TERM PREMIA IN EXPONENTIAL-AFFINE MODELS OF THE TERM STRUCTURE

MARK FISHER AND CHRISTIAN GILLES

Preliminary and Incomplete
Do not quote without permission

Comments welcome

Abstract. We derive expressions for various term premia in the context of the class of affine-exponential models of the term structure. In addition, we show how regression tests of the expectations hypothesis can be understood in terms of these models. In particular, we derive expressions for the regression coefficients in terms of the parameters of the models. Moreover, these expressions can used to define a GMM estimator of the parameters in a given non-Gaussian affine-exponential model.

1. Introduction

The unbiased expectations hypothesis—that forward rates are unbiased forecasts of future spot rates—has a number of implications that can be tested in a regression setting. These tests can often be structured so that under the expectations hypothesis the intercept is zero and the slope is one. Campbell and Shiller (1991), among others, have documented that the slope is *not* one. Moreover, the slope appears to vary systematically, but not monotonically, with various maturities. These findings have been presented as a challenge to term structure modeling.

In this paper, we show how to use the exponential-affine class of models of the term structure of interest rates to analyze various term premia and the expectations hypothesis. In addition, we show how regression tests of the expectations hypothesis can be understood in terms of these models. In particular, we derive expressions for the regression coefficients in terms of the parameters of the models. Moreover, these expressions can used to define GMM estimators of the parameters in a given non-Gaussian affine-exponential model. We show that simple models from finance can explain a large fraction of the so-called bias in the regression coefficients.

Date: April 1, 1996.
JEL Classification. G12.
Key words and phrases. Term structure, forward rates, expectations hypothesis, Campbell–Shiller regressions.
The views expressed herein are the authors' and do not necessarily reflect those of the Board of Governors or the Federal Reserve System.

In a sense, this paper is a generalization of Frachot and Lesne (1994), who first found closed-from expressions for the regression coefficients in a one-factor (generalized) CIR model. While the didactic power of that paper is great, the empirical importance is limited since term structures are typically driven by more than one factor.

The class of exponential-affine models of the term structure of interest rates is characterized by Duffie and Kan (1995). This class of models includes as special cases Vasicek (1977), Cox, Ingersoll, Jr., and Ross (1985, CIR), Longstaff and Schwartz (1992), and Chen (1996), among many others. The models are characterized by the feature that the log of zero-coupon bond prices are affine functions of the state variables. Duffie and Kan derive restrictions on the process for the state variables under the equivalent martingale measure. They also show that the partial differential equation that characterizes the term structure decomposes into a set of simultaneous ordinary differential equations that can be easily solved numerically without a closed-form analytic solution for zero-coupon bond prices.

The expression for the regression coefficients in terms of the parameters of the model involves two distinct parts. One part is the factor loadings—the relationship between the factors and zero-coupon bond prices described above. The other part is the unconditional variance of (stationary linear combinations of) the state variables under the physical measure. In order to handle both the physical and martingale measures, we need to generalize Duffie and Kan's notion of an exponential-affine model of the term structure.

Here is an outline of the rest of the paper: In Section 2 we describe exponential-affine term structure models. In Section 3, we derive expression for the first two conditional moments. (The unconditional variance is derived in the Appendix.) In Section 4, we derive closed-form expressions for a variety of term premia. In Section 5, we derive closed-form expression for the regression coefficients of the two sets of regressions that Campbell and Shiller (1991) run. In Section 6, we design GMM estimators based on those expressions. In Section 7, we report the results of our estimation.

2. Exponential-affine term-structure models

Let $X(t)$ be a length-d column vector of *state variables* or *factors*.[1] Let the process for the state variables under the physical measure be given by the process for the state variables under the physical measure:

$$dX(t) = \mu_X(X(t))\, dt + \sigma_X(X(t))^\top dW(t),$$

where $\mu_X(x)$ is a length-d column vector and $\sigma_X(x)$ is a $d \times d$ matrix. Let the price-of-risk vector be given by $\lambda(t) = \Lambda(X(t))$, where $\Lambda(x)$ is a length-d column vector, and the short-term riskless interest rate be given by $r(t) = R(X(t))$. It is convenient to define the following:

$$\mathcal{M}(x) = \begin{pmatrix} R(x) & \mu_X(x)^\top \end{pmatrix}$$

[1] We will use the two terms interchangeably.

and

$$\mathcal{S}(x) = \begin{pmatrix} \Lambda(x) & \sigma_X(x) \end{pmatrix},$$

where $\mathcal{M}(x)$ is $1 \times (d+1)$ and $\mathcal{S}(x)$ is $d \times (d+1)$. Then, for our purposes, we have an exponential-affine model of the term structure if

$$\mathcal{M}(x) \text{ and } \mathcal{S}(x)^\top \mathcal{S}(x) \text{ are affine in } x. \tag{2.1}$$

Duffie and Kan show that $\sigma_X(x)$ must have the following structure: $\sigma_X(x) = \sigma_X^o(x)\,\Omega$, where Ω is a constant $d \times d$ matrix and $\sigma_X^o(x)$ is a $d \times d$ diagonal matrix with typical element $\sqrt{\alpha_i + \beta_i^\top x}$. The affineness of the state variables means

$$\mu_X(x) = a + b\,x \tag{2.2}$$

and

$$\sigma_X(x)^\top \sigma_X(x) = A + \sum_{\ell=1}^{d} B_\ell\, x_\ell, \tag{2.3}$$

where a is a $d \times 1$ vector and b, A, and B_ℓ are $d \times d$ matrices.

From condition (2.1) flow the following two results. First, since $\mu_X(x)$ and $\sigma_X(x)^\top \sigma_X(x)$ are affine in x, we have expressions for (i) $E_t[X_\tau]$ and $V_t[X_\tau]$ that are affine in $X(t)$ and (ii) $E[X(t)] = \theta$ and $V[X(t)] = \mathcal{V}$ if $X(t)$ is stationary. We discuss these expressions in the next section.

Second, since $R(x)$, $\hat{\mu}_X(x) = \mu_X(x) - \sigma_X(x)^\top \Lambda(x)$, and $\sigma_X(x)^\top \sigma_X(x)$ are affine in x, we have exponential-affine model of the term structure under the equivalent martingale measure. In particular, the price at time t of a default-free zero-coupon bond that pays one unit of the numeraire at time T, is given by $p(t,T) = P(X(t), T-t)$, where

$$P(X(t), T-t) = \exp[-A(T-t) - B(T-t)^\top X(t)], \tag{2.4}$$

where $B(T-t) = (B_1(T-t), \ldots, B_d(T-t))^\top$ is a column vector of factor loadings. The yield-to-maturity on a zero-coupon bond is given by $y(t,T) = \mathcal{Y}(X(t), T-t)$ where

$$\begin{aligned} \mathcal{Y}(X(t), T-t) &= -\frac{\log[P(X(t), T-t)]}{T-t} \\ &= \left(\frac{1}{T-t}\right)\Big(A(T-t) + B(T-t)^\top X(t)\Big), \end{aligned} \tag{2.5}$$

and the instantaneous forward rate is given by $f(t,T) = \mathcal{F}(X(t), T-t)$ where

$$\begin{aligned} \mathcal{F}(X(t), T-t) &= \frac{-\partial \log[P(X(t), T-t)]}{\partial T} \\ &= A'(T-t) + B'(T-t)^\top X(t). \end{aligned} \tag{2.6}$$

Finally, note that $R(x) \equiv \mathcal{F}(x, 0) = A'(0) + B'(0)^\top x$.

Vasicek and CIR factor loadings. We will refer often in what follows to independent multi-factor Vasicek and CIR. In the Vasicek model the factor loadings are

$$B_i(T-t) = \frac{1-e^{-\kappa_i(T-t)}}{\kappa_i},$$

while in the CIR model they are

$$B_i(T-t) = \frac{2\,(e^{\gamma_i\,(T-t)}-1)}{2\,\gamma_i + (\kappa_i+\lambda_i+\gamma_i)(e^{\gamma_i\,(T-t)}-1)},$$

where $\gamma_i = \sqrt{(\kappa_i+\lambda_i)^2 + 2\,\sigma_i^2}$.

3. Conditional moments

In this section present closed-form solutions for the first and second conditional moments for affine state-variables. The solutions are general and do not depend on diagonalizability as in Duan and Simomato (1995). We show the simplifications that diagonalizability delivers, and we discuss trend-stationarity and cointegration.

General form of the first two conditional moments. Define the conditional expectation of $X(T)$ as

$$\hat{X}(t,T) := E[X(T)|X(t)].$$

For notational convenience define $\hat{X}'(t,T) := \partial\,\hat{X}(t,T)/\partial\,T$. From (2.2), we have a system of ordinary differential equations for the conditional expectation for fixed t,

$$\hat{X}'(t,T) = a + b\,\hat{X}(t,T) \tag{3.1}$$

with the initial condition

$$\hat{X}(t,t) = X(t).$$

Let[2]

$$\Phi(\tau) := e^{b\,\tau} \tag{3.2}$$

and

$$\mathcal{D}(\tau) := \int_0^\tau \Phi(s)\,ds. \tag{3.3}$$

The solution to the system (3.1) is

$$\hat{X}(t,T) = \Phi(T-t)\,X(t) + \mathcal{D}(T-t)\,a. \tag{3.4}$$

Applying Ito's lemma to (3.4) for fixed T, the dynamics of the conditional expectation are given by

$$d\hat{X}(t,T) = \hat{\sigma}_X(t,T)^\top dW(t), \tag{3.5}$$

where

$$\hat{\sigma}_X(t,T) := \sigma_X(X(t))\,\Phi(T-t)^\top.$$

[2] $\Phi(\tau)$ is called the fundamental matrix.

Let $V_t[Y]$ denote the conditional variance of Y, and let $v(t,T) := V_t[X(T)]$. We can derive an expression for $v(t,T)$ from (3.5) as follows. Note that

$$X(T) = \hat{X}(t,T) + \int_{s=t}^{T} d\hat{X}(s,T) = \hat{X}(t,T) + \int_{s=t}^{T} \hat{\sigma}_X(s,T)^\top dW(s).$$

Thus

$$v(t,T) = V_t\left[\int_{s=t}^{T} \hat{\sigma}_X(s,T)^\top dW(s)\right] = E_t\left[\int_{s=t}^{T} \hat{\sigma}_X(s,T)^\top \hat{\sigma}_X(s,T)\, ds\right], \quad (3.6)$$

where the second equality is shown in Duffie (1996).[3] Using (2.3), we can write

$$E_t\left[\hat{\sigma}_X(s,T)^\top \hat{\sigma}_X(s,T)\right] = \Phi(T-s)\, F(t,s)\, \Phi(T-s)^\top,$$

where,

$$F(t,s) = A + \sum_{\ell=1}^{d} B_\ell\, \hat{X}_\ell(t,s). \quad (3.7)$$

Therefore, we can write

$$v(t,T) = \int_t^T \Phi(T-s)\, F(t,s)\, \Phi(T-s)^\top\, ds. \quad (3.8)$$

Diagonalizability. When b is diagonalizable, we can write $b = Q\,\kappa\,Q^{-1}$, where κ is a diagonal $d \times d$ matrix whose diagonal elements, κ_i, are the eigenvalues of b and Q is an invertible $d \times d$ matrix whose columns are the eigenvectors of b. In this case $\Phi(\tau) = Q\, e^{\kappa\,\tau}\, Q^{-1}$, and $\mathcal{D}(\tau) = Q \int_0^\tau e^{\kappa\, s}\, ds\, Q^{-1}$. Note that when no element on the diagonal of κ is zero (*i.e.* $\kappa_i \neq 0$ for $i = 1$ to d), we can write $\int_0^\tau e^{\kappa\, s}\, ds = \kappa^{-1}\,(e^{\kappa\,\tau} - I)$.

When b is diagonalizable, we can write

$$v(t, t+\tau) = Q\left\{\int_0^\tau e^{-\kappa\,(\tau-s)}\, \mathcal{F}(t,s)\, e^{-\kappa\,(\tau-s)}\, ds\right\} Q^\top. \quad (3.9)$$

where $\mathcal{F}(t,s) := Q^{-1}\, F(t,s)\, (Q^{-1})^\top$. Let $p(t,\tau)$ denote the matrix in curly brackets in (3.9). Note that the (i,j)-th element of $p(t,\tau)$ is given by

$$p_{ij}(t,\tau) = \int_0^\tau e^{-(\kappa_i+\kappa_j)\,(\tau-s)}\, \mathcal{F}_{ij}(t,s)\, ds. \quad (3.10)$$

Let $p(t,\infty) := \lim_{\tau\to\infty} p(t,\tau)$. In the Appendix we show that when all the eigenvalues are real and negative (*i.e.*, $\kappa_i < 0$ for $i = 1$ to d), we can write

$$p_{ij}(t,\infty) = \frac{1}{\kappa_i + \kappa_j}\left(-\bar{A}_{ij} + \sum_{m=1}^{d} \frac{\beta_{ij}^m\, \bar{a}_m}{\kappa_m}\right),$$

[3] Equation (3.6) is a generalization of Duffie's equation (1) on page 84. He cites Protter for a proof.

where

$$\bar{a}_m = (Q^{-1}\, a)_m$$
$$\bar{A}_{ij} = (Q^{-1}\, A\, (Q^{-1})^\top)_{ij}$$

and

$$\beta_{ij}^m = \sum_{\ell=1}^{d} (Q^{-1}\, B_\ell\, (Q^{-1})^\top)_{ij}\, Q_{\ell m}.$$

Thus we have a closed-form solution for the unconditional variance $v(t, t+\infty)$ in this case.

Equilibrium points and trend-stationarity. An equilibrium point, θ, for system (3.1) is one where $\hat{X}'(t, t+\tau) = 0$ for all $\tau \geq 0$. From (3.1) we see that we must have $-b\,\theta = a$. If b has full rank, then the unique equilibrium point is given by $\theta = -b^{-1}\, a$. When an equilibrium point exists, we can write $\mu_X(X(t)) = K\,(\theta - X(t))$, where $K = -b$.

Since $\hat{X}(t,t) = X(t)$, at an equilibrium point we have $X(t) = \theta$. We can substitute $X(t) = \hat{X}(t, t+\tau) = \theta$ into solution (3.4) and write:

$$\Big(I - \Phi(\tau)\Big)\, \theta = \mathcal{D}(\tau)\, a. \tag{3.11}$$

Thus using (3.11) we can write (3.4) as

$$\hat{X}(t, t+\tau) = \theta + \Phi(\tau)\, \Big(X(t) - \theta\Big). \tag{3.12}$$

Equation (3.12) is the standard way in which the conditional expectation is written in CIR and Vasicek.

Consider the way in which a change in the current value of one of the state variables affects the conditional expectation of itself or another state variable. From solution (3.4) we have

$$\frac{d\hat{X}_i(t, t+\tau)}{dX_j(t)} = \Phi_{ij}(\tau).$$

Define

$$\Phi(\infty) = \lim_{\tau \to \infty} \Phi(\tau). \tag{3.13}$$

In no state-variable is to affect the expectation of itself or any other state variable at the infinite horizon, then we must have

$$\Phi(\infty) = 0. \tag{3.14}$$

We say that the system of state variables is *trend-stationary* when condition (3.14) holds. Note that if condition (3.14) holds *and* b has full rank, then we have, using (3.12), $\hat{X}(t, t+\infty) = \theta$.[4]

By contrast, if some element of $\Phi(\infty)$ is not zero, then the system of state variables is not trend stationary. In this case there is no unconditional expectation. As

[4]These two conditions need not go together, however. See Fisher and Gilles (1996).

an example, suppose that b is diagonalizable. Then $\Phi(\infty) = 0$ if and only if $\kappa_i < 0$ for $i = 1$ to d.

4. Term premia

Geometric zero-coupon bond returns. From equation (2.4), Ito's Lemma gives the process for the log of bond prices:

$$d \log[P(X(t), \tau)] = \tilde{\mu}_P(X(t), T-t)\, dt + \sigma_P(X(t), T-t)^\top dW(t)$$

where

$$\tilde{\mu}_P(X(t), \tau) := -B(\tau)^\top \mu_X(X(t)) + A'(\tau) + B'(\tau)^\top X(t)$$

and

$$\sigma_P(X(t), \tau) := -\sigma(X(t))\, B(\tau).$$

Using (2.6) we can write

$$\begin{aligned} \tilde{\mu}_P(X(t), \tau) &= \mathcal{F}(X(t), \tau) - B(\tau)^\top \mu_X(X(t)) \\ &= \mathcal{F}(\theta, \tau) + H(T-t)^\top (X(t) - \theta), \end{aligned} \tag{4.1}$$

where

$$H(\tau) = B'(\tau) - b^\top B(\tau),$$

In the Vasicek model, $H(\tau) \equiv 1$.

Constantinides (1992) calls $\tilde{\mu}_P(X(t), \tau)$ the geometric expected bond return. Here we generalize a finding of his. Given $E[\mu_X(X(t))] = 0$, then—taking unconditional expectations of (4.1)—we have

$$E[\tilde{\mu}_P(X(t), \tau)] = \mathcal{F}(\theta, \tau).$$

Hence, the unconditional geometric return equals the unconditional forward rate for the matching maturity.

Before proceeding, note that deriving the process for (the log of) bond prices we have allowed the maturity of the bond to decrease as time passes. Thus we have the written the process for a given zero-coupon bond. In contrast, one can derive the process for fixed-maturity "bond" prices. Ito's Lemma gives

$$\begin{aligned} d \log[P(X(t), \tau)] &= -B(\tau)^\top dX(t) \\ &= -B(\tau)^\top \mu_X(X(t))\, dt - B(\tau)^\top \sigma(X(t))^\top dW(t). \end{aligned}$$

Thus the difference between the two processes is the forward rate.

Forward rates.

The process for forward rates. Given (2.6), the process for forward rates can be written as

$$df(t,T) = \mu_f(X(t), T-t)\,dt + \sigma_f(X(t), T-t)^\top dW(t),$$

where

$$\begin{aligned}\mu_f(X(t),\tau) &= -\frac{\partial\,\tilde{\mu}_P(X(t),\tau)}{\partial\,\tau}\\ &= -\frac{\partial\,\mathcal{F}(X(t),\tau)}{\partial\,\tau} + B'(\tau)^\top \mu_X(X(t))\\ &= -\frac{\partial\,\mathcal{F}(\theta,\tau)}{\partial\,\tau} - H'(\tau)^\top (X(t)-\theta)\end{aligned} \tag{4.2}$$

and

$$\sigma_f(X(t),\tau) = -\frac{\partial\,\sigma_P(X(t),\tau)}{\partial\,\tau} = \sigma_X(X(t))\,B'(\tau).$$

Notice that the ergodic forward-rate drift equals minus the slope of the ergodic forward rate curve:

$$\mu_f(\theta,\tau) = -\frac{\partial\,\mathcal{F}(\theta,\tau)}{\partial\,\tau}.$$

Note that if we were holding the maturity τ fixed (rather than holding the maturity date T fixed), we would have $df(X(t),\tau) = -B'(\tau)^\top dX(t)$, the drift of which is different from (4.2) by the slope of the forward rate curve.

Forward rate premium. Define the forward rate premium as

$$\Psi_f(t,T) := f(t,T) - E_t[r(T)].$$

For the exponential-affine class, we can write

$$\Psi_f(t,t+\tau) = \Big(A'(\tau) + B'(\tau)^\top X(t)\Big) - \Big(A'(0) + B'(0)^\top E_t[X(t+\tau)]\Big).$$

Using (3.12) we can write the forward premium as the ergodic premium plus a term that depends on the deviations of the state variables from their means:

$$\Psi_f(t,t+\tau) = \mathcal{F}(\theta,\tau) - R(\theta) + G(\tau)^\top (X(t)-\theta), \tag{4.3}$$

where

$$G(\tau) := B'(\tau) - \Phi(\tau)^\top B'(0),$$

is the vector of factor loadings for the forward rate premium.

For the Vasicek model, $G(\tau) \equiv 0$, and the forward premium is non-random and depends only on maturity.[5] For the CIR model, however, the i-th factor loading is given by

$$G_i(\tau) = e^{-k_i\,\tau} - \frac{4\,\gamma_i^2\,e^{\gamma_i\,\tau}}{\Big(2\gamma_i + (\kappa_i+\lambda_i+\gamma_i)(e^{\gamma_i\,\tau}-1)\Big)^2}.$$

[5]Campbell (1986) derived this result in a general equilibrium setting.

Thus for CIR the forward premium is random, in that it depends on the state variables which are themselves random. In the next section we will see how this randomness helps explain some regression results.

The zero-coupon yield premium. In this section, we derive the expression for the zero-coupon yield premium:

$$\Psi_z(t,T) := y(t,T) - \frac{1}{T-t}\int_t^T E_t[R(X_v)]\,dv. \tag{4.4}$$

We can write the zero-coupon yield as

$$\begin{aligned} y(t,t+\tau) &= \frac{1}{\tau}\left(A(\tau) + B(\tau)^\top X(t)\right) \\ &= \mathcal{Y}(\theta,\tau) + \frac{1}{\tau}\,B(\tau)^\top (X(t)-\theta). \end{aligned} \tag{4.5}$$

Next, we can write the average expected future short rate conditional on information at time t:

$$\begin{aligned} \frac{1}{T-t}\int_t^T E_t[R(X(v))]\,dv &= \frac{1}{T-t}\int_t^T \left\{A'(0) + B'(0)^\top E_t[X(v)]\right\} dv \\ &= R(\theta) + \left(\frac{1}{T-t}\right) B'(0)^\top \mathcal{D}(T-t)\,(X(t)-\theta). \end{aligned} \tag{4.6}$$

Using equations (4.6) and (4.5), we can write the zero-coupon yield premium as

$$\Psi_z(t,T) = \mathcal{Y}(\theta, T-t) - R(\theta) + \psi(T-t)^\top (X(t)-\theta),$$

where

$$\psi(\tau) := \frac{1}{\tau}\left(B(\tau) - \mathcal{D}(\tau)^\top B'(0)\right).$$

The relationship between the factor loadings for the zero-coupon yield premium and the factor loadings for the forward premium is given by

$$\frac{\partial}{\partial \tau}\left(\tau \times \psi(\tau)\right) = G(\tau).$$

Finally, notice that Ψ_z is stochastic unless $\psi = 0$. Not surprisingly, in the Vasicek model $\psi = 0$, while in the CIR model $\psi \neq 0$.

A decomposition of the yield premium. Note that $\log(P(T,T)) = \log(P(t,T)) + \int_{s=t}^T d\log(P(s,T))$ and $\log(P(T,T)) = 0$ imply that

$$y(t,T) = \frac{1}{T-t}\int_{s=t}^T d\log(P(s,T)).$$

Thus, taking conditional expectations of both sides,

$$y(t,T) = \frac{1}{T-t}\int_{v=t}^T E_t[\tilde{\mu}_P(v,T)]\,dv.$$

As a result, we have

$$\Psi_z(t,T) = \frac{1}{T-t}\int_{v=t}^T E_t\left[\sigma_P(v,T)^\top \lambda(v) - \frac{1}{2}\|\sigma_P(v,T)\|^2\right] dv$$

which we can decompose into a risk premium and a convexity premium:

$$\Psi_z(t,T) = \mathcal{L}_y(t,T) + \mathcal{J}_y(t,T),$$

where

$$\mathcal{L}_y(t,T) = \frac{1}{T-t}\int_{v=t}^{T} E_t\left[\sigma_P(v,T)^\top \lambda(v)\right] dv$$

and

$$\mathcal{J}_y(t,T) = \frac{1}{T-t}\int_{v=t}^{T} E_t\left[-\frac{1}{2}\|\sigma_P(v,T)\|^2\right] dv.$$

This decomposition can be computed the following way. Let $y^0(t,T)$ be the zero-coupon yield where $\lambda(t) = 0$ but all other parameters are the same as for $y(t,T)$. Then $\mathcal{L}_y(t,T) = y(t,T) - y^0(t,T)$ and $\mathcal{J}_y(t,T) = \Psi_z(X(t), T-t) - \mathcal{L}_y(t,T)$.[6]

5. Campbell-Shiller regressions

Campbell and Shiller (1991) estimated two sets of regressions involving the yield spreads. The exposition relies on the existence of unconditional variance and covariances.[7] In particular, let $V[Y]$ denote the unconditional variance of Y and let $C[Y,Z]$ denote the unconditional covariance of Y and Z.[8]

The first set of regressions. Let $\beta_0(\tau-t, T-t)$ and $\beta_1(\tau-t, T-t)$ be regression coefficients indexed by $\tau - t$ and $T-t$ for $t < \tau < T$. In the first set of regressions, the change in a zero coupon yield is regressed on the slope of the yield curve:

$$y(\tau,T) - y(t,T) = \beta_0(\tau-t, T-t) + \beta_1(\tau-t, T-t)\, s(t,\tau,T) + \nu(t,\tau,T), \quad (5.1)$$

where

$$s(t,\tau,T) := \left(\frac{\tau-t}{T-\tau}\right)\Big(y(t,T) - y(t,\tau)\Big)$$

is the (weighted) slope of the yield curve and ν is an error such $C[s,\nu] = 0$. In this section we derive the theoretical regression coefficients—assuming pricing equation (2.4) holds—for the regressions that Campbell and Shiller (1991) estimated.[9]

We start by decomposing the future yield on a zero-coupon bond into an unbiased forecast and a forecast error that is uncorrelated with currently-available information:

$$y(\tau,T) = E_t\left[y(\tau,T)\right] + \epsilon(t,\tau,T), \quad (5.2)$$

The next ingredient is the forward rate. Let the forward price of a zero-coupon bond be

$$F(t,\tau,T) := \frac{P(t,T)}{P(t,\tau)},$$

[6] The same sort of decomposition can be applied to the forward rate premium.

[7] A future version will allow for cointegrated spreads.

[8] Recall that the unconditional variance of the state variables is given by $V[X] = \hat{v}(t, t+\infty)$.

[9] See Frachot and Lesne (1994) for related work.

and let the associated forward rate be

$$f(t,\tau,T) := -\frac{\log[F(t,\tau,T)]}{T-\tau} = \left(\frac{T-t}{T-\tau}\right) y(t,T) - \left(\frac{\tau-t}{T-\tau}\right) y(t,\tau). \tag{5.3}$$

Now add and subtract the forward rate from the right-hand side of (5.2):

$$y(\tau,T) = f(t,\tau,T) + \omega(t,\tau,T) + \epsilon(t,\tau,T), \tag{5.4}$$

where

$$\omega(t,\tau,T) := E_t\left[y(\tau,T)\right] - f(t,\tau,T) \tag{5.5}$$

is the forward-rate forecast error. Clearly the forward rate forecast error is related to the forward term premium.[10] Now subtract $y(t,T)$ from both sides of (5.4), producing:

$$y(\tau,T) - y(t,T) = \Big(f(t,\tau,T) - y(t,T)\Big) + \Big(\omega(t,\tau,T) + \epsilon(t,\tau,T)\Big).$$

Rewrite this equation as follows:

$$y(\tau,T) - y(t,T) = s(t,\tau,T) + \eta(t,\tau,T), \tag{5.6}$$

where and

$$\eta(t,\tau,T) := \omega(t,\tau,T) + \epsilon(t,\tau,T),$$

is an error term composed of the forward rate forecast error and the future spot forecast error.

Decompose the forward rate forecast error into two orthogonal parts by regressing ω on s:

$$\omega(t,\tau,T) = b_0 + b_1\, s(t,\tau,T) + \xi(t,\tau,T), \tag{5.7}$$

where

$$b_1(\tau-t, T-\tau) = \frac{C[s(t,\tau,T), \omega(t,\tau,T)]}{V[s(t,\tau,T)]}$$

and

$$b_0(t,\tau,T) = E[\omega(t,\tau,T)] - b_1\, E[s(t,\tau,T)].$$

Now we can substitute (5.7) into (5.6) to obtain the Campbell-Shiller regression:

$$y(\tau,T) - y(t,T) = b_0 + (1+b_1)\, s(t,\tau,T) + \Big(\xi(t,\tau,T) + \epsilon(t,\tau,T)\Big). \tag{5.8}$$

Comparing (5.8) and (5.1), we see that $\beta_0 = b_0$, $\beta_1 = 1 + b_1$, and $\nu = \xi + \epsilon$.

Campbell and Shiller point out that if the (yield-to-maturity version of the) expectations hypothesis held, then $\beta_1 = 1$ for all $t < \tau < T$. We see that the expectations hypothesis is an assertion that the forward rate forecast error is uncorrelated with the slope of the yield curve.[11]

[10] See below for the explicit relationship.

[11] Under the YTM-EH, $\tilde{\mu}_P(X(t), T-t) = R(X(t))$. CIR (1981) pointed out that this condition cannot be met in the affine-exponential framework.

In our state-variable framework, we can derive expressions for the regression coefficients. Let

$$s(t, \tau, T) = \rho_0(\tau - t, T - t) + \rho_1(\tau - t, T - t)^\top X(t)$$

and

$$\omega(t, \tau, T) = \pi_0(\tau - t, T - t) + \pi_1(\tau - t, T - t)^\top X(t),$$

where ρ_1 and π_1 are the factor loadings in s and ω respectively. Then we can write

$$b_1 = \frac{C[s, \omega]}{V[s]} = \frac{C[\rho_1 X,\ \pi_1 X]}{V[\rho_1 X]} = \frac{\rho_1 V[X] \pi_1}{\rho_1 V[X] \rho_1} \tag{5.9}$$

and

$$b_0 = \pi_0 - b_1 \rho_0 + (\pi_1 - b_1 \rho_1) \theta,$$

where $V[X]$ is given in equation (A.4). Note that since $V[s] > 0$, $\rho_1 \neq 0$. Thus the key element is π_1, the factor loadings in the forward rate forecast error, ω.

The factor loadings for $s(t, \tau, T)$ are found by substituting the factor representation of the zero-coupon yields into the definition of $s(t, \tau, T)$ and collecting terms:

$$\rho_0(\tau - t, T - t) = \left(\frac{1}{T - \tau}\right) \left(\left(\frac{\tau - t}{T - t}\right) A(T - t) - A(\tau - t)\right)$$

and

$$\rho_1(\tau - t, T - t) = \left(\frac{1}{T - \tau}\right) \left(\left(\frac{\tau - t}{T - t}\right) B(T - t) - B(\tau - t)\right).$$

The factor loadings for $\omega(t, \tau, T)$ are found by substituting the factor representations of $E_t(y(\tau, T))$ and $f(t, \tau, T)$ into the definition of $\omega(t, \tau, T)$ and collecting terms:

$$\pi_0(\tau - t, T - t) = \left(\frac{1}{T - \tau}\right) \Big(A(T - \tau) + A(\tau - t) - A(T - t) + B(T - \tau)^\top (I - \Phi(\tau - t))\, \theta\Big)$$

and

$$\pi_1(\tau - t, T - t) = \left(\frac{1}{T - \tau}\right) \Big(\Phi(\tau - t)^\top B(T - \tau) + B(\tau - t) - B(T - t)\Big),$$

It can be verified that for the Vasicek model, $\pi_1 \equiv 0$; therefore $b_1 = 0$ and $\beta_1 = 1$. In the CIR model, however, $\pi_1 \neq 0$ and there is a bias in general.

Note that the relationship between the factor loadings for the forward term premium and for the forward rate forecast error factor is given by

$$G(T - t) = - \left. \frac{\partial \Big((T - \tau) \times \pi_1(\tau - t, T - t)\Big)}{\partial T} \right|_{\tau = T}.$$

The second set of regressions. In the second set of regressions, Campbell and Shiller use what they call the *perfect foresight spread* as the dependent variable. Let $\Delta := (T - t)/n$ and define the spread as

$$S(X(t), \Delta, T - t) := \left(\frac{\tau - t}{T - \tau}\right) s(X(t), \Delta, T - t).$$

Then the perfect-foresight spread is defined as

$$S^*(t,T,\Delta) := \frac{1}{n}\sum_{i=1}^{n-1}\Big(y(X_{t+i\,\Delta},\Delta) - y(X(t),\Delta)\Big). \tag{5.10}$$

The second set of regressions have this form:

$$S^* = \gamma_0 + \gamma_1\, S + \varepsilon, \tag{5.11}$$

where the arguments have been suppressed for notational clarity.

As above, we begin by decomposing $S^*(t,T,\Delta)$ into a conditional expectation and a forecast error:

$$S^*(t,T,\Delta) = E_t[S^*(t,T,\Delta)] + \upsilon(t,T,\Delta). \tag{5.12}$$

Next we add and subtract the spread, S, on the right-hand side of (5.12):

$$S^*(t,T,\Delta) = S(X(t),\Delta,T-t) + \phi(t,T,\Delta) + \upsilon(t,T,\Delta), \tag{5.13}$$

where $\phi(t,T,\Delta) := E_t[S^*(t,T,\Delta)] - S(X(t),\Delta,T-t)$. Once again, we can turn (5.13) into a regression by decomposing ϕ into $\phi = d_0 + d_1\, S + \varpi$, where $C[S,\varpi] = 0$. Then we can rewrite (5.13) as

$$S^* = d_0 + (1+d_1)\, S + (\varpi + \upsilon). \tag{5.14}$$

Comparing (5.14) with (5.11), we see that $\gamma_1 = 1 + d_1$ and $\gamma_0 = d_0$. Noting that $C[S,\phi] = C[S, E_t[S^*] - S] = C[S, E_t[S^*]] - V[S]$ and $E[\phi] = E[E_t[S^*]] - E[S] = -E[S]$, we can write

$$\gamma_1 = 1 + d_1 = \frac{C[S, E_t[S^*]]}{V[S]},$$

and

$$\gamma_0 = d_0 = E[\phi] - d_1\, E[S] = -\gamma_1\, E[S].$$

We can find an expression for $E_t[S^*(t,T,\Delta)]$ by first noting that

$$\begin{aligned} E_t[y(X_{t+i\,\Delta},\Delta)] &= \frac{1}{\Delta}\Big(A(\Delta) + B(\Delta)\, E_t[X_{t+i\,\Delta}]\Big) \\ &= \frac{1}{\Delta}\Big(A(\Delta) + B(\Delta)\,(\theta + \Phi(i\,\Delta)\,(X(t)-\theta))\Big), \end{aligned}$$

so that

$$E_t[y(X_{t+i\,\Delta},\Delta)] - y(X(t),\Delta) = \frac{1}{\Delta}\, B(\Delta)\,(\Phi(i\,\Delta) - I)\,(X(t)-\theta).$$

Then we can write

$$\begin{aligned} E_t[S^*(t,T,\Delta)] &= \frac{1}{n\,\Delta}\sum_{i=1}^{n-1} B(\Delta)\,(\Phi(i\,\Delta) - I)\,(X(t)-\theta) \\ &= \frac{1}{T-t}\, B(\Delta)\, Q \left[\left(\sum_{i=1}^{n-1} e^{-\kappa\, i\,\Delta}\right) - (n-1)\, I\right] Q^{-1}\,(X(t)-\theta), \end{aligned} \tag{5.15}$$

where $e^{-\kappa\, i\,\Delta}$ is a diagonal matrix and I is the identity matrix.

Equation (5.15) gives the factor loadings for $X(t)$ in $E_t[S^*]$, which we denote $\chi_1(\Delta, T-t)$. The factor loadings for $X(t)$ in S are given by $\tilde{\rho}_1 = B(T-t)/(T-t) - B(\Delta)/\Delta$. Using these factor loadings, we can write

$$\gamma_1 = \frac{C[S, E_t[S^*]]}{V[S]} = \frac{\tilde{\rho}_1\, V[X]\, \chi_1}{\tilde{\rho}_1\, V[X]\, \tilde{\rho}_1}. \tag{5.16}$$

Note that in the Vasicek model, $\chi_1 = \tilde{\rho}_1$, so that $\gamma_1 = 1$. However, the CIR model, $\chi_1 \neq \tilde{\rho}_1$. For example,

$$\frac{\chi_{1i}(1,2)}{\rho_{1i}(1,2)} = \frac{\left(e^{\kappa_i/2} - 1\right)\left(\left(e^{\gamma_i} + 1\right)\gamma_i + \left(e^{\gamma_i} - 1\right)\left(\kappa_i + \lambda_i\right)\right)}{e^{\kappa_i/2}\left(e^{\gamma_i/2} - 1\right)\left(\left(e^{\gamma_i/2} - 1\right)\gamma_i + \left(e^{\gamma_i/2} + 1\right)\left(\kappa_i + \lambda_i\right)\right)}.$$

6. GMM section

We can construct a GMM estimator based on the first set of regressions as follows.[12] We have shown in the previous section that we can write β_0 and β_1 in (5.1) as functions of the parameters of the model. For models in which none of the factors is Gaussian[13], we can identify all of the parameters in the model from β_0 and β_1. Define $h_t := (h_{0t}, h_{1t})$, where $h_{0t} := \Delta\, y_t - (\beta_0 + \beta_1\, s_t)$ and $h_{1t} := s_t\, h_{0t}$. Then the moment conditions are given by $E[h_t] = 0$. Let $g := (1/T)\sum_{i=1}^{T} h_i$ be the sample average of h. Then we can construct a GMM estimator of the parameters that minimizes

$$\mathcal{Q} = g^\top\, \hat{S}^{-1}\, g, \tag{6.1}$$

where $\hat{S}$ is a consistent estimator of the asymptotic variance of $\sqrt{T}g$.

[This section is incomplete.]

7. Empirical section

[This section is incomplete.]

8. Summary and conclusions

[This section is incomplete.]

[12]Gibbons and Ramaswamy (1993) used GMM to estimate a one-factor CIR model using the unconditional moments

$$E\left[\frac{P(X(\tau), T-\tau)}{P(X(t), T-t)}\right]$$

and

$$E\left[\frac{P(X(\tau), T-\tau)}{P(X(t), T-t)} \times \frac{P(X(\tau'), T'-\tau')}{P(X(t'), T'-t')}\right],$$

for $t < \tau \le T$, $t' < \tau' \le T'$, and $\tau \le t'$. We would use the logs of the price ratios instead.

[13]Need to show this.

Appendix A. Conditional moments

When b is diagonalizable, we can write solution (3.4) as

$$\hat{X}(t,t+\tau) = Q\, e^{\kappa\, \tau}\, Q^{-1}\, X(t) + Q \int_0^{\tau} e^{\kappa\, s}\, ds\, Q^{-1}\, a. \tag{A.1}$$

Substituting from (A.1) and (3.7), we have

$$\mathcal{F}(s) = \bar{A} + \sum_{\ell=1}^{d} \bar{B}_\ell \left[Q\, e^{\kappa\, s} \left(Y(t) + \int_0^s e^{-\kappa\, v}\, dv\, \bar{a} \right) \right]_\ell, \tag{A.2}$$

where $\bar{A} = Q^{-1}\, A\, (Q^{-1})^\top$, $\bar{B}_\ell = Q^{-1}\, B_\ell\, (Q^{-1})^\top$, $Y(t) = Q^{-1}\, X(t)$, $\bar{a} = Q^{-1}\, a$, and $[x]_\ell$ is the ℓ-th component of x. Note that we can write

$$\left[Q\, e^{\kappa\, s} \left(Y(t) + \int_0^s e^{-\kappa\, v}\, dv\, \bar{a} \right) \right]_\ell = \sum_{m=1}^{d} Q_{\ell m}\, e^{\kappa_m\, s} \left(Y_m(t) + \int_0^s e^{-\kappa_m\, v}\, dv\, \bar{a}_m \right).$$

Using this expression, we can rewrite (A.2) as follows:

$$\begin{aligned} \mathcal{F}(s) &= \bar{A} + \sum_{\ell=1}^{d} \bar{B}_\ell \left(\sum_{m=1}^{d} Q_{\ell m}\, e^{\kappa_m\, s} \left(Y_m(t) + \int_0^s e^{-\kappa_m\, v}\, dv\, \bar{a}_m \right) \right) \\ &= \bar{A} + \sum_{m=1}^{d} \beta_{ij}^m\, e^{\kappa_m\, s} \left(Y_m(t) + \int_0^s e^{-\kappa_m\, v}\, dv\, \bar{a}_m \right), \end{aligned}$$

where $\beta_{ij}^m = \sum_{\ell=1}^{d} (\bar{B}_\ell)_{ij}\, Q_{\ell m}$. Then we can write

$$p_{ij}(h) = \bar{A}_{ij}\, \mathcal{E}_{ij}(h) + \sum_{m=1}^{d} \beta_{ij}^m \left(\mathcal{E}_{ij}^m(h)\, Y_m(t) + \bar{\mathcal{E}}_{ij}^m(h)\, \bar{a}_m \right),$$

where

$$\begin{aligned} \mathcal{E}_{ij}(h) &= e^{(\kappa_i+\kappa_j)\, h} \int_0^h e^{-(\kappa_i+\kappa_j)\, s}\, ds, \\ \mathcal{E}_{ij}^m(h) &= e^{(\kappa_i+\kappa_j)\, h} \int_0^h e^{-(\kappa_i+\kappa_j-\kappa_m)\, s}\, ds, \end{aligned}$$

and

$$\bar{\mathcal{E}}_{ij}^m(h) = e^{(\kappa_i+\kappa_j)\, h} \int_0^h e^{-(\kappa_i+\kappa_j-\kappa_m)\, s} \left(\int_0^s e^{-\kappa_m\, v}\, dv \right) ds.$$

Depending on the eigenvalues, κ_i, we can write the $\mathcal{E}_{ij}$, $\mathcal{E}_{ij}^m$, and $\bar{\mathcal{E}}_{ij}^m$ as follows:

$$\mathcal{E}_{ij}(h) = \begin{cases} h & \text{if } \kappa_i+\kappa_j = 0 \\ \frac{e^{(\kappa_i+\kappa_j)\, h}-1}{\kappa_i+\kappa_j} & \text{otherwise,} \end{cases}$$

$$\mathcal{E}_{ij}^m(h) = \begin{cases} e^{(\kappa_i+\kappa_j)\, h}\, h & \text{if } \kappa_i+\kappa_j-\kappa_m = 0 \\ \frac{e^{(\kappa_i+\kappa_j)\, h}-e^{\kappa_m\, h}}{\kappa_i+\kappa_j-\kappa_m} & \text{otherwise,} \end{cases}$$

and

$$\bar{\mathcal{E}}_{ij}^{m}(h) = \begin{cases} h^2/2 & \text{if } \kappa_i + \kappa_j = \kappa_m = 0 \\ \frac{1+e^{\kappa_m h}(\kappa_m h - 1)}{\kappa_m^2} & \text{if } \kappa_i + \kappa_j = \kappa_m \neq 0 \\ \frac{e^{(\kappa_i+\kappa_j)h} - 1 + (\kappa_i+\kappa_j)h}{(\kappa_i+\kappa_j)^2} & \text{if } \kappa_i + \kappa_j \neq \kappa_m = 0 \\ \frac{e^{\kappa_m h} - 1 - \kappa_m h}{\kappa_m^2} & \text{if } \kappa_m \neq \kappa_i + \kappa_j = 0 \\ \frac{\left(1-e^{\kappa_m h}\right)(\kappa_i+\kappa_j) - \left(1-e^{(\kappa_i+\kappa_j)h}\right)\kappa_m}{(\kappa_i+\kappa_j-\kappa_m)(\kappa_i+\kappa_j)\kappa_m} & \text{otherwise.} \end{cases}$$

We now examine the behavior of the variance-covariance matrix as $h \to \infty$. We are interested in which linear combinations have finite variance at as $h \to \infty$. In other words, for what vectors δ is

$$\lim_{h\to\infty} \delta^\top p(h)\,\delta$$

finite? The linear combinations of $X(t)$ that have finite variance then are given by $\phi = (Q^{-1})^\top \delta$.

We consider two cases of empirical importance. In the first case, all the eigenvalues are real and negative; *i.e.*, $\kappa_i < 0$ for $i = 1$ to d. In this case,

$$\lim_{h\to\infty} \mathcal{E}_{ij}^{m}(h) = 0 \qquad \text{for all } (m,i,j),$$

$$\lim_{h\to\infty} \bar{\mathcal{E}}_{ij}^{m}(h) = \frac{1}{(\kappa_i+\kappa_j)\,\kappa_m} \qquad \text{for all } (m,i,j),$$

and

$$\lim_{h\to\infty} \mathcal{E}_{ij}(h) = -\frac{1}{\kappa_i+\kappa_j} \qquad \text{for all } (i,j).$$

Thus

$$p_{ij}(\infty) := \lim_{h\to\infty} p_{ij}(h) = \frac{1}{\kappa_i+\kappa_j}\left(-\bar{A}_{ij} + \sum_{m=1}^{d} \frac{\beta_{ij}^{m}\,\bar{a}_m}{\kappa_m}\right),$$

and all linear combinations have finite variance. Let

$$\bar{\mathcal{B}} := \sum_{m=1}^{d}\sum_{\ell=1}^{d} \frac{\bar{B}_\ell\, Q_{\ell m}\,\bar{a}_m}{\kappa_m}. \tag{A.3}$$

Then we can write the unconditional variance-covariance matrix as follows:

$$v(t, t+\infty) = Q\left\{\frac{-\bar{A}+\bar{\mathcal{B}}}{\kappa^+}\right\}Q^\top, \tag{A.4}$$

where $\kappa^+ = \{\kappa_{ij}^+\} = \{\kappa_i + \kappa_j\}$ is a $d \times d$ matrix and the division in (A.4) is component-by-component.

In the second case, $\kappa_1 = 0$ and the rest of eigenvalues are real and negative *i.e.*, $\kappa_i < 0$ for $i = 2$ to d. In this case,

$$\begin{aligned}
\mathcal{E}_{11}(h) &= \mathcal{E}^1_{11}(h) = h, \\
\mathcal{E}_{ij}(h) &= \mathcal{E}^m_{ij}(h) = \frac{e^{(\kappa_i+\kappa_j)\,h} - 1}{\kappa_i + \kappa_j} \quad \text{for } (i,j) \neq (1,1), \\
\bar{\mathcal{E}}^1_{11}(h) &= h^2/2, \\
\bar{\mathcal{E}}^1_{ij}(h) &= \frac{e^{(\kappa_i+\kappa_j)\,h} - 1 + (\kappa_i + \kappa_j)\,h}{(\kappa_i + \kappa_j)^2} \quad \text{for } (i,j) \neq (1,1).
\end{aligned}$$

and otherwise

$$\bar{\mathcal{E}}^1_{ij}(h) = \frac{\left(1 - e^{\kappa_m\,h}\right)(\kappa_i + \kappa_j) - \left(1 - e^{(\kappa_i+\kappa_j)\,h}\right)\kappa_m}{(\kappa_i + \kappa_j - \kappa_m)\,(\kappa_i + \kappa_j)\,\kappa_m}.$$

It looks like we have trouble with any $p_{ij}(\infty)$ for $i = 1$ or $j = 1$. Moreover, we will have trouble with any other $p_{ij}(\infty)$ unless $\bar{a}_1 = 0$. In this case, when $\bar{a}_1 = 0$, there is an equilibrium point. For now, let us assume $\bar{a}_1 = 0$. With this assumption, any δ for which $\delta_1 = 0$ will have a finite value of $\delta^\top p(\infty)\,\delta$. Moreover, the same is true for any ϕ for which $(Q^\top\phi)_1 = (Q^\top)_1\phi = 0$. In other words, and linear combination orthogonal to the eigenvector associated with the zero eigenvalue will have a finite unconditional variance.

We can now address the issue of cointegrated zero-coupon rates and forward rates in the case where we have one zero eigenvalue, the rest negative. Zero-coupon rates are given by $(T-t)^{-1}\left(A(T-t) + B(T-t)^\top X(t)\right)$. Let

$$\phi(\nu, \tau_1, \tau_2) := B(\tau_1)/\tau_1 - \nu\,B(\tau_2)/\tau_2.$$

Then linear combinations of rates will be cointegrated when ν is chosen so that

$$(Q^\top)_1\phi(\nu, \tau_1, \tau_2) = 0.$$

Now we seek the conditions for which *rate spreads* are cointegrated; *i.e.* the conditions for which

$$(Q^\top)_1(B(\tau_1)/\tau_1 - B(\tau_2)/\tau_2) = 0 \quad \text{for all } \tau_1, \tau_2 \geq 0. \tag{A.5}$$

Condition (A.5) implies

$$(Q^\top)_1 B(\tau) = c\,\tau \quad \text{for all } \tau \geq 0 \tag{A.6}$$

for some constant c. Note that if $(Q^\top)_1 = (1, 0, \ldots, 0)$, then (A.6) implies $B_1(\tau) = c\,\tau$, which is the loading that results from a pure Brownian motion in the short rate function. Thus if the first factor is an independent Brownian motion that enters the short-rate function, then zero-coupon spreads (and forward spreads) will be trend-stationary even though zero-coupon rates (and forward rates) will not be.

[This section is incomplete.]

References

Campbell, J. Y. (1986). A defense of traditional hypotheses about the term structure of interest rates. *Journal of Finance 41*, 183–193.

Campbell, J. Y. and R. J. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*, 495–514.

Chen, L. (1996). *Interest rate dynamics, derivatives pricing, and risk management*, Volume 435 of *Lecture Notes in Economics and Mathematical Systems.* New York: Springer-Verlag.

Constantinides, G. M. (1992). A theory of the nominal term structure of interest rates. *Review of Financial Studies 5*(4), 531–552.

Cox, J. C., J. E. Ingersoll, Jr., and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*(2), 385–407.

Duan, J.-C. and J.-G. Simomato (1995). Estimating and testing exponential-affine term structure models by Kalman filter. Working paper, McGill University.

Duffie, D. and R. Kan (1995). A yield-factor model of interest rates. Working paper, Graduate School of Business, Stanford University. Forthcomming: *Mathematical Finance.*

Fisher, M. and C. Gilles (1996). Around and around: The expectations hypothesis. Finance and Economics Discussion Series 96–17, Board of Governors of the Federal Reserve System, Washington, DC. Forthcomming: *Journal of Finance.*

Frachot, A. and J.-P. Lesne (1994). Expectations hypotheses and stochastic volatilities. Working paper, Banque de France.

Gibbons, M. and K. Ramaswamy (1993). A test of the Cox, Ingersoll, and Ross model of the term structure. *Review of Financial Studies 6*(3), 619–658.

Longstaff, F. A. and E. Schwartz (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance 47*, 1259–1282.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

(Mark Fisher) Research and Statistics, Board of Governors of the Federal Reserve System, Washington, DC 20551

E-mail address: `mfisher@frb.gov`

(Christian Gilles) Monetary Affairs, Board of Governors of the Federal Reserve System, Washington, DC 20551

E-mail address: `cgilles@frb.gov`