

# Appendix of the Paper *The Role of No-Arbitrage on Forecasting: Lessons from a Parametric Term Structure Model*

Caio Almeida

calmeida@fgv.br

José Vicente

jose.valentim@bcb.gov.br

June 2008

## 1 Introduction

In this appendix we give a supplementary material to the paper *The Role of No-Arbitrage on Forecasting: Lessons from a Parametric Term Structure Model*. First we present a proof of Theorem 1. Next we study the algebraic details of the AFG and AFSV versions of the Legendre model with affine dynamics.

## 2 Proof of Theorem 1

**Theorem 1** *Assume  $Y_t$ -dynamics under a probability measure  $\mathcal{Q}$  equivalent to  $\mathcal{P}$  given by:*

$$dY_t = \mu^{\mathcal{Q}}(Y_t)dt + \sigma(Y_t)dW_t^*, \quad (1)$$

where  $W^*$  is a Brownian motion under  $\mathcal{Q}$ .

If  $\mu^{\mathcal{Q}}(Y_t)$  satisfies the restriction expressed in Equation (2),  $\mathcal{Q}$  is an equivalent martingale measure and the AF conditions hold<sup>1</sup>.

$$\begin{aligned} \sum_{j=2}^N (j-1)L_j Y_{t,j} \tau^{j-2} &= \sum_{j=1}^N L_j \mu_j^{\mathcal{Q}}(Y_t) \tau^{j-1} - \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} \Gamma_{jk}(Y_t) \frac{\tau^{j+k-1}}{k} \\ \Gamma_{jk}(Y_t) &= 0 \text{ for } j > \lfloor \frac{N}{2} \rfloor \text{ or } k > \lfloor \frac{N}{2} \rfloor \end{aligned} \quad (2)$$

---

<sup>1</sup>In addition to the drift restriction,  $\sigma(Y_t)$  should present enough regularity to guarantee that discounted bond prices that are local martingales, also become martingales. In practical problems, a bounded or a square-affine  $\sigma(Y_t)$  is enough to enforce the martingale condition.

with  $\Gamma(Y_t) = L\sigma(Y_t)\sigma(Y_t)L'$ ,  $L_j$  standing for the  $j$ th-line of an upper triangular matrix that depends only on  $\ell$ , and  $[\cdot]$  representing the integer part of a number.

**Proof**

The term structure of the Legendre polynomial model is given by:

$$R(\tau, Y_t) = G(\tau)'Y_t = \sum_{n=1}^N Y_{t,n}P_{n-1}\left(\frac{2\tau}{\ell} - 1\right), \quad (3)$$

that is, the loadings of the term structure are Legendre polynomials. Therefore, the  $\tau$ -maturity instantaneous forward rate is

$$f(\tau, Y_t) = \sum_{n=1}^N Y_{t,n}P_{n-1}\left(\frac{2\tau}{\ell} - 1\right) + \tau \left( \sum_{n=1}^N Y_{t,n} \frac{\partial P_{n-1}\left(\frac{2\tau}{\ell} - 1\right)}{\partial \tau} \right). \quad (4)$$

In the equation above, the forward rates are expressed as linear combinations of Legendre polynomials, which can be readily expressed as linear combinations of powers of  $\tau$ :

$$f(\tau, Y_t) = \sum_{n=1}^N L_n Y_t \tau^{n-1}, \quad (5)$$

where  $L_n$  is the  $n^{\text{th}}$  row of the upper triangular matrix  $L$ . In fact, (5) defines matrix  $L$ . If  $N = 6$  the matrix  $L$  is<sup>2</sup>

$$L = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & \frac{4}{\ell} & -\frac{12}{\ell} & \frac{24}{\ell} & -\frac{40}{\ell} & \frac{60}{\ell} \\ 0 & 0 & \frac{18}{\ell^2} & -\frac{90}{\ell^2} & \frac{270}{\ell^2} & -\frac{630}{\ell^2} \\ 0 & 0 & 0 & \frac{80}{\ell^3} & -\frac{560}{\ell^3} & \frac{2450}{\ell^3} \\ 0 & 0 & 0 & 0 & \frac{350}{\ell^4} & -\frac{3150}{\ell^4} \\ 0 & 0 & 0 & 0 & 0 & \frac{1512}{\ell^5} \end{bmatrix}. \quad (7)$$

Proposition 3.2 of Filipovic (1999) presents conditions on  $f(\tau, Y_t)$ , which guarantee that discounted bond prices are martingales under any specific interest rate model<sup>3</sup>. Using these conditions, Almeida (2005) proves that if the AF restrictions (2) hold, then the Legendre polynomial model is arbitrage-free.  $\square$

<sup>2</sup>Using the first six Legendre polynomials we have

$$\begin{aligned} f(\tau, Y_t) = & Y_{t,1} + Y_{t,2}x + \frac{Y_{t,3}}{2}(3x^2 - 1) + \frac{Y_{t,4}}{2}(5x^3 - 3x) + \\ & \frac{Y_{t,5}}{8}(35x^4 - 30x^2 + 3) + \frac{Y_{t,6}}{8}(63x^5 - 70x^3 + 15x) + \\ & \frac{2\tau}{\ell} \left[ Y_{t,2} + 3Y_{t,3}x + \frac{Y_{t,4}}{2}(15x^2 - 3) \right] + \\ & \frac{2\tau}{\ell} \left[ \frac{Y_{t,5}}{8}(140x^3 - 60x) + \frac{Y_{t,6}}{8}(315x^4 - 210x^2 + 15) \right]. \end{aligned} \quad (6)$$

where  $x = \frac{2\tau}{\ell} - 1$ . Collecting terms that are powers of  $\tau$  in the expression above we obtain the upper triangular matrix  $L$  for  $N = 6$ .

<sup>3</sup>Basically, Proposition 3.2 of Filipovic (1999) imposes a specific relationship between the partial derivatives of  $f(\tau, Y_t)$ .

### 3 Technical Details about the Sub-Class of Arbitrage-Free Legendre Models with Affine Dynamics.

The affine class of dynamic term structure models is composed by processes whose state vector  $Y$  is an affine diffusion<sup>4</sup>, and whose implied short term rate is affine in  $Y$ . Dai and Singleton (2000) proposed the following notation to describe the dynamics of canonical affine models under the risk neutral measure  $\mathcal{Q}$ :

$$dY_t = \mu^{\mathcal{Q}}(Y_t)dt + \Sigma\sqrt{S_t(Y_t)}dW_t^* = \kappa^{\mathcal{Q}}(\theta^{\mathcal{Q}} - Y_t)dt + \Sigma\sqrt{S_t(Y_t)}dW_t^* \quad (8)$$

where  $\kappa^{\mathcal{Q}}$  and  $\Sigma$  are  $N \times N$  matrices,  $\theta^{\mathcal{Q}}$  is a  $\mathbb{R}^N$ -vector, and  $S_t$  is diagonal matrix with elements  $S_t^{ii} = \alpha_i + \beta_i'Y_t$  for some scalar  $\alpha_i$  and some  $\mathbb{R}^N$ -vector  $\beta_i$ .

Now suppose we want to equip the affine class of models with a loadings structure composed by Legendre polynomials<sup>5</sup>. To this end, we have to impose the AF restrictions of Theorem 1.

Consider the auxiliary state space vector  $\tilde{Y}_t$  defined by

$$\tilde{Y}_t = LY_t, \quad (9)$$

where  $L$  is the upper triangular matrix of Theorem 1. This auxiliary process characterizes term structure movements when the loadings come from a power series in the maturity variable  $\tau$ . It appears as an intermediate step in calculations.

The dynamics of  $\tilde{Y}_t$  under probability measure  $\mathcal{Q}$  is given by

$$d\tilde{Y}_t = \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)dt + \tilde{\Sigma}\sqrt{\tilde{S}_t(\tilde{Y}_t)}dW_t^*, \quad (10)$$

where the parameters of this stochastic differential equations system are defined in similar way to (8) (i.e.,  $\tilde{S}_t^{ii} = \tilde{\alpha}_i + \tilde{\beta}_i'\tilde{Y}_t$  for some scalar  $\tilde{\alpha}_i$  and some  $\mathbb{R}^N$ -vector  $\tilde{\beta}_i$  and so on) and are related through (9) with the corresponding parameters in (8). It should be clear that  $\tilde{Y}_t$  is affine if, and only if,  $Y_t$  is affine, because  $L$  is invertible.

Under this particular sub-class (affine plus polynomial loadings), the first requirement of AF restrictions becomes

$$\sum_{j=2}^N (j-1)\tilde{Y}_{t,j}\tau^{j-2} = \sum_{j=1}^N \mu_j^{\mathcal{Q}}(\tilde{Y}_t)\tau^{j-1} - \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} \sum_{k=1}^{\lfloor \frac{N}{2} \rfloor} (\tilde{H}_{0,jk} + \tilde{H}_{1,jk}\tilde{Y}_t) \frac{\tau^{j+k-1}}{k}, \quad (11)$$

<sup>4</sup>This means that the drift and the squared diffusion terms of  $Y$  are affine functions of  $Y$ .

<sup>5</sup>Note that it is not possible to make use of Duffie and Kan (1996) separation arguments that lead to their pair of Ricatti equations since the Legendre polynomials do not satisfy one of the algebraic conditions stated in their main theorem.

where  $(\tilde{\Sigma}\tilde{S}_t\tilde{\Sigma}')_{ij} = \tilde{H}_{0ij} + \tilde{H}_{1ij}\tilde{Y}_t$ , with  $\tilde{H}_{0ij} \in \mathbb{R}$  and  $\tilde{H}_{1ij} \in \mathbb{R}^N$ .

In particular, by matching coefficients on the maturity variable  $\tau$  in (11), we obtain an explicit expression for the drift of the auxiliary process:

$$\tilde{\mu}^Q(\tilde{Y}_t)_i = i\tilde{Y}_{t,i+1} + \sum_{j=\text{Max}\{1, i-\lfloor \frac{N}{2} \rfloor\}}^{\text{Min}\{i-1, \lfloor \frac{N}{2} \rfloor\}} \frac{\tilde{H}_{0,j(i-j)} + \tilde{H}_{1,j(i-j)}\tilde{Y}_t}{i-j}. \quad (12)$$

This expression can be readily translated to a similar expression for the drift of the original state vector  $Y$  with the use of (9).

In the empirical section of our paper, we compare a three factor CS version with corresponding AF versions that present three stochastic factors with non-null diffusions. By Theorem 1, a natural way to implement this application, is to work with AF versions driven by six factors (three stochastic, three conditionally deterministic). In the next lines, we provide the restrictions that should be implemented to generate affine models with *polynomial* loadings, and how to translate those restrictions to generate affine models with *Legendre polynomial* loadings. After that, we explain in details the two AF versions chosen to be implemented in this work: the Arbitrage-Free Gaussian (AFG) version, in which the volatility of  $Y$  is deterministic and time independent, and the Arbitrage-Free Stochastic Volatility (AFSV) version, in which only one stochastic factor determines the volatility of  $Y$ .

When  $N = 6$  the dynamics of  $\tilde{Y}_t$  has the following form:

$$\begin{aligned} \tilde{S}_t^{ii}(\tilde{Y}_t) &= \begin{cases} \tilde{\alpha}_i + \tilde{\beta}'_i \tilde{Y}_t & \text{if } i \leq 3 \\ 0 & \text{if } i > 3, \end{cases} \\ \tilde{\Sigma}_{i,j} &= 0 \quad i, j > 3, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_1 &= \tilde{Y}_{t,2}, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_2 &= 2\tilde{Y}_{t,3} + \tilde{H}_{0,11} + \tilde{H}_{1,11}\tilde{Y}_t, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_3 &= 3\tilde{Y}_{t,4} + \frac{\tilde{H}_{0,12}}{2} + \frac{\tilde{H}_{1,12}}{2}\tilde{Y}_t + \tilde{H}_{0,21} + \tilde{H}_{1,21}\tilde{Y}_t, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_4 &= 4\tilde{Y}_{t,5} + \frac{\tilde{H}_{0,13}}{3} + \frac{\tilde{H}_{1,13}}{3}\tilde{Y}_t + \frac{\tilde{H}_{0,22}}{2} + \frac{\tilde{H}_{1,22}}{2}\tilde{Y}_t + \tilde{H}_{0,31} + \tilde{H}_{1,31}\tilde{Y}_t, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_5 &= 5\tilde{Y}_{t,6} + \frac{\tilde{H}_{0,23}}{3} + \frac{\tilde{H}_{1,23}}{3}\tilde{Y}_t + \frac{\tilde{H}_{0,32}}{2} + \frac{\tilde{H}_{1,32}}{2}\tilde{Y}_t, \\ \tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t)_6 &= \frac{\tilde{H}_{0,33}}{3} + \frac{\tilde{H}_{1,33}}{3}\tilde{Y}_t. \end{aligned} \tag{13}$$

The dynamics of the term structure movements  $Y$  under the original Legendre polynomial parameterization can then be obtained by solving (9). To that end, let us rewrite the drift  $\tilde{\mu}^{\mathcal{Q}}$  in matrix notation, as an affine transformation of  $\tilde{Y}$ :

$$\tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t) = M + U\tilde{Y}_t, \tag{14}$$

where  $U = U_1 + U_2$ , and  $U_1$ ,  $U_2$  and  $M$  are given by:

$$M = \begin{bmatrix} 0 \\ \tilde{H}_{0,11} \\ \frac{\tilde{H}_{0,12}}{2} + \tilde{H}_{0,21} \\ \frac{\tilde{H}_{0,13}}{3} + \frac{\tilde{H}_{0,22}}{2} + \tilde{H}_{0,31} \\ \frac{\tilde{H}_{0,23}}{3} + \frac{\tilde{H}_{0,32}}{2} \\ \frac{\tilde{H}_{0,33}}{3} \end{bmatrix}, \tag{15}$$

$$U_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (16)$$

$$U_2 = \begin{bmatrix} 0_{1 \times 6} \\ \tilde{H}_{1,11} \\ \frac{\tilde{H}_{1,12}}{2} + \tilde{H}_{1,21} \\ \frac{\tilde{H}_{1,13}}{3} + \frac{\tilde{H}_{1,22}}{2} + \tilde{H}_{1,31} \\ \frac{\tilde{H}_{1,23}}{3} + \frac{\tilde{H}_{1,32}}{2} \\ \frac{\tilde{H}_{1,33}}{3} \end{bmatrix}. \quad (17)$$

Finally, the drift and diffusion of process  $Y$  are given by:

$$\mu^{\mathcal{Q}}(Y_t) = L^{-1}\tilde{\mu}^{\mathcal{Q}}(\tilde{Y}_t) = L^{-1}\tilde{\mu}^{\mathcal{Q}}(LY_t) = L^{-1}M + L^{-1}ULY_t \quad (18)$$

and

$$\sigma(Y_t) = L^{-1}\tilde{\Sigma}\sqrt{\tilde{S}_t(LY_t)}. \quad (19)$$

In our empirical application, the maximum maturity is equal to  $\ell = 10$  years. Then, matrix  $L$  is given by:

$$L = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 0.4 & -1.2 & 2.4 & -4 & 6 \\ 0 & 0 & 0.180 & -0.9 & 2.70 & -6.3 \\ 0 & 0 & 0 & 0.08 & -0.56 & 2.24 \\ 0 & 0 & 0 & 0 & 0.035 & -0.3158 \\ 0 & 0 & 0 & 0 & 0 & 0.0152 \end{bmatrix}. \quad (20)$$

Now we are ready to specialize the drift restriction (13) to each particular AF version implemented in this paper (AFG and AFSV), and also to obtain the corresponding restrictions for the process of interest  $Y$ , the one that drives term structure movements within the Legendre polynomial model.

### 3.1 The AFG Version

Noting that in this version the matrix controlling the diffusion structure of vector  $\tilde{Y}$ , i.e.  $\tilde{S}(\cdot)$ , is the identity matrix, we directly obtain  $\tilde{\Sigma}\tilde{\Sigma}' = \tilde{H}_0$ , and from (19) we obtain the relation between  $\tilde{H}_0$  and  $\Sigma$ :

$$\tilde{H}_0 = L\Sigma^2((L^{-1})')^{-1} = L\Sigma^2L'. \quad (21)$$

If we adopt a diagonal matrix representation for  $\Sigma^6$ , with  $\Sigma_{ii}$  as the  $i^{\text{th}}$ -diagonal term, then, in order to match the second requirement of AF restrictions we must have  $\Sigma_{ii} = 0$  for  $i \geq 4$ . Therefore, using transformation  $L$  between  $Y$  and  $\tilde{Y}$ ,  $\tilde{H}_0$  can be explicitly related to the non-null diagonal terms in  $\Sigma$ :

$$\tilde{H}_0 = \begin{bmatrix} \Sigma_{11}^2 + \Sigma_{22}^2 + \Sigma_{33}^2 & -0.4\Sigma_{22}^2 - 1.2\Sigma_{33}^2 & -0.18\Sigma_{33}^2 & 0 & 0 & 0 \\ -0.4\Sigma_{22}^2 - 1.2\Sigma_{33}^2 & 0.16\Sigma_{22}^2 + 1.44\Sigma_{33}^2 & -0.216\Sigma_{33}^2 & 0 & 0 & 0 \\ 0.18\Sigma_{33}^2 & -0.216\Sigma_{33}^2 & 0.0324\Sigma_{33}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (22)$$

Since  $U_2$  is null under the Gaussian version, we learn from (18) that the two matrices ( $L^{-1}M$  and  $L^{-1}U_1L$ ) necessary to obtain an explicit expression for the

---

<sup>6</sup>This representation for  $\Sigma$  provides exactly the same identification structure of Dai and Singleton (2002).

drift  $\mu^{\mathcal{Q}}(Y_t)$  are given by:

$$L^{-1}M = \begin{bmatrix} \frac{5}{2}\Sigma_{11}^2 + \frac{5}{6}\Sigma_{22}^2 + \frac{1}{2}\Sigma_{33}^2 \\ \frac{5}{2}\Sigma_{11}^2 + \frac{3}{2}\Sigma_{22}^2 + \frac{11}{14}\Sigma_{33}^2 \\ \frac{5}{3}\Sigma_{22}^2 + \frac{5}{7}\Sigma_{33}^2 \\ \Sigma_{22}^2 + \Sigma_{33}^2 \\ \frac{9}{7}\Sigma_{33}^2 \\ \frac{5}{7}\Sigma_{33}^2 \end{bmatrix} \quad (23)$$

and

$$L^{-1}U_1L = \begin{bmatrix} 0 & 0.4 & -0.3 & 0.56667 & -0.41667 & 0.65667 \\ 0 & 0 & 0.9 & -0.5 & 1.25 & -0.77 \\ 0 & 0 & 0 & 1.3333 & -0.58333 & 1.7833 \\ 0 & 0 & 0 & 0 & 1.75 & -0.63 \\ 0 & 0 & 0 & 0 & 0 & 2.16 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (24)$$

Note that  $Y_4$ ,  $Y_5$ , and  $Y_6$  are deterministic factors under the Gaussian case. This is a consequence of two facts: (i) their dynamics do not depend on the Brownian motion vector, and (ii) their drifts do not depend on the first three components of the state vector. With matrices  $L^{-1}M$  and  $L^{-1}U_1L$  in hands, we obtain the drift of vector  $Y$ , and in particular, the drifts of the deterministic factors  $Y_4$ ,  $Y_5$ , and  $Y_6$ :

$$\begin{aligned} \mu^{\mathcal{Q}}(Y_t)_4 &= \Sigma_{22}^2 + \Sigma_{33}^2 + 1.75Y_{t,5} - 0.63Y_{t,6}, \\ \mu^{\mathcal{Q}}(Y_t)_5 &= \frac{9}{7}\Sigma_{33}^2 + 2.16Y_{t,6}, \\ \mu^{\mathcal{Q}}(Y_t)_6 &= \frac{5}{7}\Sigma_{33}^2. \end{aligned} \quad (25)$$

By explicitly solving the ordinary differential equations implied for these factors, we have

$$Y_{t,4} = Y_{0,4} + (\Sigma_{22}^2 + \Sigma_{33}^2 + 1.75Y_{0,5} - 0.63Y_{0,6})t + (0.9\Sigma_{33}^2 + 0.189Y_{0,6})t^2 + 0.45\Sigma_{33}^2t^3, \quad (26)$$

$$Y_{t,5} = Y_{0,5} + \left(\frac{9}{7}\Sigma_{33}^2 + 2.16Y_{0,6}\right)t + \frac{27}{35}\Sigma_{33}^2t^2, \quad (27)$$

$$Y_{t,6} = Y_{0,6} + \frac{5}{7}\Sigma_{33}^2t. \quad (28)$$

Note that, under this Gaussian version, the dynamics of the state variables  $Y_{t,4}$ ,  $Y_{t,5}$  and  $Y_{t,6}$ , in addition to being deterministic, are completely determined by parameters  $\Sigma_{22}$ ,  $\Sigma_{33}$ , and the initial conditions  $Y_{0,4}$ ,  $Y_{0,5}$  and  $Y_{0,6}$ .

### 3.2 The AFSV Version

The AFSV version, presents one stochastic factor driving the stochastic volatility of the three stochastic factors ( $Y_1$ ,  $Y_2$  and  $Y_3$ ). In order to keep the risk-neutral dynamics of both  $Y_t$  and  $\tilde{Y}_t$  within the sub-class of affine models with only one factor determining the volatility, we choose factor  $Y_3$  to drive the stochastic volatility<sup>7</sup>. Specifically we set

$$\beta'_i = [0 \ 0 \ \beta_{i3} \ 0 \ 0 \ 0],$$

what gives:

$$H_{1,ij} = [0 \ 0 \ h_{ij} \ 0 \ 0 \ 0] \quad 1 \leq i, j \leq 6;$$

where  $(\Sigma S_t \Sigma')_{ij} = H_{0ij} + H_{1ij}Y_t$  with  $H_{1ij} \in \mathbb{R}^N$ . This specifications imply that

$$H_1 \cdot Y_t = Y_{t,3}H,$$

where in the right-hand side we have a tensor product, with  $H$  being a  $6 \times 6$ -matrix with elements  $h_{ij}$  (see Duffie (2001) for the tensorial notation).

From (19) and the relation  $\tilde{Y}_t = LY_t$  we obtain

$$H_0 = \Sigma (\text{diag} [\alpha_1, \dots, \alpha_6]) \Sigma',$$

$$\tilde{H}_0 = LH_0L',$$

$$H = \Sigma (\text{diag} [b_1, \dots, b_6]) \Sigma'.$$

Since  $\tilde{H}_1 \cdot \tilde{Y}_t = L (H_1 \cdot Y_t) L' = Y_{t,3}LHL'$  we have

$$\tilde{H}_{1,ij} = \left[ 0 \ 0 \ z_{ij} \ 11.25z_{ij} \ \frac{720}{7}z_{ij} \ \frac{6250}{7}z_{ij} \right], \quad (29)$$

with  $z_{ij} = (LHL')_{ij}$ .

---

<sup>7</sup>Choices of any of the two remaining factors to capture stochastic volatility could have been implemented but with higher computational costs.

Hence from (15), (16) and (17) we can express  $M$  and  $U$  as well as the drift of  $Y_t$  as functions of  $\alpha_i, \beta_{i3}$  ( $i = 1, \dots, 6$ ) and  $\Sigma$ . Finally, for identification purposes, in the empirical implementation of this version, we fix  $\Sigma$  to be a diagonal matrix (with  $\Sigma_{ii} = 0$  for  $i \geq 4$  in order to match the second requirement of AF restrictions) and  $\alpha = [1 \ 1 \ 1 \ 0 \ 0 \ 0]$ .

### 3.3 Estimation Procedures

How does one estimate CS and AF versions of the Legendre polynomial model?

For the CS version, we run cross-sectional independent regressions for each point  $t$  in time, within the sample period. In a market of zero coupon bonds, assuming that we observe yields  $R_{obs}$  with measurement error, the model is estimated with the use of the following linear regression:

$$\hat{Y}_t = (F'F)^{-1}F'R_{obs}^t, \quad (30)$$

where  $R_{obs}^t$  is a vector containing observed yields, at time  $t$ , for different maturities  $(\tau_1, \dots, \tau_k)$ , and  $F$  is the following matrix:

$$F = \begin{bmatrix} P_0(\frac{2\tau_1}{\ell} - 1) & P_1(\frac{2\tau_1}{\ell} - 1) & \dots & P_{N-1}(\frac{2\tau_1}{\ell} - 1) \\ P_0(\frac{2\tau_2}{\ell} - 1) & P_1(\frac{2\tau_2}{\ell} - 1) & \dots & P_{N-1}(\frac{2\tau_2}{\ell} - 1) \\ \vdots & \vdots & & \vdots \\ P_0(\frac{2\tau_k}{\ell} - 1) & P_1(\frac{2\tau_k}{\ell} - 1) & \dots & P_{N-1}(\frac{2\tau_k}{\ell} - 1) \end{bmatrix}.$$

For both AF versions we use the Quasi-Maximum Likelihood (QML) procedure, adopting the methodology proposed by Chen and Scott (1993), with 2-, 5-, and 10-year maturity zero-coupon bonds priced exactly and 3-, and 7-year maturity zero-coupon bonds priced with i.i.d zero-mean errors. The conditional transition densities are obtained with the use of closed-form formulas for the first and the second moments of  $Y$  within the affine framework (see for instance, Duffee, 2002, and Jacobs and Karoui, 2006). Observe that for the AFG version the QML is, in fact, a pure maximum likelihood procedure since the transitions densities come exactly from a normal distribution.

## References

- [1] Almeida, C.R. (2005). Affine Processes, Arbitrage-Free Term Structures of Legendre Polynomials, and Option Pricing, *International Journal of Theoretical and Applied Finance*, **8**, 2, 1-23.
- [2] Chen, R.R. and Scott, L. (1993). Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates. *Journal of Fixed Income*, **3**, 14-31.
- [3] Dai, Q. and Singleton, K. (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance*, **LIV**, 5, 1943-1977.
- [4] Dai, Q. and Singleton, K. (2002). Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure. *Journal of Financial Economics*, **63**, 415-441.
- [5] Duffee, G. R. (2002). Term Premia and Interest Rates Forecasts in Affine Models. *Journal of Finance*, **57**, 405-443.
- [6] Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press.
- [7] Duffie, D. and Kan, R. (1996). A Yield Factor Model of Interest Rates. *Mathematical Finance*, **6**, 4, 379-406.
- [8] Filipovic, D. (1999). A Note on the Nelson and Siegel Family. *Mathematical Finance*, **9**, 4, 349-359.
- [9] Jacobs, K. and Karoui, L. (2006). Affine Term Structure Models, Volatility and the Segmentation Hypothesis, Working Paper, McGill University.