On deterministic–shift extensions of short–rate models *

Damiano Brigo Fabio Mercurio
Product and Business Development Group
Banca IMI, SanPaolo IMI Group
Corso Matteotti 6
20121 Milano, Italy
Fax: + 39 02 76019324
E-mail: {brigo,fmercurio}@bancaimi.it

First version: October 15, 1998. This version: August 3, 2001

Abstract

In the present paper we show how to extend any time–homogeneous short–rate model to a model which can reproduce any observed yield curve, through a procedure that preserves the possible analytical tractability of the original model. In the case of the Vasicek (1977) model, our extension is equivalent to that of Hull and White (1990), whereas in the case of the Cox-Ingersoll-Ross (1985) (CIR) model, our extension is more analytically tractable and avoids problems concerning the use of numerical solutions. Our approach can also be applied to the Dothan (1978) or Rendleman and Bartter (1980) model, thus yielding a shifted lognormal short–rate model which fits any given yield curve and for which there exist analytical formulae for prices of zero coupon bonds. We also consider the extension of time–homogeneous models without analytical formulae but whose tree–construction procedures are particularly appealing, such as the exponential Vasicek's. We explain why the CIR model is the more interesting model to be extended through our procedure. We also give explicit analytical formulae for bond options, and we explain how the model can be used for Monte Carlo evaluation of European path–dependent interest–rate derivatives. We finally hint at the same extension for multifactor models and explain its strong points for concrete applications.

*We are grateful to Aleardo Adotti, our current head at the Product and Business Development Group of Banca IMI, for encouraging us in the prosecution of the most speculative side of research in mathematical finance. At different times, numerical implementations have been made feasible thanks to the cooperation of Gianvittorio Mauri, Francesco Rapisarda and Giulio Sartorelli. A fruitful discussion with Marco Avellaneda and the helpful suggestions and remarks from an anonymous referee are gratefully acknowledged. This paper is available both at http://web.tiscali.it/damianohome and at http://web.tiscali.it/FabioMercurio. A short version of this paper has appeared in Finance and Stochastics, and related contents are available in our book *Interest Rate Models: Theory and Practice*

Keywords

Short–rate models, Analytical tractability, Yield–Curve fitting, Vasicek's model, Dothan's model, Cox-Ingersoll-Ross' model, Longstaff and Schwartz's model, Monte Carlo evaluation.

1 Introduction

The issue of pricing interest-rate derivatives has been addressed by the financial literature in a number of different ways. One of the oldest approaches is based on modelling the evolution of the instantaneous spot interest rate (shortly referred to as "short rate") and goes back to Merton (1973) and especially to Vasicek (1977). In both works, this rate is assumed to be normally distributed, thus having the theoretical possibility to become negative. To overcome this drawback, Dothan (1978) and Rendleman and Bartter (1980) proposed a lognormal distribution for the instantaneous spot interest rate. Cox, Ingersoll and Ross (1985) (CIR) proposed instead a noncentral chi-square distribution. All these models are "endogenous term structure" models, in that the initial term structure of interest rates is an output of the model rather than an input as observed in the financial market. This problem has been addressed in the continuous-time limit of the Ho and Lee (1986) model derived by Dybvig (1988) or Jamshidian (1988), in the Hull and White (1990) model and in the Black and Karasinski (1991) model, among others. In particular, Hull and White (1990) proposed extensions of both the Vasicek (1977) model and the Cox, Ingersoll and Ross (1985) model. However, while the first extension is quite natural and analytically tractable, the latter is less straighforward. In fact, the Hull and White extended CIR is not always capable to exactly fit an observed term structure of interest rates, and its analytical tractability can be easily questioned.

In this paper, we propose a simple method to extend any time–homogeneous short–rate model, so as to exactly reproduce any observed term structure of interest rates while preserving the possible analytical tractability of the original model. In the case of the Vasicek (1977) model, our extension is perfectly equivalent to that of Hull and White (1990). In the case of the Cox-Ingersoll-Ross (1985) model, instead, our extension is more analytically tractable and avoids problems concerning the use of numerical solutions. We are in fact able to exactly fit any observed term structure of interest rates and to derive analytical formulae both for pure discount bonds and for European bond options. The unique drawback is that in principle we can guarantee the positivity of rates only through restrictions on the paramters which might worsen the quality of the calibration to caps/floors or swaption prices.

The CIR model is the most relevant case to which our procedure can be applied. Indeed, our extension yields the unique short–rate model, to the best of our knowledge, featuring the following three properties:

- exact fit of any observed term structure;
- analytical formulae for bond prices, bond option prices, swaptions and caps prices;
- the distribution of the istantaneous spot rate has tails which are fatter than in the Gaussian case and, through restriction on the parameters, it is always possible to guarantee positive rates without worsening the volatility calibration in most situations.

Moreover, one further property of our extended model is that the term structure is affine in the short rate. The above uniqueness is the reason why we devote more space to the CIR case. We also explain how this model can be used (and has been used by us) for Monte Carlo evaluation of european path–dependent interest–rate derivatives.

Our extension procedure is also applied to the Dothan (1978) model (equivalently the Rendleman and Bartter (1980) model), thus yielding a shifted lognormal short–rate model which fits any given yield curve and for which there exist analytical formulae for zero coupon bonds.

Though conceived for analytically tractable models, the method we propose in this paper can be employed to extend more general time-homegenous models. As a clarification, we consider the example of an original short rate process that evolves as the exponential of a time-homegenous Ornstein-Uhlenbeck process (shortly referred to as "exponential Vasicek"). The only requirement that is needed in general is a numerical procedure for pricing interest rate derivatives under the original model.

It should be remarked that our approach is similar to those of Scott (1995), Dybvig (1997), and Avellaneda and Newman (1998). However, our analysis is richer in details and in the implications being considered. Our model seems also to be related to the Schmidt (1997) framework, even though in such a framework one works with the squared Gaussian model rather than with the general CIR dynamics.

A final remark is due to the fact that modelling the instantaneous spot rate seems somehow anachronistic as far as the theory of interest rates is concerned. Most recent works dealt in fact first with the modelling of instantaneous forward rates (e.g., Heath, Jarrow and Morton (1992) (HJM)) and later on with forward Libor or swap rates (Miltersen, Sandmann and Sondermann (1997), Brace, Gatarek and Musiela (1997), and Jamshidian (1997)). However, the spot-rate approach is still quite popular among practitioners both for pricing interest rate derivatives and for risk management purposes, and represents the most commonly used type of dynamical stochastic model for interest rates. There are several reasons why this happens. In particular, the reasons for which the HJM appoach is not substantially–preferable to modelling the short rate evolution have been outlined in Rogers (1995).

The paper is structured as follows. Section 2 defines an analytically tractable time–homogenous model and describes the extension we propose in this paper. Section 3 derives zero-coupon bond prices in the new model and analyzes the issue of fitting the current term structure of interest rates. Section 4 derives zero-coupon bond option prices in the new model. Section 5 derives the forward–measure dynamics. Section 6 deals with the first application of our extension procedure, namely the Vasicek (1977) case. Section 7 considers the extension in the case of the Cox-Ingersoll-Ross (1985) model. Sections 8 and 9 deal respectively with the extensions of the Dothan (1978) and the "exponential Vasicek" models. Section 10 hints at the extension of multifactor models, and presents the most relevant advantages of such extension for concrete applications. Section 11 concludes the paper.

2 The basic assumptions

On a filtered probability space $(\Omega^x, \mathcal{F}^x, I\!\!F^x, Q^x)$, $I\!\!F^x = \{\mathcal{F}^x_t : t \geq 0\}$, we consider a given time–homogeneous stochastic process, whose dynamics is expressed by:

$$dx^\alpha_t = \mu(x^\alpha_t; \alpha)dt + \sigma(x^\alpha_t; \alpha)dW_t \ , \tag{1}$$

where x^α_0 is a given real number, $\alpha = \{\alpha_1, \ldots, \alpha_n\} \in I\!\!R^n$, $n \geq 1$, is a vector of parameters, W is a standard Brownian motion and μ and σ are sufficiently well behaved real functions. We set $\mathcal{F}^x_t$ to be the sigma-field generated by x^α up to time t.

We assume that the process x^α describes the evolution of the instantaneous spot interest rate under the risk-adjusted martingale measure, and refer to this model as to the "reference model".

We denote by $P^x(t,T)$ the price at time t of a zero–coupon bond maturing at T and with unit face value, so that

$$P^x(t,T) = E^x \left\{ \exp\left[-\int_t^T x_s^\alpha ds \right] |\mathcal{F}_t^x \right\},$$

where E^x denotes the expectation under the risk-adjusted measure Q^x.

We also assume that there exists an explicit real function Π^x, defined on a suitable subset of $I\!R^{n+3}$, such that

$$P^x(t,T) = \Pi^x(t,T,x_t^\alpha;\alpha). \tag{2}$$

The best known examples of spot–rate models satisfying our assumptions are the Vasicek (1977) model, the Dothan (1978) model and the Cox-Ingersoll-Ross (1985) model.

Let us now denote by $R^x(t,T)$ the continously compounded spot-interest rate at time t for an investment maturing at time T, i.e.,

$$R^x(t,T) = -\frac{\ln P^x(t,T)}{T-t} = -\frac{\ln \Pi^x(t,T,x_t^\alpha;\alpha)}{T-t} =: \rho^x(t,T,x_t^\alpha;\alpha).$$

Models like (1) are typical examples of models with endogenous term structure of interest rates. This means that the initial term structure $T \mapsto \rho^x(0,T,x_0^\alpha;\alpha)$ does not necessarily match the term structure of interest rates observed in the market, no matter how the parameter vector α is chosen. In practice, finding a particular α by calibration of the model to the observed term structure of interest rates can produce a very poor fit, also because some typical shapes, like that of an inverted yield curve, may not be reproduced by the model. The necessity to overcome this drawback, especially to consistently price interest-rate derivatives, led to the introduction of models with time–dependent coefficients. In particular, the Hull-White (1990) "Extended Vasicek" model extends a model like (1) by exactly fitting the observed term structure of interest rates.

In this paper, we propose a simple approach for extending a time-homogenous spot-rate model (1), in such a way that our extended version preserves the analytical tractability of the initial model.[1]

Precisely, we assume we are given a new filtered probability space $(\Omega, \mathcal{F}, I\!F, Q)$, $I\!F = \{\mathcal{F}_t : t \geq 0\}$ on which the new instantaneous short rate is defined by

$$r_t = x_t + \varphi(t;\alpha)\ , \quad t \geq 0, \tag{3}$$

where x is a stochastic process that has under Q the same dynamics as x^α under Q^x, and φ is a deterministic function, depending on the parameter vector (α, x_0), that is integrable on closed intervals. Notice that x_0 is one more parameter at our disposal: We are free to select its value as long as

$$\varphi(0;\alpha) = r_0 - x_0\ .$$

The function φ can be chosen so as to fit exactly the initial term structure of interest rates. We set $\mathcal{F}_t$ to be the sigma-field generated by x up to time t.

We notice that the process r depends on the parameters $\alpha_1, \ldots, \alpha_n, x_0$ both through the process x and through the function φ. As a common practice, we can determine $\alpha_1, \ldots, \alpha_n, x_0$ by calibrating the model to the current term structure of volatilities, fitting for example caps and floors or a few swaptions prices.

[1]We found out at a later time that a similar idea has been developed independently by Dybvig (1997), and by Avellaneda and Newman (1998). In their works, both Dybvig (1997) and Avellaneda and Newman (1998) refer to the term φ as to the "fudge" factor.

Notice that, if φ is differentiable, the stochastic differential equation for the short–rate process (3) is,

$$d\,r_t = \left[\frac{d\varphi(t;\alpha)}{dt} + \mu(r_t - \varphi(t;\alpha);\alpha)\right]dt + \sigma(r_t - \varphi(t;\alpha);\alpha)dW_t\ .$$

Since in case of time–homogeneous coefficients an affine term structure in the short–rate is equivalent to affine drift and squared diffusion coefficients (see for example Björk (1997) or Duffie (1996)), we deduce immediately that if the reference model has an affine term structure so does the extended model.

3 Fitting the initial term structure of interest rates

Definition (3) immediately leads to the following.

Theorem 3.1. *The price at time t of a zero–coupon bond maturing at T and with unit face value is*

$$P(t,T) = \exp\left[-\int_t^T \varphi(s;\alpha)ds\right]\Pi^x(t,T,r_t - \varphi(t;\alpha);\alpha). \tag{4}$$

Proof. Denoting by E the expectation under the measure Q, we simply have to notice that

$$\begin{aligned} P(t,T) &= E\left\{\exp\left[-\int_t^T (x_s + \varphi(s;\alpha))ds\right]|\mathcal{F}_t\right\} \\ &= \exp\left[-\int_t^T \varphi(s;\alpha)ds\right]E\left\{\exp\left[-\int_t^T x_s ds\right]|\mathcal{F}_t\right\} \\ &= \exp\left[-\int_t^T \varphi(s;\alpha)ds\right]\Pi^x(t,T,x_t;\alpha), \end{aligned}$$

where in the last step we use the equivalence of the dynamics of x under Q and x^α under Q^x.
□

In the following, when we refer to zero–coupon bonds, we always assume a unit face value.

Let us now assume that the term structure of discount factors that is currently observed in the market is given by the sufficiently smooth function $t \mapsto P^M(0,t)$.

If we denote by $f^x(0,t;\alpha)$ and $f^M(0,t)$ the instantaneous forward rates at time 0 for a maturity t associated respectively to the bond prices $\{P^x(0,t) : t \geq 0\}$ and $\{P^M(0,t) : t \geq 0\}$, i.e.,

$$\begin{aligned} f^x(0,t;\alpha) &= -\frac{\partial \ln P^x(0,t)}{\partial t} = -\frac{\partial \ln \Pi^x(0,t,x_0;\alpha)}{\partial t}, \\ f^M(0,t) &= -\frac{\partial \ln P^M(0,t)}{\partial t}, \end{aligned}$$

we then have the following.

Corollary 3.2. *The model (3) fits the currently observed term structure of discount factors if and only if*

$$\varphi(t;\alpha) = \varphi^*(t;\alpha) := f^M(0,t) - f^x(0,t;\alpha), \tag{5}$$

i.e., if and only if

$$\exp\left[-\int_t^T \varphi(s;\alpha)ds\right] = \Phi^*(t,T,x_0;\alpha) := \frac{P^M(0,T)}{\Pi^x(0,T,x_0;\alpha)}\frac{\Pi^x(0,t,x_0;\alpha)}{P^M(0,t)}. \tag{6}$$

Moreover, the corresponding zero-coupon-bond prices at time t are given by $P(t,T) = \Pi(t,T,r_t;\alpha)$, where

$$\Pi(t,T,r_t;\alpha) = \Phi^*(t,T,x_0;\alpha)\,\Pi^x(t,T,r_t - \varphi^*(t;\alpha);\alpha) \tag{7}$$

Proof. From the equality

$$P^M(0,t) = \exp\left[-\int_0^t \varphi(s;\alpha)ds\right]\Pi^x(0,t,x_0;\alpha),$$

we obtain (5) by taking the natural logarithm of both members and then differentiating. From the same equality, we also obtain (6) by noting that

$$\exp\left[-\int_t^T \varphi(s;\alpha)ds\right] = \exp\left[-\int_0^T \varphi(s;\alpha)ds\right]\exp\left[\int_0^t \varphi(s;\alpha)ds\right] = \frac{P^M(0,T)}{\Pi^x(0,T,x_0;\alpha)}\frac{\Pi^x(0,t,x_0;\alpha)}{P^M(0,t)},$$

which, combined with (4), gives (7). □

Notice that by choosing $\varphi(t;\alpha)$ as in (5), our model exactly fits the observed term structure of interest rates, no matter which values of α and x_0 are chosen.

4 Explicit formulas for European options

In the previous section, we have seen that if an analytical expression for Π^x is available, then defining φ as in (5) yields a model which fits exactly the observed term structure of interest rates and still implies analytical formulas for bond prices.

The extension (3) is even more interesting when the reference model (1) allows for analytical formulae for zero-coupon-bond options as well. It is easily seen that the extended model preserves the analytical tractability for option prices by means of analytical correction factors that are defined in terms of φ.

To this end, we note that under the model (1), the price at time t of a European call option with maturity T, strike K and written on a zero-coupon bond maturing at time τ is

$$V_C^x(t,T,\tau,K) = E^x\left\{\exp\left[-\int_t^T x_s^\alpha ds\right](P^x(T,\tau)-K)^+|\mathcal{F}_t^x\right\}.$$

We then assume there exists an explicit real function Ψ^x defined on a suitable subset of $\mathbb{R}^{n+5}$, such that

$$V_C^x(t,T,\tau,K) = \Psi^x(t,T,\tau,K,x_t^\alpha;\alpha). \tag{8}$$

The best known examples of models (1) for which this holds are again the Vasicek (1977) model and the Cox-Ingersoll-Ross (1985) model.

Straightforward algebra leads to the following.

Theorem 4.1. *Under the model (3), the price at time t of a European call option with maturity T, strike K and written on a zero-coupon bond maturing at time τ is*

$$V_C(t,T,\tau,K) = \exp\left[-\int_t^\tau \varphi(s;\alpha)ds\right] \Psi^x\left(t,T,\tau,K\exp\left[\int_T^\tau \varphi(s;\alpha)ds\right], r_t - \varphi(t;\alpha);\alpha\right). \quad (9)$$

Proof. We simply have to notice that

$$\begin{aligned}
V_C(t,T,\tau,K) &= E\left\{\exp\left[-\int_t^T (x_s + \varphi(s;\alpha))ds\right](P(T,\tau)-K)^+|\mathcal{F}_t\right\} \\
&= \exp\left[-\int_t^T \varphi(s;\alpha)ds\right] E\left\{\exp\left[-\int_t^T x_s ds\right]\right. \\
&\qquad \left.\cdot\left(\exp\left[-\int_T^\tau \varphi(s;\alpha)ds\right]\Pi^x(T,\tau,x_T;\alpha) - K\right)^+ |\mathcal{F}_t\right\} \\
&= \exp\left[-\int_t^\tau \varphi(s;\alpha)ds\right] E\left\{\exp\left[-\int_t^T x_s^\alpha ds\right]\right. \\
&\qquad \left.\cdot\left(\Pi^x(T,\tau,x_T;\alpha) - K\exp\left[\int_T^\tau \varphi(s;\alpha)ds\right]\right)^+ |\mathcal{F}_t^x\right\} \\
&= \exp\left[-\int_t^\tau \varphi(s;\alpha)ds\right]\Psi^x\left(t,T,\tau,K\exp\left[\int_T^\tau \varphi(s;\alpha)ds\right], x_t;\alpha\right),
\end{aligned}$$

where in the last step we use the equivalence of the dynamics of x under Q and x^α under Q^x. □

The price of a European put option can be obtained through the put-call parity for bond options. Indeed, denoting by $V_P(t,T,\tau,K)$ the price at time t of a European put option with maturity T, strike K and written on a zero-coupon bond maturing at time τ, we have

$$V_P(t,T,\tau,K) = V_C(t,T,\tau,K) - P(t,\tau) + KP(t,T). \quad (10)$$

We then remark that, under our model (3), caps and floors can be priced analytically as well, since they can be viewed respectively as portfolios of put options and portfolios of call options on zero coupon bonds.

The previous formulas for European call and put options hold for any specification of the function φ. In particular, when exactly fitting the initial term structure of interest rates, the equality (5) must be used to produce the right formulas for option prices, i.e., $V_C(t,T,\tau,K) = \Psi(t,T,\tau,K,r_t;\alpha)$, where

$$\Psi(t,T,\tau,K,r_t;\alpha) = \Phi^*(t,\tau,x_0;\alpha)\Psi^x\left(t,T,\tau,K\Phi^*(\tau,T,x_0;\alpha), r_t - \varphi^*(t;\alpha);\alpha\right). \quad (11)$$

To this end, we notice that, if prices are to be calculated at time 0, we need not explicitly compute $\varphi^*(t;\alpha)$ since the relevant quantities are the discount factors at time zero.

Moreover, if Jamshidian (1989)'s decomposition for valuing coupon bearing bond options, and hence swaptions, can be applied to the model (1), the same decomposition is still feasible under (3) through straightforward modifications, so that also in the extended model we can price analytically coupon bearing bond options and swaptions. Indeed, consider a coupon bearing bond with unit face value, paying the cash flows $\mathcal{C} = [c_1, c_2, \ldots, c_n]$ at maturities $\mathcal{T} = [T_1, T_2, \ldots, T_n]$. Let $T \leq T_1$. The price of our coupon–bearing bond in T is given by

$$B_C(T, \mathcal{C}, \mathcal{T}) = \sum_{i=1}^{n} c_i P(T, T_i) = \sum_{i=1}^{n} c_i \exp\left[-\int_T^{T_i} \varphi(s;\alpha) ds\right] \Pi^x(T, T_i, r_T - \varphi(T;\alpha); \alpha).$$

Assume we need to price at time t a European put option on the coupon–bearing bond with strike price K and maturity T. The option payoff is

$$[K - B_C(T, \mathcal{C}, \mathcal{T})]^+ \ .$$

Jamshidian's decomposition is based on a decomposition of this payoff obtained through the solution x^* of the following equation:

$$\sum_{i=1}^{n} c_i \exp\left[-\int_T^{T_i} \varphi(s;\alpha) ds\right] \Pi^x(T, T_i, x^*; \alpha) = K \ .$$

The payoff can be easily rewritten as

$$\left[\sum_{i=1}^{n} c_i \exp\left[-\int_T^{T_i} \varphi(s;\alpha) ds\right] (\Pi^x(T, T_i, x^*; \alpha) - \Pi^x(T, T_i, r_T - \varphi(T;\alpha); \alpha))\right]^+ \ .$$

We now assume that the basic model (1) satisfies the following assumption:

$$\frac{\partial \Pi^x(t, s, x; \alpha)}{\partial x} < 0 \ \text{ for all } 0 < t < s, \text{ and all } \alpha \ .$$

It is easy to see that both the Vasicek and the CIR model satisfy this assumption. Under this assumption, the payoff can be rewritten as

$$\sum_{i=1}^{n} c_i \exp\left[-\int_T^{T_i} \varphi(s;\alpha) ds\right] (\Pi^x(T, T_i, x^*; \alpha) - \Pi^x(T, T_i, r_T - \varphi(T;\alpha); \alpha))^+ \ ,$$

so that we have now to value a portfolio of put options on zero–coupon bonds. If we take the risk–neutral expectation of the discounted payoff, we obtain the price at time t of the coupon–bearing bond option with maturity T and strike K:

$$V_P^C(t, T, \mathcal{C}, \mathcal{T}, K) = \sum_{i=1}^{n} c_i \exp\left[-\int_T^{T_i} \varphi(s;\alpha) ds\right] \ V_P(t, T, T_i, \Pi^x(T, T_i, x^*; \alpha)) \ , \tag{12}$$

where V_P can be computed via the formula given previously. This technique can be used to price European swaptions as well.

5 Forward Measure dynamics

In this section we derive the T-forward measure dynamics of a given short–rate model (1) from which the T-forward measure dynamics of the associated model (3) can be readily obtained through the deterministic shift $\varphi(t)$.

We recall the following possible formulation of the change-of-numeraire technique. Assume we are given two numeraire assets S and U, whose dynamics under a common measure which is equivalent to both the numeraire measures associated with S and U (say the risk-neutral measure) is

$$dS_t = x_t^\alpha S_t dt + \sigma_t^S \ \ dW_t,$$
$$dU_t = x_t^\alpha U_t dt + \sigma_t^U \ \ dW_t,$$

with W a standard Brownian motion. Then the relationship between the standard Brownian motions W^S and W^U corresponding to the martingale measures associated respectively to the numeraire assets S and U is known to be

$$dW_t^S = dW_t^U - \left(\frac{\sigma_t^S}{S_t} - \frac{\sigma_t^U}{U_t} \right) \ dt.$$

We now consider the T-forward adjusted measure, corresponding to the bond-price numeraire $S_t = P(t,T)$, and the risk-free measure, corresponding to the bank-account numeraire $U_t = \exp(\int_0^t x_t^\alpha ds)$, whose volatility is zero, i.e. $\sigma_t^U = 0$. The volatility of S can be obtained through Itô's formula applied to $S_t = P^x(t,T) = \Pi^x(t,T,x_t^\alpha;\alpha)$. One finds

$$\frac{\sigma_t^S}{S_t} = \sigma(x_t^\alpha;\alpha) \ \frac{\partial \ln \Pi^x}{\partial x},$$

so that, denoting by W_t^T the standard Brownian motion W_t^S under the T-forward adjusted measure and by W_t the standard Brownian motion W_t^U under the risk neutral measure, we have

$$dW_t = dW_t^T + \sigma(x_t^\alpha;\alpha) \ \frac{\partial \ln \Pi^x}{\partial x} \ dt.$$

Substituting this last relationship into equation (1) yields

$$dx_t^\alpha = \mu(x_t^\alpha;\alpha)dt + \sigma^2(x_t^\alpha;\alpha) \ \frac{\partial \ln \Pi^x}{\partial x} dt + \sigma(x_t^\alpha;\alpha) dW_t^T \ . \tag{13}$$

Equation (13) gives the dynamics of the short rate under the T-forward adjusted measure for the basic model x_t^α. Of course the only instances where this equation is really useful are in cases where the bond-price formula Π^x is explictly known.[2]

In the following sections we will develop some specific examples of our general model (3).

6 The Vasicek case

The first application we consider is based on the Vasicek (1977) model. In this case, the basic time–homogeneous model evolves according to

$$dx_t^\alpha = k(\theta - x_t^\alpha)dt + \sigma \ dW_t, \tag{14}$$

where the parameter vector is $\alpha = (k,\theta,\sigma)$, with k, θ, σ positive constants. This model yields a term structure which is affine in the short rate.

[2]See also Davis (1998).

Extending this model through (3) amounts to have the following short–rate dynamics:

$$dr_t = \left[k\theta + k\varphi(t;\alpha) + \frac{d\varphi(t;\alpha)}{dt} - kr_t\right] dt + \sigma\, dW_t \ . \tag{15}$$

Moreover, under the specification (5), $\varphi(t;\alpha) = \varphi^{VAS}(t;\alpha)$, where

$$\varphi^{VAS}(t;\alpha) = f^M(0,t) + (e^{-kt} - 1)\frac{k^2\theta - \sigma^2/2}{k^2} - \frac{\sigma^2}{2k^2}e^{-kt}(1 - e^{-kt}) - x_0 e^{-kt} \ .$$

The price at time t of a zero-coupon bond maturing at time T is

$$P(t,T) = \frac{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}} A(t,T)\exp\{-B(t,T)[r_t - \varphi^{VAS}(t;\alpha)]\},$$

and the price at time t of a European call option with strike K, maturity T and written on a zero-coupon bond maturing at time τ is

$$\begin{aligned} V_C(t,T,\tau,K) = & \frac{P^M(0,\tau)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,\tau)\exp\{-B(0,\tau)x_0\}} \\ & \cdot\Psi^{VAS}\left(t,T,\tau,K\frac{P^M(0,T)A(0,\tau)\exp\{-B(0,\tau)x_0\}}{P^M(0,\tau)A(0,T)\exp\{-B(0,T)x_0\}}, r_t - \varphi^{VAS}(t;\alpha);\alpha\right), \end{aligned}$$

where

$$\begin{aligned} \Psi^{VAS}(t,T,\tau,X,x;\alpha) &= A(t,\tau)\exp\{-B(t,\tau)x\}N(h) - XA(t,T)\exp\{-B(t,T)x\}N(h-\bar\sigma) \ , \\ h &= \frac{1}{\bar\sigma}\ln\frac{A(t,\tau)\exp\{-B(t,\tau)x\}}{XA(t,T)\exp\{-B(t,T)x\}} + \frac{\bar\sigma}{2} \ , \\ \bar\sigma &= \sigma B(T,\tau)\sqrt{\frac{1-e^{-2k(T-t)}}{2k}} \\ A(t,T) &= \exp\left[\frac{(B(t,T) - T + t)(k^2\theta - \sigma^2/2)}{k^2} - \frac{\sigma^2 B(t,T)^2}{4k}\right] \ , \\ B(t,T) &= \frac{1-e^{-k(T-t)}}{k} \ , \end{aligned}$$

and $N(\cdot)$ denotes the standard normal cumulative distribution function.

We now notice that by defining

$$\vartheta(t) = \theta + \varphi(t;\alpha) + \frac{1}{k}\frac{d\varphi(t;\alpha)}{dt},$$

model (15) can be written as

$$dr_t = k(\vartheta(t) - r_t)dt + \sigma\, dW_t \tag{16}$$

which is one of the particular extensions of the Vasicek (1977) model that has been proposed by Hull and White (1990, 1994a) (other possibilities include adopting time–varying coefficients k and σ). Viceversa, from the model (16), one can obtain our extension (15) by setting

$$\varphi(t;\alpha) = e^{-kt}\varphi(0;\alpha) + k\int_0^t e^{-k(t-s)}\vartheta(s)ds - \theta(1 - e^{-kt}).$$

Therefore, our extension of the Vasicek (1977) model is perfectly equivalent to that by Hull and White (1990, 1994a) with constant k and σ, as we can also verify by further expliciting our bond and option pricing formulae. This equivalence is basically due to the linearity of the reference-model equation (14). In fact our extra parameter θ turns out to be quite redundant since it is absorbed by the time-dependent function ϑ that is completely determined through the fitting of the current term structure of interest rates.

We notice that the function φ^{VAS}, for $\theta = 0$, is related to the function $\alpha(\cdot)$ in Hull and White (1994a), whose analytical expression has been derived by Pelsser (1996) or Kijima and Nagayama (1994). The only difference is that r_0 is absorbed in $\alpha(\cdot)$, whereas we let r_0 be "absorbed" partly by the reference model initial condition x_0 and partly by $\varphi^{VAS}(0;\alpha)$.

Notice that in the Vasicek case keeping a general x_0 adds no further flexibility to the extended model, due to linearity of the short–rate equation. We might as well set once and for all $x_0 = r_0$ and $\varphi^{VAS}(0;\alpha) = 0$ in the Vasicek case without affecting the caps/floors fit quality.

We finally derive the T-forward adjusted dynamics of the Vasicek model (14) by applying (13). We obtain

$$dx_t^\alpha = [k\theta - B(t,T)\sigma^2 - kx_t^\alpha]dt + \sigma dW_t^T.$$

It is easy to compute the distribution of the short rate under the forward adjusted measure Q^T and to show that it is still Gaussian. Precisely, the transition distribution of x_t^α conditional on x_s^α is given by

$$\begin{aligned} x_t^\alpha | x_s^\alpha &\sim^T \mathcal{N}\left(M^T(s,t), \sigma^2 \frac{1-e^{-2k(t-s)}}{2k}\right), \\ M^T(s,t) &= \theta + e^{-k(t-s)}(x_s^\alpha - \theta) - \frac{\sigma^2}{k^2}\left(1 - e^{-k(t-s)}\right) + \frac{\sigma^2}{2k^2}\left[e^{-k(T-t)} - e^{-k(T+t-2s)}\right], \end{aligned}$$

with $\mathcal{N}(m,v)$ denoting the normal distribution with mean m and variance v.

7 The CIR++ model

The most relevant application of the results of the previous sections is the extension of the Cox-Ingersoll-Ross (1985) model. In this case, the process (1) is given by

$$dx_t^\alpha = k(\theta - x_t^\alpha)dt + \sigma\sqrt{x_t^\alpha}dW_t,$$

where the parameter vector is $\alpha = (k, \theta, \sigma)$, with k, θ, σ positive constants. The condition

$$2k\theta > \sigma^2$$

ensures that the origin is inaccessible to the reference model, so that the process x^α remains positive. As is well known, this process x^α features a noncentral *chi-square* distribution, and yields an affine term-structure of interest rates. Accordingly, analytical formulae for prices of zero-coupon bond options, caps and floors, and, through Jamshidian's decomposition, coupon-bearing bond options and swaptions, can be derived.

We can therefore consider the CIR++ model, consisting of our extension (3), and calculate the analytical formulas implied by such a model, by simply retrieving the explicit expression for Π^x and Ψ^x as given in Cox et al. (1985).

Then, assuming exact fitting of the initial term structure of discount factors, we have that $\varphi(t;\alpha) = \varphi^{CIR}(t;\alpha)$ where

$$\begin{aligned}
\varphi^{CIR}(t;\alpha) &= f^M(0,t) - f^{CIR}(0,t;\alpha), \\
f^{CIR}(0,t;\alpha) &= 2k\theta \frac{(\exp\{th\}-1)}{2h+(k+h)(\exp\{th\}-1)} \\
&+ x_0 \frac{4h^2 \exp\{th\}}{[2h+(k+h)(\exp\{th\}-1)]^2}
\end{aligned}$$

with

$$h = \sqrt{k^2 + 2\sigma^2}.$$

Moreover, the price at time t of a zero-coupon bond maturing at time T is

$$P(t,T) = \frac{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}} A(t,T)\exp\{-B(t,T)[r_t - \varphi^{CIR}(t;\alpha)]\},$$

where

$$\begin{aligned}
A(t,T) &= \left[\frac{2h\ \exp\{(k+h)(T-t)/2\}}{2h+(k+h)(\exp\{(T-t)h\}-1)}\right]^{2k\theta/\sigma^2}, \\
B(t,T) &= \frac{2(\exp\{(T-t)h\}-1)}{2h+(k+h)(\exp\{(T-t)h\}-1)}.
\end{aligned}$$

The spot interest rate at time t for the maturity T is therefore

$$R(t,T) = \frac{1}{T-t}\left[\ln \frac{P^M(0,t)A(0,T)\exp\{-B(0,T)x_0\}}{P^M(0,T)A(0,t)\exp\{-B(0,t)x_0\}} - \ln(A(t,T)) - B(t,T)\varphi^{CIR}(t;\alpha) + B(t,T)r_t\right]$$

which is still affine in r_t.

The price at time t of a European call option with maturity $T > t$ and strike price K on a zero-coupon bond maturing at $\tau > T$ is

$$\begin{aligned}
V_C(t,T,\tau,K) = &\frac{P^M(0,\tau)A(0,t)\exp\{-B(0,t)x_0\}}{P^M(0,t)A(0,\tau)\exp\{-B(0,\tau)x_0\}} \\
&\cdot \Psi^{CIR}\left(t,T,\tau,K\frac{P^M(0,T)A(0,\tau)\exp\{-B(0,\tau)x_0\}}{P^M(0,\tau)A(0,T)\exp\{-B(0,T)x_0\}}, r_t - \varphi^{CIR}(t;\alpha);\alpha\right),
\end{aligned}$$

where

$$\begin{aligned}
\Psi^{CIR}(t,T,\tau,X,x;\alpha) &= A(t,\tau)\exp\{-B(t,\tau)x\}\chi^2\left(\mu;\frac{4k\theta}{\sigma^2},\frac{2\rho^2 x\exp\{h(T-t)\}}{\rho+\psi+B(T,\tau)}\right)\\
&- XA(t,T)\exp\{-B(t,T)x\}\chi^2\left(2\bar{r}[\rho+\psi];\frac{4k\theta}{\sigma^2},\frac{2\rho^2 x\exp\{h(T-t)\}}{\rho+\psi}\right),\\
\mu &= 2\bar{r}[\rho+\psi+B(T,\tau)],\\
\rho &= \rho(T-t) = \frac{2h}{\sigma^2\ (\exp\{h(T-t)\}-1)},\\
\psi &= (k+h)/\sigma^2,\\
\bar{r} &= \bar{r}(\tau-T) = [\ln(A(T,\tau)/X)]\,/B(T,\tau)\ ,
\end{aligned}$$

and $\chi^2(\cdot;r,\rho)$ denotes the noncentral chi–squared distribution function with r degrees of freedom and non-centrality parameter ρ, whose density is denoted by $p_{\chi^2(r,\rho)}$. By further simplifying these last formulae we obtain

$$\begin{aligned}
V_C(t,T,\tau,K) &= P(t,\tau)\chi^2\left(2\hat{r}[\rho+\psi+B(T,\tau)];\frac{4k\theta}{\sigma^2},\frac{2\rho^2[r_t-\varphi^{CIR}(t;\alpha)]\exp\{h(T-t)\}}{\rho+\psi+B(T,\tau)}\right)\\
&- KP(t,T)\chi^2\left(2\hat{r}[\rho+\psi];\frac{4k\theta}{\sigma^2},\frac{2\rho^2[r_t-\varphi^{CIR}(t;\alpha)]\exp\{h(T-t)\}}{\rho+\psi}\right)\ ,
\end{aligned}$$

with

$$\hat{r} = \frac{\ln\frac{A(T,\tau)}{K} - \ln\frac{P^M(0,T)A(0,\tau)\exp\{-B(0,\tau)x_0\}}{P^M(0,\tau)A(0,T)\exp\{-B(0,T)x_0\}}}{B(T,\tau)}\ .$$

The analogous put option price is obtained easily through put–call parity, so that also caps and floors can be valued analytically. Finally, formula (12) can be used to compute coupon–bearing bond option and swaption prices.

We now derive the T-forward adjusted dynamics and distribution of the short rate, by applying (13). We obtain

$$dx_t^\alpha = [k\theta - (k+B(t,T)\sigma^2)x_t^\alpha]dt + \sigma\sqrt{x_t^\alpha}dW_t^T. \tag{17}$$

By differentiating the call option price with respect to the strike price and by suitable decompositions, it can be shown that the transition distribution of the short rate x_t^α conditional on x_s^α under the forward adjusted measure Q^T $(s \le t \le T)$ is given by

$$\begin{aligned}
p^T_{x_t^\alpha|x_s^\alpha}(x) &= p_{\chi^2(v,\ \delta(t,s))/q(t,s)}(x) = q(t,s)p_{\chi^2(v,\ \delta(t,s))}(q(t,s)x), \qquad (18)\\
q(t,s) &= 2[\rho(t-s)+\psi+B(t,T)],\\
\delta(t,s) &= \frac{4\rho(t-s)^2 x_s^\alpha e^{h(t-s)}}{q(t,s)},\quad v = \frac{4k\theta}{\sigma^2}.
\end{aligned}$$

7.1 The positivity of rates and fitting quality

The use of a pricing model such as the CIR++ model concerns mostly non-standard interest rate derivatives. Indeed, for standard derivative products such as for example bond options, caps, floors, and swaptions, one can often find quoted "Black-like" implied volatilities prices. The analytical formulae given in this section are not used to price these standard instruments, but to determine the model parameters α such that the model prices are as close as possible to the relevant subset of the market prices above. This procedure is often referred to as "calibration" of the model to the market. An important issue for the CIR++ model which we now address is whether calibration to caps/floors prices is feasible while imposing positive rates. We know that the CIR++ rates are always positive if

$$\varphi^{CIR}(t;\alpha) > 0 \quad \text{for all } t \geq 0.$$

In turn, this condition is satisfied if

$$f^{CIR}(0,t;\alpha) < f^M(0,t) \;\; \text{for all } t \geq 0. \tag{19}$$

Studying the behaviour of the function $t \mapsto f^{CIR}(0,t;\alpha)$ will be helpful in our analysis. In particular, we are interested in its supremum

$$f^*(\alpha) := \sup_{t\geq 0} f^{CIR}(0,t;\alpha)\ .$$

There are three possible cases of interest:

i) $x_0 \leq \theta h/k$: In this case $t \mapsto f^{CIR}(0,t;\alpha)$ is monotonically increasing and the supremum of all its values is

$$f_1^*(\alpha) = \lim_{t\to\infty} f^{CIR}(0,t;\alpha) = \frac{2k\theta}{k+h};$$

ii) $\theta h/k < x_0 < \theta$: In this case $t \mapsto f^{CIR}(0,t;\alpha)$ takes its maximum value in

$$t^* = \frac{1}{h}\ln\frac{(x_0h+k\theta)(h-k)}{(x_0h-k\theta)(h+k)} > 0$$

and such a value is given by

$$f_2^*(\alpha) = x_0 + \frac{(x_0-\theta)^2k^2}{2\sigma^2 x_0};$$

iii) $x_0 \geq \theta$: In this case $t \mapsto f^{CIR}(0,t;\alpha)$ is monotonically decreasing for $t > 0$ and the supremum of all its values is

$$f_3^*(\alpha) = f^{CIR}(0,0;\alpha) = x_0.$$

We can try and enforce positivity of φ^{CIR} analytically in a number of ways. To ensure (19) one can for example impose that the market curve $t \mapsto f^M(0,t)$ remains above the corresponding CIR curve by requiring

$$f^*(\alpha) \leq \inf_{t\geq 0} f^M(0,t)\ .$$

This condition ensures (19), but appears to be too restrictive. To fix ideas, assume that when we calibrate the model parameters to market caps and floors prices we constrain the parameters to satisfy

$$x_0 > \theta\ .$$

Then we are in case iii) above. All we need to do to ensure positivity of φ^{CIR} in this case is

$$x_0 < \inf_{t \geq 0} f^M(0,t) \ .$$

This amounts to say that the initial condition of the time–homogeneous part of the model has to be placed below the whole market forward curve. On the other hand, since $\theta < x_0$ and θ is the mean–reversion level of the time–homogeneous part, this means that the time–homogeneous part will tend to decrease, and indeed we have seen that in case iii) $t \mapsto f^{CIR}(0,t;\alpha)$ is monotonically decreasing. If, on the contrary, the market forward curve is increasing (as is happening in the most liquid markets nowadays) the "reconciling role" of φ^{CIR} between the time–homogeneous part and the market forward curve will be stronger. Part of the flexibility of the time–homogeneous part of the model is then lost in this strong "reconciliation", thus subtracting freedom which the model can otherwise use to improve calibration to caps and floors prices.

In general, it turns out in real applications that the above requirements are too strong: Constraining the parameters to satisfy

$$\theta < \ x_0 < \inf_{t \geq 0} f^M(0,t)$$

leads to a caps/floors calibration whose quality is much lower than we have with less restringent constraints. For practical purposes it is in fact enough to impose weaker restrictions. Such restrictions, though not guaranteeing positivity of φ^{CIR} analytically, work well in all the market situation we tested. First consider the case of a monotonically increasing market curve $t \mapsto f^M(0,t)$: We can choose a similarly increasing $t \mapsto f^{CIR}(0,t;\alpha)$ (case i)) starting from below the market curve, $x_0 < r_0$, and impose that its asymptotic limit

$$\frac{2k\theta}{k+h} \approx \theta$$

be below the corresponding market curve limit. If we do this, calibration results are satisfactory and we obtain usually positive rates. Of course we have no analytical certainty: The f^{CIR} curve might increase quicker than the market one, f^M, for small t's, cross it, and then increase more slowly so as to return below f^M for large $t's$. In such a case φ^{CIR} would be negative in the in–between interval. However, this situation is very unlikely and never occurred in the real market situations we tested.

Next, consider a case with a decreasing market curve $t \mapsto f^M(0,t)$. In this case we take again $x_0 < r_0$ and we impose the same condition as before on the terminal limit for $t \to \infty$.

Similar considerations apply in the case of an upwardly humped market curve $t \mapsto f^M(0,t)$, which can be reporduced qualitatively by the time–homogenous model curve $t \mapsto f^{CIR}(0,t;\alpha)$. In this case one makes sure that the initial point, the analytical maximum and the asymptotic value of the time–homogeneous CIR curve remain below the corresponding points of the market curve.

Finally, the only critical situation is the case of an inverted yield curve, as was observed for example in the Italian market in the past years. The forward curve $t \mapsto f^{CIR}(0,t;\alpha)$ of the CIR model cannot mimic such a shape. Therefore, either we constrain it to stay below the inverted market curve by choosing a decreasing CIR curve (case iii)) starting below the market curve, i.e. $x_0 < r_0$, or we make the CIR curve start from a very small x_0 and increase, though not too steeply. The discrepancy between f^{CIR} and f^M becomes very large for large t's in the first case and for small t's in the second one. This feature lowers the quality of the caps/floors fitting in the case of an inverted yield curve if one wishes to maintain positive rates. However, highly inverted curves are rarely observed in liquid markets, so that this problem can be generally avoided, or has at most little effects. As far as the quality of fitting for the caps–volatility curve is concerned, we notice the following. If the initial point in the market caps–volatility curve is much smaller than the second one, so as to produce a large hump, the CIR++ model has difficulties in fitting the volatility structure. We observed through numerical simulations that, ceteris paribus, lowering the initial point of the caps volatility structure implied by the model roughly amounts to lowering the parameter x_0^α. This is consistent with the following formula

$$\sigma\sqrt{x_0}\ \frac{2h\ \exp(Th)}{[2h+(k+h)(\exp\{Th\}-1)]^2}$$

for the volatility of instantaneous forward rates $f^{CIR}(t,T;\alpha)$ at the initial time $t=0$ for the maturity T. Therefore, a desirable fitting of the caps volatility curve can require low values of x_0, in agreement with one of the conditions needed to preserve positive rates.

We conclude this section by presenting some numerical results on the quality of the model calibration to real-market cap prices, under the constraint that φ be a positive function. To this end, we have used the EURO cap volatility curve as of December 6th, 1999 at 2 p.m. The resulting fitting is displayed in Figure 1, where the volatility curve implied by the CIR++ model is compared with the market volatility curve. In Figure 2, we plot instead the graph of the function φ corresponding to the calibrated parameters.

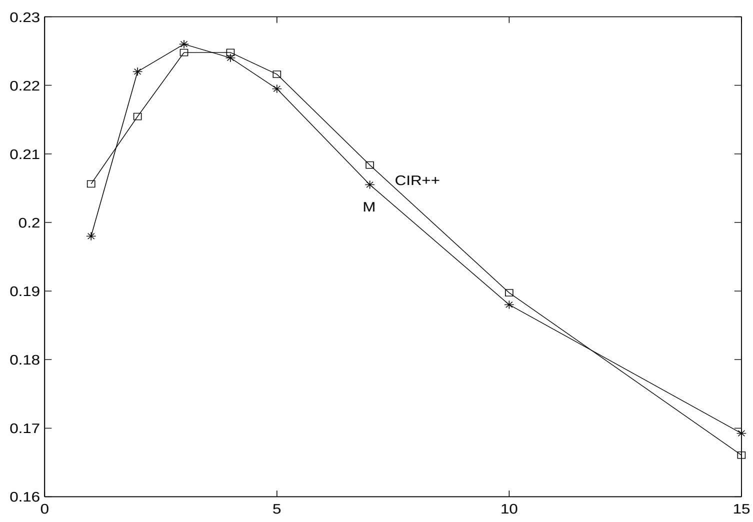

Figure 1: Comparison between the volatility curve implied by the CIR++ model and that observed in the EURO market (M) on December 6th, 1999 at 2 p.m.

7.2 The Monte Carlo simulation

This section comments on the use of the CIR++ model for Monte Carlo pricing of path–dependent interest–rate derivatives. Indeed, this model has been successfully applied to a number of practical cases.

Once we have found the α's such that the model is calibrated to the "relevant" part of the market, we may be willing to price a path-dependent payoff with European exercise features. The payoff is a function of the values of the underlying istantaneous interest rate at preassigned time

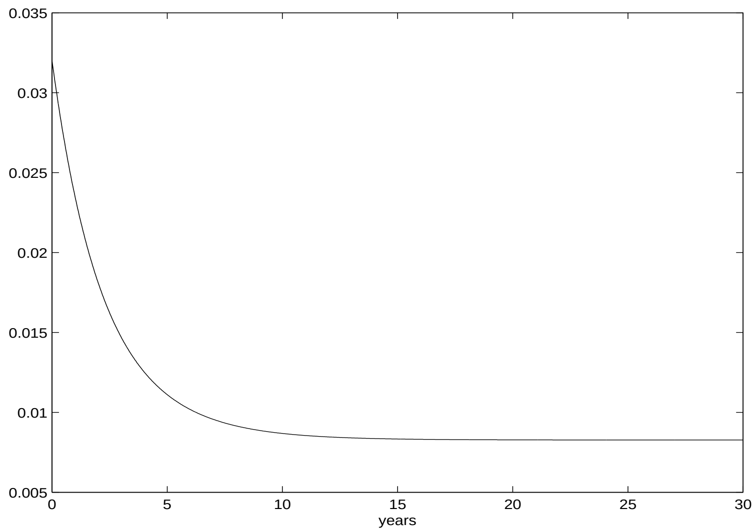

Figure 2: Graph of the function φ corresponding to the previously calibrated parameters.

instants $t = t_1, t_2, \ldots, t_m = T$. The current time is denoted by t, and usually one takes $t = 0$. The final maturity is T. Let us denote the given discounted payoff (payments of different additive components occur at different times and are discounted accordingly) by

$$\sum_{j=1}^{m} \exp\left[-\int_0^{t_j} r_s\, ds\right] H(t_1, \ldots, t_j; r_{t_1}, \ldots, r_{t_j}) \,. \tag{20}$$

Usually the payoff depends on the instantaneous rates r_{t_i} through the simple spot rates $R(t_i, t_i + \tau, r_{t_i}; \alpha) := [1/P(t_i, t_i + \tau) - 1]/\tau$ (typically six–month rates, i.e. $\tau = 0.5$). The advantage of analytical formulae for such $R's$ simplifies considerably the Monte Carlo pricing. Indeed, without analytical formulae one would compute the $R's$ from bond prices P obtained through trees or through numerical approximations of the solution of the bond price PDE. This should be done for every path, so as to increase dramatically the computational burden.

Pricing the generic additive term in the payoff (20) by a Monte Carlo method involves simulation of p paths (typically ranging from $p = 100,000$ to $p = 1,000,000$) of the short rate process r and computing the arithmetic mean of the p values assumed by the discounted payoff along each path. In order to avoid simulation of the integrals in the discounting terms

$$\exp\left[-\int_0^{t_j} r_s ds\right]$$

one can work under the t_m-forward adjusted measure, since

$$E\left\{\exp\left[-\int_0^{t_j} r_s ds\right] H(t_1, \ldots, t_j; r_{t_1}, \ldots, r_{t_j}) | \mathcal{F}_0\right\} = P(0, t_m) E^{t_m}\left\{\frac{H(t_1, \ldots, t_j; r_{t_1}, \ldots, r_{t_j})}{P(t_j, t_m)} | \mathcal{F}_0\right\}. \tag{21}$$

Notice that given the CIR analytical formula for bond prices, the term $P(t_j, t_m)$ is determined analytically from the simulated r_{t_j}, so that no further simulation is required to compute it. Here E^T denotes expected value under the T-forward adjusted measure.

In our case, simulating r is equivalent to simulating x, since φ is a known deterministic function. In detail:

- Define the sampling times s_i, $i = 0, 1, 2, \ldots, q$ up to the final maturity date of the contract, $s_q = T = t_m$. Of course the dates t_i have to be included in the s_j's.

  Now, in order to obtain the simulated paths, two approaches are possible:

  1. Sampling at each step the exact noncentral chi–squared transition density from $x^\alpha_{t_{i-1}}$ to $x^\alpha_{t_i}$ which is known analitically from (18) with $s = t_{i-1}$ and $t = t_i$;
  2. Discretizing the SDE for x^α via the Milstein scheme (see Klöden and Platen (1995)):

  $$\begin{aligned} x_{s_i+\Delta t} &= x_{s_i} + [k\theta - (k + B(s_i, t_m)\sigma^2)x_{s_i}]\,\Delta t - \frac{\sigma^2}{4}\left((W^{t_m}_{s_i+\Delta t} - W^{t_m}_{s_i})^2 - \Delta t\right) \\ &\quad + \sigma\sqrt{x_{s_i}}\left(W^{t_m}_{s_i+\Delta t} - W^{t_m}_{s_i}\right) . \end{aligned}$$

     In doing so, one has simply to sample at each step Δt the distribution associated to $W_{s_i+\Delta t} - W_{s_i}$, i.e. a normal distribution with mean 0 and variance Δt. Since the increments of the Brownian motion are independent, these normal samples are drawn from variables which are independent in different time intervals, so that one can generate a priori $p \times q$ independent realizations of such Gaussian random variables.

- Compute the discounted payoff (20) through (21) along each simultated path for r (deduced from the paths for x by adding φ) and average all the obtained payoff values.

7.3 Alternative extensions

A final issue we would like to consider, is the possible inclusion of other "time–varying coefficients", with the possible purpose of allowing exact calibration to the caps/floors (at-the-money) market.

One of the first directions in this sense has been proposed by Hull and White (1990) as a possibility to exactly fit the initial term-structure of interest rates, while keeping constant volatility parameters. Hull and White proposed the following extension of the CIR model:

$$dr_t = k[\vartheta(t) - r_t]dt + \sigma\sqrt{r_t}dW_t,$$

to exactly fit the initial term-structure of interest rates.

Such extension however is not analytically tractable. Indeed, to our knowledge, an analytical expression for $\vartheta(t)$ in terms of the observed yield curve is not available in the literature. Furthermore, there is no guarantee that a numerical approximation of $\vartheta(t)$ would keep the rate r positive, hence that the diffusion coefficient would always be well defined. On the contrary, our extension is always well defined and never leads to numerical problems, since the basic square root process x remains time–homogeneous. The only drawback is that we cannot guarantee positive interest rates analytically without influencing the caps/floors fitting quality, since $\varphi^{CIR}(t; \alpha)$ can be negative in principle. However, as noticed earlier, through simulations based on the calibration of the model to real market data, we can argue that negative rates are hardly ever observed in practical situations.

More generally, Maghsoodi (1996) studied the general case with time–varying volatility parameters

$$dr_t = k(t)[\vartheta(t) - r_t]dt + \sigma(t)\sqrt{r_t}dW_t.$$

If one chooses a constant ϑ and then extends the model through our procedure by adding a suitable φ, then one can produce an exact fitting to the term structures both of interest rates and of volatility. However, this task is not without numerical difficulties: The bond price and option price formulas in Maghsoodi's formulation rely on numerical integration and no formula for caps/floors prices is available in terms of the model "coefficients" $k(t)$ and $\sigma(t)$. Consequently, no formula for φ would be available, thus rendering the extension of too little analytical tractability. If the model loses its tractability the advantages over log-normal short–rate models or even market models become easily questionable. Finally the possibility of drawbacks related to "overfitting" are far from being remote. Indeed, forcing a one–factor model to match perfectly the caps/floors prices with a time–dependent "coefficient" leaves no synthesis to a parametric form and can lead to quite unrealistic and uncontrolled behaviour of the *future* term structure of volatility, as resulting from the calibrated model dynamics.

For further details on this extension of the CIR model, we refer to Brigo et al. (1998).

8 The Dothan case

The third example we consider is the extension of the Dothan (1978) model. This extension yields a "quasi" lognormal short–rate model which fits any given yield curve and for which there exist analytical formulae for zero coupon bonds. We offer its extension simply as an academic analytically–tractable case different from the classical affine tractable models (Vasicek and CIR). In his original paper, Dothan starts from a driftless geometric Brownian motion as short–rate process under the objective probability measure:

$$dx_t = \sigma x_t \, d\widetilde{W}_t \; .$$

Subsequently, he introduces a constant market price of risk which is equivalent to directly assuming a risk–neutral dynamics of type

$$dx_t^\alpha = \; a \; x_t^\alpha \; dt + \sigma \; x_t^\alpha \; dW_t \; , \quad \alpha = (a, \sigma), \tag{22}$$

thus yielding the Rendleman and Bartter (1980) model.

The zero coupon bond prices derived by Dothan are then given by

$$P(t,T) = \frac{\bar{x}^p}{\pi^2} \int_0^\infty \sin(2\sqrt{\bar{x}} \sinh y) \int_0^\infty f(z) \sin(yz) dz dy + \frac{2}{\Gamma(2p)} \bar{x}^p K_{2p}(2\sqrt{\bar{x}})$$

where

$$\begin{aligned} f(z) &= \exp\left[\frac{-(4p^2+z^2)s}{4}\right] z \left|\Gamma\left(-p + i\frac{z}{2}\right)\right|^2 \cosh\frac{\pi z}{2} \; , \\ \bar{x} &= \frac{2x_t}{\sigma^2} \; , \\ s &= \frac{\sigma^2(T-t)}{2} \; , \\ p &= \frac{1}{2} - a \; , \end{aligned}$$

and K_q denotes the modified Bessel function of the second kind of order q. Though somehow explicit, this formula for bond prices is rather complex since it depends on two integrals of functions involving hyperbolic sines and cosines. A double numerical integration is needed so that the advantage of having an "explicit" formula is dramatically reduced. In particular, as far as computational issues are concerned, implementing an approximating tree for the process x may be conceptually easier and not necessarily more time consuming.

We need to remark that the extension we propose in this paper, though particularly meaningful when the original model is time-homogenous and analytically tractable, is quite general and can be in principle applied to any endogenous term structure model, as we shall also see in the following section.

The construction of a binomial tree for pricing interest-rate derivatives under our extension (3) of (22) is rather straightforward. In fact, we just have to build a tree for x, the time–homogeneous part of the process r, and then shift the tree nodes at each time-period by a quantity that is defined by the corresponding value of φ.

Since x is a geometric Brownian motion, the well consolidated Cox-Ross-Rubinstein (1979) procedure can be used here (see also Rendleman and Bartter (1980)), thus rendering the tree construction extremely simple while contemporarily ensuring well known results of convergence in law. To this end, we denote by N the number of time-steps in the tree and with T a fixed maturity. We then define the following coefficients

$$u = e^{\sigma\sqrt{\frac{T}{N}}}, \tag{23}$$

$$d = e^{-\sigma\sqrt{\frac{T}{N}}}, \tag{24}$$

$$p = \frac{e^{a\frac{T}{N}} - d}{u - d}, \tag{25}$$

and build inductively the tree for x starting from x_0. Denoting by $\hat{x}$ the value of x at a certain node of the tree, the value of x in the subsequent period can either go up to $\hat{x}u$ with probability p or go down to $\hat{x}d$ with probability $1-p$. Notice that the probabilities p and $1-p$ are always well defined for a sufficiently large N, both tending to $\frac{1}{2}$ for N going to infinity.

The tree for the short-rate process r is finally constructed by displacing the previous nodes through the function φ. When exactly fitting the current term-structure of interest rates, the displacement of the tree nodes can be done, for instance, through a similar procedure to the one illustrated by Hull and White (1994a).

Our extended Dothan (1978) model, therefore, can be considered as a valid alternative to the Black-Karasinsky (1991) model, especially as far as implementation issues are concerned. However, the major drawback of our model is that the positivity of the short rate process cannot be guaranteed any more since its distribution is obtained by shifting a lognormal distribution by a possibly negative quantity.

9 The exponential Vasicek case

The last particular case we consider in this paper is the extension of what we call the exponential Vasicek model. The exponential Vasicek model assumes that the instantaneous short rate process x^α evolves as the exponential of an Ornstein-Uhlenbeck process y

$$dy_t = [\theta - a y_t]dt + \sigma dW_t, \quad \theta, a, \sigma > 0, \tag{26}$$

or equivalently, in the notation of Section 2, that

$$dx_t^\alpha = x_t^\alpha \left[\theta + \frac{\sigma^2}{2} - a \ln x_t^\alpha\right] dt + \sigma x_t^\alpha dW_t \ ,$$

with $x_0^\alpha > 0$ and $\alpha = \{\theta, a, \sigma\}$. The process x^α thus defined is therefore lognormally distributed and contrary to the Dothan (1979) process (22), always mean reverting since

$$\lim_{t\to\infty} E(x_t^\alpha) = \exp\left(\frac{\theta}{a} + \frac{\sigma^2}{4a}\right),$$

$$\lim_{t\to\infty} \mathrm{Var}(x_t^\alpha) = \exp\left(\frac{2\theta}{a} + \frac{\sigma^2}{2a}\right)\left[\exp(\frac{\sigma^2}{2a}) - 1\right].$$

The exponential Vasicek model does not imply explicit formulae for pure discount bonds. We have seen in the previous section however that analytical tractability does not mean ease of implementation and that our extension procedure can be also applied to more general reference models. To this end, similarly to the Dothan (1978) case, we can construct an approximating tree for the process x^α.

Let us define the process z as the process (26) with $\theta = 0$, i.e.,

$$dz_t = -az_t dt + \sigma dW_t,$$

with $z_0 = 0$. Since, by Ito's lemma,

$$x_t^\alpha = \exp\left(z_t + \left(\ln x_0^\alpha - \frac{\theta}{a}\right)e^{-at} + \frac{\theta}{a}\right),$$

it is then enough to construct an approximating tree for the process z. For example, the trinomial tree proposed by Hull and White (1994a) can be employed to this purpose.

Our extended exponential Vasicek (EEV) model leads to a fairly good calibration to cap prices in that the resulting model prices lie within the band formed by the market bid and ask prices. Indeed, this model clearly outperforms the classical lognormal short-rate model by Black and Karasinski (1991) (BK). An example of the quality of the model fitting to at-the-money cap prices is shown in Figure 3, where we have used the same market data as for the CIR++ case. In such a figure, it is also displayed the corresponding best-fit volatility curve implied by the BK model.

A final word has to be spent on a further advantage of the EEV model. In fact, the assumption of a shifted lognormal distribution, as implied by the EEV model, can be quite helpful when fitting volatility smiles. Notice, indeed, that the Black formula for cap prices is based on a lognormal distribution for the forward rates, and that a practical way to obtain implied volatility smiles is by shifting the support of this distribution by a quantity to be suitably determined.

10 Extensions of multifactor models

So far we have considered explicitly only the yield-curve-fitting extension in the single factor case. However, the generalization to two or more factors is indeed straightfoward. Analytical formulas for zero-coupon bond are readily available especially in the case of independent factors, when they are available for the single factor model. Indeed, in the table we refer to the case where the short–rate is obtained by adding independent reference processes. For example, the two factor case is obtained as follows:

$$r_t = \varphi(t;\alpha) + x_t^{\alpha_1} + y_t^{\alpha_2},$$

where x^{α_1} and y^{α_2} are two independent processes, both evolving similarly to (1), and $\alpha = (\alpha_1, \alpha_2)$. In particular, in the Gaussian case one can even allow for instantaneously correlated factors, thus

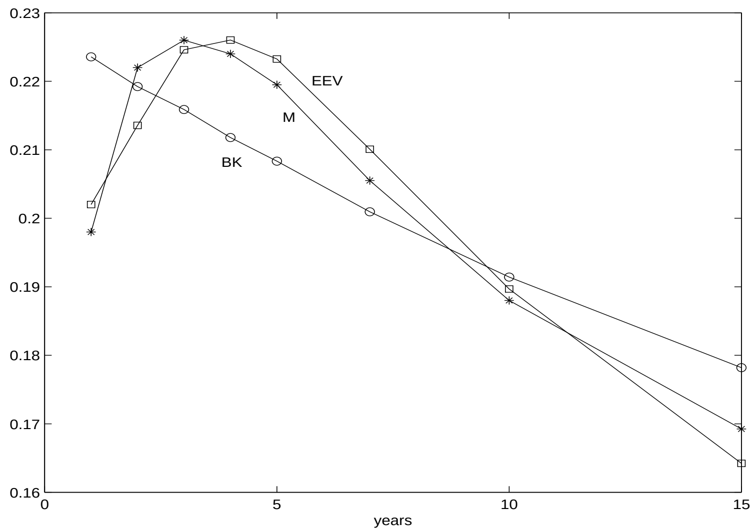

Figure 3: Comparison between the volatility curves implied by the EEV and the BK models and that observed in the EURO market (M) on December 6th, 1999 at 2 p.m.

having one more parameter, and preserve analytical tractability. This model can then be shown to be equivalent (by a change of variable) to the Hull and White (1994b) two-factor model. This "two-factor extended Vasicek", besides fitting the observed yield curve, features several other advantages, among which we recall, also in comparison with market models:

Low dimensionality; Analytical tractability (bonds, zero-coupon-bond options, caps/floors); Easy and recombining lattices (feasible calibration to swaptions); Known Gaussian transition densities both under risk-neutral and forward measures; Easy and quick Monte Carlo implementation; Humped volatility structures $T \mapsto \sqrt{\mathrm{Var}[d\ f(t,T)]/dt}$ are possible, and consequently a usually good fit to swaption or cap prices; Straightforward extension to three factors; Good interaction with log-normal FX-rate models for Monte Carlo pricing of **Quanto**-like products (e.g. quanto constant maturity swap with cap/floor features); Suitable for modelling two correlated yield curves.

The analytical tractability and the good fitting of the volatility term structure render this model a good choice, especially in areas such as risk management where the possibility to compute hedge parameters in a reasonably short time is fundamental.

Another potentially good multi-factor choice is the extended CIR model, where both x_1^α and y_2^α evolve according to the time–homogeneous CIR dynamics. This can be shown to be also equivalent to extending the Longstaff and Schwartz (1992) model to exactly fit the initial term-strucure, while using the basic parameters to fit the volatility structure. On one side, this extension is an improvement over the Vasicek case, since through restriction on the parameter values we can usually preserve positive rates. On the other hand, modelling correlated factors is not as straightforward as in the Gaussian case, and, as a result, the humped–volatility structure is not possible. It is also difficult to model two different yield curves while preserving analytical tractability if there is a non-zero correlation involved. We also recall that for pricing options on zero-coupon bonds in the extended multifactor CIR model we have to resort to numerical procedures for the calculation of multiple integrals. See also Chen and Scott (1992).

Reference Model	Tails	ABP	AOP	Multif.: ABP	Multifactor: AOP
Vasicek	$\mathcal{N}$	Yes	Yes	Yes	Yes
CIR	n.c. χ^2	Yes	Yes	Yes	Yes (Mult. Int.)
Dothan	$e^{\mathcal{N}}$	Yes (Mult. Int.)	No	No	No
Exp. Vasicek	$e^{\mathcal{N}}$	No	No	No	No

Table 1: Analytical bond prices (ABP) and analytical bond–option prices (AOP)

11 Concluding remarks

In this paper we obtained an extension of any given time-homogenous (reference) model through a simple procedure. Our extension exactly fits the initial term structure of interest rates, and has different impacts according to the nature of the reference model. In Table 1 we give a short list of reference models and the corresponding features of the extended model. "Mult. Int" in the table denotes the need for a multiple numerical integration.

When looking at Table 1, it is important to notice that the CIR model is the only reference model with analytical formulae for both bonds and bond options and with an istantaneous–rate featuring tails which are fatter than in the Gaussian case. This is the reason why the ideal choice for the reference model to which our result can be applied is the CIR model, at least in the single factor case. Our extended CIR model retains all these properties.

Concerning the pricing of American type derivatives, a binomial tree can be constructed for each possible choice of the reference model, using for example the procedure outlined by Nelson and Ramaswamy (1990). In this case, the function φ must be explicitly computed through (5), so that additional (but not too severe) assumptions on the smoothness of the initial term structure of interest rates are required (the analytical expression for φ is also needed for pricing path-dependent claims through, for example, Monte Carlo simulations). A valuable alternative is constructing a trinomial tree as suggested by Hull and White (1994a). In this case, in fact, we can avoid the differentiation of the initial yield curve and produce a more accurate pricing.

Our result can be also applied to reference models which are not analytically tractable. In fact, it can be convenient to do so when the basic model features an easy tree–implementation or other numerically desirable properties, which are usually retained in the extended model. As an example, we analysed the Dothan (1979) and the "exponential Vasicek" cases.

Finally, in the paper we have considered explicitly only the extension in the single factor case. However, the generalization to two or more factors is indeed straightforward, and we hinted at its advantages in concrete situations.

References

[1] AVELLANEDA, M., and NEWMAN, J. (1998). Positive interest rates and nonlinear term structure models. Preprint, Courant Institute of Mathematical Sciences, New York University.

[2] BJÖRK, T. (1997). Interest rate Theory, in *Lecture Notes In Mathematics*, Bressanone 1996, Berlin Heidelberg. Springer-Verlag.

[3] BLACK, F. and KARASINSKI, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, July-August 1991, pp .52-59.

[4] BRACE, A., GATAREK, D., and MUSIELA, M. (1997). The market model of interest rate dynamics. *Mathematical Finance*, Vol. 7, pp. 127–154.

[5] BRIGO, D. , MAURI, G., MERCURIO, F., SARTORELLI, G. (1998). *The CIR++ model: A new extension of the Cox-Ingersoll-Ross model with analytical calibration to bond and option prices*. Internal Report, Banca IMI, Milan.

[6] BRIGO, D. , MERCURIO, F. (2001). A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics* 5, 369-388.

[7] BRIGO, D. , MERCURIO, F. (2001). *Interest Rate Models: Theory and Practice*, Springer Verlag.

[8] CHEN, R, and SCOTT, L. (1992). Pricing interest rate options in a two–factor Cox-Ingersoll-Ross model of the term structure. *Review of Financial Studies*, 5, pp. 613–636.

[9] COX, J.C., INGERSOLL, J.E. and ROSS, S. (1985). A theory of the term structure of interest rates. *Econometrica*, 2 pp. 385-407.

[10] COX, J.C., ROSS, S., and RUBINSTEIN, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics*, 7 pagg. 229-263, 1979.

[11] DAVIS, M. (1998). A note on the forward measure. *Finance and Stochastics*, 2, pp. 19-28.

[12] DOTHAN, L.U. (1978). On the term structure of interest rates. *Journal Of Financial Economics* 6 pp. 59-69.

[13] DUFFIE, D. (1996). *Dynamic Asset Pricing Theory*, Second Edition. Academic press, New York.

[14] DYBVIG, P.H. (1988). Bond and bond option pricing based on the current term structure, *Washington University working paper*.

[15] DYBVIG, P.H. (1997). Bond and Bond Option Pricing Based on the Current Term Structure. *Mathematics of Derivative Securities*, Michael A. H. Dempster and Stanley R. Pliska, eds. Cambridge: Cambridge University Press, pp. 271-293.

[16] HEATH, D., JARROW, R., and MORTON, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. *Econometrica* 1, pp. 77-105.

[17] HULL, J. and WHITE, A. (1990). Pricing interest-rate-derivative securities. *The Review Of Financial Studies*, 4 pp. 573-592.

[18] HULL, J. and WHITE, A. (1994a). Numerical procedures for implementing term structure models I: single-factor models. *The Journal Of Derivatives*, Fall 1994 pp. 7-16, 1994.

[19] HULL, J. and WHITE, A. (1994b). Numerical procedures for implementing term structure models II: two-factor models. *The Journal Of Derivatives*, Winter 1994, 37-47, 1994.

[20] JAMSHIDIAN, F. (1988). The one–factor Gaussian interest rate model: Theory and implementation. *Working paper, Merril Lynch Capital Markets*.

[21] JAMSHIDIAN, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics*, 4, pp. 293–330.

[22] KIJIMA, M., and NAGAYAMA, I. (1994). Efficient numerical procedures for the HULL–WHITE extended Vasicek model. *The journal of financial engineering*, 3,4, pp. 275–292.

[23] KLÖDEN, P.E., and PLATEN, I. (1995). Numerical Solutions of Stochastic Differential Equations. Springer, Berlin Heidelberg New York.

[24] LONGSTAFF, F.A. and SCHWARTZ, E.S. (1992). Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. The Journal of Finance 47 , 1259-1282.

[25] MAGHSOODI, Y. (1996). Solution of the Extended CIR Term Structure and Bond Option Valuation. *Mathematical Finance*, 6, pp. 89–109.

[26] MILTERSEN K.R., SANDMANN K., and SONDERMANN D. (1997). Closed form solutions for term structure derivatives with log–normal interest rates. *Journal of Finance*, 1, pp. 409–430.

[27] MUSIELA, M., and RUTKOWSKI, M. (1997). Continuous–time term structure models: Forward measure approach. *Finance and Stochastics*, 4, pp. 261–292.

[28] NELSON, D. B. , and RAMASWAMY, K. (1990). Simple binomial processes as Diffusion Approximations in Financial Models. *The Review of Financial Studies*, 3, pagg 393–430, 1990.

[29] PELSSER, A.A.J. (1996). Efficient Methods for Valuing and Managing Interest Rate and other Derivative Securities. *PhD Dissertation*. Erasmus University Rotterdam, The Netherlands.

[30] RENDLEMAN, R. and BARTTER, B. (1980). The pricing of options on debt securities. *Journal of Financial and Quantitative Analysis*, 15, pp. 11-24.

[31] ROGERS, L.C.G. (1995). Which model for the term structure of interest rates should one choose? *Mathematical Finance*, IMA Volume 65, pp. 93–116, Springer, New York.

[32] SCHMIDT, W.M. (1997). On a general class of non-factor models for the term structure of interest rates. *Finance and Stochastics*, 4, pp. 3–24.

[33] SCOTT, L (1995). The valuation of interest rate derivatives in a multi-factor term-structure model with deterministic components. University of Georgia. Working paper.

[34] VASICEK, O.A. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5 pp .177-188.