# Modelling Multivariate Interest Rates using Time-Varying Copulas and Reducible Non-Linear Stochastic Differential Equation by

Ruijun Bu, Ludovic Giet, Kaddour Hadri and Michel Lubrano

**Workshop on Financial Econometrics**,
**Field Institute, Toronto University**

April 22-23 April 2010

# Outline

- Motivations
- Succint literature review
- Non-linear multivariate modelling
- The CEV specification for the marginals
- The conditional transformation
- Bivariate modelling of US and UK short term interest rates
- Conclusions

# Motivations

- Account for non-linear features observed in short-term interest rates and other financial assets
- At the same obtain exact discretization and closed form likelihood functions
- We seek flexible functional form over time by allowing the transformation to be time varying
- Obtain multivariate models via copulas to account for possible dependence across financial assets

- **Parametric,** assume the drift and the diffusion functions known: Merton (1973), Brennan and Schwartz (1979), Vasicek (1977), Cox (1975), Dothan (1978), Cox, Ingersoll, and Ross (1980, 1985), Courtadon (1982), Constantinides and Ingersoll (1984), Constantinides (1992), Duffie and Kan (1996) and Aït-Sahalia (1996b). Majority of SDEs do not lead to closed form expressions except: Black and Scholes (1973), Vasicek (1977), and Cox, Ingersoll, and Ross (1985). Considerable energy has been employed in developing computationally and statistically efficient approximation schemes. Examples include Lo (1988), Pedersen (1995), Brandt and Santa-Clara (2002), Shoji and Ozaki (1998), Kessler (1997), Elerian et al (2001). Durham and Gallant (2002) provides a survey on existing numerical techniques.

- **Non-parametric and semi-parametric,** do not constrain the drifts and diffusion to be within a parametric class:Aït-Sahalia (1996a,b) and Stanton (1997). Aït-Sahalia (1999, 2002, 2008) developed methods for generating closed form approximation of likelihood functions for univariate and multivariate diffusions.

We start first by deriving the marginal processes employing the reduction technique. We assume that the basic dynamic model for an interest rate process $\{r_t, t \geq 0\}$ is described by a SDE

$$dr_t = \mu\left(r_t, \boldsymbol{\phi}\right) dt + \sigma\left(r_t, \boldsymbol{\phi}\right) dW_t, \tag{1}$$

The functions $\mu\left(r_t, \boldsymbol{\phi}\right)$ and $\sigma^2\left(r_t, \boldsymbol{\phi}\right)$, typically non-linear in $r_t$, are respectively the drift and the diffusion functions of the process, and $\boldsymbol{\phi} \subset R^K$ is a vector of unknown parameters. The only assumption we impose at this stage of the analysis is that (1) belongs to the class of the so called *reducible* SDEs as defined in Kloeden and Platen (1992)

The class of SDEs that have closed form transition probability density functions can be represented by

$$dx_t = (a_1 x_t + a_2)\, dt + b x_t^{\alpha} dW_t. \tag{2}$$

Special cases of interest that arise in the finance literature include: (i) the Ornstein-Uhlenbeck (OU) process $(\alpha = 0)$ which has both the transition probability density and the marginal density normally distributed; (ii) the Cox, Ingersoll and Ross (1985) (CIR) process $(\alpha = 1/2)$ which has a non-central $\chi^2$ transition density with fractional degrees of freedom and its marginal density follows a Gamma distribution; and (iii) the *Geometric Brownian motion* $(\alpha = 1)$ and $(a_2 = 0)$ which leads to a *log-normal* transition density function.

If there exists an appropriate transformation function $U(\cdot, \boldsymbol{\phi})$ such that the process $\{x_t = U(r_t, \boldsymbol{\phi}), t \geq 0\}$ follows a SDE that is solvable analytically, then the process $\{r_t, t \geq 0\}$ governed by (1) is said to be reducible. A process that is reducible to OU is called "**OU-reducible**" and a process that can be reduced to CIR is called "**CIR-reducible**"[1]. It can be shown that under minor conditions such processes would possess an explicit analytic likelihood function via a trivial transformation of the distribution. If $\partial U(r_t, \boldsymbol{\phi}) / \partial r_t \neq 0$, the *Inverse Function Theorem* ensures the existence of a local inverse $r_t = U^{-1}(x_t, \boldsymbol{\phi})$.

**The Transformation Function**

For ease of exposition, we re-write (1), with no loss of generality, in the following way:

$$dr_t = \mu\left(r_t, \boldsymbol{\phi}\right) dt + \sigma_0 \sigma\left(r_t, \boldsymbol{\theta}\right) dW_t, \tag{3}$$

where $\boldsymbol{\phi} = \left(\boldsymbol{\theta}, \sigma_0\right)'$ and $\sigma_0$ is a normalizing scalar.
Define an analytic *transformation function* $U\left(\cdot, \boldsymbol{\phi}\right)$, where typically $U\left(\cdot, \boldsymbol{\phi}\right)$ only depends on a subset of $\boldsymbol{\phi}$, and let $x_t = U\left(r_t, \boldsymbol{\phi}\right)$.
Then according to Itô's lemma, we obtain the following dynamics for $\left\{x_t, t \geq 0\right\}$:

$$\begin{aligned} dx_t &= \left[\mu\left(r_t, \boldsymbol{\phi}\right) \frac{\partial U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t} + \frac{\sigma_0^2 \sigma^2\left(r_t, \boldsymbol{\theta}\right)}{2} \frac{\partial^2 U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t^2}\right] dt \\ &\quad + \sigma_0 \sigma\left(r_t, \boldsymbol{\theta}\right) \frac{\partial U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t} dW_t. \end{aligned} \tag{4}$$

It follows that the non-linear SDEs in (3) that are reducible to (2) via transformation function $x_t = U\left(r_t, \boldsymbol{\phi}\right)$ must satisfy the following two equations:

$$\sigma_0 \sigma\left(r_t, \boldsymbol{\theta}\right) \frac{\partial U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t} = bU^{\alpha}\left(r_t, \boldsymbol{\phi}\right) \tag{5}$$

$$\mu\left(r_t, \boldsymbol{\phi}\right) \frac{\partial U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t} + \frac{1}{2}\sigma_0^2 \sigma^2\left(r_t, \boldsymbol{\theta}\right) \frac{\partial^2 U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t^2} = a_1 U\left(r_t, \boldsymbol{\phi}\right) + a_2 \tag{6}$$

It should be noted that the three unknown functional forms $\mu\left(r_t, \boldsymbol{\phi}\right)$, $\sigma\left(r_t, \boldsymbol{\theta}\right)$ and $U\left(r_t, \boldsymbol{\phi}\right)$ cannot be uniquely identified from only two equations (5) and (6) unless an additional assumption is imposed on them.

There are three approaches for dealing with this identification problem. The first approach is to start with a desired drift function $\mu\left(r_t, \boldsymbol{\phi}\right)$. This is the most difficult route as it involves solving a higher-order differential equation for either $\sigma\left(r_t, \boldsymbol{\theta}\right)$ or $U\left(r_t, \boldsymbol{\phi}\right)$. An analytic solution is hardly obtained except in very rare cases. The most general approach is to make assumptions directly on $U\left(r_t, \boldsymbol{\phi}\right)$. Then the specification of (3) or equivalently (1) can be uniquely determined under minor identification conditions. However, it is not always straightforward to formulate such a specification without prior knowledge on the desired features that the resulting SDEs should possess. A slightly less ambitious but substantially simplified approach is to start with a desired specification of the volatility function $\sigma\left(r_t, \boldsymbol{\theta}\right)$. Then, finding $U\left(r_t, \boldsymbol{\phi}\right)$ will only involve solving the first-order differential equation in (5). The drift function $\mu\left(r_t, \boldsymbol{\phi}\right)$ can then be trivially inferred from equation (6). Given the significance of the volatility properties in financial applications, this approach appears to be fairly reasonable.

For a chosen standardized volatility function $\sigma\left(r_t, \boldsymbol{\theta}\right)$ and letting $b = \sigma_0$, the transformation function $U\left(\cdot, \boldsymbol{\phi}\right)$ can be found by solving the following ordinary differential equation:

$$\frac{\partial U\left(r_t, \boldsymbol{\phi}\right)}{\partial r_t} = U^{\alpha}\left(r_t, \boldsymbol{\phi}\right)\frac{1}{\sigma\left(r_t, \boldsymbol{\theta}\right)},$$

which yields for $\alpha \neq 1$

$$U\left(r_t, \boldsymbol{\phi}\right) = U\left(r_t, \boldsymbol{\theta}\right) = \left\{(1-\alpha)\left[\int \frac{1}{\sigma\left(r_t, \boldsymbol{\theta}\right)} dr_t + c\right]\right\}^{\frac{1}{1-\alpha}},$$

where $c$ is the *constant of integration*. Note here that the transformation function only depends on $\boldsymbol{\theta}$. When $\alpha = 0$, the original process is reducible to the OU process and the required transformation is given by

$$U\left(r_t, \boldsymbol{\theta}\right) = \int \frac{1}{\sigma\left(r_t, \boldsymbol{\theta}\right)} dr_t + c. \tag{7}$$

When $\alpha = 1/2$, the original process is reducible to CIR and the corresponding transformation is

$$U\left(r_t, \boldsymbol{\theta}\right) = \left\{ \frac{1}{2} \left[ \int \frac{1}{\sigma\left(r_t, \boldsymbol{\theta}\right)} dr_t + c \right] \right\}^2 . \tag{8}$$

Replacing the transformation function and its first and second derivatives reveals the non-linear drift function $\mu\left(r_t, \boldsymbol{\phi}\right)$. The complete specification of the process can then be written as

$$dr_t = \frac{1}{\frac{\partial U\left(r_t, \boldsymbol{\theta}\right)}{\partial r_t}} \left[ a_1 U\left(r_t, \boldsymbol{\theta}\right) + a_2 - \frac{1}{2} b^2 \sigma^2\left(r_t, \boldsymbol{\theta}\right) \frac{\partial^2 U\left(r_t, \boldsymbol{\theta}\right)}{\partial r_t^2} \right] dt + b\sigma\left(r_t, \boldsymbol{\theta} \right. \tag{9}$$

where $U\left(r_t, \boldsymbol{\theta}\right)$ is given by either (7) or (8). Note that the unknown parameter vector $\boldsymbol{\phi}$ is in fact identified as $\boldsymbol{\phi} = \left(\boldsymbol{\theta}', a_1, a_2, b\right)'$.

.Using reducible SDEs as a modelling tool has the following advantages. Firstly, since the non-linear diffusion process in (9) is a transformed process of a basic process, either OU or CIR, via a transformation function (7) or (8), many useful mathematical and statistical properties of the basic processes are preserved after the transformation. For instance, since both OU and CIR processes have exact discretization, the process in (9) also has exact discretization as a result of straight forward mapping by function $r_t = U^{-1}\left(x_t, \boldsymbol{\phi}\right)$.

For OU-reducible and CIR-reducible processes, the Jacobians of the transformations are given by

$$\mathbf{J}_{OU} = \left|\partial U\left(r_t\right) / \partial r_t\right| = 1 / \sigma\left(r_t, \theta\right),$$

and

$$\mathbf{J}_{CIR} = \left|\partial U\left(r_t\right) / \partial r_t\right| = \left[1 / 2\sigma\left(r_t, \theta\right)\right] \left|\int 1 / \sigma\left(r_t, \theta\right) dr_t\right|,$$

respectively. The transition density for the proposed model can be easily obtained by the standard transformation method of the distribution. Monotonicity in $U\left(r_t, \boldsymbol{\theta}\right)$ ensures that the transformation is unique. The corresponding marginal density function can be obtained by taking the step length $\Delta$ to the limit providing that the process is stationary and therefore the limit exists.

Focusing on the bivariate case, we denote $F_X(x)$ and $F_Y(y)$ as the continuous marginal distribution functions of $X$ and $Y$, and $F_{XY}(x,y)$ the joint distribution function. Also let $f_x(x)$ and $f_y(y)$ be the marginal density functions, and $f_{xy}(x,y)$ the joint probability density function. The Sklar's Theorem states:

$$
\begin{aligned}
F_{XY}(x,y) &= C(F_X(x), F_Y(y)) \\
f_{xy}(x,y) &= f_x(x) \cdot f_y(y) \cdot c(F_X(x), F_Y(y)),
\end{aligned}
$$

where $C : [0,1]^2 \longrightarrow [0,1]$ is the copula function for the bivariate random vector $(X, Y)$, and $c$ is the corresponding copula density. The procedure employed to construct the joint distribution is a two-step method of estimation. In the first stage we estimate the two marginal distribution models separately, and in the second stage we estimate the copula model. The estimates obtained in two-steps are consistent and asymptotically normal ( see Patton (2006b) for more details).

The constant elasticity volatility model was introduced by Chan et al. (1992). It was further studied by Aït-Sahalia (1996b) who promoted the use of a non-linear drift function to provide a better mean-reversion effect. The CEV specification of the diffusion is given by $\sigma(r_t, \gamma) = r_t^{\gamma}$, where $\gamma \in (0, 1) \cup (1, \infty)$. It follows from (7) that for a non-linear CEV process that is reducible to OU, henceforth denoted as OU-CEV, the transformation is given by

$$x_t = U(r_t, \gamma) = r_t^{1-\gamma} / (1 - \gamma) \,. \tag{10}$$

Aït-Sahalia (1999) suggests to define

$$x_t = U(r_t, \gamma) = r_t^{1-\gamma} / (\gamma - 1) \,, \tag{11}$$

for $\gamma > 1$.

It is easily verified that $\partial U\left(r_t, \gamma\right) / \partial r_t = r_t^{-\gamma}$. Since $r_t \in \mathbb{R}^+$, the above transformation is always strictly monotonic, which ensures identification of all parameters. It follows from (9) that the dynamics of OU-CEV process is governed by the following diffusion.

$$dr_t = \left(\frac{1}{2} b^2 \gamma r_t^{2\gamma-1} - a_2 r_t^{\gamma} \operatorname{sgn}\left(\gamma - 1\right) + \frac{a_1}{1-\gamma} r_t\right) dt + b r_t^{\gamma} dW_t, \tag{12}$$

where $\operatorname{sgn}(\cdot)$ is the sign function.

For a non-linear CEV model that is reducible to CIR, henceforth denoted by CIR-CEV, we have

$$x_t = U(r_t) = (1/4)\left[r_t^{1-\gamma}/(1-\gamma)\right]^2 . \tag{13}$$

It is easily verified that $\partial U(r_t)/\partial r_t = r_t^{1-2\gamma}/(2-2\gamma)$. For $r_t \in \mathbb{R}^+$, the above transformation is also strictly monotonic. The dynamics of the CIR-CEV process is therefore given by

$$dr_t = \left\{\left[2a_2(1-\gamma) + \frac{1}{2}b^2(2\gamma-1)\right] r_t^{2\gamma-1} + \frac{a_1 r_t}{(2-2\gamma)}\right\} dt + b r_t^{\gamma} dW . \tag{14}$$

The SDEs defined in (12) and (14) encompasses a number of interest rate processes that are known to have closed form likelihood functions. These models can be obtained from (12) and (14) by simply placing the appropriate restrictions on the four parameters, $a_1$, $a_2$, $b$, and $\gamma$. *Table 1* provides the specifications of nested models and the corresponding restrictions.

Table 1: Models Nested in OU-CEV and CIR-CEV Specifications

| | Models Nested in OU-CEV | |
|---|---|---|
| Model | Specification | Restriction(s) |
| 1. Merton | $dr_t = a_2 dt + bdW_t$ | $\gamma = 0, a_1 = 0$ |
| 2. OU (Vasicek) | $dr_t = (a_1 r_t + a_2)\, dt + bdW_t$ | $\gamma = 0$ |
| | **Models Nested in CIR-CEV** | |
| Model | Specification | Restriction(s) |
| 3. CIR (SR) | $dr_t = (a_1 r_t + a_2)\, dt + b\sqrt{r_t} dW_t$ | $\gamma = 1/2$ |
| 4. CIR (VR) | $dr_t = br_t^{3/2} dW_t$ | $\gamma = 3/2, a_1 = 0, a_2 = b^2$ |
| 5. CEV | $dr_t = a_1 r_t dt + br_t^{\gamma} dW_t$ | $a_2 = b^2 (2\gamma - 1) / 4 (\gamma - 1)$ |
| 6. AG | $dr_t = \left[ (b^2 - a_2)\, r_t^2 - a_1 r_t \right] dt + br_t^{3/2} dW$ | $\gamma = 3/2$ |

## Analysis of the Distributions
## CIR-CEV

We apply the methodology of Aït-Sahalia (1996b) concerning the constraints on the drift and the diffusion to the CIR-CEV model in (14) to derive the sufficient conditions for stationarity and unattainability of 0 and $\infty$ in finite expected time. The results are given in the following proposition.

**Proposition 1** *Let* $\{r_t, t \geq 0\}$ *be a CIR-CEV process defined in (14). The necessary and sufficient conditions for stationarity and unattainability of* $0$ *and* $\infty$ *in finite expected time are: (i)* $a_1 < 0$ *and* $4a_2 / b^2 > (2\gamma - 1) / (\gamma - 1)$ *if* $\gamma > 1$*; (ii)* $a_1 < 0$ *and* $4a_2 / b^2 > 1/ (1 - \gamma)$ *if* $\gamma < 1$.

**Proof.** See Appendix ■

For the CIR-CEV process, according to the transformation in (13) the transition density of the process is given by

$$f\left(r_t | r_{t-\Delta}\right) = \frac{1}{2} \frac{r_t^{1-2\gamma}}{|1-\gamma|} c e^{-u-v} \left(\frac{v}{u}\right)^{q/2} I_q \left[2\left(uv\right)^{1/2}\right],$$

where

$$\begin{aligned} c &= 2a_1 \Big/ \left[b^2\left(e^{a_1\Delta}-1\right)\right], u = \left(ce^{a_1\Delta}/4\right)\left[r_{t-\Delta}^{1-\gamma}/(1-\gamma)\right]^2 \\ v &= (c/4)\left[r_t^{1-\gamma}/(1-\gamma)\right]^2, q = 2a_2 \big/ b^2 - 1. \end{aligned}$$

and $I_q(\cdot)$ is the modified Bessel function of the first kind of order $q$.

The CIR-CEV process also permits a closed form expression for its conditional distribution function which can be written as

$$F\left(r_t|r_{t-\Delta}\right)=\left\{\begin{array}{c} D\left(2cx_t;2q+2,2u\right) \text{ for } \gamma<1 \\ 1-D\left(2cx_t;2q+2,2u\right) \text{ for } \gamma>1 \end{array}\right. ,$$

where $x_t$ is defined by (13) and $D\left(\cdot;2q+2,2u\right)$ is the non-central $\chi^2$ distribution function with $2q+2$ degrees of freedom and non-centrality parameter $2u$.

Straightforward calculation yields the $m$th conditional moments for $r_t$ following the CIR-CEV process:

$$E\left(r_t^m|r_{t-\Delta}\right)$$

$$=\left[2\left|1-\gamma\right|\right]^{\frac{m}{1-\gamma}} c^{-\frac{m}{2(1-\gamma)}} e^{-u}\frac{\Gamma\left(q+\frac{m}{2(1-\gamma)}+1\right)}{\Gamma(1+q)}$$

$${}_1F_1\left(q+\frac{m}{2(1-\gamma)}+1,1+q,u\right)$$

where ${}_1F_1\left(\cdot,\cdot,\cdot\right)$ is the *confluent hypergeometric function* and $\Gamma(\cdot)$ is the gamma function.

Since the CIR-CEV process displays mean reversion, then as $\Delta \to \infty$, its distribution is well defined. It can be shown that the steady-state density function is given by

$$\pi\left(r_t\right) = \frac{1}{2}\frac{r_t^{1-2\gamma}}{\left|1-\gamma\right|}\frac{\left(-\frac{2a_1}{b^2}\right)^{\frac{2a_2}{b^2}}}{\Gamma\left(\frac{2a_2}{b^2}\right)} x_t^{\frac{2a_2}{b^2}-1}\exp\left(\frac{2a_1}{b^2}x_t\right),$$

and the $m$th unconditional moments are

$$E\left[r_t^m\right] = \left[2\left|1-\gamma\right|\right]^{\frac{m}{1-\gamma}}\left(-\frac{2a_1}{b^2}\right)^{-\frac{m}{2(1-\gamma)}}\frac{\Gamma\left(\frac{2a_2}{b^2}+\frac{m}{2(1-\gamma)}\right)}{\Gamma\left(\frac{2a_2}{b^2}\right)}.$$

## OU-CEV

The analysis of the OU-CEV process is less straightforward than that of the CIR-CEV process. For the OU-CEV process, according to the transformations defined in (10) and (11) the probability density of the interest rate $r_t$ conditional on $r_{t-\Delta}$, where $\Delta$ is the step length, is given by the following

$$f\left(r_t \middle| r_{t-\Delta}\right) = r_t^{-\gamma} \frac{1}{\sqrt{2\pi\sigma_{ou}^2}} \exp\left[-\frac{1}{2}\left(\frac{x_t-\mu_{ou}}{\sigma_{ou}}\right)^2\right],$$

where

$$\mu_{ou} = e^{a_1\Delta} x_{t-\Delta} - \frac{a_2}{a_1}\left(1-e^{a_1\Delta}\right) \text{ and} \tag{15}$$

$$\sigma_{ou}^2 = \frac{b^2\left(e^{2a_1\Delta}-1\right)}{2a_1}. \tag{16}$$

Since the OU-CEV process is a continuous and monotonic transformation of the OU process, it has a closed form expression for its conditional distribution function which is given by

$$F\left(r_t | r_{t-\Delta}\right) = \begin{cases} \Phi\left(x_t; \mu_{ou}, \sigma_{ou}\right) - \Phi\left(0; \mu_{ou}, \sigma_{ou}\right) & \text{for } \gamma < 1 \\ 1 - \Phi\left(x_t; \mu_{ou}, \sigma_{ou}\right) & \text{for } \gamma > 1 \end{cases},$$

where $\Phi\left(\cdot; \mu_{ou}, \sigma_{ou}\right)$ is the distribution function for a normally distributed random variable with mean $\mu_{ou}$ and standard deviation $\sigma_{ou}$. Here $x_t$ is defined by (10) for $\gamma < 1$ and (11) for $\gamma > 1$, respectively.

Straightforward calculation also yields the $m$th conditional moments for $r_t$ following the OU-CEV process

$$E\left(r_t^m | r_{t-\Delta}\right) = \frac{\left[\sqrt{2}\sigma_{ou}\left|1-\gamma\right|\right]^v}{\sqrt{\pi}} \frac{1}{2} e^{-w^2} \left[\Gamma\left(\frac{v+1}{2}\right) \times {}_1F_1\left(\frac{v+1}{2},\right.\right.$$

$$\left.\left. + wv\Gamma\left(\frac{v}{2}\right) {}_1F_1\left(1+\frac{v}{2}, \frac{3}{2}, w^2\right)\right],$$

where

$$v = \frac{m}{1-\gamma}, \ w = \frac{\mu_{ou}}{\sqrt{2}\sigma_{ou}}.$$

Since the unconditional distribution of the OU process is also normal, the marginal density $\pi\left(r_t\right)$ and the unconditional moments $E\left(r_t^m\right)$ have similar expressions to their conditional counterparts. The only difference is that we will have to replace the conditional mean and variance in (15) and (16) by their corresponding limits as $\Delta \to \infty$.

For both OU-CEV and CIR-CEV processes, the transformation functions depend on a single parameter $\gamma$. A natural extension is to specify an equation describing the evolution over time of the parameter $\gamma_t$. We propose the following:
To reflect the $\gamma < 1$ case in the time-homogenous transformation design, we can define

$$\gamma_t = \Lambda\left(\omega + \sum_{i=1}^{p} \alpha_j \gamma_{t-i} + \sum_{j=1}^{q} \beta_i x_{t-j}\right) \tag{17}$$

where $\Lambda(x) \equiv \left(1 + e^{-x}\right)^{-1}$ is the logistic transformation, used to keep $\gamma_t$ in $(0, 1)$ all the time. Similarly, we can let

$$\gamma_t = \widetilde{\Lambda}\left(\omega + \sum_{i=1}^{p} \alpha_j \gamma_{t-i} + \sum_{j=1}^{q} \beta_i x_{t-j}\right) \tag{18}$$

where $\widetilde{\Lambda}(x) \equiv (1 + e^{-x})$ is the reciprocal of the the logistic transformation, used to keep $\gamma_t$ inside the range $(1, \infty)$ all the time. In practice, the choice of $p$ and $q$ can be decided by some

**Data**

We measure the US and UK short term interest rates by 1-Month Eurodollar Rate (EDR) and 1-Month London Interbank Offered Rate (LIBOR) in British Sterling. Two different frequencies, monthly and weekly, of the two rates are employed in this study. The EDR data are collected from the H.15 release of the Federal Reserve website and the data of LIBOR are obtained from BBA (British Banking Association) database. For each of the two rates, we use the longest sample period for which data are available.

Table 2: Descriptive Statistics of EDR and LIBOR Data

| | Eurodollar Rate | | LIBOR | |
|---|---|---|---|---|
| Sample period | 1971.01-2007.12 | | 1986.01-2007.12 | |
| Frequency | Monthly | Weekly | Monthly | Weekly |
| Sample size | 444 | 1930 | 264 | 1148 |
| Mean | 6.778 | 6.781 | 7.434 | 7.413 |
| Std. Dev. | 3.541 | 3.550 | 3.237 | 3.214 |
| Skewness | 1.081 | 1.099 | 1.013 | 1.003 |
| Kurtosis | 4.719 | 4.809 | 2.875 | 2.872 |
| Jarque-Bera Statistic | 141.149* | 651.979* | 45.330* | 193.280* |
| $\rho_1$ | 0.978 | 0.996 | 0.987 | 0.997 |
| $\rho_2$ | 0.944 | 0.990 | 0.971 | 0.995 |
| $\rho_3$ | 0.916 | 0.983 | 0.955 | 0.992 |
| $\rho_4$ | 0.891 | 0.975 | 0.940 | 0.989 |
| $\rho_5$ | 0.869 | 0.967 | 0.927 | 0.986 |
| $\rho_6$ | 0.851 | 0.959 | 0.909 | 0.982 |

An asterisk (*) indicates a rejection of the null hypothesis at the 0.01 level.

Figure 1: Time Series of Monthly and Weekly Eurodollar Rate and LIBOR

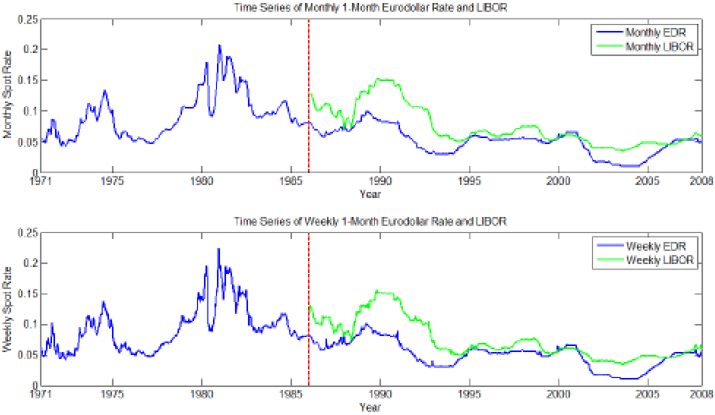


Table 3: Maximum Likelihood Estimates of Different Interest Rate Models

| | | Monthly | | | | | Weekly | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | OU | CIR | AG | OU-CEV | CIR-CEV | OU | CIR | AG | OU-CEV | CIR-CEV |
| EDR | $a_1$ | -0.259 | -0.148 | -0.105 | -0.090 | -0.100 | -0.197 | -0.120 | -0.080 | -0.070 | -0.079 |
| | | (0.120) | (0.090) | (0.076) | (0.070) | (0.074) | (0.104) | (0.080) | (0.067) | (0.062) | (0.066) |
| | $a_2$ | 1.745 | 0.992 | 0.023 | 0.355 | 0.389 | 1.301 | 0.782 | 0.018 | 0.226 | 0.206 |
| | | (0.916) | (0.509) | (0.013) | (0.308) | (0.382) | (0.796) | (0.452) | (0.012) | (0.200) | (0.172) |
| | $b$ | 2.555 | 0.782 | 0.117 | 0.193 | 0.193 | 2.245 | 0.698 | 0.104 | 0.164 | 0.163 |
| | | (0.087) | (0.026) | (0.004) | (0.017) | (0.017) | (0.036) | (0.011) | (0.002) | (0.007) | (0.007) |
| | $\gamma$ | $(\gamma = 0)$ | $(\gamma = 0.5)$ | $(\gamma = 1.5)$ | 1.184 | 1.186 | $(\gamma = 0)$ | $(\gamma = 0.5)$ | $(\gamma = 1.5)$ | 1.218 | 1.219 |
| | | | | | (0.047) | (0.047) | | | | (0.024) | (0.024) |
| | *AIC* | 983.82 | 718.99 | 599.88 | 554.35 | 554.13 | 971.78 | -131.20 | -693.36 | -841.98 | -842.18 |
| | *BIC* | 996.11 | 731.28 | 612.17 | 570.73 | 570.51 | 988.47 | -114.51 | -676.67 | -819.72 | -819.92 |
| LIBOR | $a_1$ | -0.142 | -0.174 | -0.098 | -0.124 | -0.096 | -0.115 | -0.147 | -0.090 | -0.113 | -0.096 |
| | | (0.087) | (0.091) | (0.081) | (0.083) | (0.082) | (0.083) | (0.087) | (0.080) | (0.081) | (0.080) |
| | $a_2$ | 0.740 | 0.981 | 0.020 | 0.086 | 0.012 | 0.590 | 0.829 | 0.018 | 0.198 | 0.075 |
| | | (0.707) | (0.625) | (0.012) | (0.066) | (0.014) | (0.674) | (0.596) | (0.012) | (0.143) | (0.067) |
| | $b$ | 1.315 | 0.442 | 0.058 | 0.051 | 0.051 | 1.257 | 0.423 | 0.057 | 0.079 | 0.079 |
| | | (0.058) | (0.019) | (0.003) | (0.013) | (0.013) | (0.026) | (0.009) | (0.001) | (0.009) | (0.009) |
| | $\gamma$ | $(\gamma = 0)$ | $(\gamma = 0.5)$ | $(\gamma = 1.5)$ | 1.565* | 1.565* | $(\gamma = 0)$ | $(\gamma = 0.5)$ | $(\gamma = 1.5)$ | 1.331 | 1.330 |
| | | | | | (0.132) | (0.133) | | | | (0.060) | (0.060) |
| | *AIC* | 239.90 | 170.95 | 109.54 | 111.08 | 111.31 | -748.07 | -1047.58 | -1234.46 | -1240.55 | -1240.39 |
| | *BIC* | 250.63 | 181.67 | 120.27 | 125.38 | 125.61 | -732.94 | -1032.44 | -1219.33 | -1220.37 | -1220.21 |

*Only cases where the hypothesis $\gamma = 1.5$ cannot be rejected. In all other cases, $\gamma = 0$, 0.5, or 1.5 are all rejected.

Table 4: ML Estimates of Time-varying Transformation Models

| | | Monthly | | Weekly | |
|---|---|---|---|---|---|
| | | TV-OU-CEV | TV-CIR-CEV | TV-OU-CEV | TV-CIR-CEV |
| EDR | | $(p=1)$ | $(p=1)$ | $(p=1)$ | $(p=1)$ |
| | $a_1$ | -0.100 | -0.114 | -0.091 | -0.107 |
| | | (0.074) | (0.079) | (0.070) | (0.076) |
| | $a_2$ | -0.398 | 0.441 | -0.282 | 0.258 |
| | | (0.295) | (0.300) | (0.218) | (0.178) |
| | $b$ | 0.187 | 0.186 | 0.159 | 0.159 |
| | | (0.006) | (0.006) | (0.003) | (0.003) |
| | $\omega$ | -2.105 | -2.087 | 3.522 | 3.520 |
| | | (0.697) | (0.693) | (0.468) | (0.465) |
| | $\alpha$ | 3.184 | 3.158 | -1.667 | -1.666 |
| | | (0.592) | (0.587) | (0.383) | (0.380) |
| | $\beta$ | 0.004 | 0.004 | 0.003 | 0.003 |
| | | (0.001) | (0.001) | (0.000) | (0.000) |
| | *AIC* | 526.90 | 526.67* | -930.19 | -930.49* |
| | *BIC* | 551.48 | 551.25* | -896.80 | -897.10* |
| LIBOR | | $(p=5)$ | $(p=5)$ | $(p=8)$ | $(p=8)$ |
| | $a_1$ | -0.228 | -0.223 | -0.208 | -0.198 |
| | | (0.141) | (0.146) | (0.122) | (0.123) |
| | $a_2$ | -0.124 | 0.017 | -0.336 | 0.131 |
| | | (0.072) | (0.009) | (0.190) | (0.074) |
| | $b$ | 0.046 | 0.046 | 0.077 | 0.077 |
| | | (0.002) | (0.002) | (0.002) | (0.002) |
| | $\omega$ | -1.768 | -1.766 | 1.924 | 1.903 |
| | | (0.310) | (0.311) | (1.566) | (1.715) |
| | $\alpha$ | 1.288 | 1.286 | -0.662 | -0.645 |
| | | (0.201) | (0.202) | (1.154) | (1.264) |
| | $\beta$ | 0.025 | 0.026 | 0.006 | 0.006 |
| | | (0.006) | (0.006) | (0.003) | (0.003) |
| | *AIC* | 95.06 | 95.00* | -1264.98* | -1264.95 |
| | *BIC* | 116.51 | 116.46* | -1234.71* | -1234.67 |

# Empirical analysis based on the time-varying SJC copula.

Table 5: ML Estimates of Time-varying SJC Copula

| | Time-varying SJC copula model ($p = q = 8$) | |
|---|---|---|
| | Unrestricted model | Restricted model |
| $\omega_U$ | 3.646 | 4.464 |
| | (1.554) | (0.540) |
| $\alpha_U$ | 0.628 | |
| | (1.303) | |
| $\beta_U$ | -35.565 | -40.197 |
| | (9.030) | (7.208) |
| $\omega_L$ | 3.126 | 1.845 |
| | (1.282) | (1.261) |
| $\alpha_L$ | -2.512 | |
| | (1.838) | |
| $\beta_L$ | -31.991 | -26.447 |
| | (8.723) | (8.511) |
| $AIC$ | -123.96 | -126.44 |
| $BIC$ | -93.70 | -106.26 |

Figure 2: Conditional Upper and Lower Tail Dependences in the SJC Copula

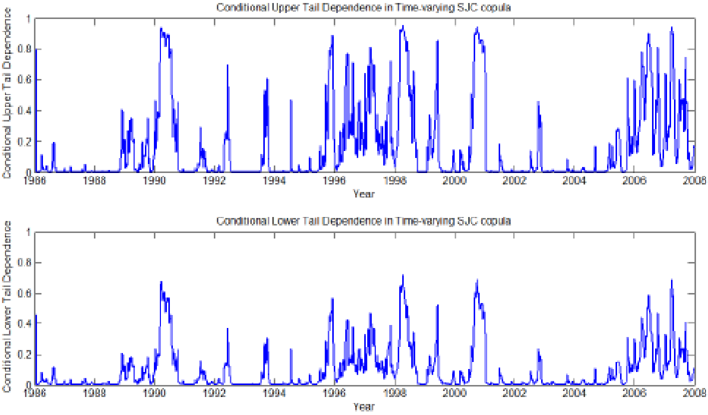
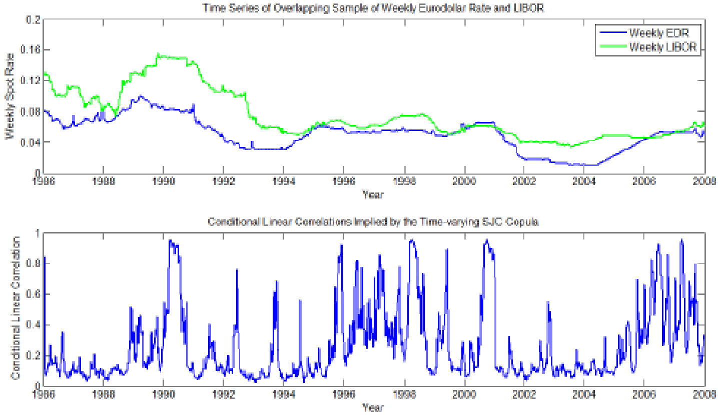


Figure 3: Overlapping Series and Conditional Linear Correlations in the Time-varying SJC Copula



We develop a copula-based non-linear multivariate interest rates models that account simultaneously for observed non-linearities and correlation across short-term interest rates. The dynamics of the marginal processes are governed by a special type of SDEs, called reducible SDEs. The use of reducible SDEs for modelling financial variables, such as the short term interest rates, has a number of advantages: exact discretization, closed form transition density functions and use of copula based multivariate modelling. The sufficient conditions for the stationarity of the CIR-CEV process are provided and the same issue for the OU-CEV process is discussed. We focused our attention on the OU-CEV and CIR-CEV models. These simple specifications encompass most existing parametric models that have closed form likelihood functions: OU, CIR and the Ahn and Gao (1999) model. The transition density, the conditional distribution function, the steady-state density function are derived in closed form as well as the conditional and unconditional moments for both processes.

In our empirical studies of monthly and weekly US and UK short term interest rates, we found that simple parametric models like OU and CIR are strongly rejected by the data under their more general CEV frameworks. The AG model is also rejected by all but the monthly LIBOR data. Hence, our new models outperform, in most cases, existing parametric models endowed with closed form likelihood functions. To generate more flexible dynamics, we extended our theory to allow for conditioning variables in the transformation functions. We found that in all four cases the time-varying effects of the transformation parameter are significant.

The dependence of the US and UK short rates were studied via a conditional copula. We found that the time-varying effect in the conditional SJC copula is significant. Also significant is the asymmetry in the tail dependence implied by the copula. From the fitted tail dependence coefficients, we found that the evolution of the conditional tail dependencies appear to coincide with that of interest rates themselves. That is, the tail dependencies tend to be higher when the interest rates are relatively high, and lower in the opposite situation. Similar relationship is also found in the conditional linear correlation coefficients implied by the conditional copula.