

Wishart Quadratic Term Structure Models

Gourieroux, C.¹, and R., Sufana²

(First version: December 2003)

¹CREST, CEPREMAP and University of Toronto.

²University of Toronto.

Wishart Quadratic Term Structure Models

Abstract

This paper reveals that the class of affine term structure models introduced by Duffie and Kan (1996) is much larger than it has been usually considered in the literature. We study "fundamental" risk factors, which represent multivariate risk aversion of the consumer or the volatility matrix of the technological activity returns, and argue that they can be defined as symmetric positive matrices. For such matrices we introduce a dynamic affine process called the Wishart autoregressive (WAR) process; this process is used to reveal the associated term structure. In this framework:

- i) we derive very simple restrictions on the parameters to ensure positive yields at all maturities;
- ii) we observe that the usual constraint that the volatility matrix of an affine process be diagonal up to a path independent linear invertible transformation can be considerably relaxed.

The Wishart Quadratic Term Structure Model is the natural extension of the one-dimensional Cox-Ingersoll-Ross model and of the quadratic models introduced in the literature.

Keywords: Affine Term Structure, Quadratic Term Structure, CAR Process, Affine Process, WAR Process.

JEL:G13, C51

1 Introduction

A term structure model requires a coherent specification of both the historical and risk-neutral properties of interest rates. Indeed, while the risk-neutral distribution is used to determine the term structure pattern and to compare the current prices of interest rate derivatives, the historical distribution is needed to predict the future term structures, the future derivative prices and for instance to determine the Value at Risk of a portfolio of bonds. Moreover, such a coherent specification is also needed on estimation. Indeed there are very few fixed income bonds or interest rate derivatives, which are liquid on a given day; therefore, much more information is contained in the time series dimension than in the cross-sectional one. While the number of liquid residual maturities for T-bonds in a given day is between 10 and 20, there are several hundreds trading days for which such observations are available. By writing a coherent specification, we establish some links between historical and risk-neutral representations, which allow us to exploit the time series information in order to improve the cross sectional analysis [Dai and Singleton (2003b)].

The most popular specification proposed in the literature and satisfying these requirements is the so-called affine term structure model, which defines the yields as affine functions of underlying state variables, with affine dynamics. The first general presentation of this class appeared in Duffie and Kan (1996), where the state variables x_t are defined in continuous time and satisfy a (multidimensional) diffusion equation with linear drift and volatility. This specification includes as special cases some well known one-factor models such as the Vasicek model [Vasicek(1977)], the CIR model [Cox, Ingersoll, Ross (1985)b], or multifactor models such as the two-factor CIR model [Chen, Scott (1992)] or the Longstaff-Schwartz model [Longstaff, Schwartz (1992)]³. The continuous-time diffusion affine processes are locally gaussian.

A unified presentation of these basic processes has been developed by [Duffie, Filipovic, Schachermayer (2003)] and includes jump processes. They considered continuous-time processes with an exponential affine conditional Laplace transform. This extended class of continuous-time affine processes includes the affine diffusion processes, but also some jump components with affine intensity.

However, the flexibility of this class of continuous-time affine models has been questioned in both the applied and theoretical literature. First, several applied studies have rejected different special cases of continuous-time affine models.⁴ Second, Duffie et alii (2003) mentioned that this class "essentially" include some mixture of Ornstein-Uhlenbeck, CIR and bifurcation processes. These remarks generate an incentive to extend the basic model. Two directions have been followed in the literature. The first consists in considering the so-called quadratic term structure models, where the yields are quadratic functions

³See also Langetieg (1980), Brown and Dybvig (1986), Hull and White (1987), Fong, Vasicek (1991), Chen and Scott (1992), Chen and Scott (1993), Hull and White (1993), Brown and Schaefer (1994), Balduzzi, Das, Foresi and Sundaram (1996), Duffie and Singleton (1997), Balduzzi, Das and Foresi (1998), Duffie and Liu (2001), Collin-Dufresne and Goldstein (2001).

⁴See e.g. Dai and Singleton (2000), Duffie (2002).

of a (multivariate) Ornstein-Uhlenbeck process ⁵. The second one introduces affine processes in discrete time (called compound autoregressive processes), defined by the condition [see Darolles, Gourieroux, Jasiak (2001)]:

$$E[\exp u' x_{t+1} | x_t] = \exp [a(u)' x_t + b(u)], \forall u. \quad (1)$$

Since the condition on the Laplace transform is written for the unitary horizon only, instead of being written for any real positive horizon as in the continuous time approach, we get an infinitely much larger class of affine processes in the discrete than in the continuous time [see Gourieroux, Monfort and Polimenis (2002) for a discussion].

It has been observed recently that the standard QTSM is a special case of an affine model obtained by stacking the factor values and their squares [Cheng, Scaillet (2002), Dai, Singleton (2003b)]. Also the QTSM has the following limitations:

- i) while the condition for positivity of the short term yield are easily derived, the positivity conditions for yields at any maturity are generally not given.
- ii) the joint distribution of the yields can be degenerate, which creates some difficulty both for interpretation and statistical inference [see Appendix 1 for a description of these models and of their properties].

The aim of this paper is to improve the quadratic term structure models (QTSM) in the framework of the affine class. The limitations of the QTSM (and of the affine term structure model) are likely due to a lack of structural interpretation of the factors introduced in the affine models. Indeed the fundamental factors have to represent the components of preferences or of the production activities related to the risk. These components can be measured by the second-order derivative of a utility function or by the volatility matrix of the activity return, and are naturally represented by symmetric negative or positive matrices. This idea is developed in this paper.

In Section 2 we introduce the Wishart autoregressive (WAR) process, which can be used for specifying the (discrete-time) dynamics of stochastic symmetric positive matrices. In the simplest case, the WAR process can be seen as the cross-sectional second-order moment of several multivariate Ornstein-Uhlenbeck processes. We compute its conditional moment generating function and show that it is an exponential affine function of the lagged values. Thus the process is compound autoregressive [see Darolles, Gourieroux, and Jasiak (2001)] and can be used to derive an affine term structure [Gourieroux, Monfort and Polimenis (2002)]. This term structure is given in Section 3, with the associated risk-neutral distribution of the factors. We also provide a simple restriction on the parameters which ensures the positivity of the yields at all maturities. At a first sight the results seem to be contradictory with the restrictive form of the

⁵See Longstaff (1989), Beaglehole and Tenney (1991, 1992), Constantinides (1992), Lu (1999), Ahn, Dittmar and Gallant (2002), Leippold and Wu (2001, 2002), Cheng and Scaillet (2002).

volatility matrix of an affine process emphasized in the literature [see Duffie and Kan (1996) for the determination of the restriction, Dai, Liu, and Singleton (1997), Dai and Singleton (2000) for a detailed discussion of the restrictions]. The purpose of Section 4 is to reconcile the two results by explaining why these "standard" constraints can be relaxed. The parallel specification in continuous time is presented in Section 5 and statistical inference in Section 6. Section 7 concludes. The proofs are gathered in Appendices.

2 The Wishart Autoregressive Process

The Wishart autoregressive (WAR) process is a process for symmetric positive definite matrices (Y_t) of dimension (n, n) . Thus it allows us to describe the dynamics of volatility-covolatility matrices, and also of the multivariate risk aversion coefficient (related to the Hessian of a utility function). The WAR process has been initially defined from gaussian VAR(1) processes as follows [see Gouriéroux, Jasiak, and Sufana (2003)]:

Definition 1 *The matricial process (Y_t) is a WAR process if it can be written as:*

$$Y_t = \sum_{k=1}^K x_{kt} x'_{kt}, \quad (2)$$

where $K \geq n$ and the n -dimensional processes (x_{kt}) , $k = 1, \dots, K$, are independent such that:

$$x_{k,t} = M x_{k,t-1} + \varepsilon_{k,t}, \quad k = 1, \dots, K,$$

with independent gaussian error terms: $\varepsilon_{k,t} \sim N(0, \Sigma)$. K is the degree of freedom, M the latent autoregressive coefficient and Σ the latent variance of the innovation.

The distribution of Y_{t+1} conditional on the values $x_{k,\tau}$, $k = 1, \dots, K$, $\tau \leq t$, is easily characterized by means of its conditional Laplace transform (or moment generating function)⁶:

$$\begin{aligned} \Psi_t(\Gamma) &= E_t [\exp Tr(\Gamma Y_{t+1})] \\ &= E_t \left[\exp \left(\sum_{i=1}^n \sum_{j=1}^n \gamma_{ij} Y_{ij,t+1} \right) \right], \end{aligned}$$

where Tr denotes the trace operator, $\Gamma = (\gamma_{ij})$ is a (n, n) symmetric matrix and E_t denotes conditional expectation. The conditional Laplace transform is derived in Appendix 2.

⁶Note that for symmetric matrices: $Tr(\Gamma Y_{t+1}) = \sum_{i=1}^n \sum_{j=1}^n \gamma_{ij} Y_{ij,t+1}$. Moreover in our framework the Laplace transform is defined on a neighbourhood of zero and characterizes the distribution of Y . This characterization is a consequence of the positivity of the matrix Y , and the existence of an expansion around $\Gamma = 0$ is a consequence of the existence and rate of increase of power moments of Y .

Proposition 2 *The conditional Laplace transform of the matricial process (Y_t) is:*

$$\Psi_t(\Gamma) = \frac{\exp \text{Tr} \left[M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M Y_t \right]}{[\det (Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2})]^{K/2}}. \quad (3)$$

In particular:

i) The process (Y_t) is a Markov process, since the conditional distribution depends on the information set through Y_t only.

ii) The conditional Laplace transform is an exponential affine function of Y_t , that is the process is compound autoregressive (CAR).

The conditional distribution is an extension of the Wishart distribution corresponding to the case of no serial dependence ($M = 0$). It is denoted $W_n(K, M, \Sigma)$.

Finally note, from Definition 1, that the matricial process $\Sigma^{-1/2}Y_t\Sigma^{-1/2} = Y_t(\Sigma^{-1/2})$, say, is also a WAR process with transition distribution $W_n(K, \Sigma^{-1/2}M\Sigma^{1/2}, Id)$.

At first sight, the construction of the WAR process looks similar to the construction of factors in a standard quadratic term structure model [see e.g. Longstaff (1989), Constantinides (1992), Beaglehole and Tenney (1992), Leippold and Wu (2001, 2002), Ahn, Dittmar and Gallant (2002)]. However it differs in two respects:

- i) First, each underlying variable $x_{k,t}$ satisfies a gaussian vector autoregressive model without intercept. This explains why the transformed process Y_t itself is a Markov process, that is why the values x_t have an effect on Y_t through the terms $x_t x_t'$ only.
- ii) Second, the Wishart process is obtained by summing several independent squared gaussian VAR's. The condition $K \geq n$ ensures that there is no deterministic relationship among the elements of the symmetric matrix Y_t and that the extended Wishart distribution is a continuous distribution on the domain of symmetric positive semidefinite matrices⁷. This eliminates the degenerate distribution which arises for the standard quadratic term structure model, where $K = 1$ and $\text{Rank}(Y_t) = \text{Rank}(x_t x_t') = 1$ [see Appendix 1 for a discussion].

The Wishart autoregressive process can be extended to allow for a noninteger degree of freedom K , in the same way as the chi-square distribution can be extended to the gamma distribution. When K is not an integer, the expression of the conditional Laplace transform remains identical, but the interpretation as a sum of squared gaussian VAR is no longer valid.⁸

⁷See Bilodeau and Brenner (1999) and references therein for a discussion of degenerate Wishart distributions.

⁸Since the term structure is quadratic with respect to $y_t^{1/2}$ (in the one dimensional case), the WAR process appears contradictory with the characterization of quadratic term structure derived in Leippold, Wu (2002) [see section 4].

3 The Wishart quadratic term structure

As noted in the introduction, it is important to develop a coherent approach to term structure analysis, which specifies both the historical and risk-neutral distributions. Following a recent literature, we focus on the law of the state price density [see Rogers (1997) or Dijkstra, Yao (2001) for a discussion]. The specification of the term structure model requires the following assumptions.

Assumption A1. The bond prices depend on some underlying factors (Y_t), say.

Assumption A2. The factors have fundamental interpretations in terms of preferences and technologies.

These assumptions are compatible with standard general equilibrium theory, in which the states of the economy are exogenous and the equilibrium prices are derived from these states (Assumption A1) [see Cox, Ingersoll, Ross (1985a,b) for an example of general equilibrium model]. In particular, they do not admit a priori interpretations in terms of asset prices. As shown later in the discussion, Assumption A2 is important to build unconstrained models.

Assumption A3. The historical distribution of the factor process is specified.

Assumption A4. A stochastic discount factor (sdf) $M_{t,t+1}$, which explains how to correct for both time and randomness in period $(t, t + 1)$, is specified as a function of Y_{t+1} .

Then the price of a zero-coupon bond with residual maturity h is:

$$B(t, h) = E_t(M_{t,t+1} \dots M_{t+h-1,t+h}). \quad (4)$$

It is important to note that the "scope for generating interest rate models (in this way) is immense" [Rogers (1997)]. Indeed for "fundamental" factors the only constraint imposed by the no-arbitrage condition is the strict positivity of the sdf⁹. In some sense the price of risk in the sdf can be chosen arbitrarily, which implies little relationship between the historical and risk-neutral distributions of the fundamental factors (see Leippold and Wu (2002) for the discussion of the choice of the price of risk in quadratic models).

In the remainder of this section we specify the sdf and the historical distribution of the factors to get a special affine term structure (Section 3.1). The model is based on the Wishart factor process to ensure in a simple way the positivity of the yields (Section 3.2). Then in Section 3.3 we discuss the risk-neutral probability and the change of probability.

⁹More restrictions have to be imposed if some factors correspond to yields, since the pricing formula (4) has to be satisfied for the associated zero-coupon bonds.

3.1 The term structure

In discrete time, affine term structures can be derived from compound autoregressive factor processes [see Gourieroux, Monfort and Polimenis (2002)]. We consider the following particular forms of Assumptions A3 and A4:

Assumption A3'. We assume Wishart autoregressive factors (Y_t), with $\Sigma = Id$.

Indeed the factors are defined up to an invertible linear mapping; therefore, a WAR factor (Y_t) with matrix Σ can always be replaced by a WAR factor $Y_t (\Sigma^{-1/2})$, with matrix $\Sigma = Id$.

Assumption A4'. The sdf is selected as an exponential affine function of the factor:

$$M_{t,t+1} = \exp [Tr (CY_{t+1}) + d], \quad (5)$$

where C is a (n, n) symmetric matrix and d is a scalar.

At first sight it might seem surprising that factors are represented by a symmetric positive definite matrix, and thus are subject to rather complicated constraints. However, the choice of the factor becomes clear in the general equilibrium framework. For instance, let us consider the underlying equilibrium model introduced by Cox, Ingersoll, Ross (1985a). At equilibrium, the interest rate is a function of different fundamentals [see Cox, Ingersoll, Ross (1985b), equation (6)], which include several symmetric positive (resp. negative) definite matrices such as the volatility matrix of the rate of return of production activities (resp. the second-order derivatives of the indirect utility functions). These symmetric matrices represent the fundamental risks and the way they are perceived by the consumer¹⁰. These fundamental risks justify the positivity constraint, which is to be satisfied by the yields. More precisely, in a structural model the impact of risk is generally explicit in the second-order term in the expansion of an objective function. Loosely speaking, it is related to a term of the type [see e.g. Karni (1979), equation (3.3)]:

$$\sum_{i=1}^n \sum_{j=1}^n x_i x_j \frac{\partial^2 u}{\partial x_i \partial x_j} = Tr \left(\frac{\partial^2 u}{\partial x \partial x'} x x' \right).$$

For instance, if the Hessian $\frac{\partial^2 u}{\partial x \partial x'} = C$ is fixed and x is random the expected effect becomes $Tr (CE (x x')) = Tr (CY)$, and has a negative sign. This structural interpretation explains why it is easy to derive in Section 3.2 the positivity restrictions on yields for such matricial factors Y_t .

Then the price at date t of a zero-coupon bond with residual maturity h is:

$$\begin{aligned} B(t, h) &= E_t (M_{t,t+1} \dots M_{t+h-1,t+h}) \\ &= E_t [M_{t,t+1} B(t+1, h-1)], \end{aligned} \quad (6)$$

by the law of iterated expectations. These prices are easily computed recursively [see Appendix 3].

¹⁰See e.g. the definition of multivariate risk aversion in Duncan (1977) and Karni (1979).

Proposition 3 *We have:*

$$B(t, h) = \exp \{ \text{Tr} [A(h) Y_t] + b(h) \}, h = 0, 1, \dots, \quad (7)$$

where the matrix $A(h)$ and the scalar $b(h)$ satisfy the recursive equations:

$$\begin{cases} A(h) = M' [C + A(h-1)] \{ Id - 2 [C + A(h-1)] \}^{-1} M, h \geq 0, \\ b(h) = d + b(h-1) - \frac{K}{2} \log \det \{ Id - 2 [C + A(h-1)] \}, h \geq 0, \end{cases} \quad (8)$$

with the initial conditions: $A(0) = 0, b(0) = 0$.

Thus the geometric yields $r(t, h) = -\frac{1}{h} \log B(t, h)$ are affine functions of the factor components (Y_t) and we get the expected affine term structure model. Note that, for integer K , the geometric yields are also quadratic functions of the latent processes ($x_{k,t}$), $l = 1, \dots, K$, which justifies the terminology quadratic term structure. The recursive equations (8) are easily solved numerically and allows us to avoid the use of the solution of the Ricatti differential equations involved in the continuous-time framework [see Section 5]. For instance, the first two yields correspond to:

$$\begin{aligned} A(1) &= M' C (Id - 2C)^{-1} M, \quad b(1) = d - \frac{K}{2} \log \det (Id - 2C), \\ A(2) &= M' \left[C + M' C (Id - 2C)^{-1} M \right] \left\{ Id - 2 \left[C + M' C (Id - 2C)^{-1} M \right] \right\}^{-1} M, \\ b(2) &= 2d - \frac{K}{2} \log \det \left\{ (Id - 2C) \left\{ Id - 2 \left[C + M' C (Id - 2C)^{-1} M \right] \right\} \right\}. \end{aligned}$$

3.2 Positivity of the yields

An interesting property of the Wishart quadratic term structure is related to the positivity of the yields [see Appendix 4]. It is a consequence of the following equivalence: $\text{Tr} [CY] \leq 0$ for any $Y \gg 0$ if and only if $C \ll 0$.

Proposition 4 *Let us assume that symmetric matrix C is negative semidefinite and that $-d + \frac{K}{2} \log \det (Id - 2C)$ is nonnegative. Then*

- i) $A(h)$ is negative semidefinite;*
- ii) $\text{Tr} [A(h) Y_t]$ is nonpositive;*
- iii) The sequence $A(h)$ is decreasing;*
- iv) The domain for the yield $r(t, h)$ is $[-\frac{1}{h} b(h), \infty)$. The lower bound $-\frac{1}{h} b(h)$ is nonnegative and increases with h .*

In practice, it seems natural to fix the constant d in the sdf to ensure that the domain of the short-term yield $r(t, 1)$ is $(0, \infty)$. Thus $d = \frac{K}{2} \log \det (Id - 2C)$ and the sdf becomes:

$$M_{t,t+1} = \exp \left[\text{Tr} (CY_{t+1}) + \frac{K}{2} \log \det (Id - 2C) \right]. \quad (9)$$

The conditions for positivity of the yields are extremely simple, especially when they are compared to the conditions usually derived for the standard Duffie-Kan model [see e.g. Duffie and Kan (1996), Dai and Singleton (2000) and the discussion in Section 4] or for the standard QTSM.

Moreover, the negativity condition on $A(h)$ [and C] allows for some interpretation of the components of the yields and of the sdf. Let us first consider the sdf. We have:

$$M_{t,t+1} = \exp [Tr (CY_{t+1}) + d] \leq \exp (d) .$$

The maximum value corresponds to a flat correction for risk and is reached when $Y_{t+1} = 0$ (absence of risk). Similarly we have:

$$r(t, h) = -\frac{1}{h} Tr [A(h) Y_t] - \frac{b(h)}{h} > -\frac{b(h)}{h} .$$

The lower bound is reached when $Y_t = 0$ (absence of risk) and corresponds to a deterministic, time-independent term structure. When Y_t increases (that is the perceived fundamental risk increases), the state prices (that is the sdf) decrease and the risk premium on interest rates increases.

3.3 Risk-neutral probability

The sdf can be decomposed as:

$$M_{t,t+1} = \frac{M_{t,t+1}}{B(t,1)} B(t,1) = \frac{M_{t,t+1}}{B(t,1)} \exp -r(t,1) ,$$

providing the density at t of the risk-neutral measure with respect to the historical one. We get:

$$\frac{M_{t,t+1}}{B(t,1)} = \exp \left[Tr (CY_{t+1}) - Tr \left[M' C (Id - 2C)^{-1} M Y_t \right] + \frac{K}{2} \log \det (Id - 2C) \right] . \quad (10)$$

Alternatively, the risk-neutral distribution can be characterized by its conditional Laplace transform. Let us denote by a * the computations under the risk-neutral distribution. We have:

$$\begin{aligned} \Psi_t^* (\Gamma) &= E_t^* [\exp Tr (\Gamma Y_{t+1})] \\ &= E_t \left[\frac{M_{t,t+1}}{B(t,1)} \exp Tr (\Gamma Y_{t+1}) \right] \\ &= \exp \left[-Tr \left[M' C (Id - 2C)^{-1} M Y_t \right] + \frac{K}{2} \log \det (Id - 2C) \right] \Psi_t (C + \Gamma) \\ &= \exp \left\{ Tr \left[M' \left\{ (C + \Gamma) [Id - 2(C + \Gamma)]^{-1} - C (Id - 2C)^{-1} \right\} M Y_t \right] \right. \\ &\quad \left. - \frac{K}{2} \log \det [Id - 2(Id - 2C)^{-1} \Gamma] \right\} . \end{aligned}$$

3.4 First and second-order conditional moments of the yields

In an affine term structure model the first and second-order conditional moments of the yields are affine functions of the lagged factor values. Appendix 5 proves that these moments are:

i) under the historical distribution

$$\begin{aligned} E_t[r(t+1, h)] &= -\frac{1}{h} \text{Tr} [A(h) M Y_t M'] + c(h), \\ V_t[r(t+1, h)] &= \frac{1}{h^2} \text{Tr} [A(h)^2 M Y_t M'] + d(h), \end{aligned}$$

where $c(h)$ and $d(h)$ are terms independent of Y_t .

ii) under the risk-neutral distribution

$$\begin{aligned} E_t^*[r(t+1, h)] &= -\frac{1}{h} \text{Tr} [A(h) (Id - 2C)^{-1} + 2C (Id - 2C)^{-1} A(h) (Id - 2C)^{-1} M Y_t M'] \\ &\quad + c^*(h), \\ V_t^*[r(t+1, h)] &= -\frac{4}{h} \text{Tr} [(Id - 2C)^{-1} A(h) (Id - 2C)^{-1} A(h) (Id - 2C)^{-1} M Y_t M'] \\ &\quad + d^*(h). \end{aligned}$$

4 Restrictions on the volatility matrix of an affine (CAR) process

Since the pioneering paper of Duffie and Kan (1996), it is usually considered in the literature that the volatility matrix of an affine process y_t , say, is essentially of the type¹¹:

$$V_t(y_{t+1}) = Q \begin{pmatrix} \alpha_1 + \beta_1' y_t & & 0 \\ & \ddots & \\ 0 & & \alpha_n + \beta_n' y_t \end{pmatrix} Q'.$$

Thus, up to a path independent linear transformation Q , the volatility matrix is diagonal [see Duffie and Kan (1996), Section 4].

This restriction is clearly contradictory with the results presented in Sections 2 and 3. We see that the factor process Y_t and the yield process are CAR

¹¹See e.g. Dai and Singleton (2000), the introduction and Section 2, Piazzesi (2003), Dai and Singleton (2003a), Section 3, Levendorskiy (2003) for presentation of this condition as necessary and sufficient in recent surveys on affine term structure models, or e.g. Andersen, Benzoni, Lund (2003) for its use in applied studies.

(affine)¹², but they clearly admit more complicated covariance structures (see Section 3.4). These results are reconciled after a careful look at the proof in Duffie and Kan (1996), page 398. Indeed the diagonal condition is necessary and sufficient if "the state space is an intersection of non parallel half spaces". It is no longer necessary if this assumption is relaxed or, equivalently, the so-called maximally flexible model introduced in Dai and Singleton (2000) is less than maximally flexible. The diagonality assumption is the reason for the apparent lack of generality of the "standard" affine model. Let us consider for expository purpose the dimension $n = 2$, which corresponds to a three-factor model. In our framework the factors y_{11} , y_{12} , y_{22} satisfy the constraints: $y_{11} > 0$, $y_{11}y_{22} - y_{12}^2 > 0$, which do not correspond to the intersection of half planes. The same type of remark is valid for the yields $r(t, 1)$, $r(t, 2)$, $r(t, 3)$, say. Since the underlying factors can be written as an affine function of the yields by inverting formula (7), they satisfy some constraints of the type: $\alpha'_1 r_t + \beta_1 > 0$, $(\alpha'_1 r_t + \beta_1)(\alpha'_2 r_t + \beta_2) - (\alpha'_3 r_t + \beta_3)^2 > 0$, say, which provide a nonlinear state space.

A similar assumption appears in Duffie, Filipovic, and Schachermayer (2003) where the state space is assumed $D = R_+^m \times R^n$ in the main part of the text, which automatically restricts the set of affine processes which are considered. The question of the choice of the state space is discussed in more details in Section 12 of Duffie, Filipovic, and Schachermayer (2003), when some examples of other state spaces are provided. They formulate a "rather bold conjecture" on the interesting state spaces. The present paper provides a partial answer to their conjecture¹³.

Loosely speaking, it seems natural to call an affine process a diffusion process with affine drift and volatility, or to call affine term structure a model in which the yields are affine functions of the factors. But the term "affine" model can be misleading. Indeed, the affine model may include nonlinear features. Typically, the state space can feature some nonlinearities, as well as a number $k < n$ of yields can satisfy a (stochastic) nonlinear relationship. This latter case is well illustrated by the basic quadratic model [see Appendix 1].

The results of section 3 are also contradictory with the necessary and sufficient conditions for a quadratic term structure model derived by Leippold, Wu (2002). A careful look at the proof shows a misuse of the so-called "principle of matching" (Appendix A). In the WAR one dimensional framework for instance we have

$$r(t, h) = C_1^*(h) + b_1^*(h)y_t = C_1^*(h) + b_1^*(h)(y_t^{1/2})^2.$$

The model is simultaneously affine and quadratic. This explains, why the WAR process which provides a QTSM and cannot be written as the square of an Ornstein-Uhlenbeck (when K is not integer) has not been found.

¹²The yield process is CAR, since it is an affine transformation of a CAR process [Darolles, Gouriéroux, and Jasiak (2001)].

¹³A complete answer in the two factor case is given in Gouriéroux, Sufana (2004)a.

Finally note that the nonlinear restrictions on the factor components due to structural interpretations can have an effect on the correlations between factors. Indeed it is known from the copula theory that a nonlinear domain of factors can imply some restriction on the range of admissible correlations. Thus we do not have necessarily to look for factors with an arbitrary correlation matrix. In some sense the concept of correlation is also a linear notion which is not necessarily appropriate in the framework of the so-called "affine" models, which include nonlinear features.

5 Continuous time specification

A similar analysis can be performed in continuous time. It is briefly presented in this section for comparison with the major part of the literature on term structure which is written in continuous time. In Section 5.1 we first define the continuous time Wishart process. The specification of the sdf in continuous time and the prices of zero-coupon bonds are discussed in Section 5.2. Then the risk-neutral distribution is derived in Section 5.3. Section 5.4 introduces a general framework

5.1 The continuous time Wishart process

Under some restrictive conditions on the parameter, a WAR process can be considered as a time discretized affine process written in continuous time.

Definition 5 *A continuous time process of symmetric positive definite matrices is a continuous time (CT) Wishart process if its drift is given by:*

$$E_t dY_t = (K\Omega + QY_t + Y_t Q') dt, \quad (11)$$

and its volatility-covolatility function is:

$$\text{cov}_t \left(dY_t^i, dY_t^j \right) = \left(\omega_{ij} Y_t + Y_t^j (\Omega^i)' + \Omega^j (Y_t^i)' + Y_{ij,t} \Omega \right) dt, \quad (12)$$

where Y_t^i is the i^{th} column of matrix Y_t and Ω^i is the i^{th} column of matrix Ω . The matrix Ω is constrained to be symmetric positive definite, the degree of freedom K is a real number larger than n and Q is a (n, n) matrix of "latent" autoregressive coefficients.

The drift and volatility functions are affine in the elements of Y_t . Thus Y_t is a continuous time affine process [see Duffie and Kan (1996), Duffie, Filipovic, and Schachermayer (2003)]. The stochastic differential system will involve a n^2 -dimensional brownian motion, but since the dY_t matrix is symmetric, the volatility matrix of (the volatility) Y_t is singular and only $\frac{n(n+1)}{2}$ brownian motions matter. However, it is more tractable to repeat the elements $Y_{ij,t}$ and $Y_{ji,t}$ in the description of the process. When the CTWAR is used as factor

process, it is always possible to assume $\Omega = Id$ for identification purposes. This simplifies the expression of the volatility-covolatility function.

To understand why the stochastic differential system (11)-(12) provides a process of symmetric positive definite matrices, it is useful to study the drift and volatility of a quadratic transformation $\alpha'Y_t\alpha$, where α is a n -dimensional vector. We get:

$$\begin{aligned} E_t(\alpha'dY_t\alpha) &= (K\alpha'\Omega\alpha + \alpha'QY_t\alpha + \alpha'Y_tQ'\alpha) dt, \\ V_t(\alpha'dY_t\alpha) &= (4K\alpha'\Omega\alpha\alpha'Y_t\alpha) dt. \end{aligned} \quad (13)$$

Let us consider what is arising when $\alpha'Y_t\alpha$ comes close to zero (or equivalently when $Y_t\alpha$ is close to zero). The conditional variance comes also close to zero, whereas the drift is close to $K\alpha'\Omega\alpha dt > 0$ (since $\alpha'Y_t\alpha = 0$ implies $Y_t\alpha = 0$ for a positive definite matrix), which is a reverting effect towards positivity. The degree of freedom K measures the magnitude of the reflection.

Finally note that any time discretized continuous-time Wishart process is a WAR process¹⁴, but there exist a lot of WAR processes without continuous-time counterpart [see [Gourieroux, Jasiak, and Sufana \(2003\)](#) for a discussion], that is which do not satisfy the so-called embeddability condition. They provide additional term structure patterns.

Example 1 : When $n = 2$, we get a three factor model, with factor components $Y_{11t}, Y_{22t}, Y_{12t}$. The factor satisfies the three dimensional diffusion system (for $\Omega = Id$) :

$$\begin{aligned} \begin{bmatrix} dY_{11t} \\ dY_{22t} \\ dY_{12t} \end{bmatrix} &= \begin{bmatrix} 2(q_{11}Y_{11t} + q_{12}Y_{12t}) + K \\ 2[q_{21}Y_{12t} + q_{22}Y_{22t}] + K \\ q_{21}Y_{11t} + q_{12}Y_{22t} + (q_{11} + q_{22})Y_{12t} \end{bmatrix} dt \\ &+ \begin{bmatrix} 4Y_{11t} & 0 & 2Y_{12t} \\ 0 & 4Y_{22t} & 2Y_{12t} \\ 2Y_{12t} & 2Y_{12t} & Y_{11t} + Y_{22t} \end{bmatrix}^{1/2} dW_t. \end{aligned}$$

It is easily checked that the volatility-covolatility matrix is positive definite since the three principal determinants are :

$$4Y_{11t}, 16Y_{11t}Y_{22t}, 16(Y_{11t} + Y_{22t})(Y_{11t}Y_{22t} - Y_{12t}^2),$$

and are all positive, when Y_t is positive definite.

¹⁴Since any time discretized Ornstein-Uhlenbeck process is a gaussian VAR(1) process.

5.2 The prices of zero-coupon bonds

Let us consider a continuous time sdf:

$$M_t = \exp \left\{ \int_0^t [Tr(CY_u) + d] du + Tr(C^*Y_t) \right\}, \quad (14)$$

where C and C^* are (n, n) symmetric matrices.

The sdf is exponential affine with respect to future factor values. Equivalently, it can be defined by:

$$d \log M_t = Tr(CY_t + d) dt + Tr(C^*dY_t).$$

Thus, the log sdf satisfies a stochastic differential equation and is predetermined in the special case $C^* = 0$. Under this assumption, the yields are also affine functions of the factor [see Appendix 6].

Proposition 6 *For a continuous time Wishart factor process and exponential affine sdf, the prices of zero-coupon bonds are:*

$$B(t, h) = \exp \{Tr[A(h)Y_t] + b(h)\},$$

where the functions A and b satisfy the multivariate Ricatti equation:

$$\begin{aligned} & Tr \left[\frac{dA(h)}{dh} Y_t + \frac{db(h)}{dh} \right] \\ = & Tr(CY_t) + d + Tr[(A(h) + C^*)(K\Omega + QY_t + Y_tQ')] \\ & + 2Tr\{[A(h) + C^*]\Omega[A(h) + C^*]Y_t\} \\ = & Tr(CY_t) + d + \frac{1}{dt} Tr[(A(h) + C^*)E_t dY_t] + \frac{1}{2} \frac{1}{dt} V_t [Tr[(A(h) + C^*)dY_t]], \end{aligned}$$

with initial conditions $A(0) = 0$, $b(0) = 0$ corresponding to $B(t, 0) = 1$.

Since the relation is valid for any level of risk factors Y_t , we can deduce the differential equations satisfied by $A(h)$ and $b(h)$. We have:

$$\begin{aligned} \frac{dA(h)}{dh} &= C + [A(h) + C^*]Q + Q'[A(h) + C^*] \\ &\quad + 2Tr\{(A(h) + C^*)\Omega(A(h) + C^*)\}, \\ \frac{db(h)}{dh} &= d + KTr[(A(h) + C^*)\Omega], \end{aligned}$$

with initial conditions $A(0) = 0$, $b(0) = 0$. In particular, the instantaneous interest rate is:

$$r_t = \lim_{h \rightarrow 0} -\frac{1}{h} \log B(t, h) = Tr \left(\frac{dA(0)}{dh} Y_t + \frac{db(0)}{dh} \right).$$

By applying the Ricatti equation for $h = 0$, we deduce:

$$\begin{aligned} r_t &= Tr(CY_t) + d + Tr[C^*(K\Omega + QY_t + Y_tQ')] \\ &\quad + 2Tr[C^*\Omega C^*Y_t] \\ &= Tr(CY_t) + d + \frac{1}{dt}Tr(C^*E_t dY_t) + \frac{1}{2}\frac{1}{dt}V_t[Tr(C^*dY_t)]. \end{aligned}$$

As in the discrete time framework, it can be verified that:

- i) the yields are nonnegative if the symmetric matrices C and C^* are negative definite;
- ii) the sequence $A(h)$ is a decreasing sequence of negative semidefinite matrices.

5.3 The risk-neutral distribution

The risk-neutral distribution does not depend on the predetermined component of the sdf. The properties below are proved in Appendix 6.

Proposition 7 *Under the risk-neutral probability:*

- i) *the factor process satisfies a stochastic system with a volatility equal to the historical volatility:*

$$V_t^*Tr(\Gamma dY_t) = V_tTr(\Gamma dY_t),$$

and a modified drift:

$$E_t^*(dY_t) = E_t(dY_t) + cov_t[Tr(C^*dY_t), dY_t].$$

- ii) *The conditional Laplace transform can be written as:*

$$\begin{aligned} \Psi^*(t, h)(\Gamma) &= E_t^*[\exp Tr(\Gamma Y_{t+h})] \\ &= \exp\{Tr[A(\Gamma, h)Y_t] + b(\Gamma, h)\}, \end{aligned}$$

where the coefficients $A(\Gamma, h)$, $b(\Gamma, h)$ satisfy the same Ricatti equation as in Proposition 6, with initial conditions: $A(\Gamma, 0) = \Gamma$, $b(\Gamma, 0) = 0$.

The result on the Laplace transform extends the formulas derived for the prices of zero-coupon bonds, which correspond to the special case $\Gamma = 0$. It explains how to price easily any derivative with a payoff, which is an exponential affine function of the fundamental risk factors. The differential equation satisfied by the sensitivity coefficients $A(\Gamma, h)$ and $b(\Gamma, h)$ does not depend on Γ and the computations differ only by the initial conditions, which is a standard result for affine models [see e.g. Duffie, Filipovic, and Schachermayer (2003) for continuous time, Gourieroux, Monfort and Polimenis (2002) for discrete time].

5.4 A unified framework

In fact it is easily seen that the risk neutral (and historical) dynamics of the factor process can be written under the following more concise form :

$$dY_t = \left(\tilde{\Omega}\tilde{\Omega}' + \tilde{A}Y_t + Y_t\tilde{A}' \right) dt + Y_t^{1/2}d\tilde{W}_tQ + Q'd\tilde{W}_t'Y_t^{1/2},$$

where \tilde{W}_t is a (n, n) stochastic matrix, whose components are independent brownian motions (see e.g. Gouriéroux, Sufana (2004)b). This representation can be useful for some computations, but it can be misleading since the information generated by the $n(n+1)/2$ elements of Y_t is strictly smaller than the information generated by the n^2 elements of \tilde{W}_t . This class of processes contains as special case (when $\tilde{\Omega} = 0, \tilde{A} = 0, Q = Id$) a subclass of Wishart processes completely studied in Bru (1989), (1991), O'Connell (2003), Donati-Martin et alii (2003).

6 Statistical inference

The estimation of an affine model is rather easy to perform if the number of observed yields is equal to the number of underlying factors [that is $\frac{n(n+1)}{2}$ in our framework]¹⁵. Indeed due to the affine term structure there is in general a one-to-one linear relationship between the observed yields and the factors. Since the Wishart factors follow a nondegenerate continuous distribution¹⁶ if $K \geq n$, the distribution of the yields is easily deduced by a standard change of variable.

We focus on the method of moments, since the transition density of the WAR has no simple analytical form, and sketch the estimation principles. For expository purposes, we assume $n = 2$ and thus $\frac{n(n+1)}{2} = 3$ factors denoted by $Y_{11t}, Y_{12t}, Y_{22t}$.

6.1 Mimicking factors

Let us consider three observed yields with residual maturities h_1, h_2, h_3 . We have:

$$r(t, h_j) = -\frac{1}{h_j}Tr[A(h_j)Y_t] - \frac{1}{h_j}b(h_j), j = 1, 2, 3,$$

¹⁵When the number of factors is smaller than the number of observed yields, the affine model implies linear affine deterministic relationships between the yields, which are systematically rejected with probability 1 from the available data.

When the number of factors is strictly larger than the number of observed yields, the exact method of moments described in this section can be replaced by a simulated method of moments.

¹⁶Note that the quadratic term structure models considered in the literature correspond to $K = 1$ and are degenerate. This degeneracy creates a lot of problems for estimation [see Appendix 1].

or equivalently:

$$r(t, h_j) = \left[-\frac{1}{h_j} a_{11}(h_j), -\frac{2}{h_j} a_{12}(h_j), -\frac{1}{h_j} a_{22}(h_j) \right] \begin{bmatrix} Y_{11t} \\ Y_{12t} \\ Y_{22t} \end{bmatrix} - \frac{1}{h_j} b(h_j), j = 1, 2, 3.$$

Thus, we can write:

$$r_t = [r(t, h_1), r(t, h_2), r(t, h_3)]' = A^* \begin{bmatrix} Y_{11t} \\ Y_{12t} \\ Y_{22t} \end{bmatrix} + b^*, \text{ say,}$$

and deduce expressions of the factors in terms of yields (called mimicking factors in the literature):

$$\begin{bmatrix} Y_{11t} \\ Y_{12t} \\ Y_{22t} \end{bmatrix} = (A^*)^{-1} (r_t - b^*). \quad (15)$$

6.2 Seemingly unrelated regressions (SUR)

The method of moments can be based on the first order conditional moments of the yields. We have [see Gouriou, Jasiak, and Sufana (2003)]:

$$E_t(Y_{t+1}) = MY_tM' + KId, \quad (16)$$

or equivalently:

$$E_t \begin{bmatrix} Y_{11,t+1} \\ Y_{12,t+1} \\ Y_{22,t+1} \end{bmatrix} = M^* \begin{bmatrix} Y_{11t} \\ Y_{12t} \\ Y_{22t} \end{bmatrix} + \begin{bmatrix} K \\ 0 \\ K \end{bmatrix},$$

where

$$M^* = \begin{pmatrix} m_{11}^2 & 2m_{11}m_{12} & m_{12}^2 \\ m_{11}m_{21} & m_{12}m_{21} + m_{11}m_{22} & m_{12}m_{22} \\ m_{21}^2 & 2m_{21}m_{22} & m_{22}^2 \end{pmatrix}.$$

We deduce:

$$\begin{aligned} E_t(r_{t+1}) &= A^* E_t \begin{bmatrix} Y_{11,t+1} \\ Y_{12,t+1} \\ Y_{22,t+1} \end{bmatrix} + b^* \\ &= A^* M^* \begin{bmatrix} Y_{11t} \\ Y_{12t} \\ Y_{22t} \end{bmatrix} + b^* + A^* \begin{bmatrix} K \\ 0 \\ K \end{bmatrix}, \end{aligned}$$

or:

$$E_t(r_{t+1}) = A^* M^* (A^*)^{-1} r_t + b^* - A^* M^* (A^*)^{-1} b^* + A^* \begin{bmatrix} K \\ 0 \\ K \end{bmatrix}. \quad (17)$$

This is a seemingly unrelated regression (SUR) model, where the parameters are nonlinear functions of the structural parameters of interest. Thus the system (17) can be used to estimate the structural parameters (if identifiable) by nonlinear least squares.

The number of structural parameters is equal to $\frac{n(n+1)}{2} + n^2 + 2$ ($\frac{n(n+1)}{2}$ for C , n^2 for M , 1 for d , 1 for K) while the number of reduced-form parameters in the SUR is $\frac{n(n+1)}{2} + \left[\frac{n(n+1)}{2}\right]^2$. Thus we will have a large degree of (first order) overidentification.

Table 1 : Overidentification.

Dimension	n	1	2	3
Number of factors	$\frac{n(n+1)}{2}$	1	3	6
Reduced parameters	$\frac{n(n+1)}{2} + \left[\frac{n(n+1)}{2}\right]^2$	2	12	42
Structural parameters	$\frac{n(n+1)}{2} + n^2 + 2$	4	9	17

6.3 The Laplace transform

Other moment conditions can be deduced from the conditional Laplace transform itself. Indeed we have:

$$E_t \exp Tr (\Gamma Y_{t+1}) = \frac{\exp Tr \left[M' \Gamma (Id - 2\Sigma \Gamma)^{-1} M Y_t \right]}{\left[\det (Id - 2\Sigma^{1/2} \Gamma \Sigma^{1/2}) \right]^{K/2}},$$

and both sides of the equation can be expressed in terms of observed yields. For instance:

$$E_t \exp Tr (\Gamma Y_{t+1}) = E_t \exp \left[(\gamma_{11}, 2\gamma_{12}, \gamma_{22}) (A^*)^{-1} (r_{t+1} - b^*) \right].$$

Thus a GMM estimator can be deduced from these moment conditions applied to a set of values of arguments $\Gamma_1, \dots, \Gamma_L$, say. It has also been proposed in Singleton (2001) the use of complex arguments Γ , that is to calibrate on the (multivariate) empirical characteristic function. The calibration on the moment generating function or on the characteristic function both provide asymptotically efficient estimators when the number of moments increase, since both functions characterize the joint conditional distribution. But more moments can be necessary in finite sample with the complex arguments, especially if the transition density features fat tails. Indeed it is easier to approach the tail by decreasing exponential functions (real Γ arguments) than by sine and cosine functions (complex Γ arguments) [see Darolles, Gouriou, and Jasiak (2001)].

7 Concluding remarks

For structural reasons it is natural to represent the fundamental factors of a term structure of interest rates by means of a (stochastic) symmetric positive

definite matrix. In this paper we consider the factor dynamics corresponding to a Wishart autoregressive process. Since the WAR process admits a conditional Laplace transform which is exponential affine with respect to lagged factor values, we get an affine term structure, which is easy to implement. Moreover due to the structural interpretation, a condition for the positivity of yields at all maturities has been easily derived. Finally the example of the Wishart quadratic term structure model shows that the restrictions usually introduced on the volatility matrix of an affine process are very constraining and explain largely the lack of flexibility of affine models mentioned in the applied literature.

APPENDIX

Appendix 1 The factors in the standard quadratic term structure model and degenerate model

A.1.1 Linear and nonlinear factors

The standard quadratic term structure model is such that the yields $r(t, h)$ are quadratic functions of a gaussian VAR with constant term. More precisely, we have:

$$r(t, h) = \mu(h)' x_t + x_t' \Lambda(h) x_t + \nu(h), \quad (18)$$

where the p -variate process (x_t) satisfies:

$$x_{t+1} = m_0 + Mx_t + \varepsilon_{t+1},$$

and (ε_t) is a sequence of iid gaussian variables $\varepsilon_t \sim N(0, \Sigma)$.

It is easily seen that the yield process is a special case of the CAR (affine) process. Indeed, it is a linear combination of x_t , $x_t x_t'$ and it can be shown that the process obtained by stacking the components of x_t , $x_t x_t'$ is also a CAR (affine), that is admits a conditional Laplace transform which is exponential affine with respect to the components of x_t , $x_t x_t'$ [see e.g. Dai and Singleton (2003b), Cheng and Scaillet (2002) for this remark. In fact the affine representation seems to be a consequence of a general result on finite dimensional term structure models obtained in Filipovic, Teichmann (2002)]. Thus equation (18) can be considered as either a nonlinear (quadratic) factor representation of the yields in terms of the p basic factors (x_t) , or a linear factor representation of the yields involving the $p + \frac{p(p+1)}{2}$ "linear" factors $x_t, x_t x_t'$.

A.1.2 Quadratic factors or jumps

The factor representation does not admit the same interpretation, according to the linear and nonlinear interpretation of equation (18). It can even become much more complicated to understand if we consider carefully the effect of the quadratic term. Let us assume for illustration $p = 1$. We get:

$$r(t, h) = \mu(h) x_t + \lambda(h) x_t^2 + \nu(h), h \geq 1. \quad (19)$$

In particular, the short-term interest rate satisfies:

$$r(t, 1) = \mu(1) x_t + \lambda(1) x_t^2 + \nu(1). \quad (20)$$

This equation does not allow to recover the factor value in a unique way. Indeed this equation admits two real solutions: $\frac{-\mu_0(1) + \epsilon_t \sqrt{\mu_0(1)^2 - 4\lambda(1)(\nu(1) + r(t, 1))}}{2\lambda(1)}$, where $\epsilon_t = \pm 1$, with only one of them being the true factor value. The solution corresponding to the factor value is obtained for $\epsilon_t = +1$, if $2\lambda(1) x_t - \mu_0(1) > 0$, or for $\epsilon_t = -1$, otherwise. Replacing x_t by its expression in formula (19), it is seen that the yields $r(t, h)$ and $r(t, 1)$ satisfy one of two nonlinear

deterministic relationships:

$$\begin{aligned}
 & r(t, h) \tag{21} \\
 = & \mu(h) \left[\frac{-\mu(1) + \epsilon_t \sqrt{\mu(1)^2 - 4\lambda(1)(\nu(1) + r(t, 1))}}{2\lambda(1)} \right] \\
 & + \lambda(h) \left[\frac{-\mu(1) + \epsilon_t \sqrt{\mu(1)^2 - 4\lambda(1)(\nu(1) + r(t, 1))}}{2\lambda(1)} \right]^2 + \nu(h),
 \end{aligned}$$

$h \geq 2$, where $\epsilon_t = \pm 1$ is independent of residual maturity h , but depends on the factor value x_t .

In these nonlinear relationships $r(t, h)$ is a rather complicated function of $r(t, 1)$, and is clearly not quadratic. Note also that $r(t, h)$ appears as a linear function of the regime indicator ϵ_t for curve selection. This explains the affine relationship between the yields which stems from the affine interpretation of the quadratic term structure model.

Moreover, the relation (21) shows that the quadratic term structure model can also be considered as a nonlinear two-factor model where the factors are the short-term interest rate (a quantitative process) and the regime indicator ϵ_t (a binary process), that is as a factor model with regime shifts. It is interesting to note that this regime indicator can be recovered as a nonlinear function of two different yields. Thus some endogenous switching regimes reported in the applied and theoretical term structure literature¹⁷ could arise as the consequence of some omitted (quadratic) nonlinearity. Note that this nonlinear feature arises while the model is a special affine term structure model.

When the dimension p increases, the results are similar: the number of regimes increases, and more yields are needed to hedge the higher number of regime indicators.

A.1.3 Degenerate conditional distribution

As mentioned above, the quadratic term structure model is a special case of an affine term structure model. Thus in a quadratic term structure model with p (nonlinear) factors, $p + \frac{p(p+1)}{2} + 1$ yields of different maturities satisfy deterministic linear relationships, and their joint distribution is degenerate.

We could expect that this conditional distribution is no longer degenerate if we only consider $p + 1$ yields, due to the additional noise generated by the quadratic terms. However, this is not true, as can be easily seen in the one factor case $p = 1$. As noted above, a yield with given maturity will satisfy one of two deterministic relationships with the short-term yield. As a consequence, the joint distribution of $[r(t, 1), r(t, h)]$ will also be degenerate, which will render impossible to find a correct fit to the data and use it for statistical inference.

¹⁷See e.g. Ahn and Thompson (1988), Gray (1996), Evans (2000), Ang and Bekaert (2001), Piazzesi (2001), Das (2002), Bansal and Zhou (2002), Dai, Singleton and Yang (2003).

Appendix 2 : Proof of Proposition 2

i) Let us first establish a preliminary lemma.

Lemma 8 For any symmetric positive semidefinite matrix Ω and any vector $\mu \in R^n$, we get:

$$\int_{R^n} \exp(-x' \Omega x + \mu' x) dx = \frac{\pi^{n/2}}{(\det \Omega)^{1/2}} \exp\left(\frac{1}{4} \mu' \Omega^{-1} \mu\right).$$

Proof. Indeed the integral on the left hand side is equal to:

$$\begin{aligned} & \int_{R^n} \exp\left[-\left(x - \frac{1}{2} \Omega^{-1} \mu\right)' \Omega \left(x - \frac{1}{2} \Omega^{-1} \mu\right)\right] \exp\left(\frac{1}{4} \mu' \Omega^{-1} \mu\right) dx \\ &= \frac{\pi^{n/2}}{(\det \Omega)^{1/2}} \exp\left(\frac{1}{4} \mu' \Omega^{-1} \mu\right), \end{aligned}$$

since the gaussian multivariate distribution with mean $\frac{1}{2} \Omega^{-1} \mu$ and covariance matrix $2\Omega^{-1}$ admits unit mass. ■

ii) We now prove Proposition 2. Let us first consider the case $K = 1$. Thus, the stochastic process (Y_t) is defined by $Y_t = x_t x_t'$, $x_{t+1} = M x_t + \Sigma^{1/2} \xi_{t+1}$ and $\xi_{t+1} \sim IIN(0, Id)$. The conditional Laplace transform of the process (Y_t) is:

$$\begin{aligned} \Psi_t(\Gamma) &= E[\exp(x_{t+1}' \Gamma x_{t+1}) | x_t] \\ &= E\left[\exp\left(\left(M x_t + \Sigma^{1/2} \xi_{t+1}\right)' \Gamma \left(M x_t + \Sigma^{1/2} \xi_{t+1}\right)\right) | x_t\right] \\ &= \exp(x_t' M' \Gamma M x_t) E\left[\exp\left(2 x_t' M' \Gamma \Sigma^{1/2} \xi_{t+1} + \xi_{t+1}' \Sigma^{1/2} \Gamma \Sigma^{1/2} \xi_{t+1}\right) | x_t\right]. \end{aligned}$$

Using the pdf of the standard normal,

$$f(\xi_{t+1}) = \frac{1}{2^{n/2} \pi^{n/2}} \exp\left(-\frac{1}{2} \xi_{t+1}' \xi_{t+1}\right),$$

and Lemma 8, we get:

$$\begin{aligned} & \Psi_t(\Gamma) \\ &= \frac{\exp(x_t' M' \Gamma M x_t)}{2^{n/2} [\det(\frac{1}{2} Id - \Sigma^{1/2} \Gamma \Sigma^{1/2})]^{1/2}} \\ & \exp\left[\frac{1}{4} \left(2 x_t' M' \Gamma \Sigma^{1/2}\right) \left(\frac{1}{2} Id - \Sigma^{1/2} \Gamma \Sigma^{1/2}\right)^{-1} \left(2 \Sigma^{1/2} \Gamma M x_t\right)\right] \\ &= \frac{\exp\left(x_t' M' \Gamma M x_t + 2 x_t' M' \Gamma (\Sigma^{-1} - 2\Gamma)^{-1} \Gamma M x_t\right)}{[\det(Id - 2 \Sigma^{1/2} \Gamma \Sigma^{1/2})]^{1/2}} \end{aligned}$$

$$\begin{aligned}
&= \frac{\exp \left[x_t' M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M x_t \right]}{[\det (Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2})]^{1/2}} \\
&= \frac{\exp \text{Tr} \left[M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M Y_t \right]}{[\det (Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2})]^{1/2}},
\end{aligned}$$

since we can commute within the trace operator. This formula is valid whenever $Id - 2\Sigma\Gamma$ is a positive definite matrix.

iii) Let us now consider the case of an arbitrary positive integer K , $K \geq 1$. We can always write: $Y_t = \sum_{k=1}^K Y_{kt}$, where the matricial processes (Y_{kt}) , $k = 1, \dots, K$, are independent WAR processes with $K = 1$. We deduce that:

$$\begin{aligned}
\Psi_t(\Gamma) &= \prod_{k=1}^n \frac{\exp \text{Tr} \left[M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M Y_{kt} \right]}{[\det (Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2})]^{1/2}} \\
&= \frac{\exp \text{Tr} \left[M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M Y_t \right]}{[\det (Id - 2\Sigma^{1/2}\Gamma\Sigma^{1/2})]^{n/2}}.
\end{aligned}$$

Appendix A.3 : Recursive computation of the zero-coupon prices

We have:

$$\begin{aligned}
B(t, h) &= E_t [M_{t,t+1} B(t+1, h-1)] \\
&= E_t [\exp \{ \text{Tr} (C Y_{t+1}) + d \} \exp \{ \text{Tr} [A(h-1) Y_{t+1}] + b(h-1) \}] \\
&= E_t [\exp \text{Tr} [(C + A(h-1)) Y_{t+1}] \exp [d + b(h-1)]] \\
&= \Psi_t [C + A(h-1)] \exp [d + b(h-1)] \\
&= \exp \text{Tr} \left[M' [C + A(h-1)] \{ Id - 2[C + A(h-1)] \}^{-1} M Y_t \right] \\
&\quad \exp \left[d + b(h-1) - \frac{K}{2} \log \det \{ Id - 2[C + A(h-1)] \} \right].
\end{aligned}$$

Finally, the initial conditions $A(0) = 0$, $b(0) = 0$ ensure that $B(t, 0) = 1$.

Appendix A.4 : Constraints on the yields

i) The proof is by recursion. Let us recall that $C \ll 0$. If $A(h-1) \ll 0$, we have also

$$[C + A(h-1)] \{ Id - 2[C + A(h-1)] \}^{-1} \ll 0,$$

and then $A(h) \ll 0$.

ii) If $A(h)$ is symmetric negative semidefinite, it can be written as $A(h) = \sum_{i=1}^n \lambda_i m_i m_i'$, where the eigenvalues λ_i are nonnegative and m_i are associated

eigenvectors. Thus we get:

$$\begin{aligned}
& Tr [A(h) Y_t] \\
&= Tr \left(\sum_{i=1}^n \lambda_i m_i m_i' Y_t \right) \\
&= \sum_{i=1}^n \lambda_i Tr (m_i m_i' Y_t) \\
&= \sum_{i=1}^n \lambda_i Tr (m_i' Y_t m_i) \\
&= \sum_{i=1}^n \lambda_i m_i' Y_t m_i \leq 0,
\end{aligned}$$

where the last two equalities follows since we can commute within the trace operator and since $\lambda_i \leq 0$ and $Y_t \gg 0$.

iii) Let us denote $D(h) = 2(C + A(h))$. It can be seen that $A(h)$ is a decreasing sequence if and only if $D(h)$ is a decreasing sequence. By the recursive formula for $A(h)$ in equation (8), we get:

$$D(h) = 2C + M' D(h-1) [Id - D(h-1)]^{-1} M,$$

which implies:

$$\begin{aligned}
& D(h) - D(h-1) \\
&= M' \left\{ D(h-1) [Id - D(h-1)]^{-1} - D(h-2) [Id - D(h-2)]^{-1} \right\} M \\
&= M' \sum_{p=1}^{\infty} [D(h-1)^p - D(h-2)^p] M.
\end{aligned}$$

The recursion hypothesis $D(h-1) \ll D(h-2)$ implies that $D(h-1)^p \ll D(h-2)^p$, for all $p \geq 1$ [see Gouriou and Monfort (1995), vol. 2, Proposition A-8.]. The result that $D(h)$ is a decreasing sequence follows since $D(1) - D(0) = 2M'C [Id - 2C]^{-1} M$ is negative semidefinite by the negativity assumption on matrix C and the fact that $M'BM \ll 0$ for any $B \ll 0$.

iv) The domain is immediately deduced since the Wishart distribution is continuous on the set of symmetric semidefinite matrices. Moreover, we have:

$$\begin{aligned}
-b(h) &= -d - b(h-1) + \frac{K}{2} \log \det \{Id - 2[C + A(h-1)]\} \\
&\geq -d - b(h-1) + \frac{K}{2} \log \det (Id - 2C).
\end{aligned}$$

Appendix A.5 : Expansion of the log-Laplace transform and conditional moments of the yields

A.5.1 : Expansion of the log-Laplace transform

The first and second order conditional moments of any linear combination of the elements of volatility-covolatility matrix are easily deduced from the log-Laplace transform. Indeed let us assume that we are interested in a linear combination $Tr(AY)$, say. The log-Laplace transform of this variable is:

$$\begin{aligned} \log E \exp Tr(\gamma AY) &= \log \Psi(\gamma A) \\ &\simeq \gamma E[Tr(AY)] + \frac{\gamma^2}{2} V[Tr(AY)], \end{aligned}$$

in a neighborhood of $\gamma = 0$.

A.5.2 : Conditional moments of the yields under the historical distribution

We have:

$$\log \Psi_t(\Gamma) = Tr \left[M' \Gamma (Id - 2\Sigma\Gamma)^{-1} M Y_t \right] - \frac{K}{2} \log \det (Id - 2\Gamma).$$

The second-order expansion with respect to Γ provides:

$$\log \Psi_t(\Gamma) \sim Tr [M' \Gamma M Y_t] + 2Tr [M' \Gamma^2 M Y_t] - \frac{K}{2} \log \det (Id - 2\Gamma).$$

We deduce that:

$$\begin{aligned} E_t Tr(\Gamma Y_{t+1}) &= Tr [M' \Gamma M Y_t] + ct \\ &= Tr [\Gamma M Y_t M'] + ct, \\ V_t Tr(\Gamma Y_{t+1}) &= 4Tr [M' \Gamma^2 M Y_t] + ct \\ &= 4Tr [\Gamma^2 M Y_t M'] + ct, \end{aligned}$$

where the constants are deduced from the expansion of $-\frac{K}{2} \log \det (Id - 2\Gamma)$.

Thus, under the historical distribution, the conditional moments of the yields are such that:

$$\begin{aligned} E_t [r(t+1, h)] &= -\frac{1}{h} Tr [A(h) M Y_t M'] + c(h), \\ V_t [r(t+1, h)] &= \frac{1}{h^2} Tr [A(h)^2 M Y_t M'] + d(h), \end{aligned}$$

where $c(h)$ and $d(h)$ are terms independent of Y_t .

A.5.3 : Conditional moments of the yields under the risk-neutral distribution

The risk-neutral log-Laplace transform is:

$$\begin{aligned} \log \Psi_t^*(\Gamma) &= Tr \left[M' \left\{ (C + \Gamma) [Id - 2(C + \Gamma)]^{-1} - C (Id - 2C)^{-1} \right\} M Y_t \right] \\ &\quad - \frac{K}{2} \log \det [Id - 2(Id - 2C)^{-1} \Gamma]. \end{aligned}$$

The second-order expansion gives:

$$\begin{aligned} \log \Psi_t^* (\Gamma) &\sim \text{Tr} \left[M' \left[2 (Id - 2C)^{-1} \Gamma (Id - 2C)^{-1} \Gamma (Id - 2C)^{-1} \right] M Y_t \right] \\ &\quad - \frac{K}{2} \log \det \left[Id - 2 (Id - 2C)^{-1} \Gamma \right]. \end{aligned}$$

We deduce:

$$\begin{aligned} &E_t^* [\text{Tr} (\Gamma Y_{t+1})] \\ &= \text{Tr} \left\{ \left[\Gamma (Id - 2C)^{-1} + 2C (Id - 2C)^{-1} \Gamma (Id - 2C)^{-1} \right] M Y_t M' \right\} + ct, \\ &V_t^* [\text{Tr} (\Gamma Y_{t+1})] \\ &= 4 \text{Tr} \left\{ \left[(Id - 2C)^{-1} \Gamma (Id - 2C)^{-1} \Gamma (Id - 2C)^{-1} \right] M Y_t M' \right\} + ct. \end{aligned}$$

The two first order conditional moments of the yields under the risk-neutral distribution follow from the above results.

Appendix A.6 : Computations in continuous time

A.6.1 : The prices of zero-coupon bonds

Let us assume that $B(t, h) = \exp \{ \text{Tr} [A(h) Y_t] + b(h) \}$, and determine the differential system satisfied by the coefficients $A(h)$ and $b(h)$. Since $M_{t, t+dt} = M_{t+dt}/M_t$, we have:

$$\begin{aligned} &B(t, h + dt) \\ &= E_t [M_{t, t+dt} B(t + dt, h)] \\ &\simeq E_t \exp \{ \text{Tr} (C Y_t + d) dt + \text{Tr} (C^* d Y_t) + \text{Tr} [A(h) Y_{t+dt}] + b(h) \} \\ &= E_t \exp \{ \text{Tr} (C Y_t + d) dt + \text{Tr} [(A(h) + C^*) Y_{t+dt}] - \text{Tr} (C^* Y_t) + b(h) \} \\ &= \exp \{ \text{Tr} (C Y_t + d) dt - \text{Tr} (C^* Y_t) + b(h) \} E_t \exp \{ \text{Tr} [(A(h) + C^*) Y_{t+dt}] \} \\ &= \exp \{ \text{Tr} (C Y_t + d) dt - \text{Tr} (C^* Y_t) + b(h) \} \exp \left\{ E_t \left\{ \text{Tr} [(A(h) + C^*) Y_{t+dt}] \right\} \right. \\ &\quad \left. + \frac{1}{2} V_t \left\{ \text{Tr} [(A(h) + C^*) Y_{t+dt}] \right\} \right\}. \end{aligned}$$

Thus if $A^i(h)'$ denotes the i^{th} row of matrix $A(h)$, $(C^{*i})'$ the i^{th} row of matrix C^* , and Y_t^i the i^{th} column of Y_t , we get:

$$\begin{aligned} &B(t, h + dt) \\ &\simeq \exp \{ \text{Tr} (C Y_t + d) dt - \text{Tr} (C^* Y_t) + b(h) \} \exp \left\{ \sum_{i=1}^n [A^i(h) + C^{*i}]' E_t (Y_{t+dt}^i) \right. \\ &\quad \left. + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [A^i(h) + C^{*i}]' \text{cov}_t (Y_{t+dt}^i, Y_{t+dt}^j) [A^j(h) + C^{*j}] \right\} \end{aligned}$$

$$\begin{aligned}
&= \exp \{Tr (CY_t + d) dt - Tr (C^* Y_t) + b (h)\} \exp \left\{ \sum_{i=1}^n [A^i (h) + C^{*i}]' (Y_t^i + m_t^i dt) \right. \\
&\quad \left. + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [A^i (h) + C^{*i}]' \Lambda_{ij,t} [A^j (h) + C^{*j}] dt \right\} \\
&= \exp \{Tr (CY_t + d) dt + b (h)\} \exp \{Tr [A (h) Y_t] + Tr [(A (h) + C^*) M_t] dt \\
&\quad + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [A^i (h) + C^{*i}]' \Lambda_{ij,t} [A^j (h) + C^{*j}] dt \}
\end{aligned}$$

where M_t and $\Lambda_{ij,t}$ are the appropriate drift and volatilities from Definition 5 and m_t^i is the i^{th} column of M_t . Comparing the expression above with the expression:

$$B(t, h + dt) = \exp \{Tr [A (h + dt) Y_t] + b (h + dt)\},$$

and taking $dt \rightarrow 0$, we deduce the differential equation:

$$\begin{aligned}
Tr \left[\frac{dA (h)}{dh} Y_t + \frac{db (h)}{dh} \right] &= Tr (CY_t) + d + Tr [(A (h) + C^*) M_t] \\
&\quad + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [A^i (h) + C^{*i}]' \Lambda_{ij,t} [A^j (h) + C^{*j}].
\end{aligned}$$

Replacing M_t and $\Lambda_{ij,t}$ by the expressions in terms of Y_t from Definition 5, we get:

$$\begin{aligned}
&Tr \left[\frac{dA (h)}{dh} Y_t + \frac{db (h)}{dh} \right] \\
&= Tr (CY_t) + d + Tr [(A (h) + C^*) (K\Omega + QY_t + Y_t Q')] \\
&\quad + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [A^i (h) + C^{*i}]' \left(\omega_{ij} Y_t + Y_t^j (\Omega^i)' + \Omega^j (Y_t^i)' + Y_{ij,t} \Omega \right) [A^j (h) + C^{*j}].
\end{aligned}$$

It is easy, but cumbersome, to check that the term involving the double summation is equal to :

$$2Tr \{A(h) + C^*\} \Omega [A(h) + C^*] Y_t \}.$$

Since this condition is valid for any Y_t , we deduce the equation in Proposition 6. The same proof can be followed for the conditional Laplace transform with $\Psi^*(t, h) (\Gamma)$ instead of $B(t, h)$.

A.6.2 : Risk-neutral distribution

By considering the local expression of the risk-neutral Laplace transform, we can derive the stochastic system satisfied by the factor under the risk-neutral

probability. We get:

$$\begin{aligned}
& E_t^* [\exp Tr (\Gamma dY_t)] \\
= & \frac{E_t [\exp Tr [(CY_t + d) dt + (C^* + \Gamma) dY_t]]}{E_t [\exp Tr [(CY_t + d) dt + C^* dY_t]]} \\
= & \frac{E_t [\exp Tr [(C^* + \Gamma) dY_t]]}{E_t [\exp Tr (C^* dY_t)]} \\
= & \exp \left\{ E_t Tr [(C^* + \Gamma) dY_t] - E_t Tr (C^* dY_t) \right. \\
& \left. + \frac{1}{2} V_t Tr [(C^* + \Gamma) dY_t] - \frac{1}{2} V_t Tr (C^* dY_t) \right\} \\
= & \exp \left\{ Tr [\Gamma E_t (dY_t)] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\Gamma^i)' cov_t (dY_t^i, dY_t^j) C^{*j} \right. \\
& \left. + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (C^{*i})' cov_t (dY_t^i, dY_t^j) \Gamma^j + \frac{1}{2} V_t Tr (\Gamma dY_t) \right\} \\
= & \exp \left\{ Tr [\Gamma E_t (dY_t)] + Tr [\Gamma cov_t [Tr (C^* dY_t), dY_t]] + \frac{1}{2} V_t Tr (\Gamma dY_t) \right\}.
\end{aligned}$$

Thus we observe that the quadratic term with respect to Γ is the same as the quadratic term corresponding to the historical distribution. Thus the volatilities are the same under the historical and risk-neutral distributions. The drift is modified and is given by:

$$\begin{aligned}
E_t^* (dY_t) &= \frac{E_t (M_{t+dt} dY_t)}{E_t (M_{t+dt})} \\
&= \frac{E_t \exp [Tr (C^* dY_t) dY_t]}{E_t \exp [Tr (C^* dY_t)]} \\
&\simeq \frac{E_t \{ [1 + Tr (C^* dY_t)] dY_t \}}{E_t [1 + Tr (C^* dY_t)]} \\
&\simeq \{ E_t (dY_t) + E_t [Tr (C^* dY_t) dY_t] \} \{ 1 - E_t [Tr (C^* dY_t)] \} \\
&\simeq E_t (dY_t) + E_t [Tr (C^* dY_t) dY_t] - E_t (dY_t) E_t [Tr (C^* dY_t)] \\
&= E_t (dY_t) + cov_t [Tr (C^* dY_t), dY_t].
\end{aligned}$$

References

- [1] Ahn, C., and H., Thompson (1988): "Jump Diffusion Processes and Term Structure of Interest Rates," *Journal of Finance*, 43, 155-174.
- [2] Ahn, D., Dittmar, R., and A., Gallant (2002): "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243-288.

- [3] Andersen, T., Benzoni, L., and J., Lund (2003): "Stochastic Volatility, Mean Drift and Jumps in the Short Rate Diffusion : Sources of Steepness, Level and Curvature", DP Northwestern Univ.
- [4] Ang A., and G., Bekaert (2001): "Regime Switches in Interest Rates," forthcoming, *Journal of Business and Economic Statistics*.
- [5] Backus, D., Foresi, S., and C., Telmer (1998): "Discrete Time Models of Bond Pricing," Working Paper, New York University.
- [6] Backus, D., Foresi, S., and C., Telmer (2001): "Affine Term Structure Models and the Forward Premium Anomaly", *Journal of Finance*, 56, 279-304.
- [7] Balduzzi, P., Das, S., Foresi, S., and R., Sundaram (1996): "A Simple Approach to Three Factor Affine Term Structure Models," *Journal of Fixed Income*, 6, 43-53.
- [8] Balduzzi, P., Das, S., and S., Foresi (1998): "The Central Tendency: A Second Factor in Bond Yields," *Review of Economics and Statistics*, 80, 62-72.
- [9] Bansal, R., and H., Zhou (2002): "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 57, 1997-2043.
- [10] Beaglehole, D., and M., Tenney (1991): "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations," *Journal of Fixed Income*, 1, 69-83.
- [11] Beaglehole, D., and M., Tenney (1992): "A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions," *Journal of Financial Economics*, 32, 345-354.
- [12] Bilodeau, M., and D., Brenner (1999): "Theory of Multivariate Statistics", Springer-Verlag.
- [13] Black, F., Derman, E., and W., Toy (1990): "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal*, 33-39.
- [14] Brown, R., and S., Schaefer (1994): "The Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross Model," *Journal of Financial Economics*, 35, 3-42.
- [15] Brown, S., and P., Dybvig (1986): "Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates," *Journal of Finance*, 41, 143-172.
- [16] Bru, M.F. (1989): "Diffusions of Perturbed Principal Component Analysis", *Journal of Multivariate Analysis*, 29, 127-136.

- [17] Bru, M.F. (1991): "Wishart Processes", *J. Theo. Probab.*,4, 725-751.
- [18] Chacko, G., and S., Das (2002): "Pricing Interest Rate Derivatives : A General Approach", *Review of Financial Studies*, 15, 195-241.
- [19] Chen, L, and H., Poor (2002): "A General Characterization of Quadratic Term Structure Models", Princeton University, DP.
- [20] Chen, R., and L., Scott (1992): "Pricing Interest Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure," *Review of Financial Studies*, 5, 613-636.
- [21] Chen, R., and L., Scott (1993): "Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14-31.
- [22] Chen, R., and L., Scott (1995): "Interest Rate Options in Multifactor Cox-Ingersoll-Ross Models of the Term Structure," *Journal of Fixed Income*, Winter, 53-72.
- [23] Cheng, P., and O., Scaillet (2002): "Linear Quadratic Jump Diffusion Modelling with Application to Stochastic Volatility", HEC Geneva DP.
- [24] Collin-Dufresne, P., and R., Goldstein (2001): "Do Bonds Span the Fixed Income Markets ? Theory and Evidence for Unspanned Stochastic Volatility," Working Paper, GSIA, Carnegie Mellon University and Ohio-State University.
- [25] Constantinides, G. (1992): "A Theory of the Nominal Term Structure of Interest Rates. *Review of Financial Studies*, 5, 531-552.
- [26] Cox, J., Ingersoll, J., and S., Ross(1985a): "An Intertemporal General Equilibrium Model of Asset Prices", *Econometrica*, 53, 363-384.
- [27] Cox, J., Ingersoll, J., and S., Ross(1985b): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385-407.
- [28] Dai, Q., Liu, J., and K., Singleton (1997): "Admissibility and Identification of Affine Asset Pricing Models", Working Paper, Stanford University.
- [29] Dai, Q., and K., Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-1978.
- [30] Dai, Q., and K., Singleton (2003a): "Term Structure Dynamics in Theory and Reality," *Review of Financial Studies*, 16, 631-678.
- [31] Dai, Q., and K., Singleton (2003b): "Fixed Income Pricing," *Handbook of Economics and Finance* , Chapter 20, Constantinides., C., Harris, M., and R., Stulz, eds, North Holland.

- [32] Dai, Q., Singleton, K., and W., Yang (2003): "Are Regime Shifts Priced in U.S. Treasury Markets?" Working Paper, Stanford University.
- [33] Darolles, S., Gouriéroux, C., and J., Jasiak (2001): "Structural Laplace Transform and Compound Autoregressive Models", CREST DP.
- [34] Das, S. (2002): "The Surprise Element: Jumps in Interest Rates," *Journal of Econometrics*, 106, 27-65.
- [35] Dijkstra, T., and Y., Yao (2002): "Moment Generating Function Approach to Pricing Interest Rate and Foreign Exchange Rate Claims", *Insurance: Mathematics and Economics*, 31, 163-178.
- [36] Donati-Martin, C., Doumerc, Y., Matsumoto, H., and M., Yor (2000) : "Some Properties of the Wishart Processes and a Matrix Extension of the Hartman-Watson Laws", Preprint University Paris VI.
- [37] Duan, J., and J., Simonato (1999): "Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter," *Review of Quantitative Finance and Accounting*, 13.
- [38] Duffee, G. (2002): "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405-443.
- [39] Duffie, D., Filipovic, D., and W., Schachermayer (2003): "Affine Processes and Applications in Finance," *Annals of Applied Probability*, Volume 13, pp. 984-1053.
- [40] Duffie, D., and R., Kan (1996): "A Yield Factor Model of Interest Rates," *Mathematical Finance*, 6, 379-406.
- [41] Duffie, D., and J., Liu (2001): "Floating-Fixed Credit Spreads," *Financial Analysts Review*, 57 (3), 76-87.
- [42] Duffie, D., and K., Singleton (1997): "An Econometric Model of the Term Structure of Interest Rate Swap Yields," *Journal of Finance*, 52, 1287-1321.
- [43] Duncan, G. (1977): "A Matrix Measure of Multivariate Local Risk Aversion", *Econometrica*, 45, 895-903.
- [44] El Karoui, N., R., Myneni, and R., Viswanathan (1992): "Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: I. Theory," DP, Paris VI University and Stanford University.
- [45] Evans, M. (2000): "Regime Shifts, Risk, and the Term Structure," Working Paper, Georgetown University.
- [46] Filipovic, D. (2001): "A General Characterization of One Factor Affine Term Structure Models", *Finance and Stochastics*, 5, 389-412.

- [47] Filipovic, D., and J., Teichmann (2002) : "On Finite Dimensional Term Structure Models", Princeton Univ. DP.
- [48] Fisher, M., and C., Gilles (1996): "Estimating Exponential Affine Models of the Term Structure," Working Paper, Federal Reserve Board, Washington DC.
- [49] Fong, H., and O., Vasicek (1991): " Fixed Income Volatility Management", Journal of Portfolio Management, 17, 177-191.
- [50] Gibson, R., Lhabitant, F., and D., Talay (2001): "Modeling the Term Structure of Interest Rates: A Review of the Literature," Working Paper, HEC.
- [51] Gouriéroux, C., and A., Monfort (1995): "Statistics and Econometric Models," Cambridge University Press.
- [52] Gouriéroux, C., Monfort, A., and V., Polimenis (2002): "Affine Term Structure Models".
- [53] Gouriéroux, C., Jasiak, J., and R., Sufana (2003): "A Dynamic Model for Multivariate Stochastic Volatility: The Wishart Autoregressive Process", University of Toronto, DP.
- [54] Gouriéroux, C., and R., Sufana (2004)a: "A Comparison of Two Factor Affine Models", CREST DP.
- [55] Gouriéroux, C., and R., Sufana (2004)b: "Derivative Pricing with Multivariate Stochastic Volatility", CREST DP.
- [56] Gray, S. (1996): "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," Journal of Financial Economics 42, 27-62.
- [57] Heston, S. (1993): "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", The Review of Financial Studies, 6, 327-343.
- [58] Hull, J., and A., White (1987): "The pricing of Options on Assets with Stochastic Volatilities", Journal of Finance 42, 281-300.
- [59] Hull, J., and A., White (1993): "One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities," Journal of Financial and Quantitative Analysis, 28, 235-254.
- [60] Jamshidian, F. (1996): "Bond, Futures and Options Evaluation in the Quadratic Interest Rate Model", Applied Mathematical Finance, 3, 93-115.
- [61] Johannes, M. (2000): "A Nonparametric View of the Role of Jumps to Interest Rates," Working Paper, Columbia University.

- [62] Karni, E. (1979): "On Multivariate Risk Aversion", *Econometrica*, 47, 1391-1401.
- [63] Kimmel, R. (2002): "Modeling the Term Structure of Interest Rates : A New Approach", *Journal of Financial Economics*, forthcoming.
- [64] Langetieg, T. (1980): "A Multivariate Model of the Term Structure," *Journal of Finance*, 35, 71-97.
- [65] Leippold, M., and L., Wu (2001): "Design and Estimation of Quadratic Term Structure Models," DP, Zurich University.
- [66] Leippold, M., and L., Wu (2002): "Asset Pricing Under the Quadratic Class", *Journal of Financial and Quantitative Analysis*, 37, 271-295.
- [67] Levendorskiy, S. (2003): "Consistency conditions for affine term structure models", *Stochastic Processes and Their Applications*, forthcoming.
- [68] Longstaff, F. (1989): "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates", *Journal of Financial Economics*, 23, 195-224.
- [69] Longstaff, F., and E., Schwartz (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *Journal of Finance*, 47, 1259-1282.
- [70] Lu, B. (1999): "An Empirical Analysis of the Constantinides Model of the Term Structure," Working Paper, University of Michigan.
- [71] O'Connell, N. (2003): "Random Matrices, Non Colliding Particle System and Queues", *Seminaire de Probabilités, XXXVI, Lecture Notes in Maths.*, 1801, 165-182, Springer Verlag.
- [72] Oldfield, G., and R., Rogalski (1987): "The Stochastic Properties of Term Structure Movements," *Journal of Monetary Economics*, 19, 229-254.
- [73] Pearson, N., and T., Sun (1994): "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model," *Journal of Finance*, 49, 1279-1304.
- [74] Piazzesi, M. (2003): "Affine Term Structure Models", UCLA DP.
- [75] Rogers, L. (1997): "The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates", *Mathematical Finance*, 7, 157-176.
- [76] Singleton, K. (2001): "Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function," *Journal of Econometrics*, 102, 111-141.
- [77] Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, 177-188.
- [78] Yan, H. (2001): "Dynamic Models of the Term Structure," *Financial Analysts Journal*, July/August, 60-75.