(^6\) In particular, for a given $n \times n$ matrix $F$ it can be shown that $x_t x_t' = tr(\Sigma x_t x_t') = tr(\Sigma x_t) = vec(F)' vec(\Sigma x_t)$; and given that $x_t x_t'$ is an $n \times n$ symmetric matrix then $vec(x_t x_t') = D_n vec(\Sigma x_t)$.

\(^7\) See Brennan and Xia (2006), Dong (2006) and Diez de los Rios (2009) for the use of this model in an international set-up.
4. Conclusions

This note presents a set of conditions that extends the tractability of the single-country ATSM to the multi-country case. In particular, the main contribution of the present paper is to provide conditions to obtain an expected rate of depreciation that is affine on the set of state variables. As shown in Diez de los Rios (2009), this result can be exploited to estimate ATSMs in a multi-country setting, and to study the exchange rate forecasting ability of such models. Finally, two main families of dynamic term structure models are shown to satisfy these conditions.

References