Les Cahiers du GERAD ISSN: 0711–2440

A Dynamic Programming Approach for Pricing Options Embedded in Bonds

H. Ben-Ameur, M. Breton
L. Karoui, P. L'Écuyer

G–2004–35

April 2004

A Dynamic Programming Approach for Pricing Options Embedded in Bonds

Hatem Ben-Ameur
GERAD and HEC Montréal

Michèle Breton
GERAD and HEC Montréal

Lotfi Karoui
McGill University

Pierre L'Écuyer
GERAD and Université de Montréal

April, 2004

Les Cahiers du GERAD

G–2004–35

Abstract

We propose a Dynamic Programming (DP) approach combined with approximation for pricing options embedded in bonds, the focus being on call and put options with advance notice. An efficient approximation method is developed in detail for the cases where the interest rate process follows the Vasicek, Cox-Ingersoll-Ross (CIR), or generalized Vasicek models. Our DP methodology uses the exact joint distribution of the interest rate and integrated interest rate at a future date, conditional on the current value of the interest rate, for these models. We provide numerical illustrations, for the Vasicek and CIR models, comparing DP with finite difference methods. The results indicate that DP compares quite favorably in terms of both efficiency and accuracy. An important advantage of the DP approach is that it can be applied to more general models calibrated to capture the term structure of interest rates, such as the generalized Vacicek model, for example, where the only available methods are the (rather inaccurate) trinomial trees.

Résumé

Nous proposons une approche de programmation dynamique et d'approximation pour la tarification d'options implicites dans des obligations, plus précisément les options de rachat et de remboursement avec délai de notification. Une méthode d'approximation efficace est développée en détail pour les cas où la dynamique du taux d'intérêt est décrite par les processus de Vasicek, Cox-Ingersoll-Ross (CIR) et Vasicek généralisé. Notre méthode de programmation dynamique utilise pour ces trois modèles les distributions jointes exactes du taux d'intérêt et de l'intégrale du taux d'intérêt à une date future, conditionnelles à la valeur courante de ce taux. Nous présentons des illustrations numériques, pour les modèles de Vasicek et CIR, comparant les résultats fournis par la programmation dynamique et les différences finies. Nos résultats indiquent que la programmation dynamique se compare avantageusement en termes d'efficacité et de précision. Un avantage important de l'approche par la programmation dynamique est qu'elle peut être appliquée à des modèles plus généraux calibrés à la structure par terme des taux d'intérêts, comme par exemple le modèle de Vasicek généralisé, où les seules méthodes actuellement disponibles sont des arbres trinomiaux.

Acknowledgments: Research supported by NSERC and IFM2.

1 Introduction

A *bond* is a contract that pays its holder a known amount, the *principal*, at a known future date, called the *maturity*. The bond may also pay periodically to its holder fixed cash dividends, called the *coupons*. When it gives no dividend, it is known as a *zero-coupon bond*. A bond can be interpreted as a loan with a known principal and interest payments equal to the coupons (if any). The borrower is the *issuer* of the bond and the lender, i.e., the holder of the bond, is the *investor*.

Several bonds contain one or several options coming in various flavors. The *call option* gives the issuer the right to purchase back its debt for a known amount, the *call price*, during a specified period within the bond's life. Several government bonds contain a call feature [see Bliss and Ronn (1995) for the history of callable US Treasury bonds from 1917]. The *put option* gives the investor the right to return the bond to the issuer for a known amount, the *put price*, during a specified period within the bond's life. These options are an integral part of a bond, and cannot be traded alone, as is the case for call and put options on stocks (for example). They are said to be *embedded* in the bond. In general, they are of the *American-type* and, thus, allow for early exercising, so that the bond with its embedded options can be interpreted as an American-style interest rate derivative. This paper focuses on call and put options embedded in bonds with *advance notice*, that is, options with exercise decisions prior to exercise benefits.

As it is usually the case in practice, we assume that exercise of the call and put options is limited to the coupon dates posterior to a known *protection period* and that there is a *notice period* of fixed duration Δt. Thus, consider a coupon date t_m where the exercise is possible, and let C_m and P_m be the *call* and *put prices* at t_m. The decision to exercise or not by the issuer and the investor must be taken at $t_m - \Delta t$. If the issuer calls back the bond at $t_m - \Delta t$, he pays C_m to the investor at t_m, and, similarly, if the investor puts the bond at $t_m - \Delta t$, he receives P_m from the issuer at t_m. If no option is exercised at t_m, by the no-arbitrage principle of asset-pricing (Elliott and Kopp, 1999), the value of the bond is equal to the expected value of the bond at the next decision date, discounted at the risk-free interest rate. This expectation is taken under the so-called *risk-neutral probability measure*, where the uncertainty lies in the future trajectory of the short-term risk-free interest rate.

There are no analytical formulas for valuing American derivatives, even under very simplified assumptions. Numerical methods, essentially trees and finite differences, are usually used for pricing. Recall that trees are numerical representations of discrete-time and finite-space models and finite differences are numerical solution methods for partial differential equations. The pricing of American financial derivatives can also be formulated as a Markov Decision process, that is, a stochastic *Dynamic Programming* (DP) problem, as pointed out by Barraquand and Martineau (1995). Here the DP value function, that is, the value of the bond with its embedded options, is a function of the current time and of the current short term risk-free interest rate, namely the *state variable*. This value function verifies a DP recurrence via the no-arbitrage principle of asset pricing, the solution of which yields both the bond value and the optimal exercise strategies of its embedded options at all time during the bond's life. For an extensive coverage of stochastic DP, see Bertsekas (1995).

The pricing of options embedded in bonds can be traced back to Brennan and Schwartz (1977). They used a standard finite differences approach, which is however known to be instable in this context. Much later, Hull and White (1990a) proposed trinomial trees in the context of generalized versions with time-dependent parameters. Trees crudely approximate the dynamics of the underlying asset and convergence is not easy to attain. Büttler and Waldvogel (1996) (denoted here by BW) proposed an approach based on Green's functions, which turns out to be a DP procedure combined with finite elements, where the transition parameters are obtained by numerical integration. Recently, d'Halluin et al. (2001) (which we denote by DFVL) were able to stabilize the finite differences approach via flux limiters. They compared their results with those of Büttler and Waldvogel (1996), concluding that their method showed better accuracy than other methods. Here, we propose a Dynamic Programming approach where the transition parameters are derived in closed-form. As BW and DFVL, we use the Vasicek and CIR specification of the interest rate dynamics for our numerical experiments. Results are compared to those of BW and DFVL and show the efficiency and robustness of DP, with an accuracy comparable to DFVL and superior to BW. An interesting advantage of our approach is that it can be applied to generalized models calibrated to capture the term structure of interest rates, such as the ones proposed by Hull and White (1990a).

The rest of the paper is organized as follows. Section 2 presents a short review of the interest rate dynamics that have been proposed in the literature. Section 3 presents the model, the general DP formulation and the approximation procedure. Section 4 derives the exact joint distribution of the interest rate and integrated interest rate at a future date, conditional on the current value of the interest rate, for the Vasicek, CIR, and generalized Vasicek models. These joint distributions are used to obtain closed-form formulas for some constants used in our DP procedure. This turns out to be a key ingredient for its efficiency. Section 5 reports our numerical experimentations for the Vasicek and CIR models. Section 6 contains a conclusion.

2 Dynamics of the short term interest rate

As pointed out by Chan et al (1992), most of the alternative dynamics for the short term risk-free interest rate are described by the general stochastic differential equation

$$dR(t) = \kappa\left(\overline{r} - R(t)\right) dt + \sigma R(t)^{\theta} dB(t), \quad \text{for } 0 \le t \le T, \tag{1}$$

where $\{B(t),\, t \geq 0\}$ is a standard Brownian motion whereas κ (the *reverting rate*), $\overline{r}$ (the *reverting level*), σ (the *volatility*), and θ, are real-valued parameters. Table 1 summarizes various versions of (1) found in the literature in which a blank entry indicates a non-specified constant.

Table 1: Models for the short-term risk-free interest rate (on the last line, $\overline{r}(t)$ represents a deterministic function of t)

Model	$\overline{r}$	κ	θ
1. Vasicek (1977)			0
2. Brennan-Schwartz (1977)		0	1
3. Brennan-Schwartz (1980)			1
4. Marsh-Rosenfeld (1983)	0		
5. Cox-Ingersoll-Ross (1985a, b)			1/2
6. Hull-White (1990a)	$\overline{r}(t)$		0 or 1/2

Vasicek (1977) used a mean-reverting Ornstein-Uhlenbeck process. This model gives nice distributional results and closed-form solutions for zero-coupon bonds and for several European-style interest rate derivatives. But it has the undesirable property of allowing negative interest rates, although with very low probabilities. Several authors took advantage of its properties to price various interest rate derivatives, often in closed-form. Examples include Jamshidian (1989) and Rabinovitch (1989).

Brennan and Schwartz (1977, 1980) were pioneers on the modeling of options embedded in bonds. They let the interest rate move as a geometric Brownian motion without a drift to price the call and the put options (Model 2) and as a mean-reverting proportional process to price the conversion option (Model 3). Model 2 was also used by Dothan (1978) to price bonds in closed-form, and Model 3 by Courtadon (1982) to price several European as well as American options on bonds. Notice that Model 2 is a special case of Model 3 and that the latter includes the geometric Brownian motion of Black and Scholes (1973).

Model 4 corresponds to the so-called constant-elasticity-of-variance process (Cox 1996). It was considered by Marsh and Rosenfeld (1983), among others, as an alternative process for the interest rate.

Cox, Ingersoll, and Ross (1985a, b) (CIR) used the mean-reverting square-root process to handle the interest rate movements. This model is extensible to several factors, ensures strictly positive interest rates, and gives closed-form solutions for zero-coupon bonds and for some European-style interest rate derivatives. Several authors used the CIR model to price various interest rate derivatives (e.g., Richard 1978, Ananthanarayanan and Schwartz 1980, and Schaefer and Schwartz 1984). Notice that the Vasicek and the CIR models are special cases of the so-called single-factor affine term structures models. Duffie and Kan (1996) provide a characterization of multifactor affine term structure models. For the models described by (1), matching all theoretical bond values with their market counterparts is unfeasible because this gives much more equations than the number of parameters to estimate. A remedy, proposed by Hull and White (1990a), is to consider time-dependent parameters to better adjust the model to the observed yield curves. This leads to the generalized Vasicek and CIR models. Hull and White (1990b, 1993, 1994a, 1994b, 1996) interpreted the finite differences method as a trinomial tree, and priced several interest rate derivatives within their extended models. See also the note by Carverhill (1995) and the response by Hull and White (1995) for a discussion about the performance of these models.

3 A dynamic programming approach

In this section, we first provide a precise formulation of the model considered for a bond with its embedded call and put options. We then develop the DP equations and an approximation procedure to solve them. The pricing is made under a standard *no-arbitrage* assumption.

3.1 Model and formulation

Let $t_0, \ldots, t_n$ be a sequence of dates, where $t_0 = 0$ is the initial time, $t_1, \ldots, t_{n-1}$ are the *coupon dates*, and $t_n = T$ is the *maturity date* of the bond, i.e., the time at which the principal and the last coupon are due. The principal is scaled to 1 and the corresponding coupon is denoted by c. The periods $t_{m+1} - t_m$, for $m = 1, \ldots, n-1$, are equal except perhaps for the the first one, $t_1 - t_0$, which can be different. Assume also that the exercise decisions of the call and put options are made at times $\tau_m = t_m - \Delta t > t_{m-1}$, for $m = n^*, \ldots, n$. The benefits of an exercise decision at τ_m are obtained at the coupon date t_m. The lag $\Delta t = t_m - \tau_m$ is the *notice period* and the time increment $\tau_{n^*} - t_0$ is the *protection period* (against early exercising). All these dates t_m, and the values of Δt and n^*, are known in advance.

Let C_m and P_m be the *call* and *put prices* at t_m, for $m = n^*, \ldots, n$, respectively. Thus, if the issuer calls back the bond at τ_m, he pays C_m to the investor at t_m, and, similarly, if the investor puts the bond at τ_m, he receives P_m from the issuer at t_m. We assume that $0 \leq P_m \leq C_m$, as is usual in practice, and that $C_n = P_n = 1$.

Let $E^Q[\cdot \mid R(\tau) = r]$ represent the conditional expectation operator, conditional on $R(\tau) = r$ where $R(\tau)$ is the interest rate at time τ, under the so-called *risk-neutral probability measure* Q. Specific models for the dynamics of the interest rate process under Q will be considered in Sections 4 and 5. For $\delta \geq 0$, define

$$\rho(r, \tau, \delta) = E^Q\left[\exp\left(-\int_\tau^{\tau+\delta} R(t)dt\right) \,\middle|\, R(\tau) = r\right],$$

which represents the discount factor over the period $[\tau, \tau + \delta]$ when $R(\tau) = r$. Equivalently, this is the price at τ of a zero-coupon bond with maturity date $\tau + \delta$. Let $v(t, r)$ denote the value of the bond with its embedded options at time t if $R(t) = r$ and assuming that both parties always make their decisions in an optimal way.

We adopt the convention that $\tau_{n+1} = t_n = T$ and $\tau_0 = t_0$, and use the short notation $E_{m,r}[\cdot] = E^Q[\cdot \mid R(\tau_m) = r]$. We also denote $v(\tau_m, r)$ by $v_m(r)$ and $\rho(r, \tau_m, t_m - \tau_m)$ by $\rho_m(r)$. The ex-coupon holding value at time τ_m when $R(\tau_m) = r$, denoted $v_m^h(r)$, can be written as

$$v_m^h(r) = E_{m,r}\left[v_{m+1}\left(R(\tau_{m+1})\right)\exp\left(-\int_{\tau_m}^{\tau_{m+1}} R(t)dt\right)\right] \tag{2}$$

for $m = n^*, \ldots n-1$, and

$$v_0^h(r) = E_{0,r}\left[v_{n^*}\left(R(\tau_{n^*})\right)\exp\left(-\int_0^{\tau_{n^*}} R(t)dt\right)\right]. \tag{3}$$

For $m = n^*, \ldots, n$, by the no-arbitrage principle of asset pricing, the expected benefit from exercising or holding the bond with its embedded options at time τ_m is given by

$$\begin{array}{ll} \rho_m(r)(C_m + c) & \text{if the issuer calls,} \\ c\rho_m(r) + v_m^h(r) & \text{if the investor holds,} \\ \rho_m(r)(P_m + c) & \text{if the investor puts.} \end{array}$$

The optimal exercising strategies are as follows. The issuer should call the bond at τ_m if its holding value exceeds the exercise benefit, i.e., if

$$v_m^h(r) > C_m\rho_m(r);$$

the investor should put the bond at τ_m if the exercise benefit exceeds the holding value, i.e., if

$$v_m^h(r) < P_m\rho_m(r);$$

and the bond should be held for at least another period otherwise, i.e., if

$$P_m\rho_m(r) \le v_m^h(r) \le C_m\rho_m(r).$$

Notice that it cannot be optimal for the issuer to call and for the investor to put simultaneously, since $0 \le P_m \le C_m$.

Assembling all these ingredients, we find that

$$\begin{aligned} v_n(r) &= 1 + c, && (4) \\ v_m(r) &= c\rho_m(r) + \max\left\{P_m\rho_m(r),\, \min\left(C_m\rho_m(r),\, v_m^h(r)\right)\right\} && (5) \\ &\qquad \text{for } m = n^*, \ldots, n-1, \\ v_0(r) &= v_0^h(r) + c\sum_{m=1}^{n^*}\rho(r, 0, t_m). && (6) \end{aligned}$$

Equations (2) to (6) are the *DP recurrence equations*. Solving these equations backwards from the maturity date to the origin (i.e., computing the functions v_m^h and v_m successively for $m = n-1, n-2, \cdots, 0$) yields both the initial value of the bond and the optimal exercise strategies of its embedded options at all exercise dates, via the functions v_m^h. However, these functions do not have an analytical form, so they must be approximated numerically. An approximation method is detailed in the next section.

For all commonly encountered versions of model (1), conditionally on $R(\tau) = r$ and for $\delta > 0$, the two random variables $R(\tau + \delta)$ and $\int_\tau^{\tau+\delta} R(t)dt$ can be formally defined in a way that they are continuous in r. Combining this property with backwards induction on m, one can easily show that the functions v_m^h and v_m are also continuous in r, for $m = 0, \ldots, n$.

3.2 Approximation procedure

The general idea is to start from the function v_n, which is a known constant, use it in (2) to compute $v_{n-1}^h(r)$ at a finite number of values of r, use (5) to get $v_{n-1}(r)$ at these values of r, and interpolate these values to obtain an approximation of $\widehat{v}_{n-1}^h$ of v_{n-1}^h. This can be repeated iteratively for $m = n-2, \ldots, 1$: for a finite number of values of r, use the approximation $\widehat{v}_{m+1}$ of v_{m+1} in (2) to compute an approximation of $v_m^h(r)$, put it in (5) to get an approximation $\widehat{v}_m(r)$ of $v_m(r)$, and interpolate these values to obtain an approximation $\widehat{v}_m$ of v_m. Do it again for $m = 0$, but using (3) and (6) instead of (2) and (5). Note that the functions v_m^h and v_m are continuous for all m. In general, the expectation (or integral) in (2) can be approximated numerically for each r of interest, but we shall introduce shortcuts that can dramatically reduce the computing time at that stage by avoiding the explicit numerical integration.

There are of course many ways of defining the interpolations $\widehat{v}_m$ for each m (e.g., Hermite interpolation, splines, wavelets, etc.). Here, we use one of the simplest possible methods: *piecewise-linear interpolation* between the evaluation points. The main advantage of this simple method is that it allows us, under the dynamics presented here, to express the expectation in (2) in closed form when v_{m+1} is replaced by its piecewise-linear approximation.

To be more specific, let $a_0 < a_1 < \ldots < a_p < a_{p+1} = \infty$ be a fixed set of points forming a partition of $\mathbb{R}$. These are the values of r at which the functions v_m will be evaluated for the interpolation. The piecewise-linear approximation of v_m has the form

$$\begin{aligned}\widehat{v}_m(r) &= \begin{cases} 0 & \text{for } r < a_0; \\ \alpha_i^m + \beta_i^m r & \text{for } a_i \le r \le a_{i+1},\ i = 0, \ldots, p; \end{cases} \\ &= \sum_{i=0}^{p} (\alpha_i^m + \beta_i^m r)\, I\,(a_i \le r < a_{i+1})\end{aligned}$$

for some real-valued coefficients α_i^m and β_i^m, for $i = 0, \ldots, p$, where I denotes the indicator function. After computing the value of $\widehat{v}_m$ at $a_0, \ldots, a_p$, the $2(p+1)$ coefficients α_i^m and

β_i^m are obtained by solving the following system of $2(p+1)$ linear equations:

$$\begin{aligned}
\widehat{v}_m(a_0) &= \alpha_0^m + \beta_0^m a_0, && (7)\\
\widehat{v}_m(a_i) &= \alpha_i^m + \beta_i^m a_i \;=\; \alpha_{i-1}^m + \beta_{i-1}^m a_i \quad \text{for } i = 1, \ldots, p, && (8)\\
0 &= \beta_p^m. && (9)
\end{aligned}$$

The resulting function $\widehat{v}_m$ is continuous everywhere except perhaps at a_0. Now, if we use in (2) the piecewise-linear approximation of $\widehat{v}_{m+1}$, we obtain the following approximation for $v_m^h(a_k)$ at a grid point a_k:

$$\begin{aligned}
\widehat{v}_m^h(a_k) &= E_{m,a_k}\left[\exp\left(-\int_{\tau_m}^{\tau_{m+1}} R(t)dt\right)\widehat{v}_{m+1}\left(R(\tau_{m+1})\right)\right] && (10)\\
&= \sum_{i=0}^{p}\left(\alpha_i^{m+1}A_{k,i}^m + \beta_i^{m+1}B_{k,i}^m\right) && (11)
\end{aligned}$$

for $m = n^*, \ldots n-1$ and similarly for $m = 0$, where

$$\begin{aligned}
A_{k,i}^m &= E_{m,a_k}\left[\exp\left(-\int_{\tau_m}^{\tau_{m+1}} R(t)dt\right) I\left(a_i \le R(\tau_{m+1}) < a_{i+1}\right)\right], && (12)\\
A_{k,i}^0 &= E_{0,a_k}\left[\exp\left(-\int_{\tau_0}^{\tau_{n^*}} R(t)dt\right) I\left(a_i \le R(\tau_{n^*}) < a_{i+1}\right)\right], && (13)\\
B_{k,i}^m &= E_{m,a_k}\left[\exp\left(-\int_{\tau_m}^{\tau_{m+1}} R(t)dt\right) R(\tau_{m+1}) I\left(a_i \le R(\tau_{m+1}) < a_{i+1}\right)\right], && (14)\\
B_{k,i}^0 &= E_{0,a_k}\left[\exp\left(-\int_{\tau_0}^{\tau_{n^*}} R(t)dt\right) R(\tau_{n^*}) I\left(a_i \le R(\tau_{n^*}) < a_{i+1}\right)\right]. && (15)
\end{aligned}$$

In the next section, we obtain closed-form formulas for the constants or *transition parameters* $A_{k,i}^m$ and $B_{k,i}^m$ for the Vasicek, generalized Vasicek, and CIR models. For the Vasicek and CIR models, these constants depend on m only through $\tau_{m+1} - \tau_m$, and the discount factor $\rho_m(a_k)$ (for which a closed-form formula is also available) depends on τ_m and t_m only through $t_m - \tau_m$, which are constant for $m \ge n^*$ in our models. The availability of these closed-form formulas and the fact that the constants do not depend on m has a strong impact on the efficiency of the DP procedure. The piecewise-linear approximation of the function v_m was chosen precisely for that reason.

4 Conditional distributions and expectations for specific models

To derive formulas for the constants $A_{k,i}^m$ and $B_{k,i}^m$ defined in (12)-(15), for the Vasicek, generalized Vasicek, and CIR models, we obtain the distribution of the random vector

$$\left(R(\tau+\delta), \int_{\tau}^{\tau+\delta} R(t)dt\right) \tag{16}$$

conditional on the value of $R(\tau)$, for $0 \leq \tau \leq \tau + \delta \leq T$. This conditional distribution can also be used to price derivatives on the process $R(\cdot)$ over some time interval $[\tau,\, \tau+\delta]$. The formulas for the coefficients $A^m_{k,i}$, $B^m_{k,i}$ and $\rho(a_k, \tau_m, \delta)$ follow as corollaries.

4.1 The Vasicek model

Under the risk-neutral probability measure, the short rate process is the solution to the following stochastic differential equation

$$dR(t) = \kappa\left(\overline{r} - R(t)\right)dt + \sigma dB(t), \quad \text{for } 0 \leq t \leq T. \tag{17}$$

Closed form solutions for prices of zero-coupon bonds and European bond options are available in this model (see, e.g., Jamshidian 1989).

Theorem 1 *For the Vasicek model, the distribution of the random vector (16) conditional on $R(\tau) = r$ is bivariate normal with mean*

$$\mu(r,\delta) = \left(\mu_1(r,\delta), \mu_2(r,\delta)\right) = \left(\overline{r} + e^{-\kappa\delta}\left(r - \overline{r}\right), \overline{r}\delta + \frac{1 - e^{-\kappa\delta}}{\kappa}\left(r - \overline{r}\right)\right) \tag{18}$$

and covariance matrix

$$\Sigma(\delta) = \begin{bmatrix} \sigma_1^2(\delta) & \sigma_{12}(\delta) \\ \sigma_{21}(\delta) & \sigma_2^2(\delta) \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\delta}\right) & \frac{\sigma^2}{2\kappa^2}\left(1 - 2e^{-\kappa\delta} + e^{-2\kappa\delta}\right) \\ \sigma_{12}(\delta) & \frac{\sigma^2}{2\kappa^3}\left(-3 + 2\kappa\delta + 4e^{-\kappa\delta} - e^{-2\kappa\delta}\right) \end{bmatrix}. \tag{19}$$

Proof. By applying Itô's lemma to the process $\{\xi\left(t, R(t)\right) = e^{\kappa t}R(t),\ 0 \leq t \leq T\}$, we obtain that for $u \geq t$,

$$R(u) = \overline{r} + e^{-\kappa(u-t)}\left(r - \overline{r}\right) + \sigma \int_\tau^u e^{-\kappa(u-t)} dB\left(t\right),$$

and consequently that

$$\begin{aligned} \int_\tau^{\tau+\delta} R(u)du &= \overline{r}\delta + \frac{1 - e^{-\kappa\delta}}{\kappa}\left(r - \overline{r}\right) + \sigma \int_\tau^{\tau+\delta}\left(\int_\tau^u e^{-\kappa(u-t)} dB\left(t\right)\right) du \\ &= \overline{r}\delta + \frac{1 - e^{-\kappa\delta}}{\kappa}\left(r - \overline{r}\right) + \sigma \int_\tau^{\tau+\delta}\left(\int_t^{\tau+\delta} e^{-\kappa(u-t)} du\right) dB\left(t\right), \end{aligned}$$

where the last equality comes from Lemma 6 of the appendix. Conditioning on the information available at time τ, we can decompose $R(\tau+\delta)$ and $\int_\tau^{\tau+\delta} R(t)dt$ into a deterministic part and a random part. The random part turns out to be a limit of linear combinations of the same standard Brownian motion taken at different points in time. The random variables $R(\tau+\delta)$ and $\int_\tau^{\tau+\delta} R(t)dt$, conditional on $R(\tau)$, are thus jointly normal.

The mean vector and covariance matrix of the normally distributed random vector

$$\left(R(\tau+\delta),\ \int_{\tau}^{\tau+\delta} R(t)dt\right),$$

conditional on $R(\tau)$, can then be derived from the basic properties of stochastic integrals (e.g., Øksendal 1995). This is done in Lemma 7 of the appendix. □

Corollary 2 *For the Vasicek model, the transition parameters and discount factor are given by*

$$\rho(a_k,\tau_m,\delta) = \exp(-\mu_2\,(r,\delta)+\sigma_2^2(\delta)/2),$$
$$A_{k,i}^m = e^{-\mu_2(a_k,\delta)+\sigma_2^2(\delta)/2}\left[\Phi\left(x_{k,i}\right)-\Phi\left(x_{k,i-1}\right)\right]$$

and

$$B_{k,i}^m = e^{-\mu_2(a_k,\delta)+\sigma_2^2(\delta)/2}\Bigg[\left(\mu_1\left(a_k,\delta\right)-\sigma_{12}(\delta)\right)\left(\Phi\left(x_{k,i}\right)-\Phi\left(x_{k,i-1}\right)\right) - \frac{\sigma_1(\delta)}{\sqrt{2\pi}}\left(e^{-x_{k,i}^2}-e^{-x_{k,i-1}^2}\right)\Bigg],$$

where

$$\begin{aligned} x_{k,i} &= \left(a_i-\mu_1\left(a_k,\delta\right)+\sigma_{12}(\delta)\right)/\sigma_1(\delta) \quad \textit{for } i=0,\ldots,p \\ x_{k,-1} &= -\infty, \end{aligned}$$

the $\mu_i(a_k,\delta)$ and the $\sigma_{ij}(\delta)$ are given in (18) and (19) with $\delta=\tau_{m+1}-\tau_m$ for $m=n^*,\ldots n-1$ and $\delta=\tau_{n^*}-\tau_0$ for $m=0$, and Φ is the standard normal distribution function.

Proof. Using Theorem 1, the derivation is straightforward. □

4.2 The generalized Vasicek model

We now consider a special case of the *generalized Vasicek model* with time dependent long run mean, proposed by Hull and White (1990). Under the risk-neutral probability measure, the short rate process is the solution to the stochastic differential equation

$$dR(t) = \kappa\left(\overline{r}\,(t)-R(t)\right)dt+\sigma dB(t), \quad \text{for } 0\le t\le T. \tag{20}$$

This model requires the use of market data by fitting the zero-coupon yield curve. This can easily be done using, for instance, the Nelson and Siegel (1987) model. Once this is done, it provides an exact fit to the initial term structure and has a closed-form solution for the long run mean $\overline{r}\,(t)$, given by

$$\overline{r}\,(t) = -\frac{1}{\kappa}\frac{\partial^2\log(P(t))}{\partial t^2}-\frac{\partial\log(P(t))}{\partial t}+\frac{\sigma^2}{2\kappa^2}\left(1-e^{-2\kappa t}\right),$$

where $P(t)$ denotes the function fitted to the observed prices at time 0 of zero coupon bonds with maturity t.

Closed form solutions for prices of zero-coupon bonds and European bond options are also available in this model (see, e.g., Hull and White 1990a).

Theorem 3 *For this generalized Vasicek model, the distribution of the random vector (16) conditional on $R(\tau) = r$ is bivariate normal with mean*

$$\mu^{\star}(r,\delta) = \left(\mu_1^{\star}(r,\delta),\, \mu_2^{\star}(r,\delta)\right)$$

where

$$\begin{aligned}
\mu_1^{\star}(r,\delta) &= re^{-\kappa\delta} - \left.\frac{\partial \log(P(t))}{\partial t}\right|_{\tau+\delta} + e^{-\kappa\delta}\left.\frac{\partial \log(P(t))}{\partial t}\right|_{\tau} \\
&\quad + \frac{\sigma^2}{2\kappa^2}\left(1-e^{-\kappa\delta}\right) + \frac{\sigma^2}{2\kappa^2}e^{-\kappa\tau}\left(e^{-\kappa(\tau+\delta)} - e^{-\kappa\tau}\right), \\
\mu_2^{\star}(r,\delta) &= r\frac{\left(1-e^{-\kappa\delta}\right)}{\kappa} - \log\left(\frac{P(\tau+\delta)}{P(\tau)}\right) + \left.\frac{\partial \log(P(t))}{\partial t}\right|_{\tau}\frac{\left(1-e^{-\kappa\delta}\right)}{\kappa} \\
&\quad + \frac{\sigma^2}{2\kappa^2}\delta - \frac{\sigma^2}{2\kappa^3}\left(1-e^{-\kappa\delta}\right) + \frac{\sigma^2}{4\kappa^3}\left(e^{-\kappa(\tau+\delta)} - e^{-\kappa\tau}\right)^2,
\end{aligned} \tag{21}$$

and the same covariance matrix $\Sigma(\delta)$ as in Theorem 1.

Proof. The joint normality of $\left(R(\tau+\delta), \int_{\tau}^{\tau+\delta} R(t)dt\right)$ conditional on $R(\tau)$ can be proved by a similar argument as in Theorem 1. Stochastic integration then yields the mean and covariance matrix of this conditional distribution.

More specifically, by applying Itô's lemma to the process $\{\xi\left(t, R(t)\right) = e^{\kappa t}R(t),\, 0 \le t \le T\}$, we obtain that

$$R(u) = R(\tau)e^{-\kappa(u-\tau)} + \int_{\tau}^{u} \kappa\overline{r}\left(t\right)e^{-\kappa(u-s)}dt + \sigma\int_{\tau}^{u} e^{-\kappa(u-t)}dB(t).$$

This implies that

$$\begin{aligned}
\mu_1^{\star}(r,\delta) &= re^{-\kappa\delta} + \int_{\tau}^{\tau+\delta} \kappa\overline{r}\left(t\right)e^{-\kappa(\tau+\delta-t)}dt \\
&= re^{-\kappa\delta} + \int_{\tau}^{\tau+\delta}\Bigg(\kappa - \frac{1}{\kappa}\frac{\partial^2 \log(P(t))}{\partial t^2} - \frac{\partial \log(P(t))}{\partial t} \\
&\qquad\qquad + \frac{\sigma^2}{2\kappa^2}\left(1-e^{-2\kappa t}\right)e^{-\kappa(\tau+\delta-t)}\Bigg)dt \\
&= re^{-\kappa\delta} - \left.\frac{\partial \log(P(t))}{\partial t}\right|_{\tau+\delta} + e^{-\kappa\delta}\left.\frac{\partial \log(P(t))}{\partial t}\right|_{\tau} \\
&\quad + \frac{\sigma^2}{2\kappa^2}\left(1-e^{-\kappa\delta}\right) + \frac{\sigma^2}{2\kappa^2}e^{-\kappa(\tau+\delta)}\left(e^{-\kappa(\tau+\delta)} - e^{-\kappa\tau}\right).
\end{aligned}$$

The conditional mean $\mu_2^\star(r,\delta)$ can then be computed as follows:

$$\begin{aligned}
\mu_2^\star(r,\delta) &= E\left[\int_\tau^{\tau+\delta} R(t)dt \mid R(\tau)=r\right] \\
&= \int_\tau^{\tau+\delta} E\left[R(t)\mid R(\tau)=r\right]dt \\
&= \int_\tau^{\tau+\delta} \left(re^{-\kappa(t-\tau)} - \frac{\partial \log(P(t))}{\partial t} + e^{-\kappa(t-\tau)} \left.\frac{\partial \log(P(t))}{\partial t}\right|_\tau \right. \\
&\quad \left. + \frac{\sigma^2}{2\kappa^2}\left(1-e^{-\kappa(t-\tau)}\right) + \frac{\sigma^2}{2\kappa^2} e^{-\kappa t}\left(e^{-\kappa t} - e^{-\kappa\tau}\right)\right) dt \\
&= r\frac{\left(1-e^{-\kappa\delta}\right)}{\kappa} - \log\left(\frac{P(\tau+\delta)}{P(\tau)}\right) + \left.\frac{\partial \log(P(t))}{\partial t}\right|_\tau \frac{\left(1-e^{-\kappa\delta}\right)}{\kappa} \\
&\quad + \frac{\sigma^2}{2\kappa^2}\delta - \frac{\sigma^2}{2\kappa^3}\left(1-e^{-\kappa\delta}\right) + \frac{\sigma^2}{4\kappa^3}\left(e^{-\kappa(\tau+\delta)} - e^{-\kappa\tau}\right)^2.
\end{aligned}$$

The conditional covariance matrix is the same as in Theorem 1 and can be derived in the same way. □

Corollary 4 *For the generalized Vasicek model, the discount factor and the transition parameters are given by the same expressions as in Corollary 2, but with μ replaced by $\mu^\star$.*

Proof. The derivation is straightforward from Theorem 3 and we omit the details. □

4.3 The CIR model

Under the risk-neutral probability measure, the short rate process is the solution to the following stochastic differential equation:

$$dR(t) = \kappa\left(\overline{r} - R(t)\right)dt + \sigma\sqrt{R(t)}dB(t), \quad \text{for } 0 \le t \le T. \tag{22}$$

A key ingredient in this model is the joint distribution of the random vector (16) conditional on $R(\tau)=r$. This distribution can be characterized by its Laplace transform as follows (see, e.g., Feller 1951):

$$\begin{aligned}
& E\left[\exp\left(-\omega\int_\tau^{\tau+\delta} R(t)dt - \upsilon R(\tau+\delta)\right) \mid R(\tau)=r\right] \\
={}& \exp\left(X(r,\delta,\omega,\upsilon) - Y(r,\delta,\omega,\upsilon)r\right)
\end{aligned} \tag{23}$$

where

$$\begin{aligned}
\gamma(\omega) &= \sqrt{\kappa^2 + 2\omega\sigma^2}, \\
X(\delta,\omega,\upsilon) &= \frac{2\kappa\overline{r}}{\sigma^2}\log\left[\frac{2\gamma(\omega)e^{(\gamma(\omega)+\kappa)\delta/2}}{\left(\upsilon\sigma^2+\gamma(\omega)+\kappa\right)\left(e^{\gamma(\omega)\delta}-1\right)+2\gamma(\omega)}\right],
\end{aligned} \tag{24}$$

$$Y(\delta,\omega,\upsilon) \quad = \quad \frac{\upsilon\left(\gamma\left(\omega\right)+\kappa+e^{\gamma(\omega)\delta}\left(\gamma\left(\omega\right)-\kappa\right)\right)+2\omega\left(e^{\gamma(\omega)\delta}-1\right)}{\left(\upsilon\sigma^{2}+\gamma\left(\omega\right)+\kappa\right)\left(e^{\gamma(\omega)\delta}-1\right)+2\gamma\left(\omega\right)}.$$

When $\omega = 1$ and $\upsilon = 0$, the Laplace transform gives the price of a zero coupon bond. We also use this characterization to obtain the following proposition, whose proof is given in the appendix.

Proposition 5 *For the CIR model, the discount factor and transition parameters are given by*

$$\rho(a_k, \tau_m, \delta) = \exp\left(X(a_k, \delta, 1, 0) - Y(a_k, \delta, 1, 0)r\right),$$

$$A_{k,i}^m = \rho(a_k, \tau_m, \delta) \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^n}{t!} \left(F_{d+2t}\left(\frac{a_{i+1}}{\eta}\right) - F_{d+2t}\left(\frac{a_i}{\eta}\right)\right),$$

and

$$\begin{aligned} B_{k,i}^m \quad = \quad & \rho(a_k, \tau_m, \delta)\eta \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^t}{t!} \left[-2\left(a_{i+1} f_{d+2t}(a_{i+1}/\eta) - a_i f_{d+2t}(a_i/\eta)\right)\right. \\ & \left. +(d+2t)\left(F_{d+2t}(a_{i+1}/\eta) - F_{d+2t}(a_i/\eta)\right)\right] \end{aligned}$$

for $i = 0, \ldots, p$, where F_{d+2t} and f_{d+2t} are the distribution function and the density of a chi-square random variable with $d + 2t$ degrees of freedom, $X(\cdot), Y(\cdot)$ are defined in (24),

$$\begin{aligned} \delta &= \tau_{m+1} - \tau_m \text{ for } m = n^*, \ldots n-1, \\ \delta &= \tau_{n^*} - \tau_0 \text{ at } m = 0, \\ \gamma &= \sqrt{\kappa^2 + 2\sigma^2}, \\ \eta &= \frac{\sigma^2(e^{\gamma\delta} - 1)}{2(\gamma+\kappa)(e^{\gamma\delta} - 1) + 2\gamma}, \\ d &= \frac{4\kappa\overline{r}}{\sigma^2}, \text{ and} \\ \lambda_k &= \frac{8\gamma^2 e^{\gamma\delta} a_k}{\sigma^2\left[(\gamma+\kappa)(e^{\gamma\delta}-1)+2\gamma\right](e^{\gamma\delta}-1)}. \end{aligned}$$

5 Numerical illustrations

In this section, we compare the performance of DP with that of the methods proposed by Büttler and Waldvogel (1996) (denoted BW) and by d'Halluin et al. (2001) (denoted DFVL), on a numerical example taken from Büttler and Waldvogel (1996) and also considered by d'Halluin et al. (2001). The comparisons are made with both the Vasicek and CIR interest rate models.

The security to be priced is a 4.25% callable bond issued by the Swiss Confederation, over the period 1987–2012. At the pricing date $t_0 = 0$, December 23, 1991, the time to maturity was $T = t_n = 20.172$ years, with $n = 21$, a principal scaled to 1, a coupon $c = 0.0425$ coming once a year with the first coupon coming at time $t_1 = 0.172$, a notice period of 2 months, i.e., $\Delta t = t_m - \tau_m = 0.1666$, and a protection period of $t_{n^*} = 10.172$ years, with $n^* = 11$. The call prices are $C_{11} = 1.025$, $C_{12} = 1.020$, $C_{13} = 1.015$, $C_{14} = 1.010$, $C_{15} = 1.005$, and $C_{16} = \cdots = C_{21} = 1$.

The parameters $\overline{r}^*$, κ^*, σ^* given in Table 2 below were estimated by fitting the theoretical term structure at the pricing date. They were then adjusted via the *price of risk*, denoted here by q, to describe the dynamics of the interest rate under the risk-neutral probability measure required by no-arbitrage pricing. For the Vasicek model, one has $\overline{r} = \overline{r}^* + q\sigma^*/\kappa^*$, $\kappa = \kappa^*$, $\sigma = \sigma^*$, and $\theta = 0$. For the CIR model, one has $\overline{r} = \overline{r}^*\kappa^*/(\kappa^* + q)$, $\kappa = \kappa^* + q$, $\sigma = \sigma^*$, and $\theta = 1/2$.

Table 2: Input data for the Vasicek and CIR models

	Vasicek	CIR
$\overline{r^*}$	0.0348468515	0.0348468515
κ^*	0.44178462	0.54958046
σ^*	0.13264223	0.38757496
q	0.21166329	−0.40663675

The grid points $a_1, \ldots, a_p$ are selected to be equally spaced with $a_0 = -\infty$, $a_1 = \overline{r} - 6\sigma_1$ and $a_p = \overline{r} + 6\sigma_1$ for the Vasicek model, whereas $a_0 = 0$, $a_1 = 10^{-6}$ and $a_p = 3$, for the CIR model.

We first compare the accuracies of the three methods on a simple case where a closed-form solution is available for the exact value: the *straight bond*, without any embedded options. In other words, we assume that $n^* = n$ and keep the values given above for all other parameters. Tables 3 and 4 report the results for the Vasicek and CIR models, respectively. The column "Formula" gives the exact value obtained from the closed-form formula. The last two columns report the values obtained by BW and DFVL in their papers. Columns 2 to 4 give the approximations obtained by our DP procedure, for grid sizes with $p = 600$, 1200, and 2400. We find that the DP approximations converge very nicely to the exact values as p increases. The precision of DP compares advantageously with DFVL and BW. BW reported their results with only 4 digits.

Table 5 reports the value of the callable bond, with its embedded options and $n^* = 11$, obtained by the three approximation methods. The DP results are for a grid size with $p = 1200$. The results for $p = 2400$ are practically the same as those reported in the table (with absolute error less than 0.00003, comparable to the differences observed in Tables 3 and 4). This indicates that convergence has occurred. The results of DP and DFVL are close to each other, but the results of BW differ significantly for the Vasicek model.

Table 3: Convergence of the DP procedure for the Vasicek model

	Values of the Straight Bonds for Vasicek Model					
r	$p = 600$	$p = 1200$	$p = 2400$	Formula	BW	DFVL
0.01	0.92746	0.92743	0.92742	0.92742	0.9274	0.92739
0.02	0.90899	0.90896	0.90896	0.90895	0.9089	0.90892
0.03	0.89091	0.89089	0.89088	0.89088	0.8908	0.89084
0.04	0.87322	0.87319	0.87319	0.87318	0.8731	0.87315
0.05	0.85590	0.85588	0.85587	0.85587	0.8558	0.85583
0.06	0.83895	0.83893	0.83892	0.83892	0.8389	0.83887
0.07	0.82236	0.82233	0.82233	0.82233	0.8223	0.82228
0.08	0.80612	0.80610	0.80609	0.80609	0.8060	0.80604
0.09	0.79022	0.79020	0.79020	0.79019	0.7901	0.79014
0.10	0.77466	0.77464	0.77464	0.77464	0.7746	0.77458

Table 4: Convergence of the DP procedure for the CIR model

	Values of the Straight Bonds for CIR Model					
r	$p = 600$	$p = 1200$	$p = 2400$	Formula	BW	DFVL
0.01	0.95537	0.95528	0.95526	0.95525	0.9552	0.95527
0.02	0.93166	0.93157	0.93154	0.93154	0.9315	0.93155
0.03	0.90858	0.90848	0.90846	0.90845	0.9084	0.90846
0.04	0.88610	0.88601	0.88599	0.88598	0.8859	0.88599
0.05	0.86422	0.86414	0.86411	0.86411	0.8641	0.86411
0.06	0.84292	0.84284	0.84282	0.84281	0.8428	0.84281
0.07	0.82219	0.82211	0.82208	0.82208	0.8220	0.82207
0.08	0.80200	0.80192	0.80190	0.80189	0.8018	0.80188
0.09	0.78235	0.78227	0.78225	0.78224	0.7822	0.78223
0.10	0.76322	0.76314	0.76312	0.76311	0.7631	0.76309

Table 5: Bond values obtained by the three methods for the Vasicek and CIR models

	Values of Callable Bonds					
	Vasicek model			CIR model		
r	BW	DFVL	DP	BW	DFVL	DP
0.01	0.8556	0.84282	0.84285	0.9392	0.93926	0.93921
0.02	0.8338	0.82627	0.82630	0.9159	0.91598	0.91595
0.03	0.8223	0.81010	0.81009	0.8933	0.89333	0.89330
0.04	0.8062	0.79420	0.79423	0.8712	0.87127	0.87125
0.05	0.7904	0.77868	0.77871	0.8498	0.84980	0.84978
0.06	0.7749	0.76348	0.76351	0.8289	0.82890	0.82888
0.07	0.7598	0.74860	0.74862	0.8085	0.80855	0.80854
0.08	0.7450	0.73403	0.73406	0.7887	0.78874	0.78873
0.09	0.7305	0.71977	0.71980	0.7694	0.76945	0.76945
0.10	0.7163	0.70578	0.70583	0.7507	0.75067	0.75067

Our code is written in C, compiled with `gcc`, and executed on a 2.0 GHz Pentium 4 processor running under Windows XP. Our DP procedure takes between 2 and 3 seconds of CPU time, including all precomputations, to compute one entry of Table 5 for a grid size of $p = 1200$. As an indication, DFVL report 12 to 14 seconds and more than 200 seconds, respectively, for the DFVL and BW methods, on a Sun Ultra Sparc.

The value of the corresponding embedded call option can be obtained by

$$v_0^{\text{call option}}(r) = v_0^{\text{straight bond}}(r) - v_0^{\text{callable bond}}(r),$$

where $r = R(t_0)$ is the current interest rate at t_0.

6 Conclusion

In this paper, we model the valuation of options embedded in bonds as a stochastic dynamic programming problem. We propose a simple and efficient approximation for the case of three affine models of the term structure. Our method is based on a piecewise linear approximation of the value function and is quite easy to implement. It differs from the usual lattice-based and finite-difference methods in that the exact conditional probability distribution of the short rate is used in order to compute the transition probabilities. We provide numerical results for the Vasicek and CIR models. These results demonstrate convergence and an efficiency that compares favorably with that of the other available methods. In addition, our methodology readily extends to handle the Generalized Vasicek model with time-dependent long-run mean, provided that a function can be fitted to the observed zero-coupon bond prices. In that context, our pricing method can be used to investigate empirical features of the yield curve implied by bonds with embedded options.

Extensive empirical studies of the term structure generated by callable and puttable bonds have been lacking in the literature, mainly due to the lack of a robust numerical procedure for pricing these bonds.

Our solution technique is flexible and can be extended to other problems involving different continuous-time as well as discrete-time processes, whenever the transition parameters can be obtained efficiently.

7 Appendix

Lemma 6 *If f and g are two real-valued functions continuously differentiable in $[\tau, \tau + \delta]$, for $0 \leq \tau \leq \tau + \delta \leq T$, and $\{B(t), t \geq 0\}$ is a standard Brownian motion, then*

$$\begin{aligned} & \int_{\tau}^{\tau+\delta} \left(\int_{\tau}^{u} f(t) g(u) \, dB(t) \right) du \\ = & \int_{\tau}^{\tau+\delta} \left(\int_{\tau}^{\tau+\delta} f(t) g(u) I(t \in [\tau, u]) \, dB(t) \right) du \\ = & \int_{\tau}^{\tau+\delta} \left(\int_{\tau}^{\tau+\delta} f(t) g(u) I(u \in [t, \tau+\delta]) \, du \right) dB(t) \\ = & \int_{\tau}^{\tau+\delta} \left(\int_{t}^{\tau+\delta} f(t) g(u) \, du \right) dB(t) . \end{aligned}$$

Proof. One can use the integration by parts theorem of stochastic calculus (e.g., Øksendal 1995). Define the function $h(t) = f(t) \int_t^{\tau+\delta} g(u) \, du$ and transform the last expression in the theorem's statement as

$$\begin{aligned} & \int_{\tau}^{\tau+\delta} \left(\int_{t}^{\tau+\delta} f(t) g(u) \, du \right) dB(t) \\ = & \int_{\tau}^{\tau+\delta} h(t) \, dB(t) \\ = & h(\tau+\delta) B(\tau+\delta) - h(\tau) B(\tau) - \int_{\tau}^{\tau+\delta} \frac{\partial h(t)}{\partial t} B(t) \, dt \\ = & -\int_{\tau}^{\tau+\delta} f(t) g(u) B(\tau) \, du + \int_{\tau}^{\tau+\delta} f(t) g(t) B(t) \, dt \\ & -\int_{\tau}^{\tau+\delta} \left(\int_{t}^{\tau+\delta} \frac{\partial f(t)}{\partial t} g(u) B(t) \, du \right) dt. \end{aligned}$$

The same argument can be used to transform $\int_{\tau}^{u} f(t) g(u) \, dB(t)$ in order to obtain

$$\begin{aligned}\int_{\tau}^{\tau+\delta}\left(\int_{\tau}^{u} f(t)\, g(u)\, dB(t)\right) du &= \int_{\tau}^{\tau+\delta} f(u)\, g(u)\, B(u)\, du - \\ &\quad \int_{\tau}^{\tau+\delta} f(\tau)\, g(u)\, B(\tau)\, du - \\ &\quad \int_{\tau}^{\tau+\delta}\left(\int_{\tau}^{u} \frac{\partial f(t)}{\partial t} g(u)\, B(t)\, dt\right) du.\end{aligned}$$

The final result then comes from the basic properties of multidimensional real integrals. □

Lemma 7 *In Theorem 1, the mean vector and covariance matrix of the random vector* $\left(R(\tau+\delta), \int_{\tau}^{\tau+\delta} R(t)dt\right)$*, conditional on* $R(\tau) = r$*, are given by (18) and (19).*

Proof. The conditional mean is

$$E\left[\left(R(\tau+\delta), \int_{\tau}^{\tau+\delta} R(t)dt\right) \mid R(\tau) = r\right] = \left(\overline{r} + e^{-\kappa\delta}(r-\overline{r}), \overline{r}\delta + \frac{1-e^{-\kappa\delta}}{\kappa}(r-\overline{r})\right),$$

since the centered random vector

$$\left(\int_{\tau}^{\tau+\delta} e^{-\kappa(u-t)} dB(t), \int_{\tau}^{\tau+\delta}\left(\int_{t}^{\tau+\delta} e^{-\kappa(u-\tau)} du\right) dB(t)\right)$$

is independent of $\{R(t),\ t \leq \tau\}$.

The conditional variance of $R(\tau+\delta)$ is

$$\begin{aligned}\operatorname{Var}\left[R(\tau+\delta) \mid R(\tau) = r\right] &= E\left[\left(\int_{\tau}^{\tau+\delta} \sigma e^{-\kappa(\tau+\delta-t)} dB(t)\right)^2 \mid R(\tau) = r\right] \\ &= \sigma^2 \int_{\tau}^{\tau+\delta} e^{-2\kappa(\tau+\delta-t)} dt = \frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa\delta}\right).\end{aligned}$$

The conditional variance of $\int_{\tau}^{\tau+\delta} R(t)dt$ is

$$\begin{aligned}&\operatorname{Var}\left[\int_{\tau}^{\tau+\delta} R(t)dt \mid R(\tau) = r\right] \\ =\ & E\left[\left(\int_{\tau}^{\tau+\delta}\left(\int_{t}^{\tau+\delta} \sigma e^{-\kappa(u-\tau)} du\right) dB(t)\right)^2 \mid \mathcal{F}(\tau)\right] \\ =\ & \sigma^2 \int_{\tau}^{\tau+\delta}\left(\int_{\tau}^{\tau+\delta} e^{-\kappa(u-t)} du\right)^2 dt = \frac{\sigma^2}{2\kappa^3}\left(-3 + 2\kappa\delta + 4e^{-\kappa\delta} - e^{-2\kappa\delta}\right).\end{aligned}$$

The conditional covariance between $R(\tau+\delta)$ and $\int_\tau^{\tau+\delta} R(t)dt$ is

$$\begin{aligned}
& \operatorname{Cov}\left[R(\tau+\delta), \int_\tau^{\tau+\delta} R(t)dt \mid R(\tau)=r\right] \\
= \; & E\left[\int_\tau^{\tau+\delta} \sigma e^{-\kappa(\tau+\delta-t)} dB\left(t\right) \int_\tau^{\tau+\delta}\left(\int_\tau^{\tau+\delta} \sigma e^{-\kappa(u-t)} du\right) dB\left(t\right) \mid R(\tau)=r\right] \\
= \; & \sigma^2 \int_\tau^{\tau+\delta} e^{-\kappa(\tau+\delta-t)} \left(\int_\tau^{\tau+\delta} e^{-\kappa(u-t)} du\right) dt \\
= \; & \frac{\sigma^2}{2\kappa^2}\left(1-2e^{-\kappa\delta}+e^{-2\kappa\delta}\right).
\end{aligned}$$

□

Proof of Proposition 5. We start by defining a new probability measure $\widetilde{Q}$, known as the forward neutral measure (see, e.g., Geman, El Karoui and Rochet 1995), such that:

$$\frac{d\widetilde{Q}}{dQ} = \frac{e^{-\int_\tau^{\tau+\delta} R(t)dt}}{E\left[e^{-\int_\tau^{\tau+\delta} R(t)dt} \mid R(\tau)=r\right]}$$

Under this new measure, it can be proved that the conditional distribution of $R(\tau+\delta)$ is proportional to a non-central chi-square random variable. For a detailed discussion on the properties of square-root processes, see Jamshidian (1995,1996). More precisely, under $\widetilde{Q}$, $R(\tau+\delta)/\eta$ has the non-central chi-square distribution with non-centrality parameter λ and with d degrees of freedom, where

$$\lambda = \frac{8\gamma^2 e^{\gamma\delta} r}{\sigma^2\left[(\gamma+\kappa)(e^{\gamma\delta}-1)+2\gamma\right]\left(e^{\gamma\delta}-1\right)},$$

and where γ, η, and d are defined as in the theorem's statement. Now, we can rewrite $A_{k,i}^m$ in terms of the new probability measure $\widetilde{Q}$ as follows:

$$\begin{aligned}
A_{k,i}^m &= E_{m,a_k}\left[e^{-\int_{\tau_m}^{\tau_{m+1}} R(t)dt} I\left(a_i \le R(\tau_{m+1}) < a_{i+1}\right)\right] && (25) \\
&= \rho\left(a_k,\tau_m,\delta\right) \Pr^{\widetilde{Q}}\left[a_i \le R(\tau_{m+1}) < a_{i+1}\right] \\
&= \rho\left(a_k,\tau_m,\delta\right)\left(\Phi_{\lambda_k,d}\left(\frac{a_{i+1}}{\eta}\right) - \Phi_{\lambda_k,d}\left(\frac{a_i}{\eta}\right)\right) && (26)
\end{aligned}$$

where $\rho\left(a_k,\tau_m,\delta\right)$, the discount factor over the period $[\tau_m,\tau_m+\delta]$ when $R(\tau_m)=a_k$ is derived using the Laplace transform given by equation (23) and $\Phi_{\lambda_k,d}$ is the distribution function of a non-central chi-squared with non-centrality parameter λ_k and d degrees of freedom, with λ_k defined as in the theorem's statement. The distribution function $\Phi_{\lambda_k,d}$ can be written as

$$\Phi_{\lambda_k,d}(z) = e^{-\lambda_k/2} \sum_{t=0}^{\infty} \frac{(\lambda_k/2)^t}{\tau!} F_{d+2t}(z),$$

where F_{d+2t} is the chi-square distribution function with $d+2t$ degrees of freedom.

Combining this result with (26), we obtain

$$A^m_{k,i} = \rho\left(a_k, \tau_m, \delta\right) \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^t}{t!} \left(F_{d+2t}\left(\frac{a_{i+1}}{\eta}\right) - F_{d+2t}\left(\frac{a_i}{\eta}\right)\right)$$

where λ_k, d and η are defined as previously.

With the same change of probability measure, we also obtain

$$\begin{aligned} B^m_{k,i} &= E_{m,a_k}\left[e^{-\int_{\tau_m}^{\tau_{m+1}} R(t)dt} R(\tau_{m+1}) I\left(a_i \le R(\tau_{m+1}) < a_{i+1}\right)\right] \\ &= \rho\left(a_k, \tau_m, \delta\right) \eta E^{\widetilde{Q}}_{m,a_k}\left[\frac{R(\tau_{m+1})}{\eta} I\left(\frac{a_i}{\eta} \le \frac{R(\tau_{m+1})}{\eta} < \frac{a_{i+1}}{\eta}\right)\right]. \end{aligned}$$

The truncated expectation in this expression can be computed explicitly as

$$\begin{aligned} & E^{\widetilde{Q}}_{m,a_k}\left[\frac{R(\tau_{m+1})}{\eta} I\left(\frac{a_i}{\eta} \le \frac{R(\tau_{m+1})}{\eta} < \frac{a_{i+1}}{\eta}\right)\right] \\ = & \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^t}{t!} \int_{a_i/\eta}^{a_{i+1}/\eta} y f_{d+2t}(y) dy. \end{aligned}$$

After integration, it comes that

$$\begin{aligned} & E^{\widetilde{Q}}_{m,a_k}\left[\frac{r_{\tau_{m+1}}}{\mu} I\left(\frac{a_i}{\mu} \le \frac{r_{\tau_{m+1}}}{\mu} < \frac{a_{i+1}}{\mu}\right)\right] \\ = & \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^t}{t!} \left[-2\left(a_{i+1} f_{d+2t}(\frac{a_{i+1}}{\mu}) - a_i f_{d+2t}(\frac{a_i}{\mu})\right)\right. \\ & \left. +(d+2t) \times \left(F_{d+2t}(\frac{a_{i+1}}{\mu}) - F_{d+2t}(\frac{a_i}{\mu})\right)\right]. \end{aligned}$$

This implies that

$$\begin{aligned} B^m_{k,i} = & \rho\left(a_k, \tau_m, \delta\right) \eta \sum_{t=0}^{\infty} e^{-\lambda_k/2} \frac{(\lambda_k/2)^t}{t!} \left[-2\left(a_{i+1} f_{d+2t}(\frac{a_{i+1}}{\eta}) - a_i f_{d+2t}(\frac{a_i}{\eta})\right)\right. \\ & \left. +(d+2t)\left(F_{d+2t}(\frac{a_{i+1}}{\eta}) - F_{d+2t}(\frac{a_i}{\eta})\right)\right]. \end{aligned}$$

□

References

[1] Ananthanarayanan, A.L. and E.S. Schwartz, "Retractable and Extendible Bonds: The Canadian Experience," *Journal of Finance*, 35 (1980), 31–47.

[2] Barraquand, J. and D. Martineau, "Numerical Valuation of High Dimensional Multivariate American Securities," *Journal of Financial and Quantitative Analysis*, 30 (1995), 383–405.

[3] Bertsekas, D. P., *Dynamic Programming and Optimal Control, Volumes I and II*, Athena Scientific, Belmont, Mass., 1995.

[4] Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (1973), 637–654.

[5] Bliss, R.R. and E.I. Ronn, "To Call or not to Call? Optimal Call Policies for Callable U.S. Treasury Bonds," *Economic Review*, 80 (1995), 1–15.

[6] Brennan, M.J. and E.S. Schwartz, "Savings Bonds, Retractable Bonds, and Callable Bonds," *Journal of Financial Economics*, 5 (1977), 67–88.

[7] Brennan, M.J. and E.S. Schwartz, "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance*, 3 (1979), 133–155.

[8] Brennan, M.J. and E.S. Schwartz, "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, 15 (1980), 907–929.

[9] Büttler, H.J. and J. Waldvogel, "Pricing Callable Bonds By Means of Green's Function," *Mathematical Finance*, 1996 (6), 53–88.

[10] Carverhill,A., "A Note on the Models of Hull and White for Pricing Options on the Term Structure," *Journal of Fixed Income*, 5 (1995), 89–96.

[11] Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47 (1992), 1209–1227.

[12] Courtadon, G., "The Pricing of Options on Default-Free Bonds," *Journal of Financial and Quantitative Analysis*, 17 (1982), 75–100.

[13] Cox, J.C., "The Constant Elasticity of Variance Option Pricing Model," *Journal of Portfolio Management*, 22 (1996), 15–17.

[14] Cox, J.C., J.E. Ingersoll, and S.A. Ross, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53 (1985a), 385–407.

[15] Cox, J.C., J.E. Ingersoll, and S.A. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53 (1985b), 385–407.

[16] de Boor, C., *A Practical Guide to Splines*, Springer-Verlag, New York, 1978.

[17] d'Halluin, Y., P.A. Forsyth, K.R. Vetzal, and G. Labahn, "A Numerical PDE Approach for Pricing Callable Bonds," *Applied Mathematical Finance*, 2001 (8), 49–77.

[18] Dothan, L.U., "On the Term Structure of Interest Rates," *Journal of Financial Economics*, 6 (1978), 59–69.

[19] Duffie, D. and R. Kan., "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6 (1996), 379–406.

[20] Elliott, R.J. and P.E. Kopp, *Mathematics of Financial Markets*, Springer-Verlag, New York, 1999.

[21] Feller, W. "Two Singular Diffusion Problems," *Annals of Mathematics*, 54 (1951), 173–182.

[22] Geman, H., N. El Karoui, and J.C. Rochet., "Changes of Numraire, Changes of Probability Measure and Option Pricing," *Journal of Applied Probability*, 32 (1995), 443–458.

[23] Hull, J. and A. White, "Pricing Interest-Rate Derivative Securities," *Review of Financial Studies*, 3 (1990a), 573–592.

[24] Hull, J. and A. White, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25 (1990b), 87–100.

[25] Hull, J. and A. White, "One Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities," *Journal of Financial and Quantitative Analysis*, 28 (1993), 235–254.

[26] Hull, J. and A. White, "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives.* 2 (1994a), 7–16.

[27] Hull, J. and A. White, "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models," *Journal of Derivatives*, 2 (1994b), 37–48.

[28] Hull, J. and A. White, "A Note on the Models of Hull and White for Pricing Options on the Term Structure: Response," *Journal of Fixed Income*, 5 (1995), 97–103.

[29] Hull, J. and A. White, "Using Hull-White Interest Rate Trees," *Journal of Derivatives*, 4 (1996), 26–36.

[30] Jamshidian, F., "An Exact Bond Option Formula," *Journal of Finance*, 44 (1989), 205–209.

[31] Jamshidian,F., "A Simple Class of Square-Root Interest Rate Models," *Applied Mathematical Finance*, 2 (1995), 61–72.

[32] Jamshidian, F., "Bond Futures and Option Evaluation in the Quadratic Interest Rate Model," *Applied Mathematical Finance*, 3 (1996), 93–115.

[33] Marsh, T.A. and E.R. Rosenfeld, "Stochastic Processes for Interest Rates and Equilibrium Bond Prices," *Journal of Finance*, 38 (1983), 635–646.

[34] Nelson, C. R., and A. F. Siegel, "Parsimonious Modeling of Yield Curves," *Journal of Business* , 60 (1987), 473–489.

[35] Øksendal, B., *Stochastic Differential Equations, An Introduction with Applications*, Fourth Edition, Springer-Verlag, Germany, 1995.

[36] Rabinovitch, R., "Pricing Stock and Bond Options when the Default-Free Rate is Stochastic," *Journal of Financial and Quantitative Analysis*, 24 (1989), 447–457.

[37] Richard, S.F., "An Arbitrage Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 6 (1978), 33–57.

[38] Schaefer, S.M. and E.S. Schwartz, "A Two-Factor Model of the Term Structure: An Appropriate Analytical Solution," *Journal of Financial and Quantitative Analysis*, 19 (1984), 413–424.

[39] Vasicek, O., "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5 (1977), 177–188.