Generalized Affine Models

Bruno Feunou[*]

Nour Meddahi[†]

Université de Montréal and CREST

Imperial College London

November 15, 2007

Abstract

Affine models are very popular in modeling financial time series as they allow for analytical calculation of prices of financial derivatives like treasury bonds and options. The main property of affine models is that the conditional cumulant function, defined as the logarithmic of the conditional characteristic function, is affine in the state variable. Consequently, an affine model is Markovian, like an autoregressive process, which is an empirical limitation. The paper generalizes affine models by adding in the current conditional cumulant function the past conditional cumulant function. Hence, generalized affine models are non-Markovian, such as ARMA and GARCH processes, allowing one to disentangle the short term and long-run dynamics of the process. Importantly, the new model keeps the tractability of prices of financial derivatives. This paper studies the statistical properties of the new model, derives its conditional and unconditional moments, as well as the conditional cumulant function of future aggregated values of the state variable which is critical for pricing financial derivatives. It derives the analytical formulas of the term structure of interest rates and option prices. Different estimating methods are discussed (MLE, QML, GMM, and characteristic function based estimation methods). Three empirical applications developed in companion papers are presented. The first one based on Feunou (2007) presents a no-arbitrage VARMA term structure model with macroeconomic variables and shows the empirical importance of the inclusion of the MA component. The second application based on Feunou and Meddahi (2007a) models jointly the high-frequency realized variance and the daily asset return and provides the term structure of risk measures such as the Value-at-Risk, which highlights the powerful use of generalized affine models. The third application based on Feunou, Christoffersen, Jacobs and Meddahi (2007) uses the model developed in Feunou and Meddahi (2007a) to price options theoretically and empirically.

Keywords: Affine models; cumulant function; option pricing; term structure of interest rates.

JEL codes: C12; C13; C14; C22; C51; E43; G12.

---

[*]Corresponding author. Email: bruno.feunou.kamkui@umontreal.ca.

[†]Email: n.meddahi@imperial.ac.uk.

1 Introduction

Affine models are often used to model the short term of interest rates because they lead to closed form of the bond prices and yields whatever the maturity. In addition, these yields are linear functions of the state variables, often the short term interest rate when one considers a one-factor model, which makes the pricing and the statistical inference simple. This approach has been introduced in continuous time by Vasicek (1977) where the short term interest rate is assumed to follow a Gaussian autoregressive process of order one and extended by Duffie and Kan (1997) to more non-negative models. Discrete time versions of affine models are studied theoretically in Darolles, Gourieroux, and Jasiak (2006) and Gourieroux, Monfort, and Polimenis (2002) among others while several papers, including Piazzesi (2005) and Ang and Piazzesi (2003), used them to characterize the term structure of interest rates and its interaction with macroeconomic variables; see Piazzesi (2003) for a survey on affine term structure models. Likewise, several authors used the affine processes for modeling the stochastic volatility of asset returns and characterized analytically the formulas of option prices; see Heston (1993) and Duffie, Pan and Singleton (2000) in continuous time and Heston and Nandi (2001) in discrete time.

A discrete time process x_t is called affine when its conditional cumulant function, denoted $\psi_t(u)$, and defined as the logarithmic of the moment generating function,[1] i.e.,

$$\psi_t(u) \equiv \log[E[\exp(ux_{t+1}) \mid x_\tau, \tau \leq t]],$$

is given by

$$\psi_t(u) = \omega(u) + \alpha(u)x_t. \tag{1.1}$$

Any autoregressive process of order one, AR(1), with i.i.d. innovations is affine. A consequence of (1.1) is that an affine process is Markovian, which could be a limitation for modeling some financial data. It is well known that financial data, like volatility of asset returns, exhibit serial correlation that the Markov ARCH models of Engle (1982) do not describe well, which led to the introduction of the GARCH models in Bollerslev (1986). Likewise, we do know that

---

[1]Instead of considering the moment generating function, one could use the characteristic function which exists for any random variable while the moment generating function does not exist for some random variables. The theory developed in this paper holds for characteristic functions. However, we decided to use the moment generating function for convenience and due to its familiarity with researchers in financial economics.

allowing for non-Markovian components in a model, like moving average (MA) components, allows one to disentangle the short-term and the long-run dynamics of the variable of interest, which could be important for some financial data like volatility of asset returns and short term of interest rates (Andersen and Lund (1997)).

Several generalizations of affine models have been introduced in order to include more memory in the basic model (1.1) and maintaining the tractability of affine models, i.e. by maintaining closed forms for the yields. Dai and Singleton (2003) and Dai, Singleton and Yang (2006) assumed that the coefficients that drive the affine model follow a Markov switching model. The authors show the empirical usefulness of this approach although filtering techniques are needed to price and estimate the model. Darolles, Gourieroux and Jasiak (2006) added lags of x_t in (1.1), i.e. they proposed an affine model of order $p > 1$. Monfort and Pegararo (2007a) successfully applied this approach to the term structure of interest rates, although one could need several lags leading to the estimation of many parameters. In a different paper, Monfort and Pegoraro (2007b) combined the two approaches describe above, i.e. they added lags and assumed that some parameters are driven by a Markov switching model. Again, such a method needs filtering techniques for both pricing and estimating the model.

In this paper, we follow a more traditional approach by including MA component in the model. The following example highlights our approach. Assume that the process x_t is an ARMA(1,1) given by

$$x_t = a + bx_{t-1} + \varepsilon_t - c\varepsilon_{t-1}, \quad \varepsilon_t \text{ i.i.d.}, \ |b| < 1, \ |c| < 1,$$

where the cumulant function of ε is denoted $\psi_\varepsilon(\cdot)$. One can show (see Section 2) that

$$\psi_t(u) = (ua + (1 - c)\psi_\varepsilon(u)) + u(b - c)x_t + c\psi_{t-1}(u),$$

which suggests the following extension of (1.1)

$$\psi_t(u) = \omega(u) + \alpha(u)x_t + \beta\psi_{t-1}(u). \tag{1.2}$$

One could view the new model as an extension of AR models to ARMA ones. Actually, it is much more than that because Eq. (1.2) implies that any power function of x_t is an ARMA process. This is the case because the conditional cumulant function of x_t is an auto-regression.

Our approach has several advantages. It involves fewer parameters than the approach in Darolles, Gourieroux and Jasiak (2006). The pricing and estimation procedures of the model are simpler than those of a model with Markov switching factors like Dai, Singleton and Yang

2

(2006). It also allows one to disentangle the short term dynamics of x_t from its long-run ones. When one considers an affine model (1.1), the function $\alpha(u)$ captures both short and long-run dynamics, which could be restrictive. We do know from the volatility literature that GARCH models allow for more persistence than ARCH models and that this is empirically important. Our empirical examples highlight this advantage.

Several dynamic term structure models with macroeconomic variables introduce latent variables in the affine state variable; see Ang and Piazzesi (2003). Such an approach is often used because current values of the macroeconomic variables do not fully explain the term structure of interest. However, it is always hard to understand what exactly these latent variables are. It is well known from the time series literature that AR models with latent variables, called structural models, imply reduced form ARMA representations for the observable variables. Consequently, one could interpret the new model as a reduced from of affine models with latent factors.

We introduce a slightly more general model than (1.2) by allowing the coefficient in front of $\psi_{t-1}(u)$ to be a function of u, i.e., we study the model defined by

$$\psi_t(u) = \omega(u) + \alpha(u)x_t + \beta(u)\psi_{t-1}(u). \tag{1.3}$$

This extension is theoretically important because Eq. (1.2) implies that the vector (x_t, m_t), where $m_t = E[x_{t+1} \mid x_\tau, \tau \le t]$, is a bivariate affine model while it is not the case for the model defined by Eq. (1.3). Likewise, we allowed for several lags of x_t and $\psi_{t-1}(u)$ in Eq. (1.3), i.e., we consider ARMA(p,q) type models.

The paper has several contributions. First, we study the statistical properties of the model and derive several conditional and unconditional moments and cumulants. We also derive the conditional cumulant function of the vector $(x_{t+1}, x_{t+2}, ..., x_{t+h})$. This function is critical when one wants to derive analytical formulas of yields and option prices. We then derive the Treasury yields when assuming that short term of interest rate is given by (1.3) under the risk neutral measure or the physical measure (the latter needs the specification of the price of the risk). Fiannly, we derive the formulas of options prices when assuming a stochastic volatility where the dynamics of the stochastic variance is given by Eq. (1.3).

One can use several methods to estimate the model. One could either characterize the likelihood of the model as in our empirical analysis, or follow Singleton (2001) by using the characteristic function of the process x_t and the instrumental variable approach of Hansen (1982). Actually, an efficient use of the whole characteristic function leads to an efficient estimation of the parameters comparable to the maximum likelihood estimators; see Carrasco

3

and Florens (2001, 2006) and Carrasco, Chernov, Florens, and Ghysels (2006). It would also be possible to use the conditional mean and variance of the process x_t combined with the Gaussian quasi-maximum likelihood approach to consistently estimate the parameters.

Three empirical applications developed in companion papers are presented. The first one based on Feunou (2007) presents a no-arbitrage VARMA term structure model with macroeconomic variables and shows the empirical importance of the inclusion of the MA component. The second application based on Feunou and Meddahi (2007a) models jointly the high-frequency realized variance and the daily asset return and provides the term structure of risk measures such as the Value-at-Risk, and highlights the powerful use of generalized affine models. The third application based on Feunou, Christoffersen, Jacobs and Meddahi (2007) uses the model developed in Feunou and Meddahi (2007a) to price options theoretically and empirically.

The rest of the paper is organized as follows. Section 2 provides the simple generalized affine model and provides its statistical properties. Section 3 provides the analytical formulas of the term structure of interest rates when the short term of interest rates is a generalized affine process under the physical or the risk-neutral measure. Likewise, section 3 provides the formulas of the option prices when the volatility of the stock returns is a generalized affine process under the physical or the risk-neutral measure. Section 4 provides three empirical examples while Section 5 concludes. Appendix A provides an example where the function $\beta(\cdot)$ is not constant. The proofs of Sections 2 and 3 are provided in Appendix B, while Appendix C provides the generalized affine model of higher order.

2   The Simple Generalized Affine Model

This section introduces and studies the simple model defined in the previous section.

Definition: Generalized Affine Process. A process x_t is called a generalized affine process of order (1,1) when the conditional cumulant function of x_{t+1} given its lagged values $x_t, x_{t-1}, ...,$ is characterized by

$$\psi_t(u) \equiv \log E[\exp(ux_{t+1}) \mid x_\tau, \tau \leq t] = \omega(u) + \alpha(u)x_t + \beta(u)\psi_{t-1}(u). \qquad (2.1)$$

4

2.1 Examples

Several well know examples in the time series and financial literatures are generalized affine. Obviously, affine models correspond to the case $\beta(u) = 0$. Other examples are given below.

2.1.1 Linear and Non-Linear ARMA(1,1) Models

Assume that x_t follows a linear ARMA(1,1) where the innovation process is i.i.d., i.e.

$$x_t = a + bx_{t-1} + \varepsilon_t - c\varepsilon_{t-1}, \quad \varepsilon_t \text{ i.i.d.}, \ |b| < 1, \ |c| < 1,$$

where the cumulant function of ε is denoted $\psi_\varepsilon(\cdot)$. Denote the conditional mean of x_{t+1} by m_t, i.e.,

$$m_t = a + bx_t - c\varepsilon_t.$$

Observe that

$$m_t = a + (b - c)x_t + cm_{t-1}.$$

Hence,

$$\psi_t(u) = \log E_t[\exp(ux_{t+1})] = um_t + \psi_\varepsilon(u) = u(a + (b - c)x_t) + \psi_\varepsilon(u) + ucm_{t-1}$$
$$= u(a + (b - c)x_t) + \psi_\varepsilon(u) + c(\psi_{t-1}(u) - \psi_\varepsilon(u))$$
$$= (ua + (1 - c)\psi_\varepsilon(u)) + u(b - c)x_t + c\psi_{t-1}(u),$$

i.e. any ARMA(1,1) process with i.i.d. innovations defined in (2.1.1) is a generalized affine process given in (2.1) where

$$\omega(u) = ua + (1 - c)\psi_\varepsilon(u), \quad \alpha(u) = u(b - c), \quad \beta(u) = c.$$

Let us now assume that the conditional mean of x_t is non-linear but still has an MA(1) structure, i.e.

$$x_t = f(x_{t-1}) + \varepsilon_t - c\varepsilon_{t-1}, \quad \varepsilon_t \text{ i.i.d.}, \ |c| < 1.$$

The condition mean of x_{t+1} denoted m_t is given by

$$m_t = f(x_t) - c\varepsilon_t = f(x_t) - cx_t + cm_{t-1}.$$

Hence,

$$\psi_t(u) = \log E_t[\exp(ux_{t+1})] = um_t + \psi_\varepsilon(u) = u(f(x_t) - cx_t) + \psi_\varepsilon(u) + ucm_{t-1}$$
$$= (1 - c)\psi_\varepsilon(u) + u(f(x_t) - cx_t) + c\psi_{t-1}(u).$$

Consequently, a non-linear ARMA(1) process with i.i.d. innovations is not a generalized affine process but belongs to the family defined by

$$\psi_t(u) = \omega(u) + \alpha(u, x_t) + \beta(u)\psi_{t-1}(u). \tag{2.2}$$

This family is studied in Feunou and Meddahi (2007b) and called generalized non-affine models.

2.1.2 GARCH(1,1) Type Models

We start the analysis by considering the model introduced in Bollerslev (1986), i.e.,

$$x_t = \mu + \varepsilon_t = \mu + \sqrt{h_{t-1}}z_t, \ z_t \text{ i.i.d. } \mathcal{N}(0,1), \ h_t = \omega + \alpha\varepsilon_t^2 + \beta h_{t-1},$$

with $\alpha \geq 0, \beta \geq 0, \alpha + \beta < 1$. By doing the same calculations as in the ARMA example, one gets

$$\psi_t(u) = \left((1-\beta)\mu u + \omega\frac{u^2}{2}\right) + \frac{1}{2}\alpha u^2(x_t - \mu)^2 + \beta\psi_{t-1}(u). \tag{2.3}$$

In other words, the GARCH(1,1) is not a generalized affine process as (2.1) but a generalized non-affine process given by (2.2).

It is well known that the GARCH(1,1) does not lead to closed forms of option prices. Heston and Nandi (2000) proposed a different specification for h_t that solved the problem where h_t is given by[2]

$$h_t = \omega + \alpha(z_t - \gamma\sqrt{h_{t-1}})^2 + \beta h_{t-1}.$$

Likewise, one can show that the conditional cumulant function of x_{t+1} is given by

$$\psi_t(u) = \left(u\mu(1-\beta-\alpha\gamma^2) + \frac{u^2}{2}(\omega + \alpha\frac{(x_t-\mu)^2}{h_{t-1}} - 2\gamma(x_t-\mu))\right) + (\beta+\alpha\gamma^2)\psi_{t-1}(u). \tag{2.4}$$

Consequently, the Heston and Nandi (2000) model is a generalized non-linear model defined by (2.2) where the function $\alpha(x_t, u)$ depends on x_t and h_{t-1}, i.e., the whole past of x_t.

Eq. (2.4) looks more non-linear than Eq. (2.3), which is puzzling given that the Heston and Nandi (2000) model leads to analytical formulas for option prices while the Bollerslev (1986) does not. As already mentioned, affine models lead to closed form of prices of derivatives. It turns out that the variance process h_t is affine when one considers the Heston and Nandi (2000) while it is not the case for the traditional GARCH model. More precisely, one has

$$\text{Heston \& Nandi}: \ \log E[\exp(uh_{t+1}) \mid h_\tau, \tau \leq t] = u\omega + \psi_{\chi^2(1)}(\alpha u) + ((\beta+\alpha\gamma^2)u - 2\gamma^2 u^2)h_t$$

$$\text{Bollerslev}: \ \log E[\exp(uh_{t+1}) \mid h_\tau, \tau \leq t] = u\omega + \psi_{\chi^2(1)}(\alpha u h_t) + \beta u h_t,$$

---

[2]There is an additional coefficient γ that appears in (2.1.2) which captures the leverage effect. One could easily add such term in Bollerslev's GARCH equation.

where $\psi_{\chi^2(1)}(\cdot)$ denotes the cumulant function of the $\chi^2(1)$ distribution. We will consider again the Heston and Nandi (2000) model when we will derive the option pricing formulas of generalized affine models.

2.1.3 ACD(1,1) type models

Engle and Russell (1997) introduced the autoregressive conditional duration (ACD) model where the duration x_i between two consecutive trades follows the process

$$x_i = \eta_{i-1}\, v_i, \ \ v_i \ \text{i.i.d.}, \ \ v_i > 0, \ \ E[v_i] = 1, \ \ \eta_i = \omega + \alpha x_i + \beta \eta_{i-1}.$$

If one assume that v_i follows an exponential distribution whose density function is $f_v(v) = \exp(-v)$, then one gets

$$\psi_i = E_i[\exp(ux_{i+1})] = \frac{1}{1 - u\eta_i}, \ \ u < \frac{1}{\eta_i},$$

which is not a generalized affine model. However, it is the case for the logarithmic duration model of Bauwens and Giot (2000) defined by

$$x_i = \exp(\eta_{i-1})\, v_i, \ \ v_i \ \text{i.i.d.}, \ \ v_i > 0, \ \ E[v_i] = 1, \ \ \log(\eta_i) = \omega + \alpha \log(x_i) + \beta \eta_{i-1}.$$

For this model, $\log(x_i)$ is an ARMA(1,1) and therefore a generalized affine process.

2.1.4 The Function $\beta(\cdot)$ is Varying

It is worth noting that in all the previous examples, the function $\beta(u)$ given in (2.1) does not depend on u. We provide in Appendix A an example where $\beta(\cdot)$ is varying.

2.2 Existence of Generalized Affine Models

Generalized affine models are defined recursively by their conditional cumulant function in (2.1). Therefore, one needs to show that the function $\psi_t(\cdot)$ in (2.1) is a proper cumulant function. The rest of this subsection focuses on the model where $\beta(\cdot)$ is constant, the other model being studied case by case as in the previous subsection.

The first important property of cumulant functions is that the sum of cumulant functions is a cumulant function. Consequently, when $\omega(u)$, $\alpha(u)x_t$, and $\beta(u)\psi_{t-1}(u)$ are cumulant functions, the function $\psi_t(u)$ defined in (2.1) is a cumulant function. Observe that often, as in

7

our empirical examples, $\omega(u) + \alpha(u)x_t$ is the cumulant function of an affine model. Therefore, the generalized affine model is well defined when $\beta(u)\psi_{t-1}(u)$ is a cumulant function.

The second important property of cumulant functions is related to infinitely divisible random variables. A random variable z whose cumulant function denoted $\psi_z(u)$, is called infinitely divisible when for any positive number c, $c\psi_z(u)$ is a cumulant function. Observe that a consequence of this definition is that $c\psi_z(u)$ is the cumulant function of an infinitely divisible random variable. Such variables appear in central limit theorems; examples of infinitely divisible random variables include normal, Poisson, and Gamma random variables. The first version of Darolles et al. (2006) provided sufficient conditions such that an affine process is infinitely divisible. In particular, popular affine models in finance, i.e. the Gaussian and the square root processes are infinitely divisible. This second property of cumulant functions is quite important for our purpose. By expanding recursively $\psi_t(u)$ given in (2.1), one gets

$$\psi_t(u) = \sum_{i=0}^{t-1} \beta^i(\omega(u) + \alpha(u)x_{t-i}) + \beta^t\psi_0(u)$$

where $\psi_0(u)$ is the unconditional cumulant function of x_0. Consequently, when $\beta > 0$ and $\omega(u) + \alpha(u)x_{t-i}$ is the cumulant function of an indivisible random variable (like some affine models derived in Darolles et al. (2006)), $\beta^i(\omega(u) + \alpha(u)x_{t-i})$ is a cumulant function of an infinitely divisible random variable. The definition of infinitely divisible random variables implies that the sum of infinitely divisible random variables is also an infinitely divisible random variable. Therefore, $\sum_{i=0}^{t} \beta^i(\omega(u) + \alpha(u)x_{t-i})$ is the cumulant function of an infinitely divisible random variable. Consequently, $\psi_t(u)$ is the cumulant function of an infinitely divisible random variable when one assumes that this is the case for $\psi_0(u)$. In other words, sufficient conditions to guarantee that $\psi_t(u)$ defined in (2.1) is a proper cumulant function are: $\beta \geq 0$, $\omega(u) + \alpha(u)(x)$ and $\psi_0(u)$ are cumulant functions of indivisible random variables.

A question not studied here is the existence of a stationary solution of (2.1). As usual, such a question is very difficult for discrete time non-linear models like GARCH models and it is left for future research. In the sequel of the paper, we assume such existence.

2.3   Cumulant and Moment Structures

In this section ww derive some moments and cumulants of the process.

2.3.1 Conditional Cumulants and Moments

Given that the process x_t is defined by its conditional cumulant function, it is more convenient to derive the conditional cumulant of x_{t+1} and then the conditional moments. The conditional cumulant of x_{t+1} of order n denoted $\kappa_{n,t}$, is given by

$$\kappa_{n,t} = \psi_t^{(n)}(0),$$

where $f^{(n)}(\cdot)$ denotes the n-th derivative function of a function $f(\cdot)$. We will also use the notation

$$\overline{\kappa_{n,t}} \equiv \left(\kappa_{1,t}, \kappa_{2,t}, ..., \kappa_{n,t}\right)^{\top}. \tag{2.5}$$

Proposition 2.1 *Let x_t be a generalized affine process defined by (2.1). Then,*

$$\kappa_{n,t} = \omega^{(n)}(0) + \alpha^{(n)}(0)x_t + \sum_{j=0}^{n-1} \begin{pmatrix} n \\ j \end{pmatrix} \beta^{(j)}(0)\kappa_{n-j,t-1} \tag{2.6}$$

and

$$\overline{\kappa}_{n,t} = \overline{\omega}_n + \overline{\alpha}_n x_t + \overline{\beta}_n \overline{\kappa}_{n,t-1} \tag{2.7}$$

where

$$\overline{\omega}_n = \begin{pmatrix} \omega^{(1)}(0) \\ \omega^{(2)}(0) \\ \vdots \\ \vdots \\ \omega^{(n)}(0) \end{pmatrix}, \ \overline{\alpha}_n = \begin{pmatrix} \alpha^{(1)}(0) \\ \alpha^{(2)}(0) \\ \vdots \\ \vdots \\ \alpha^{(n)}(0) \end{pmatrix}, \tag{2.8}$$

and

$$\overline{\beta}_n = \begin{bmatrix} \beta(0) & 0 & 0 & ... & 0 \\ \binom{2}{1}\beta^{(1)}(0) & \beta(0) & 0 & ... & 0 \\ \vdots & \vdots & \vdots & ... & \vdots \\ \vdots & \vdots & \vdots & ... & \vdots \\ \binom{n}{n-1}\beta^{(n-1)}(0) & \binom{n}{n-2}\beta^{(n-2)}(0) & ... & \binom{n}{1}\beta^{(1)}(0) & \beta(0) \end{bmatrix} \tag{2.9}$$

Consequently, the vector $\kappa_{n,t}$ is a VAR(1).

An important implication of the proposition is that any conditional cumulant $\kappa_{n,t}$ is a linear combination of x_t and its lagged values. This property is a characteristic of affine type models. One has different forms when one considers generalized non-affine models defined in (2.2). Another consequence of the VAR representation is that when $\beta(\cdot)$ is not constant, a cumulant

admits an ARMA representation of higher order. However, when $\beta(\cdot)$ is constant, one has a GARCH(1,1) type equation for $\kappa_{n,t}$

$$\kappa_{n,t} = \omega^{(n)}(0) + \alpha^{(n)}(0)x_t + \beta\kappa_{n,t-1}.$$

We will show below that x_t admits an ARMA(1,1) representation. Hence, $k_{n,t}$ admits an ARMA(2,1) representation.

There is a mapping between the cumulants and the moments of a random variable, which allows one to derive the conditional moments of x_{t+1} from its conditional cumulants. More precisely, by denoting the conditional moments by $m_{n,t}$, i.e.,

$$m_{n,t} = E[x_{t+1}^n \mid x_\tau, \tau \leq t],$$

one has

$$m_{1,t} = \kappa_{1,t}$$

$$m_{2,t} = \kappa_{2,t} + \kappa_{1,t}^2$$

$$m_{3,t} = \kappa_{3,t} + 3\kappa_{2,t}\kappa_{1,t} + \kappa_{1,t}^3$$

$$m_{4,t} = \kappa_{4,t} + 4\kappa_{3,t}\kappa_{1,t} + 3\kappa_{2,t}^2 + 6\kappa_{2,t}\kappa_{1,t}^2 + \kappa_{1,t}^4$$

$$m_{5,t} = \kappa_{5,t} + 5\kappa_{4,t}\kappa_{1,t} + 10\kappa_{3,t}\kappa_{2,t} + 10\kappa_{3,t}\kappa_{1,t}^2 + 15\kappa_{2,t}^2\kappa_{1,t} + 10\kappa_{2,t}\kappa_{1,t}^3 + \kappa_{1,t}^5$$

$$m_{6,t} = \kappa_{6,t} + 6\kappa_{5,t}\kappa_{1,t} + 15\kappa_{4,t}\kappa_{2,t} + 15\kappa_{4,t}\kappa_{1,t}^2 + 10\kappa_{3,t}^2 + 60\kappa_{3,t}\kappa_{2,t}\kappa_{1,t}$$
$$+ 20\kappa_{3,t}\kappa_{1,t}^3 + 15\kappa_{2,t}^3 + 45\kappa_{2,t}^2\kappa_{1,t}^2 + 15\kappa_{2,t}\kappa_{1,t}^4 + \kappa_{1,t}^6.$$

Therefore, by using the results of Proposition 2.1, one gets the conditional moments of x_{t+1}.

2.3.2 Unconditional first and second moments

As in affine models, we can compute unconditional moments which are useful to understand the dynamics of the model and to estimate it. We start by focusing on the covariance structure of the process x_t which will allow us to show that x_t is an ARMA(1,1) (with possibly heteroskedastic innovations).

Proposition 2.2 *Let x_t be a generalized affine process of order (1,1). Then,*

$$E\left[x_t\right] = \frac{\omega^{(1)}(u)}{1 - \left(\alpha^{(1)}(u) + \beta(0)\right)}$$

$$Var\left(x_t\right) = \left(1 + \frac{\alpha^{(1)}(0)^2}{1 - \left(\alpha^{(1)}(0) + \beta(0)\right)^2}\right) \times \left(\frac{\alpha^{(2)}(0) + 2\beta^{(1)}(0)}{1 - \beta(0)}E\left[x_t\right] + \frac{\omega^{(2)}(0)}{1 - \beta(0)}\right)$$

10

$$E\left(x_t x_{t+h}\right) = \omega^{(1)}\left(0\right) E\left[x_t\right] + \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right) E\left(x_t x_{t+h-1}\right).$$

Corollary 1

$$Cov\left(x_t, x_{t+h}\right) = \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1} cov\left(x_t, x_{t+1}\right).$$

Hence, x_t is an ARMA(1,1) whose autoregressive root equals $\alpha^{(1)}(0) + \beta(0)$. In addition, one has:

Proposition 2.3

$$corr\left(x_t, x_{t+1}\right) = \frac{cov\left(x_t, x_{t+1}\right)}{Var\left(x_t\right)} = \alpha^{(1)}\left(0\right)\left[1 + \frac{\alpha^{(1)}\left(0\right)\beta\left(0\right)}{1 - \left(\beta\left(0\right)^2 + 2\alpha^{(1)}\left(0\right)\beta\left(0\right)\right)}\right].$$

2.3.3 Unconditional moments of cumulants

Given that we know all the dynamics of the generalized affine process, it is of interest to study the dynamics of the higher moments like $cov(x_t^n, x_{t+h}^m)$ which will be useful for estimation purpose. Since $cov(x_t^n, x_{t+h}^m)$ is related to $cov(\kappa_{n,t}, \kappa_{m,t+h})$, we need first to compute unconditional first and second moments of cumulant.

By using the vector of cumulant $\overline{\kappa}_{n,t}$ given by equation (2.7), we deduce the following unconditional mean of cumulant.

$$E\left[\overline{\kappa}_{n,t}\right] \equiv \overline{\kappa}_n = \left(I - \overline{\beta}_n\right)^{-1}\left(\overline{\omega}_n + E\left(x_t\right)\overline{\alpha}_n\right).$$

Hence, one has

Proposition 2.4

$$E\left[\kappa_{1,t}\overline{\kappa}_{n,t}\right] = \left(I - \overline{\beta}_n\right)^{-1}\left[E\left(x_t\right)\overline{\omega}_n + \left(cov\left(X_t, X_{t+1}\right) - E\left(x_t\right)^2\right)\overline{\alpha}_n\right]$$

$$cov\left(\overline{\kappa}_{n,t}\ \kappa_{1,t}\right) = V\left(x_t\right)\left(I - \overline{\beta}_n\right)^{-1}\overline{\alpha}_n.$$

Consequently, one gets

Proposition 2.5

$$E\left[\kappa_{1,t+h}\overline{\kappa}_{n,t}\right] = \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^h E\left[\kappa_{1,t}\overline{\kappa}_{n,t}\right] + \left[1 - \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^h\right] E\left(x_t\right) E\left[\overline{\kappa}_{n,t}\right]$$

$$cov\left(\overline{\kappa}_{n,t}\ \kappa_{1,t+h}\right) = \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^h cov\left(\overline{\kappa}_{n,t}\ \kappa_{1,t}\right).$$

11

One needs to characterize the variance matrix of the vector of cumulants $\overline{\kappa}_{n,t}$, which is the goal of the following proposition.

Proposition 2.6 $V\left(\overline{\kappa}_{n,t}\right)$ *is the solution to the following quaternion Matrix equation*

$$V\left(\overline{\kappa}_{n,t}\right) - \overline{\beta}_n V\left(\overline{\kappa}_{n,t}\right) \overline{\beta}_n^\top = \theta_n$$

where

$$\theta_n = V\left(x_t\right)\left[\overline{\alpha}_n \overline{\alpha}_n^\top + \overline{\alpha}_n \overline{\alpha}_n^\top \left(I - \overline{\beta}_n^\top\right)^{-1} \overline{\beta}_n^\top + \overline{\beta}_n \left(I - \overline{\beta}_n\right)^{-1} \overline{\alpha}_n \overline{\alpha}_n^\top\right].$$

By using the formula

$$vec\left(AXB\right) = \left(B^\top \otimes A\right) vec\left(X\right),$$

one gets the unconditional variance-covariance matrix of of $\overline{\kappa}_{n,t}$:

$$vec\left[V\left(\overline{\kappa}_{n,t}\right)\right] = \left(I - \overline{\beta}_n \otimes \overline{\beta}_n\right)^{-1} vec\left(\theta_n\right).$$

Consequently, one gets the covariance matrix between $\overline{\kappa}_{n,t}$ and $\overline{\kappa}_{n,t+h}$:

Proposition 2.7

$$
\begin{aligned}
Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h}\right] &= Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h-1}\right]\overline{\beta}_n^\top + cov\left[\overline{\kappa}_{n,t} \, \kappa_{1,t+h-1}\right]\overline{\alpha}_n^\top \\
&= Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h-1}\right]\overline{\beta}_n^\top + V\left(x_t\right)\left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1}\left(I - \overline{\beta}_n\right)^{-1}\overline{\alpha}_n \overline{\alpha}_n^\top \\
&= V\left(\overline{\kappa}_{n,t}\right)\left(\overline{\beta}_n^\top\right)^h \\
&+ V\left(x_t\right)\left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1}\left(I - \overline{\beta}_n\right)^{-1}\overline{\alpha}_n \overline{\alpha}_n^\top \left(I - \overline{\beta}_n^\top\right)^{-1}\left[I - \left(\overline{\beta}_n^\top\right)^h\right].
\end{aligned}
$$

2.3.4 Higher order covariance

In this subsection, we use the results of the previous subsection to characterize the covariance structure of the third and fourth moments. We begin by deriving the third moments.

Proposition 2.8 *One has*

$$Cov\left(x_t^2,\ x_{t+h}\right) = \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^h Cov\left(x_t^2,\ x_t\right).$$

We need to compute $Cov[x_t^2, x_t]$. For this purpose, we will use $E\left[\kappa_{1,t}^3\right]$ who is given by

$$
\begin{aligned}
E\left[\kappa_{1,t}^3\right]\left(1-\left(\alpha^{(1)}(0)+\beta(0)\right)^3\right) &= \left[\omega^{(1)}(0)\right]^3+\left[\alpha^{(1)}(0)\right]^3 E\left[\kappa_{3,t}\right] \\
&\quad +3\left[\alpha^{(1)}(0)\right]^2\left(\beta(0)+\alpha^{(1)}(0)\right) E\left[\kappa_{1,t}\kappa_{2,t}\right] \\
&\quad +3\left[\omega^{(1)}(0)\right]^2\left(\beta(0)+\alpha^{(1)}(0)\right) E\left[x_t\right] \\
&\quad +3\omega^{(1)}(0)\beta(0)^2 E\left(\kappa_{1,t}^2\right)+3\omega^{(1)}(0)\left[\alpha^{(1)}(0)\right]^2 E\left[x_t^2\right].
\end{aligned}
$$

All the terms in the right hand side have already been computed in section 2.3.3. Consequently, one can deduce $E\left[x_t^3\right]$ and $cov\left(x_t^2,\ x_t\right)$ from

$$
E\left[x_t^3\right]=E\left[\kappa_{3,t}\right]+3E\left[\kappa_{1,t}\kappa_{2,t}\right]+E\left[\kappa_{1,t}^3\right] \quad\text{and}\quad cov\left(x_t^2,\ x_t\right)=E\left[x_t^3\right]+E\left[x_t^2\right]E\left[x_t\right]
$$

Likewise, $cov\left(x_t^2,\ \kappa_{2,t+h}\right)$ and $cov\left(x_t^2,\ \kappa_{1,t+h}^2\right)$ can be computed recursively as follows:

$$
cov\left(x_t^2,\ \kappa_{2,t+h}\right)=\beta(0)\,cov\left(x_t^2,\ \kappa_{2,t+h-1}\right)+\left(\alpha^{(2)}(0)+2\beta^{(1)}(0)\right)Cov\left(x_t^2,\ x_{t+h}\right),
$$

$$
\begin{aligned}
cov\left(x_t^2,\ \kappa_{1,t+h}^2\right) &= \left(\beta(0)+\alpha^{(1)}(0)\right)^2 cov\left(x_t^2,\ \kappa_{1,t+h-1}^2\right)+\left[\alpha^{(1)}(0)\right]^2 cov\left(x_t^2,\ \kappa_{2,t+h-1}\right) \\
&\quad +2\omega^{(1)}(0)\left(\beta(0)+\alpha^{(1)}(0)\right)Cov\left(x_t^2,\ x_{t+h}\right).
\end{aligned}
$$

We then get $cov\left(x_t^2,\ x_{t+h}^2\right)$ from

$$
cov\left(x_t^2,\ x_{t+h}^2\right)=cov\left(x_t^2,\ \kappa_{1,t+h-1}^2\right)+cov\left(x_t^2,\ \kappa_{2,t+h-1}\right).
$$

2.4 Cumulant Function of Future Aggregated Returns

An important formula used in the analytical calculation of the term structure of interest rates and option prices is the conditional distribution function of $\sum_{i=1}^h a_i x_{t+i}$ for given real numbers a_i. Affine models allow one to derive the conditional cumulant function of $(x_{t+1}, x_{t+2}, ..., x_{t+h})$ and consequently the one of $\sum_{i=1}^n a_i x_{t+i}$. It turns out that this is the case for generalized affine models.

Proposition 2.9 *The conditional cumulant function of $(x_{t+1}, x_{t+2}, ..., x_{t+h})$ is given by*

$$
\log E_t\left[\exp\left(\sum_{i=1}^h u_i x_{t+i}\right)\right]=\sum_{k=1}^h\left\{\beta(d_k)^{k-1}\Psi_t(d_k)+\frac{1-\beta(d_k)^{k-1}}{1-\beta(d_k)}\omega(d_k)\right\} \qquad (2.10)
$$

where the sequence $(d_k)_{1 \le k \le h}$ *is defined as follows:*

$$d_k = u_k + \sum_{j=k}^{h-1} \beta \, (d_{j+1})^{j-k} \, \alpha \, (d_{j+1}) \ \ for \ k \le h-1, \ \ d_h = u_h. \tag{2.11}$$

We will often use (2.10) in the next section when we derive yields and option prices.

3 Analytical Formulas of Prices of Financial Derivatives

This section analytically characterizes the yields and option prices when one assumes a generalized affine model for the interest rate and the stochastic volatility respectively. For each of them, we follow two approaches. First assume the affine model under the physical measure (P-measure) and specify the price of risk and then derive the price of the financial derivatives (bonds or options). The second approach consists on specifying the affine model under the risk neutral measure (Q-measure) and then derive the prices of the financial derivatives. We start the analysis by studying the term structure model

3.1 The Term Structure of Interest Rates

3.1.1 Generalized Affine Model Under the P-Measure

We assume that under the P-measure, the short term of interest rate denoted r_t follows a generalized affine process given in (2.1), i.e.,

$$\log E_t^P[\exp(ur_{t+1})] \equiv \psi_t^P(u) = \omega^P(u) + \alpha^P(u)r_t + \beta^P(u)\psi_{t-1}^P(u).$$

When $\beta^P(u) = 0$ one gets affine models like those of Vasicek (1977) and Duffie and Kan (1996) who derived the term structure of interest rates.

In order to derive the dynamics of r_t under the Q-measure and the yield curve, one needs to specify the stochastic discount factor denoted here $M_{t,t+1}$ or the price of risk. We follow the general approach of Gourieroux and Monfort (2007) who proposed the following formulation:

$$M_{t,t+1} = \exp(\gamma r_{t+1} + \theta_t). \tag{3.1}$$

Given the restriction

$$\exp(-r_t) = E_t^P[M_{t,t+1}], \tag{3.2}$$

one gets

$$\theta_t = -r_t - \psi_t(\gamma) \quad \text{and} \quad M_{t,t+1} = \exp(\gamma r_{t+1} - r_t - \psi_t(\gamma)). \tag{3.3}$$

In the sequel, we define $B(t,h)$ and $r_{t,h}$ as

$$B(t,h) = E_t^P\left[\prod_{i=1}^h M_{t+i-1,t+i}\right], \quad r_{t,h} = -\frac{\log(B(t,h))}{h}. \tag{3.4}$$

We are now able to derive the term structure of interest rates, i.e., the formula of $r_{t,h}$ when h varies.

Proposition 3.1

$$
\begin{aligned}
r_{t,t+h} &= \frac{1 - \beta^P(\gamma)^h}{1 - \beta^P(\gamma)}\frac{\Psi_t^P(\gamma)}{h} + \frac{1}{h}\frac{\omega^P(\gamma)}{1 - \beta^P(\gamma)}\left(h - \frac{1 - \beta^P(\gamma)^h}{1 - \beta^P(\gamma)}\right)\\
&\quad + \frac{r_t}{h} - \sum_{k=1}^h \beta^P(d_k)^{k-1}\frac{\Psi_t^P(d_k)}{h} - \sum_{k=1}^h \frac{1 - \beta^P(d_k)^{k-1}}{1 - \beta^P(d_k)}\frac{\omega^P(d_k)}{h}
\end{aligned}
$$

with

$$d_k = u_k + \sum_{j=k}^{h-1}\beta^P(d_{j+1})^{j-k}\alpha^P(d_{j+1}) \text{ for } k \leq h-1, \ d_h = u_h$$

where

$$u_1 = \gamma - \alpha^P(\gamma)\frac{1 - \beta^P(\gamma)^{h-1}}{1 - \beta^P(\gamma)}, \quad u_h = \gamma, \text{ and } u_j = \gamma - 1 - \alpha^P(\gamma)\frac{1 - \beta^P(\gamma)^{h-j}}{1 - \beta^P(\gamma)} \text{ for } 1 < j < h.$$

A major difference of the yield curve between affine and generalized affine models is the introduction of terms like $\psi_t^P(\gamma)$ and $\psi_t^P(d_k)$. These terms imply that the whole past of r_t influences the term structure of interest rates given that $\psi_t^P(u)$ is a function of r_t and its past.

Another question of interest is the characterization of the dynamics of r_t under the Q-measure. We denote $\psi_t^Q(u)$ the conditional cumulant function of r_{t+1} under the Q-measure, i.e.,

$$\psi_t^Q(u) \equiv \log E_t^Q[\exp(ur_{t+1})]. \tag{3.5}$$

Proposition 3.2 Dynamics of r_t under the Q-measure. *One has*

$$\psi_t^Q(u) = \psi_t^P(u + \gamma) - \psi_t^P(\gamma). \tag{3.6}$$

Hence,

$$\psi_t^Q(u) = \omega^Q(u) + \alpha^Q(u)r_t + \beta^Q(u)\psi_{t-1}^Q(u) + [\beta(u+\gamma) - \beta(u)]\psi_t(u), \tag{3.7}$$

15

where

$$\omega^Q(u) = \omega^P(u+\gamma) - \omega^P(\gamma), \ \alpha^Q(u) = \alpha^P(u+\gamma) - \alpha^P(\gamma), \ \beta^Q(u) = \beta^P(u+\gamma). \quad (3.8)$$

Eq. (3.6) is model free, i.e. it does not depend on our generalized affine specification. In particular, the same equation appears in affine models; see Gourieroux and Monfort (2007) and Monfort and Pegoraro (2006a). An additional term appears in (3.7) which vanishes when $\beta(\cdot)$ is constant, as in our empirical examples. When this term does not vanish, the short term of interest rate is not a generalized affine under the Q-measure. However, the following proposition characterizes the conditional cumulant of $(r_{t+1}, \psi_{t+1}^P(\gamma))$ which will allow us to understand the dynamics of r_{t+1} under the Q-measure. In the sequel, $\psi_{r,\psi(\gamma),t}^Q(u,v)$ denotes the conditional cumulant function of $(r_{t+1}, \psi_{t+1}^P(\gamma))$ under the Q-measure.

Proposition 3.3

$$\psi_{r,\psi(\gamma),t}^Q(u,v) = \omega_1^Q(u,v) + \left(\alpha_1^Q(u,v)\, r_t + \alpha_2^Q(u,v)\, \psi_t^p(\gamma)\right) + \beta_1^Q(u,v)\, \psi_{r,\psi(\gamma),t-1}^Q(u,v)$$
$$- \alpha_2^Q(u,v)\beta_1^Q(u,v)\psi_{t-1}^Q(\gamma)$$

$$(3.9)$$

where

$$\omega_1^Q(u,v) = v\omega^P(\gamma)\left[1 - \beta^P\left(u + v\alpha^P(\gamma) + \gamma\right)\right] + \omega^P\left(u + v\alpha^P(\gamma) + \gamma\right)$$
$$\alpha_1^Q(u,v) = \alpha^P\left(u + v\alpha^P(\gamma) + \gamma\right), \quad \alpha_2^Q(u,v) = v\beta^P(\gamma) - 1$$
$$\beta_1^Q(u,v) = \beta^P\left(u + v\alpha^P(\gamma) + \gamma\right).$$

While the definition of generalized affine models (2.1) is given for univariate processes and of order (1,1), the extensions to multivariate and higher order is not very difficult (see Appendix C); Eq. (3.9) means that the bivariate vector $(r_{t+1}, \psi_{t+1}^P(\gamma))$ is a generalized affine process of order (2,1). Consequently, one can characterize formulas of financial derivatives, including yields, by using the generalized affine dynamics of $(r_{t+1}, \psi_{t+1}^P(\gamma))$ under the Q-measure.

3.1.2 Generalized Affine Model Under the Q-Measure

We now assume that the short term of interest rate r_t follows a generalized affine process given in (2.1) under the Q-measure, i.e.

$$\log E_t^Q[\exp(ur_{t+1})] \equiv \psi_t^Q(u) = \omega^Q(u) + \alpha^Q(u)r_t + \beta^Q(u)\psi_{t-1}^Q(u).$$

The following proposition provides the formula of the yield curve.

Proposition 3.4 *The yield at horizon h is*

$$r_{t,t+h} = \frac{r_t}{h} - \frac{1}{h} \sum_{k=1}^{h-1} \left\{ \beta^Q \left(d_k\right)^{k-1} \Psi_t^Q \left(d_k\right) + \frac{1 - \beta^Q \left(d_k\right)^{k-1}}{1 - \beta^Q \left(d_k\right)} \omega^Q \left(d_k\right) \right\} \tag{3.10}$$

where

$$d_k = -1 + \sum_{j=k}^{h-2} \beta^Q \left(d_{j+1}\right)^{j-k} \alpha^Q \left(d_{j+1}\right) \text{ for } k \leq h - 2, \ d_{h-1} = -1. \tag{3.11}$$

On could also characterize the dynamics of r_t under the P-measure if one assumes a stochastic discount factor. We assume again that the stochastic discount factor is given by[3] (3.1). Hence, one gets

$$\psi_t^p(u) = -r_t - \theta \psi_t^Q(u - \gamma) - \psi_t^Q(-\gamma),$$

which leads to

$$\psi_t^p(u) = \psi_t^Q(u - \gamma) - \psi_t^Q(-\gamma). \tag{3.12}$$

Again, this equation is model free and appears in affine models (Gourieroux and Monfort (2007), Monfort and Pegoraro (2006a)). Likewise, r_{t+1} is not a generalized affine process under the P-measure. However, the vector $(r_{t+1}, \psi_{t+1}(-\gamma))$ is a generalized affine process of order $(2,1)$ as shown in the following proposition. In the sequel, $\psi_{r,\psi(-\gamma),t}^P$ denotes the conditional cumulant function of $(r_{t+1}, \psi_{t+1}(-\gamma))$ under the P-measure:

Proposition 3.5

$$\psi_{r,\psi(-\gamma),t}^P (u, v) = \omega_1^P (u, v) + \left(\alpha_1^P (u, v) r_t + \alpha_2^P (u, v) \psi_t^p (-\gamma) \right) + \beta_1^P (u, v) \psi_{r,\psi(-\gamma),t-1}^Q (u, v)$$
$$- \alpha_2^Q(u,v)\beta_1^P(u,v)\psi_{t-1}^P(-\gamma)$$

$$\tag{3.13}$$

where

$$\begin{aligned}
\omega_1^P (u, v) &= v\omega^Q (-\gamma) \left[1 - \beta^Q \left(u + v\alpha^Q (-\gamma) - \gamma\right)\right] + \omega^Q \left(u + v\alpha^Q (-\gamma) - \gamma\right) \\
\alpha_1^P (u, v) &= \alpha^Q \left(u + v\alpha^Q (-\gamma) - \gamma\right), \quad \alpha_2^P (u, v) = v\beta^Q (-\gamma) - 1 \\
\beta_1^P (u, v) &= \beta^Q \left(u + v\alpha^Q (-\gamma) - \gamma\right).
\end{aligned}$$

[3]Observe that when one specifies the dynamics of r_t under the Q-measure as a generalized affine process, one could allow γ in (3.1) to be time-varying and adapted to the information available at time t. A consequence is that the short term of interest rate will not be a generalized affine process under the P-measure; see Gourieroux and Monfort (2007) for the same discussion about affine models.

3.2 Option Pricing

We now consider models of stock returns where we assume that the conditional variance of the returns is time-varying and is generalized affine. In what follows r_t denotes the log-returns of the stock price, i.e.

$$r_{t+1} = \ln\left(\frac{S_{t+1}}{S_t}\right).$$

The key approach behind the analytical calculations of Heston (1993), Duffie, Pan, and Singleton (2000), and Heston and Nandi (2000) is the possibility to write the joint process (r_t, h_t) as an affine process, where h_{t+1} is the conditional variance of r_{t+1} given an information set that contains r_t and its lagged values and possibly another variable, latent or not, like in stochastic volatility models. In what follows, we will allow for both cases. We will write the joint model of (r_{t+1}, h_{t+1}). The variable h_{t+1} could be the conditional variance of r_{t+2} given $\{r_\tau, h_\tau, \tau \leq t\}$ (including GARCH type models). The variable h_{t+1} could be an observable variable like in our empirical example where it equals the high frequency realized volatility. In the rest of this section the information set I_t is the sigma algebra generated by $\{r_\tau, h_\tau, \tau \leq t\}$. The conditional expectation operator $E[\cdot \mid I_t]$ will be denoted $E_t[\cdot]$.

3.2.1 Generalized Affine process under the P-Measure

We denote the conditional cumulant function of (r_{t+1}, h_{t+1}) under the P measure by $\psi_t^P(u,v)$:

$$\psi_t^P(u,v) = \log E_t^P[\exp(ur_{t+1} + vh_{t+1})] = \omega^P(u,v) + \alpha^P(u,v)h_t + \beta^P(u,v)\psi_{t-1}^P(u,v).$$

When one assumes that h_t is exactly the conditional variance of r_{t+1}, one needs to impose the following restrictions on the cumulant function in order to guarantee this assumption:

$$\frac{\partial^2 \omega^P}{\partial u^2}(0,0) = 0, \ \frac{\partial^2 \alpha^P}{\partial u^2}(0,0) = 1, \ \frac{\partial^2 \beta^P}{\partial u^2}(0,0) = 0 \tag{3.14}$$

which implies

$$\frac{\partial^2 \psi_t}{\partial u^2}(0,0) = Var_t^P[r_{t+1}] = h_t.$$

We denote by r the short term interest rate supposed constant for simplicity. We consider the following stochastic discount factor

$$M_{t,t+1} = \exp(\gamma r_{t+1} + \lambda h_{t+1} + \theta_t). \tag{3.15}$$

Observe that both Heston and Nandi (2000) and Christoffersen et al. (2006) assumed that $\lambda = 0$. There is no theoretical foundation for such assumption other than simplicity. In other words, we allow the volatility to be priced.

In addition, one needs to impose restrictions in order to guarantee that $M_{t,t+1}$ is a stochastic discount factor, which implies that prices under the Q-measure are martingales. This is the purpose of the following proposition.

Proposition 3.6 *The parameters γ and λ are restricted by the following system of equations*

$$
\begin{aligned}
\omega\left(1+\gamma,\lambda\right) - \omega\left(\gamma,\lambda\right) &= r\left(1 - \beta\left(\gamma,\lambda\right)\right) \\
\alpha\left(1+\gamma,\lambda\right) &= \alpha\left(\gamma,\lambda\right) \\
\beta\left(1+\gamma,\lambda\right) &= \beta\left(\gamma,\lambda\right).
\end{aligned}
$$

Observe that when $\beta(\cdot)$ is a constant function, the third equation in the previous system holds, which leads to a fully identified system.

We will now characterize the dynamics of (r_{t+1}, h_{t+1}) under the Q-measure by deriving its conditional cumulant function denoted $\psi_t^Q(u,v)$.

Proposition 3.7 *We have*

$$
\Psi_t^Q(u,v) = \Psi_t^P(u+\gamma, v+\lambda) - \Psi_t^P(\gamma,\lambda) \tag{3.16}
$$

and

$$
\begin{aligned}
\Psi_t^Q(u,v) = {}& (\omega^p(u+\gamma, v+\lambda) - \omega^p(\gamma,\lambda)) \\
&+ (\alpha^p(u+\gamma, v+\lambda) - \alpha^p(\gamma,\lambda))h_t + \beta(u+\gamma, v+\lambda)\Psi_{t-1}^Q(u,v) \\
&+ (\beta(u+\gamma, v+\lambda) - \beta(\gamma,\lambda))\psi_{t-1}(\gamma,\lambda).
\end{aligned} \tag{3.17}
$$

Several remarks are in order. When one assumes that h_t is the conditional variance of r_{t+1} under the P-measure, i.e. under (3.14), Eq. (3.17) implies that h_t is also the conditional variance of r_{t+1} under the Q measure. In other words, one keeps (3.17) under the Q-measure. In addition, as for the term structure of interest rates, an additional term appears in (3.17), implying that the process (r_{t+1}, h_{t+1}) is not generalized affine process under the Q-measure. Likewise, this additional term vanishes when the function $\beta(\cdot)$ is constant. Again, one can still prove that a particular vector is a generalized affine model of higher order, which will allow us to derive option prices. More precisely, one can show that the vector $(r_{t+1}, h_{t+1}, \psi_{t+1}(\gamma,\lambda))$ is a generalized affine process of order.

We now provide the formula of the option prices.

Proposition 3.8 *The price at time t of a European call option with payoff $(S_{t+h} - X)^+$ at time $t + h$ is given by*

$$C_t = \exp(-rh)S_t C_{1,t} - \exp(-rh)X C_{2,t} \qquad (3.18)$$

where

$$C_{1,t} = \frac{\exp(rh)}{2} + \int_0^{+\infty} \frac{1}{\pi u} Im \left[\exp\left(\Psi_{t,t+h}^Q (1 + iu) - iu \ln\left(\frac{X}{S_t}\right) \right) \right] du$$

$$C_{2,t} = \frac{1}{2} + \int_0^{+\infty} \frac{1}{\pi u} Im \left[\exp\left(-iu \ln\left(\frac{X}{S_t}\right) + \Psi_{t,t+h}^Q (iu) \right) \right] du$$

and

$$\Psi_{t,t+h}^Q (u) = -\Psi_t(\gamma, \lambda) \frac{1 - \beta(\gamma, \lambda)^h}{1 - \beta(\gamma, \lambda)} - \frac{\omega(\gamma, \lambda)}{1 - \beta(\gamma, \lambda)} \left[h - \frac{1 - \beta(\gamma, \lambda)^h}{1 - \beta(\gamma, \lambda)} \right]$$

$$+ \sum_{k=1}^{h} \left\{ \beta(d_k)^{k-1} \Psi_t(d_k) + \frac{1 - \beta(d_k)^{k-1}}{1 - \beta(d_k)} \alpha(d_k) \right\}$$

with

$$d_k = (u + u_k, v_k) + \sum_{j=k}^{h-1} \beta(d_{j+1})^{j-k} (0, \alpha(d_{j+1})) \ \text{for } k \le h - 1 \qquad (3.19)$$

$$d_h = (u + u_h, v_h)$$

and

$$u_h = \gamma, \ v_h = \lambda$$

$$u_j = \gamma - \alpha(\gamma, \lambda) \frac{1 - \beta(\gamma, \lambda)^{h-j}}{1 - \beta(\gamma, \lambda)} \ \text{for } 1 \le j < h$$

$$v_j = \lambda - \alpha(\gamma, \lambda) \frac{1 - \beta(\gamma, \lambda)^{h-j}}{1 - \beta(\gamma, \lambda)} \ \text{for } 1 \le j < h.$$

This proposition uses Fourier transforms, which is a traditional approach in affine models. It is important to notice that, for this purpose, we had to use the logarithmic of the characteristic function instead of the logarithmic of the moment generating function. A simple modification of the notation is sufficient to make this change.

3.3 Generalized Affine process under the Q-Measure

This subsection specifies the dynamics of (r_{t+1}, h_{t+1}) under the Q-measure, $\Psi_t^Q(u, v)$, and derives the option prices. We assume that

$$\Psi_{t+1}^Q (u, v) = \omega(u, v) + \alpha(u, v) h_{t+1} + \beta(u, v) \Psi_t^Q(u, v). \qquad (3.20)$$

A well defined risk-neutral distribution for log-returns must satisfy

$$\exp(r) = E^Q\left[\exp(r_{t+1})\,|I_t\right]$$

where r is the risk-free rate. Thus $\Psi_t^Q(1,0)$ must satisfy

$$\Psi_t^Q(1,0) = r.$$

Proposition 3.9 *Eq. (3.20) is a valid risk-neutral model if and only if*

$$\frac{\omega(1,0)}{1 - \beta(1,0)} = r \tag{3.21}$$
$$\alpha(1,0) = 0.$$

The result is an implication of the following representation of expression of the model:

$$\Psi_{t+1}^Q(u,v) = \frac{\omega(u,v)}{1 - \beta(u,v)} + \alpha(u,v)\sum_{i=0}^{\infty}\beta(u,v)^i\,h_{t-i+1}.$$

We are now able to characterize the option prices.

Proposition 3.10 *When (3.21) holds, the price at time t of European call option with payoff $(S_{t+h} - X)^+$ at time $t+h$:*

$$C_t = \exp(-rh)C_{1,t} - \exp(-rh)XC_{2,t}$$

where

$$C_{1,t} = \frac{\exp(rh)}{2} + \int_0^{+\infty}\frac{1}{\pi u}Im\left[\exp\left(\Psi_{t,t+h}^Q(1+iu) - iu\ln\left(\frac{X}{S_t}\right)\right)\right]du$$

$$C_{2,t} = \frac{1}{2} + \int_0^{+\infty}\frac{1}{\pi u}Im\left[\exp\left(-iu\ln\left(\frac{X}{S_t}\right) + \Psi_{t,t+h}^Q(iu)\right)\right]du$$

$$\Psi_{t,t+h}^Q(u) = \sum_{k=1}^{h}\left\{c\,(d_k)^{k-1}\,\Psi_t^Q(d_k) + \frac{1 - \beta(d_k)^{k-1}}{1 - \beta(d_k)}\omega(d_k)\right\}$$

and the sequence $(d_k)_{1\leq k\leq h}$ is defined as follows:

$$d_k = (u,0) + \sum_{j=k}^{h-1}\beta(d_{j+1})^{j-k}(0,\alpha(d_{j+1}))\ \text{for}\ k \leq h-1 \tag{3.22}$$
$$d_h = (u,0).$$

We will use these formulas in the empirical section.

4   Three Empirical Examples

This section provides three empirical examples developed in companion papers.

4.1   The Term Structure of Realized Risk

This example is studied in Feunou and Meddahi (2007a) and has two goals. The first one is to model the joint dynamics of the returns and the realized variance. The second goal is to compute the term structure of the value-at-risk, i.e. to characterize the quantile function of the aggregated returns, $\sum_{i=1}^{h} r_{t+i}$, when h varies.

We consider the daily realized variance computed as the sum of squared intra-daily returns, five-minutes and thirty-minutes returns in our empirical application. The recent literature on volatility shows the importance of such measures. The basic theory on realized volatility assumes that the underlying process is in continuous time and shows that the realized variance converges to the integrated variance when the length of intra-day returns goes to zero. In our empirical analysis, we specify the model in discrete time and we do not make the formal connection between the realized variance and the daily returns. We will specify discrete models, affine or generalized affine, and allow the data to select the best model. We will, however, use some insights from continuous time when we specify the discrete model. In what follows the conditioning information is $I_t = \sigma(r_\tau, RV_\tau, \tau \leq t)$ where r_t is the daily returns.

We start our analysis by modeling the realized variance as either an affine process or a generalized one. Consider the affine model given by

$$\psi_t(u) = \log E_t[\exp(uRV_{t+1})] = \omega(u) + \alpha(u)RV_t. \tag{4.1}$$

Given the non-negativity of the realized variance process, we will consider two examples. The first one corresponds to the Inverse Gaussian case while the second one is the Gamma case, which corresponds to the exact discretization of the square-root process, studied in Gourieroux and Jasiak (2006)):

$$\text{Inverse Gaussian}: \quad \omega(u) = \nu(1 - \sqrt{1 - 2u\mu}), \ \alpha(u) = \frac{\rho}{\mu}(\exp(1 - \sqrt{1 - 2u\mu}) - 1) \tag{4.2}$$

$$\text{Gamma}: \quad \omega(u) = -\nu \log(1 - u\mu), \ \alpha(u) = \frac{\rho u}{1 - u\mu}. \tag{4.3}$$

When we extend our analysis to the generalized affine case, i.e.,

$$\psi_t(u) = \log E_t[\exp(uRV_{t+1})] = \omega(u) + \alpha(u)RV_t + \beta\psi_{t-1}(u), \tag{4.4}$$

22

we still consider the same two examples of Inverse-Gaussian and Gamma. We prove in Section 2 that this leads to a proper cumulant function.

We use the maximum likelihood method to estimate the four models. The empirical results are provided in Tables 1 and 2. The main empirical result is that the coefficient β is non-zero whatever the model or the realized volatility measure (based on five-minutes or thirty-minutes). In particular, the increase of the log-likelihood is substantial when one allows β to be non-zero. Another interesting result is that the inverse Gaussian model describes better the date for the two frequencies.

We now want to specify a joint model for the returns and the realized variance. When one considers a continuous time stochastic volatility model

$$d \log p_u = (a + b\sigma_u^2)du + \sigma_u dW_u$$

and assumes that there is no leverage effect, one gets that the daily return $r_{t+1} = \log(p_{t+1}) - \log(p_t)$ has the following distribution:

$$r_t | \sigma(p_\tau, \sigma_s, \tau \leq t, s \leq t+1) \ \sim \ \mathcal{N}(a + bIV_{t+1}, IV_{t+1}),$$

which suggests the following discrete time model that we study:

$$r_{t+1} \mid \sigma(r_\tau, RV_\tau, RV_{t+1}, \tau \leq t) \ \sim \ \mathcal{N}(a + bRV_{t+1}, c + dRV_{t+1}). \tag{4.5}$$

We assume that RV_{t+1} follows (4.4) where $\alpha(u)$ follows either (4.2) or (4.3). By denoting the joint cumulant function of (r_{t+1}, RV_{t+1}) as $\psi_{r,RV;t}(v, u)$ defined by

$$\psi_{r,RV;t}(v, u) \equiv \log E_t[\exp(vr_{t+1} + uRV_{t+1})],$$

one gets

$$\psi_{r,RV;t}(v, u) = (va + v^2 c/2) + \psi_t(vb + v^2 d/c + u).$$

Hence, the joint process r_t, RV_t is indeed a generalized affine process because one has

$$\psi_{r,RV;t}(v, u) = \tilde{\omega}(v, u) + \tilde{\alpha}(v, u)RV_t + \beta\psi_{r,RV;t-1}(v, u), \tag{4.6}$$

where

$$\tilde{\omega}(v, u) = (va + v^2 c/2)(1 - \beta) + \omega(vb + v^2 d/2 + u) \tag{4.7}$$

$$\tilde{\alpha}(v, u) = \alpha(vb + v^2 d/2 + u). \tag{4.8}$$

Tables 3 and 4 provide the results of the maximum likelihood estimators. Again, the coefficient β is non-zero and the inverse Gaussian model provided the best fit.

We compute the term structure of the Value-at-Risk, i.e., we compute the 5%-quantile of

$$\overline{r}_{t+1:t+h} \equiv \frac{1}{\sqrt{h}} \sum_{i=1}^{h} r_{t+i}.$$

For this purpose, we derive the conditional characteristic function of $\overline{r}_{t+1:t+h}$ and then we invert it to get the cumulative distribution function. This approach has been used in the affine case and continuous time by Duffie and Pan (2001).

In practice, the value at risk of $\overline{r}_{t+1:t+h}$ will depend on RV_t and its lagged values. In order to graphically present the results, one needs to choose RV_t. We proceed by taking from the data three values for RV_t: a small value (low case), a median one (median case) and a large one (high case). Then, we use the lagged values of each of them to plot the term structure of the value-at-risk (VaR).

Figures 1 to 5 present and compare Affine and Generalized affine term structures of the value-at-risk. Figure 3 shows that in a low variance day, the VaR increases with the maturity and that the affine model overestimates the VaR. In contrast, in a high or median volatility day, affine model overestimates the VaR for lowest maturity and underestimates it for longer maturities. Underestimation of the VaR could lead to important risk management problems; see Feunou and Meddahi (2007a) for more discussions. We also show in that paper that it is useful to consider realized variances, i.e., we used the same approach with the Heston and Nandi (2000) daily model and show that the model with realized volatility is the best one. We also provide in Feunou and Meddahi (2007a) the term structure of another risk measure called the expected shortfall.

4.2 A No-Arbitrage VARMA Term Structure Model with Macroeconomic Variables

This example hinges on Feunou (2007) where a no-arbitrage VARMA term structure model with macroeconomic variables is studied. Ang and Piazzesi (2003) and Ang, Piazzesi and Wei (2006) studied the term structure of interest rates by assuming that the state variable is a Vector AR process (VAR). The state variable included macroeconomic variables (inflation and real activity), denoted X_t, and financial variables (short term yield which proxies the level and term spread which proxies the slope), denoted Y_t. The state variable is $Z_t = (X_t, Y_t)'$. Here, instead of assuming that the state variable is VAR, we will assume that it is a VARMA one:

$$Z_{t+1} = \mu + \phi Z_t + \Sigma \left(\varepsilon_{t+1} - \Theta \varepsilon_t \right), \quad \varepsilon \sim i.i.d. \ \mathcal{N}(0, I).$$

24

We can show that the conditional cumulant function of Z_{t+1} denoted $\psi_t(u)$ satisfies the following recursive relation:

$$\psi_t(u) = \omega(u) + \alpha(u)'Z_t + \psi_{t-1}(\theta(u))$$

with

$$\omega(u) = u'\mu + \frac{1}{2}u'\Sigma\left(I_4 - \Theta\Theta'\right)\Sigma'u, \ \ \alpha(u) = u'\left(\phi - \Sigma\Theta\Sigma^{-1}\right), \ \ \theta(u) = \left(\Sigma\Theta\Sigma^{-1}\right)'u.$$

Feunou (2007) extends the theoretical analysis of the current paper to the multivariate case. In particular, the analytical form of the term structure is derived. Some details are given below.

The estimation by the maximum likelihood shows that the VARMA model describes better the data than the VAR model. In addition, we find that the VARMA forecasts are more accurate than those of VAR and a naive random walk model; see Table 7. We did the forecasting exercise for different horizons, and find that the VARMA model provides smaller forecasting errors in and out of sample. Since we want to assess the VARMA performance in fitting the whole yield curve, we need a coherent framework which allows to connect yields of maturity n to the state variable Z_t.

In a no-arbitrage context, we need to specify a pricing kernel. Our parametric specification of the pricing kernel is similar to the one used in the literature (see Ang and Piazzesi (2003, 2006))

$$M_{t+1} = \exp\left(-y_t^{(1)} - \frac{1}{2}\lambda_t^{\mathsf{T}}\lambda_t - \lambda_t^{\mathsf{T}}\varepsilon_{t+1}\right).$$

To maintain the tractability of the model, affine price of risk (λ_t) is often used $\lambda_t = \lambda_0 + \lambda_1 Z_t$, (again see Ang and Piazzesi (2003)). In the context of our VARMA model, we will depart from the literature by adding the expected future state variable $E_t(Z_{t+1})$, i.e.,

$$\lambda_t = \lambda_0 + \lambda_1 Z_t + \lambda_2 E_t(Z_{t+1}). \tag{4.9}$$

Several interpretation can be given to equation (4.9). First, we can reformulate it as follows

$$
\begin{aligned}
\lambda_t &= \lambda_0 + \lambda_1 Z_t + \lambda_2 \left(\mu + \left(\phi - \Sigma\Theta\Sigma^{-1}\right)Z_t + \Sigma\Theta\Sigma^{-1}E_{t-1}(Z_t)\right) \\
&= \lambda_0^* + \lambda_1^* Z_t + \lambda_2^* E_{t-1}(Z_t)
\end{aligned}
$$

where

$$\lambda_0^* = \lambda_0 + \lambda_2\mu, \ \ \lambda_1^* = \lambda_1 + \lambda_2\left(\phi - \Sigma\Theta\Sigma^{-1}\right), \ \ \text{and} \ \ \lambda_2^* = \lambda_2\Sigma\Theta\Sigma^{-1}.$$

25

Thus, the parameter λ_2 captures the past information impact on the current market price of risk. Another way of rewriting the price of risk is to express it in terms of the expected variable $E_{t-1}(Z_t)$ and the unexpected news $\Sigma \varepsilon_t$:

$$\lambda_t = \lambda_0^* + \lambda_1^* \Sigma \varepsilon_t + (\lambda_1^* + \lambda_2^*) E_{t-1}(Z_t).$$

Feunou (2007) shows that bond yields (with maturity n) are no longer affine of the state variable Z_t, but are rather affine function of the state variable Z_t and its past conditional expectation $E_{t-1}(Z_t)$, i.e.

$$y_t^{(n)} = a_n + b_{1,n}^\mathsf{T} Z_t + b_{2,n}^\mathsf{T} E_{t-1}(Z_t). \tag{4.10}$$

Another representation derived in Feunou (2007) is

$$y_t^{(n)} = a_n + (b_{1,n} + b_{2,n})^\mathsf{T} E_{t-1}(Z_t) + b_{1,n}^\mathsf{T} \Sigma \varepsilon_t, \tag{4.11}$$

where the coefficients a_n, $b_{1,n}$, and $b_{2,n}$ are given in Feunou (2007).

The estimation of the unknown parameters (parameters of the historical distribution and parameters of the price of risk) is done in two steps. The first step estimates the parameters of the historical distribution of the state vector by using the maximum likelihood method. By taking the parameters of the historical dynamic as their estimated values, we estimate in the second step the pricing kernel's parameters by minimizing the squared difference between the model implied yield and the observed yield (in practice, the maturities are 1, 2, 3, and 4 years). Since there is an endogenous component in the state vector (the term spread), we run a constrained minimization in order to guarantee that the dynamic of the yield to maturity 60 months in the VARMA model is coherent with the relation (4.10); see Ang and Piazzesi (2006) and Feunou (2007) for more details.

In order to assess the impact of the MA component, i.e., the impact of the vector Θ on the yield curve, we provide in Figures 6 to 9 the loading coefficients a_n, $b_{1,n}$, $b_{2,n}$ and $b_{1,n} + b_{2,n}$ for the different models and the different component of the state variable. These figures show clearly that there are differences between the VAR and VARMA term structure models. The likelihood estimates as well as the pricing errors are in favor of the VARMA approach, which highlights the importance of using generalized affine models.

4.3 Realized Option Pricing model

This example hinges on Feunou, Christoffersen, Jacobs and Meddahi (2007). We used the model developed in the first empirical example and used the option pricing formulas derived

in Section 3.2 where h_t equals the realized variance RV_t. We model jointly the dynamics of the return r_t and realized variance RV_t in the same way as in section 4.1, with a slight modification of the distribution of the stock log-returns r_t conditional on realized variance RV_t. Following Christoffersen et al (2006), Feunou (2006), and Feunou and Tedongap (2007), we used a skewed inverse Gaussian distribution, which nests the normal distribution. This extension is empirically important.

The model is given by

$$r_{t+1} \mid \sigma(r_\tau, RV_\tau, RV_{t+1}, \tau \leq t) \ \sim \ a + bRV_{t+1} - \eta(c + dRV_{t+1}) + \frac{1}{\eta}y_{t+1}, \qquad (4.12)$$

with $y_{t+1} \sim \mathcal{IG}(\eta^2(c + dRV_{t+1}))$. $\mathcal{IG}$ means the standard inverse gaussian distribution. The conditional cumulant function of the return r_{t+1} conditional on I_t and RV_{t+1} is given by

$$E[\exp(ur_{t+1})|RV_{t+1}, I_t] = \exp(\omega_0(u) + \alpha_0(u)RV_{t+1})$$

with

$$\omega_0(u) = u(a - c\eta) + c\eta^2 \left(1 - \sqrt{1 - \frac{2u}{\eta}}\right), \ \text{ and } \ \alpha_0(u) = u(b - d\eta) + d\eta^2 \left(1 - \sqrt{1 - \frac{2u}{\eta}}\right).$$

In the affine case, the conditional cumulant function of RV_{t+1} given I_t is given by (4.1) where $\omega(u)$ and $\alpha(u)$ are defined either by (4.2) for the inverse gaussian case or by (4.3) for the gamma case. We extend this affine case to the generalized affine of order (1,2) as follows

$$\psi_t(u) = \log E_t[\exp(uRV_{t+1})] = \omega(u) + \alpha(u)RV_t + \beta_1\psi_{t-1}(u) + \beta_2\psi_{t-2}(u). \qquad (4.13)$$

Consequently, the joint cumulant function of (r_{t+1}, RV_{t+1}) given I_t is

$$\psi_{r,RV;t}(v, u) = \omega_0(v) + \psi_t(u + \alpha_0(v)).$$

Eq. (4.13) implies that the joint process (r_t, RV_t) is a generalized affine process

$$\psi_{r,RV;t}(v, u) = \tilde{\omega}(v, u) + \tilde{\alpha}(v, u)RV_t + \beta_1\psi_{r,RV;t-1}(v, u) + \beta_2\psi_{r,RV;t-2}(v, u) \qquad (4.14)$$

with $\tilde{\omega}(v, u) = \omega_0(v)(1 - \beta_1 - \beta_2) + \omega(u + \alpha_0(v))$ and $\tilde{\alpha}(v, u) = \alpha(u + \alpha_0(v))$.

We assume that the generalized affine model is defined under the risk-neutral probability measure. The estimation is done by minimizing the MSE of the implied Black-Scholes volatility from the option (IVMSE) defined as

$$IVMSE = \frac{1}{n}\sum_{i=1}^{n}(\sigma_i - \sigma_i(\theta))^2$$

27

where the implied volatilities are obtained as

$$\sigma_i = BS^{-1}\left(C_i, T_i, X_i, S, r\right) \quad \text{and} \quad \sigma_i\left(\theta\right) = BS^{-1}\left(C_i\left(\theta\right), T_i, X_i, S, r\right),$$

with BS^{-1} being the inverse of the Black-Scholes formula, T_i the time to maturity, X_i the strike price, S the price of the underlying stocks and r the riskless interest rate.

Figures 10, 11 and 12 represent the daily implied volatility bias, option price bias and implied volatility RMSE. The generalized affine model clearly outperforms the affine model in terms of pricing errors. This result holds whatever the maturity of the moneyness; see Tables 5 and 6.

5 Conclusion

The paper extends affine models by introducing moving average type components in the conditional cumulant functions. The extension is important theoretically because important models like ARMA are not affine. The extension is also empirically important as shown in three empirical examples.

There is an alternative approach that leads to non-Markov affine processes. It uses the Laplace transform of the process x_t defined as $\mathcal{L}_t(u) = \exp(\psi_t(u))$ instead of the cumulant function. The traditional affine models are characterized by

$$\mathcal{L}_t(u) = \exp(\omega(u) + \alpha(u)x_t).$$

In a companion paper, we are currently studying the process defined by

$$\mathcal{L}_t(u) = \gamma(u) + \exp(\omega(u) + \alpha(u)x_t) + \beta(u)\mathcal{L}_{t-1}(u).$$

References

[1] Ang, A. and M., Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics 50 745-787.*

[2] Ang, A., M., Piazzesi and M., Wei (2006). What does the yield curve tell us about GDP growth. *Journal of Econometrics 131 359-403.*

[3] Barndorff-Nielsen OE. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24, 1-13.*

[4] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics 31, 307-327.*

[5] Bollerslev, T. and L., Forsberg (2002). Bridging the gap between the distribution of realized (ECU) and ARCH modelling (of the euro): The GARCH-NIG model. *Journal of Applied Econometrics 17, 535-548.*

[6] Carrasco, M. and J-P Florens (2002). Efficient GMM Estimation Using the Empirical Characteristic Function. *Working Paper; Departement of Economics: University of Rochester.*

[7] Carrasco, M., and J-P., Florens(2000).*Generalization of GMM to a continuum of moment conditions.*, Econometric Theory, 16, 2000, 797-834.

[8] Carrasco, M., M., Chernov, J-P., Florens, and E., Ghysels (2007). Efficient estimation of jump diffusions and general dynamic models with a continuum of moment conditions. *Journal of Econometrics, 140 529-573*

[9] Christoffersen, P. (1998) valuating interval forecasts. *E International Economic Review 39, 841.862.*

[10] Christoffersen, P., R., Elkamhi, B., Feunou and K., Jacobs (2006). Option Valuation with Conditional Heteroskedasticity and Non-Normality. *Working Paper, McGill University.*

[11] Christoffersen, P., S., Heston, and K., Jacobs (2006). Option valuation with conditional skewness. *Journal of Econometrics 131 253-284.*

[12] Christoffersen, P., and K. Jacobs (2004). Which GARCH Model for Option Valuation?. *Management Sci. 50. 1204-1221.*

[13] Christoffersen, P., K., Jacobs and Y., Wang (2005). Volatility Components: Affine Restrictions and Non-normal Innovations. *Working Paper; Faculty of Management, McGill University.*

[14] Cox, J.C., J.E., Ingersoll and S.A. Ross (1985).A theory of the term structure of interest rates. *Econometrica 53, 385-407.*

[15] Dai, Q. and K., Singleton, (2000). Specification Analysis of affine term structure models. *Journal of Finance 55, 1943-1978.*

[16] Diebold, F., M., Piazzesi, and D., Rudebusch (2005).Modeling Bond Yields in Finance and Macroeconomics. *American Economic Review P&P Volume 95, Issue 2, May 2005, pp. 415-420.*

[17] Diebold, F., D., Rudebusch, and B., Aruoba (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics 131 309-338.*

[18] Darolles, S., C., Gourieroux, and J., Jasiak (2006). Structural Laplace Transform and Compound Autoregressive Models. *Journal of time serie analysis Vol. 27, No. 4.*

[19] Diebold, F. X., T. A., Gunther and T. S., Tay (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review 39, 863. 883.*

[20] Duan, J. C. (1999). Conditionally Fat-Tailed Distributions and the Volatility Smile in Options. *Working paper, Hong Kong University of Science and Technology.*

[21] Duan, J. C. (1995). The GARCH option pricing model. *Math. Finance 5, 13-32.*

[22] Duan, J. C. (1996). Cracking the smile.*Risk 9 55-59.*

[23] Duan, J. C., J.G. Simonato (1998). Empirical martingale simulation for asset prices. *Management Sci. 44 1218-1233.*

[24] Duffie, D., J., Pan and K., Singleton (2000). Transform Analysis and Asset Pricing for affine jump diffusions. *Econometrica. 68, 1343-1376.*

[25] Duffie, D., and , R., Kan(1996). *A yield-Factor model of interest rates* Mathematical Finance. 6, 379-406.

[26] Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica 50 987-1008.*

[27] Feller, W. (1968). An introduction to probability theory and its applications. *John Wiley & Sons.*

[28] Feuerverger, A. and R.A., Mureika (1977). The Empirical Characteristic Function and its Applications. *The annals of Statistics. 5, 88-97.*

[29] Feuerverger, A., and P., McDunnough (1981). On the Efficiency of Empirical Characteristic Function Procedure. *Journal of the Royal Statistical Society. Series B. 43, 20-27.*

[30] Feunou, B. (2007). A no-arbitrage VARMA term structure model with macroeconomic variables. *Working Paper, Université de Montréal.*

[31] Feunou, B., and N., Meddahi (2007). Realized term structure of risk. *Work in Progress.*

[32] Feunou, B., P., Christoffersen, K., Jacobs, and N., Meddahi (2007). Realized Option pricing models. *Work in progress.*

[33] Feunou, B., and R. Tedongap (2007). Affine Stochastic Skewness Models. *Working Paper, Université de Montréal.*

[34] Gourieroux, C. (2006). Continuous Time Wishart Process for Stochastic Risk. *25, Issue 2 & 3, pages 177 - 217.*

[35] Gourieroux, C., and J.Jasiak (2002). Value at Risk . *Handbook of financial Econometric.*

[36] Gourieroux, C., and J., Jasiak (2006). Autoregressive Gamma Processes. *Journal of Forecasting. 25, 129-152.*

[37] Gourieroux, C., A., Monfort and V., Polimenis (2002). Affine term structure models. *CREST DP.*

[38] Gourieroux, C. and A., Monfort (2006). Econometric Specifications of Stochastic Discount Factor Models. *Journal of Econometrics 136 509-530.*

[39] Hansen, L. P.(1982). Large Sample properties of generalized method of moment estimators. *Econometrica 50, 1029-1053.*

[40] Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies, 6, 327-343.*

[41] Heston, S. and S. Nandi (2000). A Closed-Form GARCH Option Pricing Model. *Review of Financial Studies, 13, 585-626.*

[42] Johnson, N.L., S., Kotz and N., Balakrishnan. Continuous Univariate distributions. *Volume1. Second Edition. Wiley series in probability and mathematical statistics.*

[43] Monfort, A. and F. Pegoraro (2006). Switching VARMA Term Structure Models. *Journal of Financial Econometrics 5(1): 105-153.*

[44] Monfort, A. and F., Pegoraro (2006). Multi-Lag Term Structure Models with Stochastic Risk Premia. Working paper, CREST.

[45] Lukacs, E. (1970). Characteristic Functions.*Griffin 1820 London.*

[46] Singleton, K. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics 102, 111-141.*

[47] Vasicek, O. (1977). An equilibrium characterization of term structure. *Journal of Financial Economics 5, 177-188.*

[48] Wilhelmsson, A. (2006) Value at Risk with time varying variance, skewness and kurtosis The NIG-S&ARCH-tv model. *Working Paper Swedish School of Economics.*

[49] Yu, J. (2004). Empirical Characteristic Function Estimation and its applications. *Econometric Reviews, 23 (2) :93-123.*

Appendix A

In this appendix, we build a generalized affine model where the function $\beta(\cdot)$ varies. Let us consider a positive process X_t with conditional cumulant function Ψ_t.

$$E_t \left[\exp\left(uX_{t+1}\right)\right] = \exp\left(\Psi_t\left(u\right)\right)$$

We define $\Psi_t\left(u\right)$ recursively as follows:

$$
\begin{aligned}
\Psi_0\left(u\right) &= \omega\left(u\right) + a_0\left(u, X_0\right) \\
\Psi_1\left(u\right) &= \omega\left(u\right) + a_0\left(u, X_1\right) + a_1\left(u, X_0\right) \\
\Psi_2\left(u\right) &= \omega\left(u\right) + a_0\left(u, X_2\right) + a_1\left(u, X_1\right) + a_2\left(u, X_0\right)
\end{aligned}
$$

and generally, we have

$$\Psi_t\left(u\right) = \omega\left(u\right) + \sum_{i=0}^{t} a_i\left(u, X_{t-i}\right) \tag{A.1}$$

The first issue is to give some conditions on sequence functions $a_i(u, x)$ and $\omega\left(u\right)$ such that $\Psi_t(u)$ is a well defined cumulant function.

If $\omega\left(u\right)$ and $a_i\left(u, x\right)$ are cumulant functions $\forall\ i$, then $\Psi_t\left(u\right)$ is a well defined cumulant function. Indeed, the sum of cumulant function is a cumulant function.

Consequently we will choose $\omega\left(u\right)$ and $a_i\left(u, x\right)$ such that they will be always cumulant functions. Another consequence is the fact that we can write X_{t+1} as follows

$$X_{t+1} = \eta_{t+1} + \sum_{i=0}^{t} Z_{i,t+1}$$

where η_{t+1} and $Z_{i,t+1}$ are mutually conditionally independent with cumulant function $\omega\left(u\right)$ and $a_i\left(u, X_{t-i}\right)$. This give us a simple approach to simulate X_{t+1}.

The final goal is to rewrite definition of $\Psi_t\left(u\right)$ given by (A.1) recursively. To achieve this goal the following expression is given to cumulant function $a_i\left(u, x\right)$

$$
\begin{aligned}
a_i\left(u, x\right) &= P_i\left(x\right)\left[\exp\left(a\left(u\right)i + b\left(u\right)\right) - 1\right] \tag{A.2}\\
P_i\left(x\right) &= \exp\left(\lambda_0 + \lambda_1 i\right) x
\end{aligned}
$$

As it was the case with $\Psi_t\left(u\right)$, we need to make sure that (A.2) is a valid cumulant function. This is done using Lemma 5.4.1 of Lukacs (1970) (page 111) where it is shown that $p(g(u) - 1)$ is an infinitely divisible cumulant function whenever $g(u)$ is a characteristic function and $p > 0$.

Thus if $a\left(u\right)$ and $b\left(u\right)$ are cumulant functions and X a positive process, then $a_i\left(u, x\right)$ is a cumulant function. Since process X_t is built using cumulant generating function, it is hard to simulate. We give an answer in the following lines. Proposition 5.1 shows how a random variable with cumulant function $a_i\left(u, x\right)$ can be simulated.

Proposition 5.1 $p(g(u) - 1)$ *is the cumulant function of* Z *iff*

$$Z = \sum_{n=0}^{N} Y_n$$

where random variables N *and* Y_n *are mutually independent,* N *follows Poisson distribution of parameter* p *and the moment generating function of* Y_n *is* $g(u)$.

Since $\Psi_t(u)$ is the conditional cumulant function of X_{t+1} (which is a positive random variable), we must then choose $\omega(u)$, $a(u)$ and $b(u)$ such that $\Psi_t(u)$ is a cumulant function of a positive random variable. The following proposition addresses this issue.

Proposition 5.2 *If* $a(u)$, $b(u)$ *and* $\omega(u)$ *are cumulant functions of positive random variable, then* $\Psi_t(u)$ *is a well defined conditional cumulant function of positive random variable* X_{t+1}

We are now ready to write $\Psi_t(u)$ recursively.

$$
\begin{aligned}
\Psi_t(u) &= \omega(u) + \sum_{i=0}^{t} P_i(X_{t-i}) \left[\exp(a(u) i + b(u)) - 1 \right] \\
&= \omega(u) + \sum_{i=0}^{t} P_i(X_{t-i}) \exp(a(u) i + b(u)) - \sum_{i=0}^{t} P_i(X_{t-i}) \\
&= \omega(u) + \sum_{i=0}^{t} \exp((a(u) + \lambda_1) i + \lambda_0 + b(u)) X_{t-i} - \sum_{i=0}^{t} \exp(\lambda_0 + \lambda_1 i) X_{t-i}
\end{aligned}
$$

Proposition 5.3 $\omega(u)$ *can always be reformulated was following*

$$\omega(u) = \frac{c(u)}{1 - \exp(a(u) + \lambda_1)} - \frac{c(0)}{1 - \exp(\lambda_1)}$$

As shown below, the proof of Proposition 5.3 is a direct consequence of the fact that $\omega(u)$ is a cumulant function.

We can then rewrite $\Psi_t(u)$ as following.

$$\Psi_t(u) = f_t(u) - f_t(0)$$

with

$$
\begin{aligned}
f_t(u) &= \frac{c(u)}{1 - \exp(a(u) + \lambda_1)} + \sum_{i=0}^{t} \exp((a(u) + \lambda_1) i + \lambda_0 + b(u)) X_{t-i} \\
&= \frac{c(u)}{1 - \exp(a(u) + \lambda_1)} + \exp(\lambda_0 + b(u)) \sum_{i=0}^{t} \exp(a(u) + \lambda_1)^i X_{t-i}
\end{aligned}
$$

Proposition 5.4 $f_{t+1}(u)$ *evolves recursively as follows:*

$$f_{t+1}(u) = c(u) + \exp(\lambda_0 + b(u)) X_{t+1} + \exp(a(u) + \lambda_1) f_t(u)$$

An immediate consequence of Proposition 5.4 is the recursive formulation of $\Psi_t(u)$ given by proposition 5.5.

Proposition 5.5

$$\Psi_t(u) = \omega_0(u) + \alpha_1(u) f_t(0) + \alpha_2(u) f_{t-1}(0) + \beta(u) \Psi_{t-1}(u) \tag{A.3}$$

where

$$
\begin{aligned}
\omega_0(u) &= c(u) - c(0)e^{b(u)} \\
\alpha_1(u) &= e^{b(u)} - 1 \\
\alpha_2(u) &= e^{\lambda_1}\left[e^{a(u)} - e^{b(u)}\right] \\
\beta(u) &= e^{\lambda_1 + a(u)}
\end{aligned}
$$

Note that in the right hand side of equation (A.3), we have $f_t(0)$ instead of X_t. For this reason the conditional cumulant generating function of $f_t(0)$ is evaluated. The joint conditional moment generating function of X_{t+1} and $f_{t+1}(0)$ is:

$$
\begin{aligned}
E_t\left[\exp(uX_{t+1} + vf_{t+1}(0))\right] &= E_t\left[\exp\left(uX_{t+1} + v\left(c(0) + e^{\lambda_0}X_{t+1} + e^{\lambda_1}f_t(0)\right)\right)\right] \\
&= \exp\left(vc(0) + ve^{\lambda_1}f_t(0) + \Psi_t(u + ve^{\lambda_0})\right)
\end{aligned}
$$

Thus if we denote $\Psi_t^c(u, v) = \ln\left(E_t\left[\exp(uX_{t+1} + vf_{t+1}(0))\right]\right)$, we have

$$
\begin{aligned}
\Psi_t^c(u, v) &= vc(0) + ve^{\lambda_1}f_t(0) + \Psi_t(u + ve^{\lambda_0}) \\
&= vc(0) + ve^{\lambda_1}f_t(0) + \omega_0\left(u + ve^{\lambda_0}\right) + \alpha_1\left(u + ve^{\lambda_0}\right)f_t(0) \\
&\quad + \alpha_2\left(u + ve^{\lambda_0}\right)f_{t-1}(0) + \beta\left(u + ve^{\lambda_0}\right)\Psi_{t-1}(u + ve^{\lambda_0}) \\
&= vc(0) + ve^{\lambda_1}f_t(0) + \omega_0\left(u + ve^{\lambda_0}\right) + \alpha_1\left(u + ve^{\lambda_0}\right)f_t(0) \\
&\quad + \alpha_2\left(u + ve^{\lambda_0}\right)f_{t-1}(0) + \beta\left(u + ve^{\lambda_0}\right)\left[\Psi_{t-1}^c(u, v) - vc(0) - ve^{\lambda_1}f_{t-1}(0)\right]
\end{aligned}
$$

The whole expression of $\Psi_t^c(u, v)$ is summarized in the following proposition.

Proposition 5.6

$$\Psi_t^c(u, v) = W(u, v) + A_1(u, v) f_t(0) + A_2(u, v) f_{t-1}(0) + B(u, v) \Psi_{t-1}^c(u, v)$$

where

$$
\begin{aligned}
W(u, v) &= vc(0)\left(1 - \beta\left(u + ve^{\lambda_0}\right)\right) + \omega_0\left(u + ve^{\lambda_0}\right) \\
A_1(u, v) &= ve^{\lambda_1} + \alpha_1\left(u + ve^{\lambda_0}\right) \\
A_2(u, v) &= \alpha_2\left(u + ve^{\lambda_0}\right) - ve^{\lambda_1}\beta\left(u + ve^{\lambda_0}\right) \\
B(u, v) &= \beta\left(u + ve^{\lambda_0}\right)
\end{aligned}
$$

In conclusion the vector $(X_{t+1}, f_{t+1}(0))$ is a generalized affine of order $(2,1)$, implying a univariate generalized affine for $f_{t+1}(0)$ as stated in the following corollary.

Corollary 2 *Notice that by imposing $u = 0$ we have a generalized affine model of order $(2,1)$ for $f_t(0)$.*

Indeed

$$
\begin{aligned}
E_t\left[\exp\left(vf_t(0)\right)\right] &\equiv \exp\left(\Psi_t^f(v)\right) \\
&= \exp\left(\Psi_t^c(0,v)\right)
\end{aligned}
$$

with

$$
\Psi_t^c(0,v) = W(0,v) + A_1(0,v) f_t(0) + A_2(0,v) f_{t-1}(0) + B(0,v) \Psi_{t-1}^c(0,v)
$$

Proposition 5.7 *Generally, for any given s, $(f_t(0), f_t(s))$ is a generalized affine of order $(2,1)$*

We can restrict $f_t(0)$ to be positive by just imposing $c(0)$ to be positive and considering positive initial value $f_0(0)$. On the other hand $f_t(0)$ can take any sign if any restriction is made on $c(0)$ and $f_0(0)$. All these assertions are consequences of the recursive definition of $f_{t+1}(0)$

$$
f_{t+1}(0) = c(0) + \exp(\lambda_0) X_{t+1} + \exp(\lambda_1) f_t(0)
$$

Since X_{t+1} is a positive random variable, if $f_t(0) \geq 0$ and $c(0) \geq 0$, then $f_{t+1}(0) \geq 0$. $c(0)$ is an undetermined parameter with undetermined sign. This implies that if the sign of $c(0)$ and $f_t(0)$ are undetermined then $f_{t+1}(0)$'s sign is also undetermined.

Generalized affine of order $(1,1)$ (for $f_t(0)$) can be obtained by restricting functions a and b to satisfy $A_2(0,v) = 0$. Solving $A_2(0,v) = 0$ implies

$$
b(v) = a(v) + \ln\left(1 - ve^{\lambda_1 - \lambda_0}\right)
$$

Let us denote $e^{\lambda_1 - \lambda_0} = \mu$, $\ln(1 - v\mu)$ remind us a component of ARG of Gourieroux and Jasiak (2006). Thus the following choice of function $a(.)$,

$$
a(v) = \frac{v\rho}{1 - v\mu} - \nu \ln(1 - v\mu)
$$

with $\rho \geq 0$, $\nu \geq 1$ insures well defined generalized affine model of order $(1,1)$ for $f_t(0)$.These examples of functions $a(.)$ and $b(.)$ are cumulant functions of positive random variables which are in fact the only restrictions needed on function $a(.)$ and $b(.)$. In general, the following general function $a(.)$ is sufficient to satisfy $A_2(0,v) = 0$.

$$
a(v) = \bar{a}(v) - \nu \ln(1 - v\mu)
$$

where $\bar{a}(v)$ is the cumulant function of a positive random variable, and $\nu > 1$.

Proposition 5.8 *Let denote* $e^{\lambda_1 - \lambda_0} = \mu$, *if*

$$b\left(v\right) = a\left(v\right) + \ln\left(1 - v\mu\right)$$

and

$$a\left(v\right) = \bar{a}\left(v\right) - \nu \ln\left(1 - v\mu\right)$$

where $\bar{a}\left(v\right)$ is the cumulant function of a positive random variable, and $\nu > 1$. then $A_2\left(0, v\right) = 0$ which implies that $f_t(0)$ is a generalized affine of order $(1, 1)$.

Proof of Proposition 5.1: The proof is quite easy, in fact it is done by realizing that if G is the distribution function corresponding to characteristic function g (or moment generating function), then $F = e^{-p} \sum_0^\infty \frac{p^n}{n!} G^{n\star}$ is the distribution function corresponding to characteristic function (or moment generating function $\exp\left(p(g(u) - 1)\right)$). In this expression $G^{n\star}$ means the convolution of n identical distribution function G. The simulation of random variable corresponding to distribution function F is also easy to deal with. Let consider a sequence of *iid* random variable $(Y_i)_{i=1,2...}$, and a discrete random variable N which is independent to $(Y_i)_{i=1,2...}$ and which follows a Poisson distribution with parameter p. The following random variable X has F as distribution function:

$$Z = \sum_{n=0}^{N} Y_n$$

where Y_0 is a constant.

Proof of Proposition 5.2: The result is the consequence of the fact that $p(g(u) - 1)$ is the cumulant function of positive random variable when $g\left(u\right)$ is the moment generating function of a positive random variable. This result is deduced from the previous Proposition, indeed since $p(g(u) - 1)$ is the cumulant function of $Z = \sum_{n=0}^{N} Y_n$, and $g(u)$ the moment generating function of Y_n. $Y_n \geq 0 \Rightarrow Z \geq 0$

Proof of Proposition 5.3: In fact, for any given choice of a cumulant function of positive random variable $\omega\left(u\right)$, choose $c\left(u\right)$ as follows

$$c\left(u\right) = \left(1 - \exp\left(a\left(u\right) + \lambda_1\right)\right) \left[\omega\left(u\right) + \frac{\delta}{1 - \exp\left(\lambda_1\right)}\right]$$

for any real δ. Since $\omega\left(u\right)$ and $a\left(u\right)$ are a cumulant functions, thus $\omega\left(0\right) = a\left(0\right) = 0$, which implies that

$$c\left(0\right) = \delta$$

Proof of Proposition 5.4: Indeed

$$c\left(u\right) + \exp\left(\lambda_0 + b\left(u\right)\right) X_{t+1} + \exp\left(a\left(u\right) + \lambda_1\right) f_t\left(u\right)$$

$$= \ c\left(u\right) + \exp\left(\lambda_0 + b\left(u\right)\right) X_{t+1} + \frac{\exp\left(a\left(u\right) + \lambda_1\right) c\left(u\right)}{1 - \exp\left(a\left(u\right) + \lambda_1\right)} + \exp\left(\lambda_0 + b\left(u\right)\right) \sum_{i=0}^{t} \exp\left(a\left(u\right) + \lambda_1\right)^{i+1} X_{t-i}$$

$$= \ \frac{c\left(u\right)}{1 - \exp\left(a\left(u\right) + \lambda_1\right)} + \exp\left(\lambda_0 + b\left(u\right)\right) \sum_{i=0}^{t+1} \exp\left(a\left(u\right) + \lambda_1\right)^{i} X_{t+1-i}$$

$$= \ f_{t+1}\left(u\right)$$

Proof of Proposition 5.5:

$$f_{t+1}\left(u\right) = c\left(u\right) + \exp\left(\lambda_0 + b\left(u\right)\right) X_{t+1} + \exp\left(a\left(u\right) + \lambda_1\right) f_t\left(u\right)$$

and

$$\Psi_t\left(u\right) = f_t\left(u\right) - f_t\left(0\right)$$

imply that

$$\begin{aligned}
\Psi_t(u) + f_t(0) \ &= \ c(u) + e^{b(u)}\left[f_t(0) - c(0) - e^{\lambda_1} f_{t-1}(0)\right] \\
&\quad + e^{\lambda_1 + a(u)}\left[\Psi_{t-1}(u) + f_{t-1}(0)\right]
\end{aligned}$$

Proof of Proposition 5.7: Indeed

$$\begin{aligned}
E_t\left[\exp\left(u f_{t+1}(0) + v f_{t+1}(s)\right)\right] \ &= \ E_t\left[\exp\left(\begin{array}{c} u\left(c(0) + e^{\lambda_0} X_{t+1} + e^{\lambda_1} f_t(0)\right) \\ + v\left(c(s) + e^{\lambda_0 + b(s)} X_{t+1} + e^{\lambda_1 + a(s)} f_t(s)\right) \end{array}\right)\right] \\
&= \ \exp\left(\begin{array}{c} u c(0) + v c(s) + u e^{\lambda_1} f_t(0) \\ + v e^{\lambda_1 + a(s)} f_t(s) + \Psi_t(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}) \end{array}\right) \\
&\equiv \ \exp\left(\Psi_t^{c,s}\left(u,v\right)\right)
\end{aligned}$$

$$\begin{aligned}
\Psi_t^{c,s}\left(u,v\right) \ &= \ u c(0) + v c(s) + u e^{\lambda_1} f_t(0) + v e^{\lambda_1 + a(s)} f_t(s) + \Psi_t(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}) \\
&= \ u c(0) + v c(s) + u e^{\lambda_1} f_t(0) + v e^{\lambda_1 + a(s)} f_t(s) + \omega_0\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) \\
&\quad + \alpha_1\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) f_t(0) + \alpha_2\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) f_{t-1}(0) \\
&\quad + \beta\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) \Psi_{t-1}(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}) \\
&= \ u c(0) + v c(s) + u e^{\lambda_1} f_t(0) + v e^{\lambda_1 + a(s)} f_t(s) + \omega_0\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) \\
&\quad + \alpha_1\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) f_t(0) + \alpha_2\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right) f_{t-1}(0) \\
&\quad + \beta\left(u e^{\lambda_0} + v e^{\lambda_0 + b(s)}\right)\left[\Psi_{t-1}^{c,s}\left(u,v\right) - u c(0) - v c(s) - u e^{\lambda_1} f_{t-1}(0) - v e^{\lambda_1 + a(s)} f_{t-1}(s)\right]
\end{aligned}$$

Hence

$$\Psi_t^{c,s}(u,v) = W^s(u,v) + A_1^s(u,v)' f_t(0,s) + A_2^s(u,v)' f_{t-1}(0,s) + B^s(u,v) \Psi_{t-1}^{c,s}(u,v)$$

where

$$f_t(0,s) = \begin{pmatrix} f_t(0) \\ f_t(s) \end{pmatrix}$$

$$W^s(u,v) = (uc(0) + vc(s))\left(1 - \beta\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right)\right) + \omega_0\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right)$$

$$A_1^s(u,v) = \begin{pmatrix} ue^{\lambda_1} + \alpha_1\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right) \\ ve^{\lambda_1 + a(s)} \end{pmatrix}$$

$$A_2^s(u,v) = \begin{pmatrix} \alpha_2\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right) - ue^{\lambda_1}\beta\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right) \\ -ve^{\lambda_1 + a(s)}\beta\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right) \end{pmatrix}$$

$$B^s(u,v) = \beta\left(ue^{\lambda_0} + ve^{\lambda_0 + b(s)}\right)$$

Appendix B

This appendix provides the proofs of Section 2 and Section 3.

Proof of Proposition 2.2. Cumulant's equation implies that

$$\kappa_{1,t} = \omega^{(1)}(0) + \alpha^{(1)}(0) X_t + \beta(0) \kappa_{1,t-1} \tag{B.1}$$

by taking the unconditional expectation in both side of the equation we get the first result. i.e

$$E[X_t] = E[\kappa_{1,t}] = \frac{\omega^{(1)}(u)}{1 - \left(\alpha^{(1)}(u) + \beta(0)\right)}$$

Since

$$V(X_t) = E[\kappa_{2,t}] + V[\kappa_{1,t}]$$

the equation of $\kappa_{2,t}$ is

$$\kappa_{2,t} = \omega^{(2)}(0) + \alpha^{(2)}(0) X_t + 2\beta^{(1)}(0) \kappa_{1,t-1} + \beta(0) \kappa_{2,t}$$

this implies that

$$E[\kappa_{2,t}] = \frac{\alpha^{(2)}(0) + 2\beta^{(1)}(0)}{1 - \beta(0)} E[X_t] + \frac{\omega^{(2)}(0)}{1 - \beta(0)}$$

$$
\begin{aligned}
V[\kappa_{1,t}] &= \alpha^{(1)}(0)^2 V(X_t) + \beta(0)^2 V[\kappa_{1,t-1}] + 2\alpha^{(1)}(0)\beta(0)\left[E\left(\kappa_{1,t-1}^2\right) - E[X_t]^2\right] \\
&= \alpha^{(1)}(0)^2 V(X_t) + \left(\beta(0)^2 + 2\alpha^{(1)}(0)\beta(0)\right) V[\kappa_{1,t-1}]
\end{aligned}
$$

thus

$$V[\kappa_{1,t}] = \frac{\alpha^{(1)}(0)^2}{1 - \left(\beta(0)^2 + 2\alpha^{(1)}(0)\beta(0)\right)} V(X_t)$$

hence

$$V(X_t) = \left(1 + \frac{\alpha^{(1)}(0)^2}{1 - \left(\alpha^{(1)}(0) + \beta(0)\right)^2}\right) E[\kappa_{2,t}]$$

Thus we can the second result.

for $h \geq 2$ we have

$$
\begin{aligned}
E[X_t X_{t+h}] &= E[X_t E_{t+h-1}[X_{t+h}]] \\
&= E[X_t \kappa_{1,t+h-1}] \\
&= E\left[X_t \left[\omega^{(1)}(0) + \alpha^{(1)}(0) X_{t+h-1} + \beta(0) \kappa_{1,t+h-2}\right]\right] \\
&= \omega^{(1)}(0) E(X_t) + \alpha^{(1)}(0) E[X_t X_{t+h-1}] + \beta(0) E[X_t X_{t+h-1}] \\
&= \omega^{(1)}(0) E(X_t) + \left(\alpha^{(1)}(0) + \beta(0)\right) E(X_t X_{t+h-1})
\end{aligned}
$$

Proof of Corollary 1: From the Proposition 2.2, we have that

$$E\left(X_t X_{t+h}\right) = \omega^{(1)}\left(0\right) E\left[X_t\right] + \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right) E\left(X_t X_{t+h-1}\right)$$

this implies that

$$
\begin{aligned}
E\left(X_t X_{t+h}\right) &= \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1} E\left(X_t X_{t+1}\right) + E\left(X_t\right)^2 \left[1 - \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1}\right] \\
&= \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right)^{h-1} cov\left(X_t, X_{t+1}\right) + E\left(X_t\right)^2
\end{aligned}
$$

we then get the result.

Proof of Proposition 2.3:

$$
\begin{aligned}
cov\left(X_t, X_{t+1}\right) &= cov\left(X_t, \kappa_{1,t}\right) \\
&= cov\left(X_t, \omega^{(1)}\left(0\right) + \alpha^{(1)}\left(0\right) X_t + \beta\left(0\right)\kappa_{1,t-1}\right) \\
&= \alpha^{(1)}\left(0\right) V\left(X_t\right) + \beta\left(0\right) V\left(\kappa_{1,t-1}\right) \\
&= \alpha^{(1)}\left(0\right) V\left(X_t\right) + \frac{\beta\left(0\right)\alpha^{(1)}\left(0\right)^2}{1 - \left(\beta\left(0\right)^2 + 2\alpha^{(1)}\left(0\right)\beta\left(0\right)\right)} V\left(X_t\right)
\end{aligned}
$$

Proof of Proposition 2.4: Indeed

$$
\begin{aligned}
E\left[\kappa_{1,t}\overline{\kappa}_{n,t}\right] &= E\left[\kappa_{1,t}\left(\overline{\omega}_n + \overline{\alpha}_n x_t + \overline{\beta}_n \overline{\kappa}_{n,t-1}\right)\right] \\
&= \overline{\omega}_n E\left(x_t\right) + E\left(x_t \kappa_{1,t}\right)\overline{\alpha}_n + \overline{\beta}_n E\left[\kappa_{1,t-1}\overline{\kappa}_{n,t-1}\right]
\end{aligned}
$$

Proof of Proposition 2.5:

$$
\begin{aligned}
E\left[\kappa_{1,t+h}\overline{\kappa}_{n,t}\right] &= E\left[\left(\omega^{(1)}\left(0\right) + \alpha^{(1)}\left(0\right) X_{t+h} + \beta\left(0\right)\kappa_{1,t+h-1}\right)\overline{\kappa}_{n,t}\right] \\
&= \omega^{(1)}\left(0\right) E\left[\overline{\kappa}_{n,t}\right] + \left(\alpha^{(1)}\left(0\right) + \beta\left(0\right)\right) E\left[\kappa_{1,t+h-1}\overline{\kappa}_{n,t}\right]
\end{aligned}
$$

Proof of Proposition 2.6:

$$\overline{\kappa}_{n,t} = \overline{\omega}_n + \overline{\alpha}_n x_t + \overline{\beta}_n \overline{\kappa}_{n,t-1}$$

implies that

$$
\begin{aligned}
V\left(\overline{\kappa}_{n,t}\right) - \overline{\beta}_n V\left(\overline{\kappa}_{n,t}\right)\overline{\beta}_n^\top &= V\left(x_t\right)\overline{\alpha}_n\overline{\alpha}_n^\top + \overline{\beta}_n cov\left(\overline{\kappa}_{n,t-1}\ \kappa_{1,t-1}\right)\overline{\alpha}_n^\top + \overline{\alpha}_n cov\left(\kappa_{1,t-1}\ \overline{\kappa}_{n,t-1}\right)\overline{\beta}_n^\top \\
&= \theta_n
\end{aligned}
$$

40

Proof of Proposition 2.7:

$$
\begin{aligned}
Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h}\right] &= Cov\left[\overline{\kappa}_{n,t}, \overline{\omega}_n + \overline{\alpha}_n x_{t+h} + \overline{\beta}_n \overline{\kappa}_{n,t+h-1}\right] \\
&= Cov\left[\overline{\kappa}_{n,t}, x_{t+h}\right]\overline{\alpha}_n^\top + Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h-1}\right]\overline{\beta}_n^\top \\
&= Cov\left[\overline{\kappa}_{n,t}, \overline{\kappa}_{n,t+h-1}\right]\overline{\beta}_n^\top + cov\left[\overline{\kappa}_{n,t}\,\kappa_{1,t+h-1}\right]\overline{\alpha}_n^\top
\end{aligned}
$$

Proof of Proposition 2.8:

$$
\begin{aligned}
Cov\left(x_t^2,\ x_{t+h}\right) &= Cov\left(x_t^2,\ \kappa_{1,t+h-1}\right) \\
&= Cov\left(x_t^2,\ \omega^{(1)}(0) + \alpha^{(1)}(0)\,x_{t+h-1} + \beta(0)\,\kappa_{1,t+h-2}\right) \\
&= \alpha^{(1)}(0)\,Cov\left(x_t^2,\ x_{t+h-1}\right) + \beta(0)\,Cov\left(x_t^2,\ \kappa_{1,t+h-2}\right) \\
&= \left(\alpha^{(1)}(0) + \beta(0)\right)Cov\left(x_t^2,\ x_{t+h-1}\right)
\end{aligned}
$$

Proof of Proposition 2.9: We have

$$
\begin{aligned}
V_{t,h}(u_1, u_2, \ldots, u_h) &\equiv E_t\left[\exp\left(\sum_{j=1}^{h} u_j x_{t+j}\right)\right] \\
&= E_t\left\{\exp\left(\sum_{j=1}^{h-1} u_j x_{t+j}\right) E_{t+h-1}\left[\exp\left(u_h x_{t+h}\right)\right]\right\} \\
&= E_t\left[\exp\left(\Psi_{t+h-1}(u_h) + \sum_{j=1}^{h-1} u_j x_{t+j}\right)\right]
\end{aligned}
$$

From Eq. (2.1), one gets easily

$$
\forall j \geq 1,\ \Psi_{t+j}(u) = \beta(u)^j\,\Psi_t(u) + \omega(u)\,\frac{1-\beta(u)^j}{1-\beta(u)} + \alpha(u)\sum_{k=1}^{j}\beta(u)^{j-k}\,x_{t+k}. \tag{5.2}
$$

Hence,

$$
\begin{aligned}
\log(V_{t,h}(u_1, u_2, \ldots, u_h)) &= \beta(u_h)^{h-1}\,\Psi_t(u_h) + \omega(u_h)\,\frac{1-\beta(u_h)^{h-1}}{1-\beta(u_h)} \\
&+ \log E_t\left[\exp\left(\sum_{j=1}^{h-1}\left(u_j + \alpha(u_h)\beta(u_h)^{h-1-j}\right)X_{t+j}\right)\right] \\
&= \beta(u_h)^{h-1}\,\Psi_t(u_h) + \omega(u_h)\,\frac{1-\beta(u_h)^{h-1}}{1-\beta(u_h)} \\
&+ \log V_{t,h-1}(u_1 + \alpha(u_h)\beta(h_u)^{h-2}, u_2 + \alpha(u_h)\beta(h_u)^{h-3}, \ldots, u_{h-1} + \alpha(u_h)).
\end{aligned}
$$

The result (2.10) is then obtained by induction.

Proof of Proposition 3.4. $B(t,h)$ is the price at t of a zero-coupon bond which give 1 at $t+h$

$$
\begin{aligned}
B(t,h) &= E_t^Q \left[\exp\left(-\sum_{i=0}^{h-1} r_{t+i} \right) \right] \\
&= \exp\left[-r_t + \sum_{k=1}^{h-1} \left\{ \beta(d_k)^{k-1} \Psi_t(d_k) + \frac{1 - \beta(d_k)^{k-1}}{1 - \beta(d_k)} \omega(d_k) \right\} \right]
\end{aligned}
$$

where the last equality follows from Proposition 2.9 while the sequence $(d_k)_{1 \leq k \leq h-1}$ is defined in (3.11). Given that the yield at horizon h is $r_{t,t+h} = -\ln(B(t,h))/h$, one gets (3.10).

Proof of Proposition 3.6. One has

$$
\begin{aligned}
E_t[M_{t+1}] &= \exp(-r) \\
E_t[M_{t+1} \exp(r_{t+1})] &= 1,
\end{aligned}
$$

which leads to

$$
\begin{aligned}
\theta_t + \Psi_t(\gamma, \lambda) &= -r \\
\theta_t + \Psi_t(1 + \gamma, \lambda) &= 0.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\theta_t &= -r - \Psi_t(\gamma, \lambda) \\
\Psi_t(1 + \gamma, \lambda) - \Psi_t(\gamma, \lambda) &= r
\end{aligned}
$$

By using the following expression of the model:

$$
\Psi_{t+1}(u, v) = \frac{\omega(u, v)}{1 - \beta(u, v)} + \alpha(u, v) \sum_{i=0}^{\infty} \beta(u, v)^i h_{t-i+1}
$$

one gets,

$$
\frac{\omega(1 + \gamma, \lambda)}{1 - \beta(1 + \gamma, \lambda)} - \frac{\omega(\gamma, \lambda)}{1 - \beta(\gamma, \lambda)} + \sum_{i=0}^{\infty} \left[\beta(1 + \gamma, \lambda)^i \alpha(1 + \gamma, \lambda) - \beta(\gamma, \lambda)^i \alpha(\gamma, \lambda) \right] h_{t-i} = r
$$

which implies

$$
\begin{aligned}
\frac{\omega(1 + \gamma, \lambda)}{1 - \beta(1 + \gamma, \lambda)} - \frac{\omega(\gamma, \lambda)}{1 - \beta(\gamma, \lambda)} &= r \\
\beta(1 + \gamma, \lambda)^i \alpha(1 + \gamma, \lambda) - \beta(\gamma, \lambda)^i \alpha(\gamma, \lambda) &= 0, \ \forall i \geq 0.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\omega(1 + \gamma, \lambda)}{1 - \beta(1 + \gamma, \lambda)} - \frac{\omega(\gamma, \lambda)}{1 - \beta(\gamma, \lambda)} &= r \\
\beta(1 + \gamma, \lambda) &= \beta(\gamma, \lambda) \\
\alpha(1 + \gamma, \lambda) &= \alpha(\gamma, \lambda).
\end{aligned}
$$

Proof of Proposition 3.7.

$$E_t^Q \left[\exp\left(ur_{t+1} + v\Psi_{t+1}\left(\gamma\right)\right)\right] = E_t^Q \left[\exp\left(ur_{t+1} + v\left(\omega\left(\gamma\right) + \alpha\left(\gamma\right)r_{t+1} + \beta\left(\gamma\right)\Psi_t\left(\gamma\right)\right)\right)\right]$$

thus

$$
\begin{aligned}
\Psi_t^*\left(u,v\right) &= v\omega\left(\gamma\right) + v\beta\left(\gamma\right)\Psi_t\left(\gamma\right) + \Psi_t^Q\left(u^*\right) \\
&= v\omega\left(\gamma\right) + v\beta\left(\gamma\right)\Psi_t\left(\gamma\right) + \Psi_t\left(u^* + \gamma\right) - \Psi_t\left(\gamma\right)
\end{aligned}
$$

where

$$u^* = u + v\alpha\left(\gamma\right)$$

this implied that

$$
\begin{aligned}
\Psi_{t+1}^*\left(u,v\right) &= v\omega\left(\gamma\right) + \left(v\beta\left(\gamma\right) - 1\right)\Psi_{t+1}\left(\gamma\right) + \Psi_{t+1}\left(u^* + \gamma\right) \\
&= v\omega\left(\gamma\right) + \left(v\beta\left(\gamma\right) - 1\right)\Psi_{t+1}\left(\gamma\right) + \omega\left(u^* + \gamma\right) \\
&\quad + \alpha\left(u^* + \gamma\right)r_{t+1} + \beta\left(u^* + \gamma\right)\Psi_t\left(u^* + \gamma\right) \\
&= v\omega\left(\gamma\right) + \left(v\beta\left(\gamma\right) - 1\right)\Psi_{t+1}\left(\gamma\right) + \omega\left(u^* + \gamma\right) + \\
&\quad + \alpha\left(u^* + \gamma\right)r_{t+1} + \beta\left(u^* + \gamma\right)\left[\Psi_t^*\left(u,v\right) - v\omega\left(\gamma\right) - \left(v\beta\left(\gamma\right) - 1\right)\Psi_t\left(\gamma\right)\right]
\end{aligned}
$$

For Proposition 3.8, we provide a general proof when the joint dynamic of return and its conditional variance follow a generalized affine of order (p,q). See the proof of Proposition 5.12 for more details.

Appendix C: Generalized Affine Models of Higher Order

The generalized affine model of order (1,1) can be straightforwardly extended to order (p,q) as following.

$$\psi_t(u) = \omega(u) + \sum_{j=0}^{p} \alpha_j(u)x_{t-j} + \sum_{j=1}^{q} \beta_j(u)\psi_{t-j}(u) \tag{C.1}$$

Let note

$$\kappa_{n,t} = \psi_t^{(n)}(0)$$

In this section, we give some details on generalization of several issue which have been addressed in order (1,1).

Cumulant and moment structure

$$\kappa_{n,t} = \omega^{(n)}(0) + \sum_{j=0}^{p} \alpha_j^{(n)}(0)x_{t-j} + \sum_{j=1}^{q}\sum_{i=0}^{n-1} \binom{n-1}{i}\beta_j^{(i)}(0)\kappa_{n-i,t-j} \tag{C.2}$$

$$\overline{\kappa}_{n,t} = (\kappa_{1,t}, \kappa_{2,t}, ..., \kappa_{n,t})'$$

$$\overline{\kappa}_{n,t} = \overline{\omega}_n + \sum_{j=0}^{p} \overline{\alpha}_{j,n}x_{t-j} + \sum_{j=1}^{q} \overline{\beta}_{j,n}\overline{\kappa}_{n,t-j} \tag{C.3}$$

where

$$\overline{\omega}_n = \begin{pmatrix} \omega^{(1)}(0) \\ \omega^{(2)}(0) \\ .. \\ .. \\ \omega^{(n)}(0) \end{pmatrix}, \ \overline{\alpha}_{j,n} = \begin{pmatrix} \alpha_j^{(1)}(0) \\ \alpha_j^{(2)}(0) \\ .. \\ .. \\ \alpha_j^{(n)}(0) \end{pmatrix}$$

and

$$\overline{\beta}_{j,n} = \begin{bmatrix} \beta_j(0) & 0 & 0 & ... & 0 \\ \binom{2}{1}\beta_j^{(1)}(0) & \beta_j(0) & 0 & ... & 0 \\ \vdots & \vdots & \vdots & ... & \vdots \\ \vdots & \vdots & \vdots & ... & \vdots \\ \binom{n}{n-1}\beta_j^{(n-1)}(0) & \binom{n}{n-2}\beta_j^{(n-2)}(0) & ... & \binom{n}{1}\beta_j^{(1)}(0) & \beta_j(0) \end{bmatrix}$$

$$\overline{\kappa}_{n,t} = \overline{\omega}_n + \overline{\alpha}_n X_t + \iota'\overline{\beta}_n\overline{\overline{\kappa}}_{n,t-1}$$

where

$$\overline{\alpha}_n = [\overline{\alpha}_{0,n}, ..., \overline{\alpha}_{p,n}]$$

$$X_t = \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ \vdots \\ x_{t-p} \end{pmatrix}, \quad \underset{(n \times q)}{\iota'} = (1, ..., 1), \quad \overline{\overline{\kappa}}_{n,t-1} = \begin{pmatrix} \overline{\kappa}_{n,t-1} \\ \overline{\kappa}_{n,t-2} \\ \vdots \\ \vdots \\ \overline{\kappa}_{n,t-q} \end{pmatrix}$$

and

$$\overline{\beta}_n = \begin{bmatrix} \overline{\beta}_{1,n} & 0 & 0 & \cdots & 0 \\ 0 & \overline{\beta}_{2,n} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \overline{\beta}_{q,n} \end{bmatrix}$$

Unconditional moments

From equation (C.3), we deduce unconditional mean of cumulant,

$$E(x_t) = \frac{\omega^{(1)}(0)}{1 - \sum_{j=0}^{p} \alpha_j^{(1)}(0) - \sum_{j=1}^{q} \beta_j(0)}$$

$$E[\overline{\kappa}_{n,t}] = \left(I - \sum_{j=1}^{q} \overline{\beta}_{j,n} \right)^{-1} \left[\overline{\omega}_n + E(x_t) \sum_{j=0}^{p} \overline{\alpha}_{j,n} \right]$$

If the process is covariance stationary, then the autocovariance function $\gamma(h) = Cov(x_t, x_{t+h})$ is the solution of the following recurrence sequence

$$\gamma(h) = \sum_{j=0}^{p} \alpha_j^{(1)}(0)\gamma(h - 1 - j) + \sum_{j=1}^{q} \beta_j(0)\gamma(h - j)$$

with $\gamma(0) = V(x_t)$

Indeed

$$
\begin{aligned}
Cov(x_t, x_{t+h}) &= cov(x_t, \kappa_{1,t+h-1}) \\
&= cov\left(x_t, \omega^{(1)}(0) + \sum_{j=0}^{p} \alpha_j^{(1)}(0)x_{t+h-1-j} + \sum_{j=1}^{q} \beta_j(0)\kappa_{1,t+h-1-j} \right) \\
&= \sum_{j=0}^{p} \alpha_j^{(1)}(0)cov(x_t, x_{t+h-1-j}) + \sum_{j=1}^{q} \beta_j(0)cov(x_t, \kappa_{1,t+h-1-j})
\end{aligned}
$$

45

Let us denote the covariance between x_t and $\overline{\kappa}_{n,t+h}$ by $\delta_n(h) = cov(x_t \, ; \, \overline{\kappa}_{n,t+h})$. $\delta_n(h)$ is the solution of the following recurrence sequence

$$\delta_n(h) = \sum_{j=0}^{p} \gamma(h-j)\overline{\alpha}_{j,n} + \sum_{j=1}^{q} \delta_n(h-j)\overline{\beta}_{j,n}^{\mathsf{T}}$$

Denote the covariance between $\overline{\kappa}_{n,t}$ and $\overline{\kappa}_{n,t+h}$ by $\overline{\delta}_n(h) = cov(\overline{\kappa}_{n,t} \, ; \, \overline{\kappa}_{n,t+h})$. $\overline{\delta}_n(h)$ is the solution of the following recurrence sequence

$$\overline{\delta}_n(h) = \sum_{j=0}^{p} \delta_n(h-j)^{\mathsf{T}}\overline{\alpha}_{j,n} + \sum_{j=1}^{q} \overline{\delta}_n(h-j)\overline{\beta}_{j,n}^{\mathsf{T}}$$

Cumulant Function of Aggregated Returns.

Let $V_{t,h}(u_1,\cdots,u_h) = E_t\left[\exp\left(\sum_{i=1}^{h} u_i x_{t+i}\right)\right]$; the following proposition shows how to compute $V_{t,h}$ by recursion. Details on proofs are provided in appendix.

Proposition 5.9

$$\ln\left[V_{t,h}(u_1,\cdots,u_h)\right] = \omega_h + \sum_{j=0}^{p-1} \delta_j^h x_{t-j} + \sum_{j=0}^{q} \gamma_j^h \Psi_{t-j}(u) + \ln\left[V_{t,h-1}\left(u_1^{h-1},\cdots,u_{h-1}^{h-1}\right)\right]$$

Details on definition of sequences ω_h, γ_j^h and δ_j^h are provided in the proof.

A more general result.

$$W_{t,h}(u,v;v_0) = E_t\left[\left(v_0 + \sum_{i=1}^{h} v_i x_{t+i}\right)\exp\left(\sum_{i=1}^{h} u_i x_{t+i}\right)\right]$$

$$W_{t,h}(u,v;v_0) = v_0 V_{t,h}(u_1,\cdots,u_h) + \sum_{i=1}^{h} v_i E_t\left[x_{t+i}\exp\left(\sum_{i=1}^{h} u_i x_{t+i}\right)\right]$$

for $1\leq i \leq h$

$$E_t\left[x_{t+i}\exp\left(\sum_{i=1}^{h} u_i x_{t+i}\right)\right] = \frac{\partial}{\partial u_i}\left[V_{t,h}(u_1,\cdots,u_h)\right]$$

thus

$$W_{t,h}(u, v; v_0) = v_0 V_{t,h}(u_1, \cdots, u_h) + \sum_{i=1}^{h} v_i \frac{\partial}{\partial u_i} [V_{t,h}(u_1, \cdots, u_h)]$$

The term structure of interest rates.

We specify the generalized affine model of order (p,q) for the historical dynamic of short term rate r_{t+1}.

$$\ln E_t^P [\exp(u r_{t+1})] \equiv \psi_t(u) = \omega(u) + \sum_{j=0}^{p} \alpha_j(u) r_{t-j} + \sum_{j=1}^{q} \beta_j(u) \psi_{t-j}(u)$$

We give the exponential affine form to the pricing kernel $M_{t,t+1}$

$$M_{t,t+1} = \exp(\gamma r_{t+1} - r_t - \psi_t(\gamma))$$

The following proposition provides closed-form expression of yield to maturity h at time t (noted $r_{t,t+h}$).

Proposition 5.10

$$r_{t,t+h} = -\frac{\overline{\omega_h^*}(\gamma)}{h} - \sum_{j=1-p}^{0} \frac{\overline{\delta_j}(\gamma)}{h} r_{t+j} + \frac{r_t}{h} - \sum_{k=0}^{q} \frac{\overline{\gamma_k}(\gamma)}{h} \Psi_{t-k}(\gamma) - \frac{\ln V_{t,h}(u_1, \cdots, u_h)}{h}$$

Proof and definition of function sequences $\overline{\gamma_k}$, $\overline{\omega_h^*}(u)$ and $\overline{\delta_j}(u)$ are provided below.

Option Pricing.

As for order (1,1), we assume that the joint dynamic of log returns r_{t+1} and it conditional variance h_{t+1} follows a generalized affine model under the historical probability measure.

$$\Psi_t(u, v) = \log E_t^P [\exp(u r_{t+1} + v h_{t+1})] = \omega(u, v) + \sum_{j=0}^{p} \alpha_j(u, v) h_{t-j} + \sum_{j=1}^{q} \beta_j(u, v) \psi_{t-j}(u, v)$$

We give the following pricing kernel

$$M_{t+1} = \exp(\gamma r_{t+1} + \lambda h_{t+1} + \theta_t)$$

where

$$\theta_t = -r - \psi_t(\gamma, \lambda)$$
$$\psi_t(1 + \gamma, \lambda) - \psi_t(\gamma, \lambda) = r$$

We show below the following proposition

Proposition 5.11 *Price of risk γ are solution of the following equation*

$$\sum_{i=1}^{q} \sum_{j=0}^{\min(k,p)} \left\{ \nu_i \left(1+\gamma,\lambda\right) \alpha_j \left(1+\gamma,\lambda\right) \widehat{\beta}_i^{k-j} \left(1+\gamma,\lambda\right) - \nu_i \left(\gamma,\lambda\right) \alpha_j \left(\gamma,\lambda\right) \widehat{\beta}_i^{k-j} \left(\gamma,\lambda\right) \right\} = 0, \ \forall k \geq 0$$

where $\widehat{\beta}_j(u,v)$ for j=1,..,q, are inverse of real or complex root of $1 - \sum_{j=1}^{q} \beta_j(u,v) L^j$

Observe that sufficient conditions which guarantee these equalities are:

$$\begin{aligned}
\nu_i \left(1+\gamma,\lambda\right) &= \nu_i \left(\gamma,\lambda\right), \ \forall \, 1 \leq i \leq q \\
\alpha_j \left(1+\gamma,\lambda\right) &= \alpha_j \left(\gamma,\lambda\right), \ \ \forall \, 0 \leq j \leq p \\
\widehat{\beta}_i \left(1+\gamma,\lambda\right) &= \widehat{\beta}_i \left(\gamma,\lambda\right), \ \forall \, 1 \leq i \leq q
\end{aligned}$$

Let $\psi_{t,t+h}^{Q}$ denotes the conditional risk-neutral cumulant function of aggregated future returns $\sum_{i=1}^{h} r_{t+i}$. The following proposition give the closed form expression of $\psi_{t,t+h}^{Q}$.

Proposition 5.12

$$\begin{aligned}
\psi_{t,t+h}^{Q}(u) &= -rh - \psi_t\left(\gamma,\lambda\right) - \sum_{i=2}^{h} \omega_{i-1}^{*}\left(\gamma,\lambda\right) - \sum_{i=2}^{h} \left[\sum_{k=0}^{q} \gamma_{i-1,k}\left(\gamma,\lambda\right) \psi_{t-k}\left(\gamma,\lambda\right) \right] \\
&\quad - \sum_{j=1-p}^{0} \left[\sum_{i=j+2}^{h} \delta_{i-1-j}\left(\gamma,\lambda\right) \right] h_{t+j} + \log V_{t,h}\left(\left(u_1,v_1\right), ..., \left(u_h,v_h\right) \right)
\end{aligned}$$

See below for the proof and definition of the sequences u_i, v_i, and the functions sequence $\omega_i^{*}(u,v)$, $\delta_i(u,v)$ and $\gamma_{i,j}(u,v)$.

The price at t of a European call option which pays $\left(S_{t+h} - X\right)^{+}$ at t+h is given by

$$C_t = \exp\left(-rh\right) \left[S_t C_{1,t} - X C_{2,t} \right]$$

where

$$\begin{aligned}
C_{1,t} &= \frac{e^{rh}}{2} + \int_{0}^{+\infty} \frac{1}{\pi u} Im \left[\exp\left(\psi_{t,t+h}^{Q}\left(1+iu\right) - iu \ln\left(\frac{X}{S_t}\right) \right) \right] \\
C_{2,t} &= \frac{1}{2} + \int_{0}^{+\infty} \frac{1}{\pi u} Im \left[\exp\left(\psi_{t,t+h}^{Q}\left(iu\right) - iu \ln\left(\frac{X}{S_t}\right) \right) \right]
\end{aligned}$$

Proof of Proposition 5.9. Let note

$$\overline{\Psi_t\left(u\right)} = \begin{pmatrix} \Psi_t\left(u\right) \\ \Psi_{t-1}\left(u\right) \\ : \\ : \\ \Psi_{t-q}\left(u\right) \end{pmatrix}, \ \overline{\omega\left(u\right)} = \begin{pmatrix} \omega\left(u\right) \\ 0 \\ 0 \\ : \\ 0 \end{pmatrix}, \ \alpha\left(u\right) = \begin{pmatrix} \alpha_0\left(u\right) \\ \alpha_1\left(u\right) \\ : \\ : \\ \alpha_p\left(u\right) \end{pmatrix}, \ \overline{\alpha\left(u\right)} = \begin{pmatrix} \alpha\left(u\right)^\mathsf{T} \\ 0 \\ \scriptstyle(p+1,1) \\ : \\ : \\ 0 \\ \scriptstyle(p+1,1) \end{pmatrix}$$

$$\beta\left(u\right) = \begin{bmatrix} \beta_1\left(u\right) & \beta_2\left(u\right) & \cdots & \beta_q\left(u\right) & 0 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & : \\ : & \ddots & \ddots & 0 & : \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

$$\overline{\Psi_t\left(u\right)} = \overline{\omega\left(u\right)} + \overline{\alpha\left(u\right)}X_t + \beta\left(u\right)\overline{\Psi_{t-1}\left(u\right)}$$

$$\overline{\Psi_{t+j}\left(u\right)} = \sum_{k=0}^{j-1}\beta\left(u\right)^k\left[\overline{\omega\left(u\right)} + \overline{\alpha\left(u\right)}X_{t+j-k}\right] + \beta\left(u\right)^j\overline{\Psi_t\left(u\right)}$$

$\Rightarrow$

$$\Psi_{t+j}\left(u\right) = \sum_{k=0}^{j-1}\beta_{k,1}\left(u\right)\left[\omega\left(u\right) + \alpha\left(u\right)^\mathsf{T} X_{t+j-k}\right] + \beta_{j,\underline{1}}\left(u\right)\overline{\Psi_t\left(u\right)}$$

where

$$\beta_{k,1}\left(u\right) = \left[\beta\left(u\right)^k\right]_{(1,1)}$$
$$\beta_{j,\underline{1}}\left(u\right) = \left[\beta\left(u\right)^j\right]_{(1,:)}$$

$$\begin{aligned} \Psi_{t+j}\left(u\right) &= \omega\left(u\right)\sum_{k=0}^{j-1}\left[\beta\left(u\right)^k\right]_{(1,1)} + \sum_{k=0}^{j-1}\beta_{k,1}\left(u\right)\left[\sum_{i=0}^{p}\alpha_i\left(u\right)x_{t+j-k-i}\right] + \sum_{i=0}^{q}\left[\beta\left(u\right)^j\right]_{(1,i+1)}\Psi_{t-i}\left(u\right) \\ &= \omega\left(u\right)\sum_{k=0}^{j-1}\left[\beta\left(u\right)^k\right]_{(1,1)} + \sum_{i=0}^{p+j-1}\left[\sum_{k=0}^{i}\alpha_{i-k}\left(u\right)\beta_{k,1}\left(u\right)\right]x_{t+j-i} + \sum_{i=0}^{q}\left[\beta\left(u\right)^j\right]_{(1,i+1)}\Psi_{t-i}\left(u\right) \\ &= \omega_j^*\left(u\right) + \sum_{i=0}^{p+j-1}\delta_i\left(u\right)x_{t+j-i} + \sum_{i=0}^{q}\gamma_{j,i}\left(u\right)\Psi_{t-i}\left(u\right) \end{aligned}$$

$$\Psi_{t+j}\left(u\right) = \omega_j^*\left(u\right) + \sum_{i=0}^{p+j-1}\delta_i\left(u\right)x_{t+j-i} + \sum_{i=0}^{q}\gamma_{j,i}\left(u\right)\Psi_{t-i}\left(u\right) \tag{C.4}$$

49

where

$$\omega_j^*(u) = \omega(u)\sum_{k=0}^{j-1}\left[\beta(u)^k\right]_{(1,1)}$$

$$\delta_i(u) = \sum_{k=0}^{i}\alpha_{i-k}(u)\left[\beta(u)^k\right]_{(1,1)}$$

$$\gamma_{j,i}(u) = \left[\beta(u)^j\right]_{(1,i+1)}$$

$$
\begin{aligned}
V_{t,h}(u_1,\cdots,u_h) &= E_t\left[\exp\left(\sum_{i=1}^{h}u_i x_{t+i}\right)\right]\\
&= E_t\left[\exp\left(\sum_{i=1}^{h-1}u_i x_{t+i}+\Psi_{t+h-1}(u_h)\right)\right]\\
&= E_t\left[\exp\left(\sum_{i=1}^{h-1}u_i x_{t+i}+\sum_{k=0}^{h-2}\beta_{k,1}(u_h)\left[\omega(u_h)+\alpha(u_h)^\mathsf{T}X_{t+h-1-k}\right]+\beta_{h-1,\underline{1}}(u_h)\overline{\Psi_t(u_h)}\right)\right]\\
&= \exp\left(
\begin{array}{c}
\omega(u_h)\sum_{k=0}^{h-2}\beta_{k,1}(u_h)+\beta_{h-1,\underline{1}}(u_h)\overline{\Psi_t(u_h)}\\
+\sum_{i=1-p}^{0}\left(\sum_{k=0}^{h-1-i}\beta_{k,1}(u_h)\alpha_{h-1-k-i}(u_h)\right)x_{t+i}
\end{array}
\right)V_{t,h-1}\left(u_1^{h-1},\cdots,u_{h-1}^{h-1}\right)
\end{aligned}
$$

where

$$u_i^{h-1}=u_i+\sum_{k=0}^{h-1-i}\beta_{k,1}(u_h)\alpha_{h-1-k-i}(u_h)$$

$$\ln\left[V_{t,h}(u_1,\cdots,u_h)\right]=\omega_h+\sum_{j=0}^{p-1}\delta_j^h x_{t-j}+\sum_{j=0}^{q}\gamma_j^h\Psi_{t-j}(u)+\ln\left[V_{t,h-1}\left(u_1^{h-1},\cdots,u_{h-1}^{h-1}\right)\right]$$

where

$$\omega_h = \omega(u_h)\sum_{k=0}^{h-2}\beta_{k,1}(u_h)$$

$$\gamma_j^h = \left[\beta(u_h)^{h-1}\right]_{(1,j+1)}$$

$$\delta_j^h = \sum_{k=0}^{h-1+j}\alpha_{h-1-k+j}(u_h)\left[\beta(u_h)^k\right]_{(1,1)}$$

Proof of Proposition 5.10.

The price at t of a zero coupon bond which pay 1 at t+h is given by:

$$
\begin{aligned}
B(t, h) &= E_t^P \left[\prod_{i=1}^{h} M_{t+i-1, t+i} \right] \\
&= E_t^P \left[\exp \left(\gamma \sum_{i=1}^{h} r_{t+i} - \sum_{i=1}^{h} r_{t+i-1} - \sum_{i=1}^{h} \psi_{t+i-1}(\gamma) \right) \right]
\end{aligned}
$$

Using relation (5), we have

$$
\begin{aligned}
\sum_{i=1}^{h} \psi_{t+i-1}(\gamma) &= \sum_{i=1}^{h} \left[\omega_{i-1}^*(\gamma) + \sum_{k=0}^{p+i-2} \delta_k(\gamma) r_{t+i-k-1} + \sum_{k=0}^{q} \gamma_{i-1,k}(\gamma) \Psi_{t-k}(\gamma) \right] \\
&= \sum_{i=1}^{h} \omega_{i-1}^*(\gamma) + \sum_{j=1-p}^{h-1} \left[\sum_{i=j+1}^{h} \delta_{i-1-j}(\gamma) \right] r_{t+j} + \sum_{k=0}^{q} \left[\sum_{i=1}^{h} \gamma_{i-1,k}(\gamma) \right] \Psi_{t-k}(\gamma) \\
&= \overline{\omega_h^*}(\gamma) + \sum_{j=1-p}^{h-1} \overline{\delta_j}(\gamma) r_{t+j} + \sum_{k=0}^{q} \overline{\gamma_k}(\gamma) \Psi_{t-k}(\gamma)
\end{aligned}
$$

where

$$
\begin{aligned}
\overline{\omega_h^*}(\gamma) &= \sum_{i=1}^{h} \omega_{i-1}^*(\gamma) \\
\overline{\delta_j}(\gamma) &= \sum_{i=j+1}^{h} \delta_{i-1-j}(\gamma) \\
\overline{\gamma_k}(\gamma) &= \sum_{i=1}^{h} \gamma_{i-1,k}(\gamma)
\end{aligned}
$$

The price at t of zero-coupon bond which pay 1 at $t+h$ is defined as follows:

$$
\begin{aligned}
B(t, h) &= \exp \left(\overline{\omega_h^*}(\gamma) + \sum_{j=1-p}^{0} \overline{\delta_j}(\gamma) r_{t+j} - r_t + \sum_{k=0}^{q} \overline{\gamma_k}(\gamma) \Psi_{t-k}(\gamma) \right) \times E_t^P \left[\exp \left(\sum_{i=1}^{h} u_i r_{t+i} \right) \right] \\
&= \exp \left(\overline{\omega_h^*}(\gamma) + \sum_{j=1-p}^{0} \overline{\delta_j}(\gamma) r_{t+j} - r_t + \sum_{k=0}^{q} \overline{\gamma_k}(\gamma) \Psi_{t-k}(\gamma) \right) V_{t,h}(u_1, \cdots, u_h)
\end{aligned}
$$

where

$$
\begin{aligned}
u_i &= \gamma - 1 + \overline{\delta_i}(\gamma), \quad i \leq h-1 \\
u_h &= \gamma
\end{aligned}
$$

Thus

$$\ln B(t,h) = \overline{\omega_h^*}(\gamma) + \sum_{j=1-p}^{0} \overline{\delta_j}(\gamma) r_{t+j} - r_t + \sum_{k=0}^{q} \overline{\gamma_k}(\gamma) \Psi_{t-k}(\gamma) + \ln V_{t,h}(u_1, \cdots, u_h)$$

$$r_{t,t+h} = -\frac{\overline{\omega_h^*}(\gamma)}{h} - \sum_{j=1-p}^{0} \frac{\overline{\delta_j}(\gamma)}{h} r_{t+j} + \frac{r_t}{h} - \sum_{k=0}^{q} \frac{\overline{\gamma_k}(\gamma)}{h} \Psi_{t-k}(\gamma) - \frac{\ln V_{t,h}(u_1, \cdots, u_h)}{h}$$

Proof of Proposition 5.11. We can rewrite the model as follows:

$$\Psi_t(u,v) \left[1 - \sum_{j=1}^{q} \beta_j(u,v)L^j \right] = \omega(u,v) + \sum_{j=0}^{p} \alpha_j(u,v) h_{t-j}$$

Denote by $\widehat{\beta}_j(u,v)$ for j=1,..,q, the inverse of real or complex root of $1 - \sum_{j=1}^{q} \beta_j(u,v)L^j$. Then, there exist real numbers $\nu_1, ..., \nu_q$ such that:

$$\frac{1}{1 - \sum_{j=1}^{q} \beta_j(u,v)L^j} = \sum_{j=1}^{q} \frac{\nu_j(u,v)}{1 - \widehat{\beta}_j(u,v)L}$$

Thus

$$
\begin{aligned}
\Psi_t(u,v) &= \frac{\omega(u,v)}{1 - \sum_{j=1}^{q} \beta_j(u,v)} + \sum_{j=0}^{p} \sum_{i=1}^{q} \frac{\nu_i(u,v)\alpha_j(u,v)}{1 - \widehat{\beta}_i(u,v)L} h_{t-j} \\
&= \frac{\omega(u,v)}{1 - \sum_{j=1}^{q} \beta_j(u,v)} + \sum_{j=0}^{p} \sum_{i=1}^{q} \nu_i(u,v) \alpha_j(u,v) \sum_{k=0}^{\infty} \widehat{\beta}_i(u,v)^k h_{t-j-k} \\
&= \frac{\omega(u,v)}{1 - \sum_{j=1}^{q} \beta_j(u,v)} + \sum_{j=0}^{p} \sum_{i=1}^{q} \sum_{k=j}^{\infty} \nu_i(u,v) \alpha_j(u,v) \widehat{\beta}_i(u,v)^{k-j} h_{t-k} \\
&= \frac{\omega(u,v)}{1 - \sum_{j=1}^{q} \beta_j(u,v)} + \sum_{k=0}^{\infty} \left[\sum_{i=1}^{q} \sum_{j=0}^{\min(k,p)} \nu_i(u,v) \alpha_j(u,v) \widehat{\beta}_i(u,v)^{k-j} \right] h_{t-k}
\end{aligned}
$$

$$
\begin{aligned}
\psi_t(1+\gamma,\lambda) - \psi_t(\gamma,\lambda) =\ & \frac{\omega(1+\gamma,\lambda)}{1 - \sum_{j=1}^{q} \beta_j(1+\gamma,\lambda)} - \frac{\omega(\gamma,\lambda)}{1 - \sum_{j=1}^{q} \beta_j(\gamma,\lambda)} \\
& + \sum_{k=0}^{\infty} \left[\sum_{i=1}^{q} \sum_{j=0}^{\min(k,p)} \left\{ \begin{array}{c} \nu_i(1+\gamma,\lambda) \alpha_j(1+\gamma,\lambda) \widehat{\beta}_i^{k-j}(1+\gamma,\lambda) \\ -\nu_i(\gamma,\lambda) \alpha_j(\gamma,\lambda) \widehat{\beta}_i^{k-j}(\gamma,\lambda) \end{array} \right\} \right] h_{t-k}.
\end{aligned}
$$

This implies that

$$\frac{\omega(1+\gamma,\lambda)}{1 - \sum_{j=1}^{q} \beta_j(1+\gamma,\lambda)} - \frac{\omega(\gamma,\lambda)}{1 - \sum_{j=1}^{q} \beta_j(\gamma,\lambda)} = r$$

52

$$\sum_{i=1}^{q}\sum_{j=0}^{\min(k,p)}\left\{\nu_i\left(1+\gamma,\lambda\right)\alpha_j\left(1+\gamma,\lambda\right)\widehat{\beta}_i^{k-j}\left(1+\gamma,\lambda\right)-\nu_i\left(\gamma,\lambda\right)\alpha_j\left(\gamma,\lambda\right)\widehat{\beta}_i^{k-j}\left(\gamma,\lambda\right)\right\}=0,\ \forall k\geq 0.$$

Sufficient conditions that guarantee these equalities are:

$$
\begin{aligned}
\nu_i\left(1+\gamma,\lambda\right) &= \nu_i\left(\gamma,\lambda\right),\ \forall\, 1\leq i\leq q\\
\alpha_j\left(1+\gamma,\lambda\right) &= \alpha_j\left(\gamma,\lambda\right),\ \ \forall\, 0\leq j\leq p\\
\widehat{\beta}_i\left(1+\gamma,\lambda\right) &= \widehat{\beta}_i\left(\gamma,\lambda\right),\ \forall\, 1\leq i\leq q
\end{aligned}
$$

Proof of Proposition 5.12.

$$
\begin{aligned}
\psi_{t,t+h}^{Q}\left(u\right) &= \log\left[E_t^Q\left[\exp\left(u\sum_{i=1}^{h}r_{t+i}\right)\right]\right]\\
&= \log\left[E_t^P\left[M_{t+1}\times\cdots\times M_{t+h}\times\exp\left(u\sum_{i=1}^{h}r_{t+i}\right)\right]\right]\\
&= \log\left[E_t^P\left[\exp\left(u\sum_{i=1}^{h}r_{t+i}+\gamma\sum_{i=1}^{h}r_{t+i}+\lambda\sum_{i=1}^{h}h_{t+i}+\sum_{i=1}^{h}\theta_{t+i-1}\right)\right]\right]\\
&= \log\left[E_t^P\left[\exp\left((u+\gamma)\sum_{i=1}^{h}r_{t+i}+\lambda\sum_{i=1}^{h}h_{t+i}-rh-\sum_{i=1}^{h}\psi_{t+i-1}\left(\gamma,\lambda\right)\right)\right]\right]
\end{aligned}
$$

using the fact that

$$\Psi_{t+j}\left(u,v\right)=\omega_j^*\left(u,v\right)+\sum_{i=0}^{p+j-1}\delta_i\left(u,v\right)h_{t+j-i}+\sum_{i=0}^{q}\gamma_{j,i}\left(u,v\right)\Psi_{t-i}\left(u,v\right)$$

with $\omega_j^*\left(u,v\right)=\omega\left(u,v\right)\sum_{k=0}^{j-1}\left[\beta\left(u,v\right)^k\right]_{(1,1)}$, $\delta_i\left(u\right)=\sum_{k=0}^{i}\alpha_{i-k}\left(u,v\right)\left[\beta\left(u,v\right)^k\right]_{(1,1)}$, $\gamma_{j,i}\left(u\right)=\left[\beta\left(u,v\right)^j\right]_{(1,i+1)}$

and

$$\beta\left(u,v\right)=\begin{bmatrix}\beta_1\left(u,v\right) & \beta_2\left(u,v\right) & \cdots & \beta_q\left(u,v\right) & 0\\ 1 & 0 & \cdots & \cdots & 0\\ 0 & 1 & \ddots & & \vdots\\ \vdots & \ddots & \ddots & 0 & \vdots\\ 0 & \cdots & 0 & 1 & 0\end{bmatrix}$$

$$
\begin{aligned}
\psi^{Q}_{t,t+h}(u) \;=\;& -rh - \psi_t(\gamma,\lambda) - \sum_{i=2}^{h} \omega^{*}_{i-1}(\gamma,\lambda) - \sum_{i=2}^{h}\left[\sum_{k=0}^{q}\gamma_{i-1,k}(\gamma,\lambda)\,\psi_{t-k}(\gamma,\lambda)\right] \\
& - \sum_{j=1-p}^{0}\left[\sum_{i=j+2}^{h}\delta_{i-1-j}(\gamma,\lambda)\right] h_{t+j} + \log V_{t,h}\left((u_1,v_1),...,(u_h,v_h)\right)
\end{aligned}
$$

where

$$
V_{t,h}\left((u_1,v_1),...,(u_h,v_h)\right) = E_t^{P}\left[\exp\left(\sum_{i=1}^{h}(u_i,v_i)\begin{pmatrix}r_{t+i}\\h_{t+i}\end{pmatrix}\right)\right]
$$

with

$$
\begin{aligned}
u_i &= u + \gamma,\ \forall\ 1 \le i \le h \\
v_h &= v_{h-1} = \lambda \\
v_j &= \lambda - \sum_{i=j+2}^{h}\delta_{i-1-j}(\gamma,\lambda),\, 1 \le j \le h-2
\end{aligned}
$$

Table 1: **MLE Estimation of Generalized- Autoregressive Inverse Gaussian Process on Realized variance Data.**

The data is the Deutsche mark (DM) / US dollar (USD) exchange rate realized variance. Sample period is 1986:12:01 to 1996:12:01 with a total of 2449 observations

| | 30 min | | | | 5 min | | | |
| | Affine | | G-Affine | | Affine | | G-Affine | |
par	Est	STD	Est	STD	Est	STD	Est	STD
β			0.6044	0.0354			0.5686	0.0342
ρ	0.3438	0.0183	0.1857	0.0166	0.3860	0.0174	0.2264	0.0174
μ	0.2124	0.0091	0.1665	0.0071	0.1525	0.0059	0.1213	0.0046
ν	1.3454	0.0386	0.5447	0.0522	2.1936	0.0624	0.9156	0.0836
LIK	54.9468		121.4220		-216.4841		-143.1975	
BIC	-0.0129		-0.0368		0.0980		0.0712	

Table 2: **MLE Estimation of Generalized- Autoregressive Gamma Process on Realized variance Data.**

The data is the Deutsche mark (DM) / US dollar (USD) exchange rate realized variance. Sample period is 1986:12:01 to 1996:12:01 with a total of 2449 observations

| | 30 min | | | | 5 min | | | |
| | Affine | | G-Affine | | Affine | | G-Affine | |
par	Est	STD	Est	STD	Est	STD	Est	STD
β			0.6093	0.0341			0.5781	0.0324
ρ	0.3288	0.0191	0.1875	0.0163	0.3976	0.0184	0.2340	0.0173
μ	0.1867	0.0063	0.1527	0.0054	0.1414	0.0046	0.1156	0.0038
ν	1.5890	0.0433	0.5880	0.0577	2.3374	0.0682	0.8918	0.0835
LIK	-151.6401		-80.2328		-358.8048		-276.5893	
BIC	0.07148		0.0455		0.1561		0.1257	

Table 3: **MLE Estimation of Generalized- Autoregressive Normal Inverse Gaussian Process on returns and realized variance Data.**

The data is the Deutsche mark (DM) / US dollar (USD) exchange rate returns and realized variance. Sample period is 1986:12:01 to 1996:12:01 with a total of 2449 observations

| | 30 min | | | | 5 min | | | |
| | Affine | | G-Affine | | Affine | | G-Affine | |
par	Est	STD	Est	STD	Est	STD	Est	STD
β			0.6111	0.0396			0.5449	0.0419
ρ	0.3255	0.0203	0.1754	0.0179	0.3444	0.0193	0.2150	0.0192
μ	0.2341	0.0114	0.1834	0.0087	0.1642	0.0071	0.1328	0.0058
ν	1.2565	0.0398	0.5045	0.0545	2.0818	0.0647	0.9380	0.0961
a	0.0063	0.0139	0.0063	0.0140	0.0064	0.0180	0.0064	0.0180
c	-0.0214	0.0448	-0.0214	0.0449	-0.0180	0.0433	-0.0180	0.0434
b	1.74E-08	5.709E-06	1.44E-08	6.024E-06	4.98E-08	1.093E-05	1.54E-08	5.777E-06
d	0.9282	0.0290	0.9282	0.0290	0.7551	0.0236	0.7551	0.0236
LIK	-1600.0932		-1547.4719		-1838.6743		-1790.0531	
BIC	0.8069		0.7850		0.9234		0.9034	

Table 4: **MLE Estimation of Generalized- Autoregressive Normal Gamma Process on returns and realized variance Data.**

The data is the Deutsche mark (DM) / US dollar (USD) exchange rate returns and realized variance. Sample period is 1986:12:01 to 1996:12:01 with a total of 2449 observations

Table 2: Joint Estimation, Panel B:R-RV- DM/USD: gamma								
	30 min				5 min			
	Affine		G-Affine		Affine		G-Affine	
par	Est	STD	Est	STD	Est	STD	Est	STD
β			0.6172	0.0381			0.5696	0.0383
ρ	0.3043	0.0211	0.1744	0.0174	0.3599	0.0203	0.2206	0.0190
μ	0.2033	0.0075	0.1668	0.0065	0.1503	0.0054	0.1239	0.0046
ν	1.5165	0.0447	0.5542	0.0610	2.2383	0.0698	0.8907	0.0935
a	0.0063	0.0139	0.0063	0.0139	0.0064	0.0180	0.0064	0.0179
c	-0.0214	0.0447	-0.0214	0.0448	-0.0180	0.0433	-0.0180	0.0432
b	7.0E-09	2.701E-06	3.36E-08	7.974E-06	1.75E-08	7.241E-06	5.9E-09	3.153E-06
d	0.9282	0.0290	0.9282	0.0290	0.7551	0.0236	0.7551	0.0236
LIK	-1782.7432		-1726.3807		-1975.3616		-1915.0602	
BIC	0.8961		0.8723		0.9901		0.9644	

Table 5: **Implied volatility Root mean squared error by maturity, Implied volatility bias and option price bias by Moneyness**

We estimate the models on a total of 16, 506 contracts with an average call price of 46.05 and average implied volatility of 20.26. The estimation have been done by minimizing the Black-Scholes IVRMSE

Moneyness					
	S/X<0.975	0.975<S/X<1	1<S/X<1.025	1.025<S/X	All
Model	IVRMSE (%)				
Affine	3.8809	4.2988	4.4313	5.0642	4.3768
G-Affine	2.9471	2.9476	3.2201	3.7181	3.1915
Model	IV bias (%)				
Affine	0.2134	-0.0661	0.0346	-0.3556	-0.0166
G-Affine	0.0211	0.1873	0.3723	-0.2216	0.0694
Model	Option price bias				
Affine	0.4357	-0.3601	-0.4654	-1.2721	-0.3124
G-Affine	0.1809	0.0281	0.0990	-0.9638	-0.1342

Table 6: **Implied volatility Root mean squared error by maturity, Implied volatility bias and option price bias by Moneyness**

We estimate the models on a total of 16, 506 contracts with an average call price of 46.05 and average implied volatility of 20.26. The estimation have been done by minimizing the Black-Scholes IVRMSE

Maturity					
	DTM<30	30<DTM<90	90<DTM<180	180<DTM	All
Model	IVRMSE (%)				
Affine	5.4750	4.3179	3.9963	3.8295	4.3768
G-Affine	3.9103	3.1432	2.8824	2.9301	3.1915
Model	IV bias (%)				
Affine	0.5038	-0.2423	-0.0693	0.1737	-0.0166
G-Affine	0.8447	0.1338	-0.2344	-0.4328	0.0694
Model	Option price bias				
Affine	0.3352	-0.5526	-0.5843	0.0199	-0.3124
G-Affine	0.8570	0.3845	-0.5387	-1.8517	-0.1342

Table 7: **State variable forecasting errors: RMSE.**

we measure the difference between model forecast of state variable Z_t, for a given horizon h ($E_t[Z_{t+h}]$) and observed state variable Z_{t+h}. $RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(E_t[Z_{t+h}] - Z_{t+h}\right)^2}$ In sample period is 1952:06 to 2000:12. Out of sample exercise is conducted by successively estimating on 200+i th first observations and forecasting the 200+i+1

	$x_{1,t}$	$x_{2,t}$	$y_t^{(1)}$	$y_t^{(60)} - y_t^{(1)}$	$x_{1,t}$	$x_{2,t}$	$y_t^{(1)}$	$y_t^{(60)} - y_t^{(1)}$
	IS RMSE 1 month horizon				IS RMSE 3 months horizon			
RW	0.1368	0.2491	0.6601	0.6049	0.2880	0.5309	1.0632	0.8731
VAR	0.1312	0.2439	0.6271	0.5705	0.2651	0.4993	0.9722	0.7678
VARMA	0.1296	0.2372	0.6116	0.5586	0.2631	0.4858	0.9692	0.7615
	OS RMSE 1 month horizon				OS RMSE 3 months horizon			
RW	0.1248	0.1367	0.5424	0.5874	0.2782	0.2557	0.7559	0.7818
VAR	0.1244	0.1367	0.5161	0.5684	0.2799	0.2545	0.7264	0.7515
VARMA	0.1212	0.1375	0.5076	0.5681	0.2779	0.2493	0.6829	0.7254
	IS RMSE 6 months horizon				IS RMSE 12 months horizon			
RW	0.4461	0.9014	1.3589	1.0510	0.7319	1.4017	1.7494	1.2651
VAR	0.3915	0.8034	1.2137	0.8711	0.6142	1.1155	1.5490	0.9487
VARMA	0.3885	0.7796	1.2012	0.8586	0.6096	1.0646	1.5406	0.9437
	OS RMSE 6 months horizon				OS RMSE 12 months horizon			
RW	0.4154	0.3983	0.9770	0.9765	0.6149	0.6192	1.3765	1.2150
VAR	0.4197	0.3819	0.9556	0.9078	0.6182	0.5488	1.1998	0.9481
VARMA	0.4186	0.3722	0.8761	0.8702	0.6187	0.5272	1.1204	0.9336

Table 8: **Cross Section Root Mean Squared Errors.**

I measure the difference between model-yields $\hat{y}_t^{(n)}$ and observed yield $y_t^{(n)}$. $RMSE^{(n)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{y}_t^{(n)} - y_t^{(n)}\right)^2}$ In sample period is 1952:06 to 2000:12. Out of sample exercise is conducted by successively estimating on 200+i th first observations and forecasting the 200+i+1

	$y_t^{(12)}$	$y_t^{(24)}$	$y_t^{(36)}$	$y_t^{(48)}$	Total
In Sample					
VAR	0.3669	0.2519	0.1774	0.1774	0.2466
VARMA	0.3529	0.2433	0.1740	0.1740	0.2385
Out of Sample					
VAR	0.3867	0.2832	0.1866	0.0846	0.2607
VARMA	0.3636	0.2679	0.1781	0.0830	0.2462

Table 9: **Yield curve forecasting errors by horizon.**

I measure the difference between model forecast of yield to maturity n, for a given horizon m $(E_t[\hat{y}^{(n)}_{t+h}])$ and observed yield $y^{(n)}_{t+h}$. $RMSE^{(n)}(h) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(E_t[\hat{y}^{(n)}_{t+h}] - y^{(n)}_{t+h}\right)^2}$ In sample period is 1952:06 to 2000:12. Out of sample exercise is conducted by successively estimating on 200+i th first observations and forecasting the 200+i+1

	$y_t^{(12)}$	$y_t^{(24)}$	$y_t^{(36)}$	$y_t^{(48)}$	$y_t^{(12)}$	$y_t^{(24)}$	$y_t^{(36)}$	$y_t^{(48)}$
	IS RMSE 1 month horizon				IS RMSE 3 months horizon			
RW	0.5065	0.4395	0.4034	0.3898	0.9221	0.8201	0.7377	0.6884
VAR	0.5650	0.4624	0.4125	0.3829	0.8638	0.7619	0.6876	0.6479
VARMA	0.5493	0.4533	0.4083	0.3814	0.8524	0.7570	0.6858	0.6481
	OS RMSE 1 month horizon				OS RMSE 3 months horizon			
RW	0.3297	0.3386	0.3434	0.3487	0.6299	0.6671	0.6551	0.6472
VAR	0.5024	0.4356	0.3808	0.3464	0.7368	0.6971	0.6438	0.6117
VARMA	0.4656	0.4115	0.3675	0.3414	0.6783	0.6636	0.6239	0.6039
	IS RMSE 6 months horizon				IS RMSE 12 months horizon			
In Sample								
RW	1.2328	1.1003	0.9938	0.9375	1.6257	1.4786	1.3614	1.3059
VAR	1.1274	1.0021	0.9086	0.8562	1.4826	1.3475	1.2284	1.1645
VARMA	1.1106	0.9929	0.9031	0.8541	1.4694	1.3378	1.2202	1.1601
	OS RMSE 6 months horizon				OS RMSE 12 months horizon			
RW	0.9455	0.9785	0.9531	0.9318	1.4762	1.4753	1.4156	1.3777
VAR	1.0187	0.9651	0.8976	0.8491	1.3991	1.3413	1.2597	1.1914
VARMA	0.9440	0.9180	0.8676	0.8339	1.3237	1.2831	1.2166	1.1622
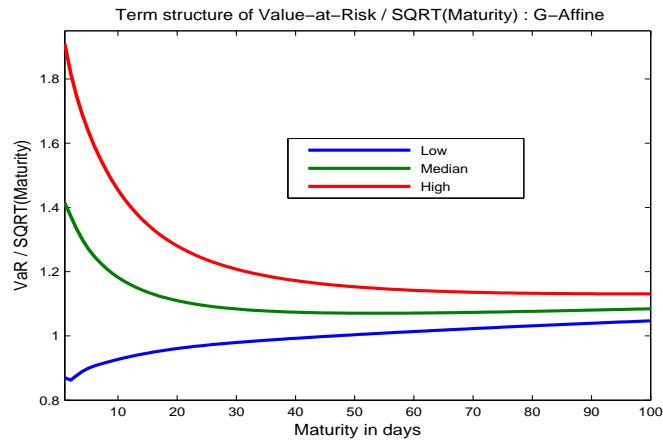
Figure 1: **Generalized Autoregressive Normal Inverse gaussian term structure of value-at-risk**

We use parameters estimated from the MLE to compute the term structure of value at risk. Several cases have been considered depending on the day where the term structure is evaluated. The cases are High volatility day (day with higher realized variance), Median volatility day and Low volatility day
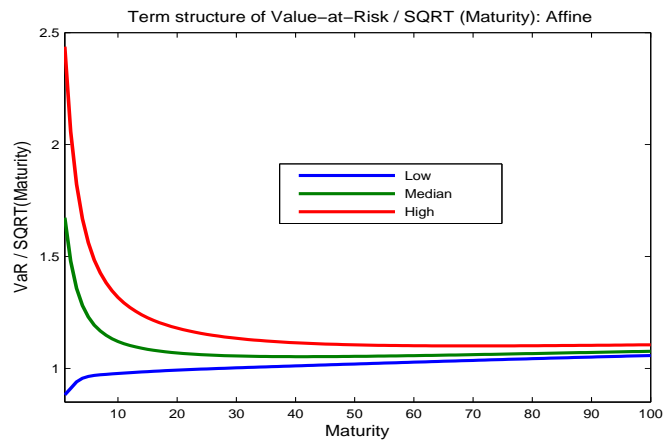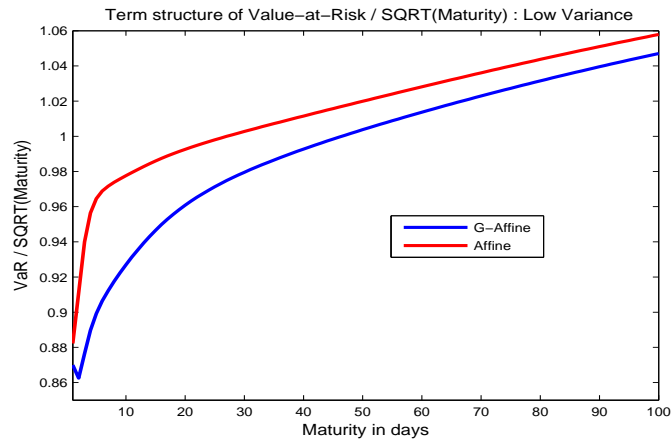

Figure 2: **Autoregressive Normal Inverse gaussian term structure of value-at-risk**

We use parameters estimated from the MLE to compute the term structure of value at risk. Several cases have been considered depending on the day where the term structure is evaluated. The cases are High volatility day (day with higher realized variance), Median volatility day and Low volatility day

Figure 3: **Term structure of Value-at-risk Conditional on a low variance day**
We use parameters estimated from the MLE to compute the term structure of value at risk. We compared Affine and Generalized affine term structure
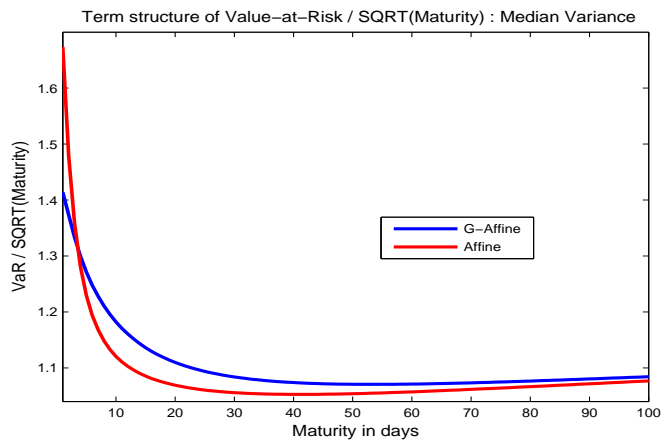

Figure 4: **Term structure of Value-at-risk Conditional on a median variance day**
We use parameters estimated from the MLE to compute the term structure of value at risk. We compared Affine and Generalized affine term structure
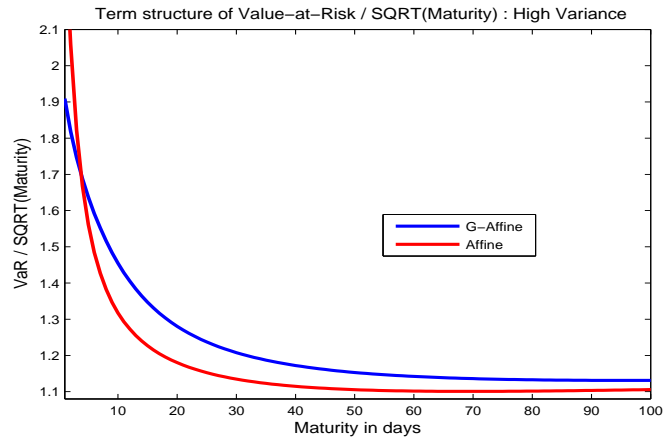
Figure 5: **Term structure of Value-at-risk Conditional on a high variance day**
We use parameters estimated from the MLE to compute the term structure of value at risk. We compared Affine and Generalized affine term structure.
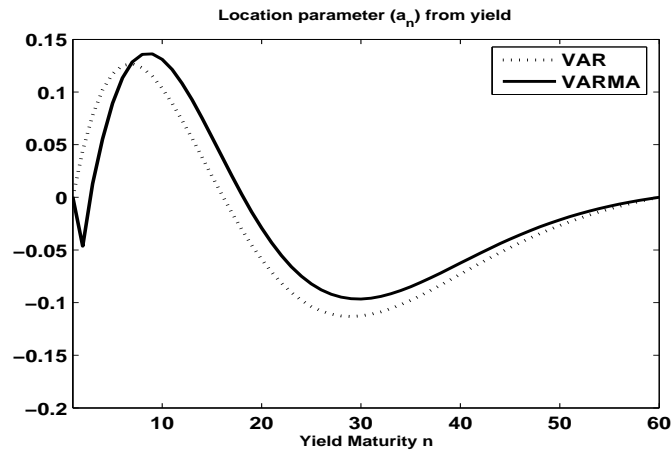
Figure 6: **Constant yield coefficient a_n for the VAR and VARMA model**
The figure displays a_n yield constant coefficient as a function of maturity n.
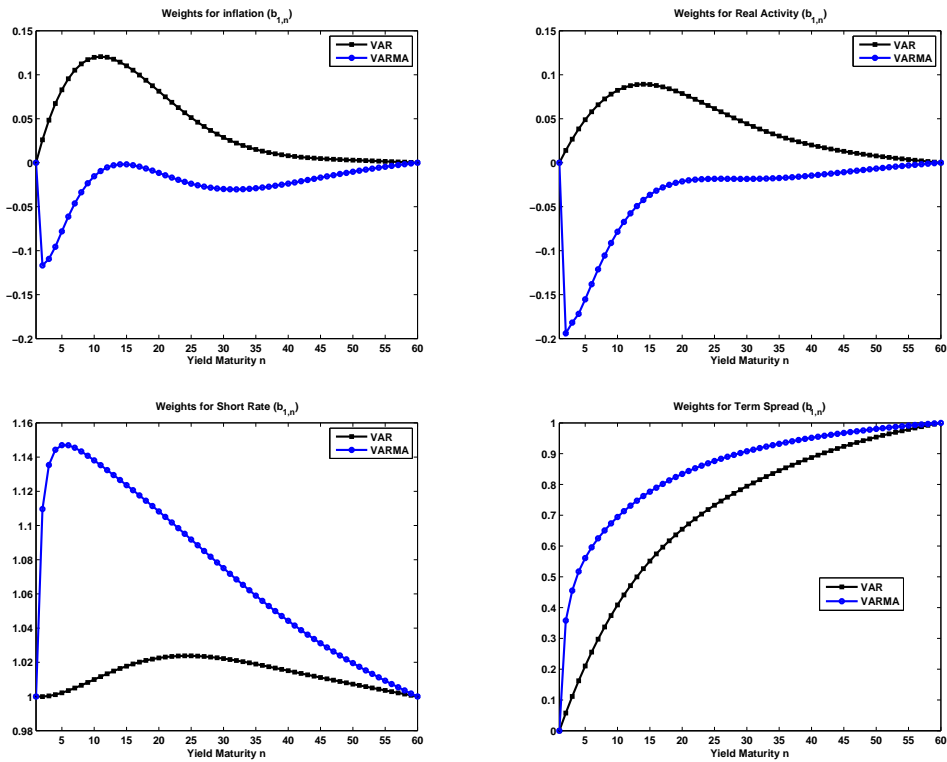
Figure 7: $b_{1,n}$ **yield weights for the VAR and VARMA model.**

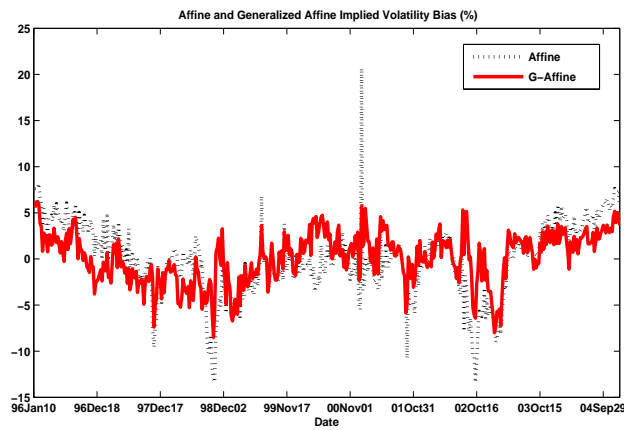The figure displays $b_{1,n}$ yield weights as a function of maturity n.
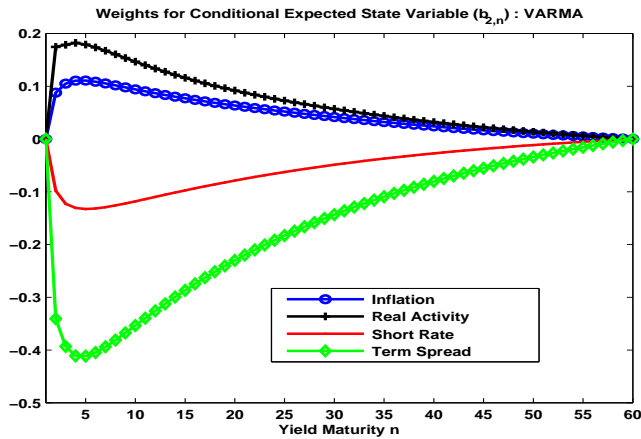
Figure 8: $b_{2,n}$ **yield weights for the VARMA model.**

The figure displays $b_{2,n}$ yield weights as a function of maturity n. Notice that these weights are zero for the VAR model.
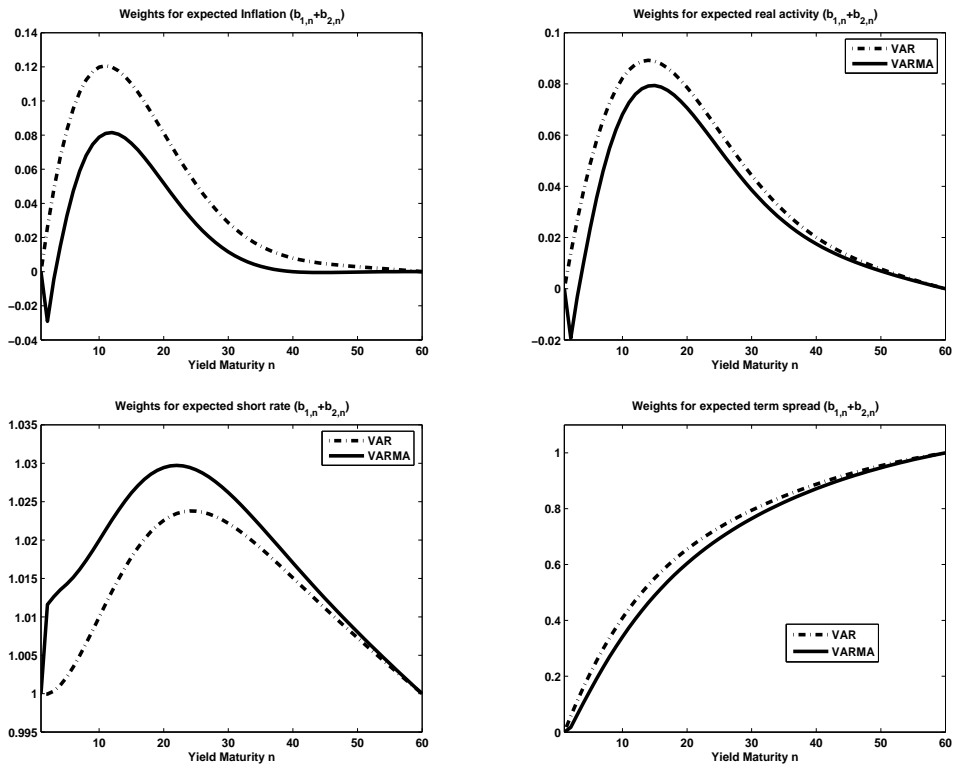
Figure 9: $b_{1,n} + b_{2,n}$ **yield weights for the VAR and VARMA model.**

The figure displays $b_{1,n} + b_{2,n}$ yield weights as a function of maturity n.

Figure 10: **Implied Volatility Bias**

The figure displays implied volatility bias as a function of the day at which option is priced. Implied volatility bias is the difference between model and observed black scholes implied volatility. For each day we compute average available Implied volatility bias
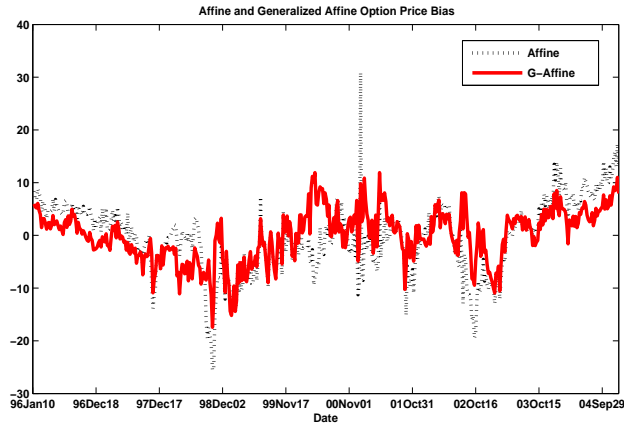
Figure 11: **Option price Bias**

The figure displays Option price bias as a function of theb day at which option is priced. Option price bias is the difference between model and observed Option price. For each day we compute average available option bias
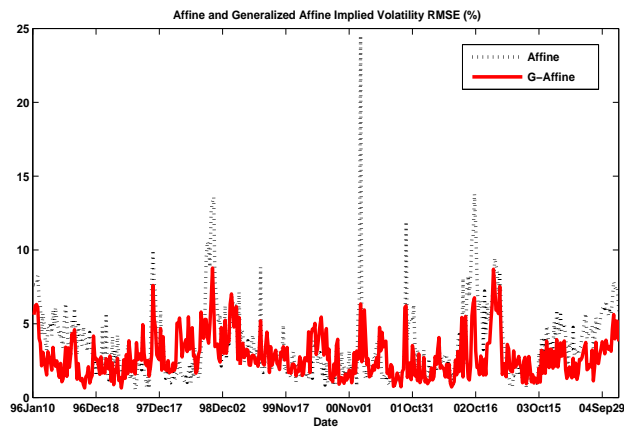

Figure 12: **Implied Volatility Root mean squared error (IVRMSE)**

The figure displays IVRMSE as a function of day at which option is priced. IVRMSE is the square-root of the average squared difference between model and observed black scholes implied volatility.