

Efficient Calibration of Trinomial Trees for One-Factor Short Rate Models *

MARKUS LEIPPOLD[†]

Swiss Banking Institute, University of Zurich

ZVI WIENER[‡]

School of Business Administration, Hebrew University of Jerusalem

May 7, 2003

*The authors thank seminar participants at the Z-Quants Workshop of the Zürcher Kantonalbank. Furthermore, we would like to thank Menachem Brenner, Paolo Vanini, Ton Vorst, and, in particular, Simon Benninga for his thoughtful comments. Markus Leippold acknowledges the financial support of the Swiss National Science Foundation (NCCR FINRISK). Zvi Wiener acknowledges the financial support of the Krueger and Rosenberg funds at the Hebrew University of Jerusalem. We welcome comments, including references to related papers we inadvertently overlooked.

[†]**Correspondence Information:** University of Zurich, ISB, Plattenstr. 14, 8032 Zurich, Switzerland; tel: +41 1-634-2951; fax: +41 1-634-4903; leippold@isb.unizh.ch.

[‡]**Correspondence Information:** Hebrew University of Jerusalem, Mount Scopus, Jerusalem, 91905, Israel; tel: +972-2-588-3049; fax: +972-2-588-1341; mwiener@mscc.huji.ac.il.

Efficient Calibration of Trinomial Trees for One-Factor Short Rate Models

ABSTRACT

In this paper we discuss the implementation of general one-factor short rate models with a trinomial tree. Taking the Hull-White model as a starting point, our contribution is threefold. First, we show how trees can be spanned using a set of general branching processes. Secondly, we improve Hull-White's procedure to calibrate the tree to bond prices by a much more efficient approach. This approach is applicable to a wide range of term structure models. Finally, we show how the tree can be adjusted to the volatility structure. The proposed approach leads to an efficient and flexible construction method for trinomial trees, which can be easily implemented and calibrated to both prices and volatilities.

JEL Classification Codes: G13, C6.

KEY WORDS: Short Rate Models, Trinomial Trees, Forward Measure.

In this paper, we elaborate on the implementation and calibration of one-factor short rate models. We contribute to the existing literature threefold. *First*, we show how spanning the tree can be generalized to allow for different alternative branching processes. This not only allows to use alternative branching processes to avoid negative interest rates as in Hull and White (1994), but also allows to obtain a “slender” tree at the edges. This can substantially reduce the computational time. Moreover, the additional flexibility in defining branching processes becomes important when pricing certain types of exotic options. For barrier options, as an example, a finer grid around the barrier helps to increase the convergence of the numerical tree method. *Second*, and this is our main contribution, we improve the procedure of Hull and White (1994) to calibrate the tree to bond prices by a computationally much more efficient approach. In particular, when pricing an interest rate derivative, one has first to perform a forward induction to match the tree, and secondly one has to do backward induction to price the instrument. The forward induction can become especially computationally intensive. With our approach, however, we are only left with backward induction. Thus, a substantial reduction in computational costs is achieved. Moreover, our approach is not only restricted to the extended Vasicek model, but it is applicable to a wide range of short rate models. *Finally*, we show how the tree can be adjusted to the volatility structure in such a manner that our approach to match the initial term structure is still applicable. This is done by exploiting the flexibility embedded in the trinomial tree model.

Our paper is structured as follows. Section 1 presents the basic setup and introduces the notation. Furthermore, this section gives a brief overview of term structure models, which can be handled with our approach. Section 2 elaborates on the modelling of the state variable

process within a trinomial tree. Section 3 presents our calibration procedures and provides some explicit examples. Section 4 concludes.

1. Model Setup

We assume that the market is complete and arbitrage opportunities are absent. Then, there exists a unique risk-neutral pricing measure \mathbb{P} on (Ω, \mathcal{F}) such that every claim discounted by the money market account is a martingale under \mathbb{P} . The state space is described by the one-dimensional process X_t with $X_0 = x$. We assume that X_t follows a time-homogeneous Markov process. Interest rates and bond prices can be expressed as functions of X_t , i.e. we assume $P(X_t, t, T)$ to be the time- t pricing functional for a zero bond in state X_t , which pays \$1 at the maturity date T . Then,

$$P(X_t, t, T) = \mathbb{E} \left[e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right],$$

where r_t is the short term interest rate and $\mathbb{E}[\cdot \mid \mathcal{F}_t]$ is the expectation operator under \mathbb{P} . We further require the zero bond to satisfy the conditions $P(X_t, t, t) = 1$ and $\lim_{T \rightarrow \infty} P(X_t, t, T) = 0$ as $T \rightarrow \infty$ for all X_t and t . To abbreviate notation, bond prices observed from the initial term structure, $P(x, 0, T)$, are denoted by $P^*(T)$. Given the above setting, specifying a short rate model can be done in two steps:

1. Defining a state variable process X_t . Here, we make the basic assumption that X follows an Itô diffusion,

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t, \quad X_0 = x, \quad (1)$$

under the measure \mathbb{P} .

2. Defining the functional form $g(X, t) = r_t$, which translates the state variable to the interest rate. Under the appropriate technical conditions, we can then write the general dynamics for the short rate as

$$dr_t = \mathcal{L}g(X, t)dt + \frac{\partial g(X, t)}{\partial X} \sigma_X(X_t) dW_t,$$

where \mathcal{L} is the extended generator.¹

Thus, depending on the choices μ_X, σ_X and $g(X, t)$, we can construct a wide class of short rate models. However, since the tree construction starts with spanning a trinomial tree for X_t by matching its moments, a simple choice for the process X_t is appropriate. When calculating examples in Section 3, we will therefore focus on a Gaussian specification for X_t .

To illustrate the generality of the above short rate specification, we briefly discuss a few possible models. We start with affine term structure models. These are obtained by setting $r_t = a + bX_t$. If X_t follows a mean-reverting Ornstein-Uhlenbeck process with $\mu_X(X_t) = -\kappa X_t$ and $\sigma_X(X_t) = \sigma$, then

$$dr_t = (a\kappa - \kappa r_t)dt + b\sigma dW_t, \tag{2}$$

which corresponds to the short rate model of Vasicek (1977). With time-dependent parameters equation (2) is often referred to as the extended-Vasicek model of Hull and White (1994). When $\mu_X(X_t) = \theta - \kappa X_t$ and $\sigma_X(X_t) = \sigma\sqrt{X_t}$, and $b > 0$, we get

$$dr_t = (\bar{\theta} - \kappa r_t)dt + \sigma\sqrt{b(r_t - a)} dW_t, \tag{3}$$

with $\bar{\theta} = b\theta + a\kappa$. Setting $a = 0$ we obtain the classical Cox, Ingersoll, and Ross (1985) term structure model. A lognormal model would be obtained by setting $\sigma_X(X_t) = \sigma X_t$.

Suppose now that $r_t = a + bX_t + cX_t^2$. Such a specification defines a short rate model which belongs to the class of quadratic models.² As a specific example of the quadratic class with Gaussian state variable X_t , take $r_t = cX_t^2$. Then, with $\mu_X(X_t) = \theta - \kappa X_t$ the interest rate process becomes

$$dr_t = (c\sigma^2 + 2(\sqrt{c}\theta\sqrt{r_t} + \kappa r_t)) dt + 2\sqrt{c}\sigma\sqrt{r_t}dW_t. \quad (4)$$

As a special case of the dynamics in (4), we can obtain a parameterized version of the Cox, Ingersoll, and Ross (1985) model, namely by setting $\theta = 0$. As another special case of the dynamics in (4) arises the double square-root interest rate model of Longstaff (1989), and Beaglehole and Tenney (1992). These authors investigate a one-factor model with the short rate process given as

$$dr_t = \hat{\kappa} \left(\frac{\hat{\sigma}^2}{4\hat{\kappa}} - \sqrt{r_t} \right) dt + \hat{\sigma}\sqrt{r_t}dW_t,$$

for some parameters $\hat{\kappa}, \hat{\sigma}$. It fits into our framework by setting $\kappa = 0$, $\theta = -\frac{1}{2}\sqrt{a}$ and $\sigma = \frac{1}{2}\hat{\sigma}/\sqrt{a}$ in (4).

In lognormal term structure models, the short rate is of the form $r_t = \exp(a + bX_t)$. A popular example is the Black and Karasinski (1991) model, which is a generalization of the

continuous-time formulation of the Black, Derman, and Toy (1990) model. The Black and Karasinski (1991) model assumes the logarithm of the interest rate to evolve according to

$$d \ln r_t = (\theta_t - \kappa_t \ln r_t) dt + \sigma_t dW_t. \quad (5)$$

To obtain the process in (5) we have to set $r_t = \exp(a_t + b_t X_t)$ with X_t following a normal distribution.³

2. State Variable Tree

The standard procedure to construct a trinomial tree approximation of the short rate process is to start spanning a tree for the state variable X_t on an equidistant tree. The process X_t is assumed to follow a time-homogeneous stochastic differential equation (SDE). Then, a tree representation can be constructed to provide a discrete-time and discrete-space Markov approximation for X_t . Usually, a trinomial tree is preferred to a binomial tree approximation, since the additional flexibility provided by the trinomial tree can be used to match not only the first, but also the second moment of the process X_t .

We make the following notational conventions. The nodes of the trinomial tree are denoted by (i, j) , where i is the vertical placement (the space axis) and j is the horizontal placement (the time axis). We define by $\pi_{i,j}^{k,h}$ the probability by which the process moves from node (i, j) to node $(i + k, j + h)$ within time interval h . For the standard branching process in the Hull and White (1994) model, it is assumed that *a*) the time interval h is constant for the whole tree, and *b*) the jump k is either 1, 0, or -1 . Hence, the transition from a node (i, j) to nodes

$\{(i-1, j+1), (i, j+1), (i+1, j+1)\}$ are equidistant both on the time as well as on the space axis. Furthermore, it is assumed that the nodes (i, j) and $(i, j+1)$ remain on the same vertical level. This branching process is illustrated in Panel (A) of Figure 2 and the resulting trinomial tree structure is illustrated in Figure 1. Having fixed the tree geometry, we have to determine the tree probabilities in such a way that the distributional properties of the state variable are mimicked at every node. However, instead of matching the variance as in Hull and White (1994), we match the second moment. This gives the same result, but simplifies the formulas somewhat. For the branching process given in Panel (A) of Figure 2, the system of equations is given by

$$\begin{aligned}
1 &= \pi_{i,j}^{-1,1} + \pi_{i,j}^{0,1} + \pi_{i,j}^{1,1}, \\
\mathbb{E}(X_{j+1}|X_{i,j}) \equiv M_1 &= \delta(\pi_{i,j}^{1,1} - \pi_{i,j}^{-1,1}) + X_{i,j}, \\
\mathbb{E}(X_{j+1}^2|X_{i,j}) \equiv M_2 &= \pi_{i,j}^{-1,1} (X_{i,j} - \delta)^2 + \pi_{i,j}^{0,1} X_{i,j}^2 + \pi_{i,j}^{1,1} (X_{i,j} + \delta)^2,
\end{aligned} \tag{6}$$

which is linear in the probabilities and hence straightforward to solve. In order to guarantee that the threesome $\{\pi_{i,j}^{-1,1}, \pi_{i,j}^{0,1}, \pi_{i,j}^{1,1}\}$ can be interpreted as probabilities for all i 's, we have to guarantee $\pi_{i,j}^{-1,1} + \pi_{i,j}^{0,1} + \pi_{i,j}^{1,1} = 1$ together with the three inequality constraints $\{\pi_{i,j}^{-1,1} \geq 0, \pi_{i,j}^{0,1} \geq 0, \pi_{i,j}^{1,1} \geq 0\}$. This can be done in several different ways. First, we can build some constraints on the number of time steps we are considering. However by doing so, we impose some severe restrictions on the depth of the tree. This method would probably fail to value either derivatives with complex payoff structures or long term instruments with intermediate payoffs, since such instruments require a reasonable depth for the tree. Another possibility is to relax the assumptions that the trinomial tree evolves to the neighbor states. In this case an

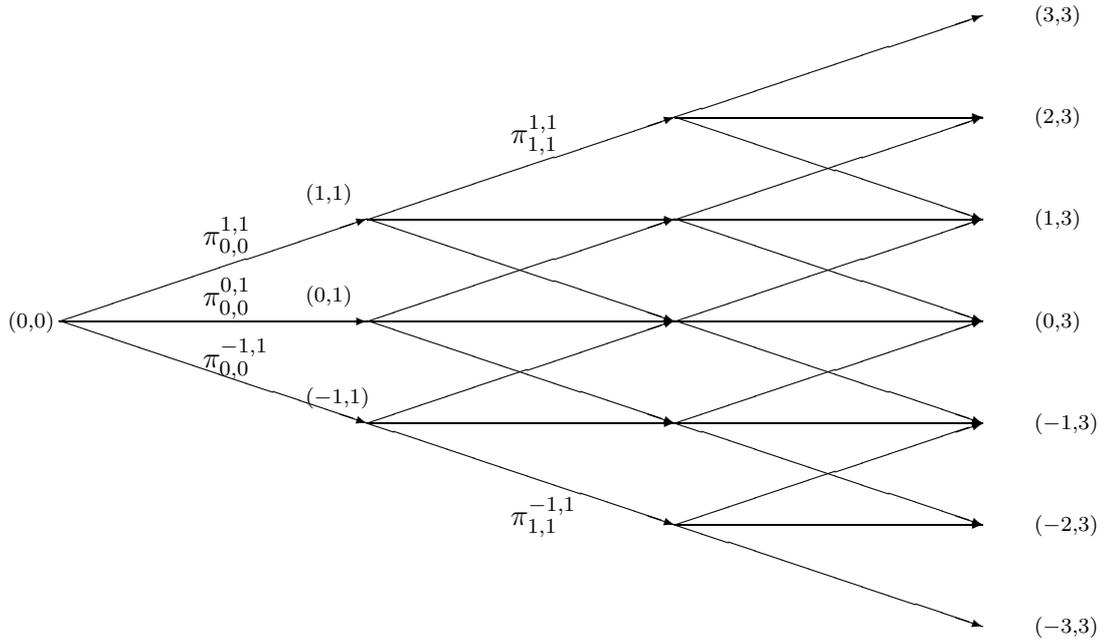


Figure 1. Trinomial Tree. The tree starts at note $(0,0)$. At each node there is a threesome $\{\pi_{i,j}^{-1,1}, \pi_{i,j}^{0,1}, \pi_{i,j}^{1,1}\}$ evolving to the neighbor nodes $(i-1, j)$, (i, j) , and $(i+1, j)$ respectively.

equation system with six variables and three equality constraints has to be solved. Thus, to uniquely select a particular threesome of transitions, we would have to impose some additional constraints. An alternative way to treat negative probabilities is by considering them as a finite difference scheme applied to the basic pricing PDE. In such a case we do not need the weights (probabilities) to be positive as soon as a specific finite difference scheme converges to the solution of the PDE.⁴

A more serious problem, in particular for Gaussian interest rate models, is the possibility of obtaining negative interest rates. As was done e.g. in Hull and White (1994), this can be avoided by altering the geometry of the tree. Of course, altering the geometry is an arbitrary manipulation of the pricing problem and thus subject to some criticism. Nevertheless, it is widely used in practice. An obvious way to avoid negative interest rates is to use the branching process illustrated in Panel (C) of Figure 2. At the same time, the branching process in

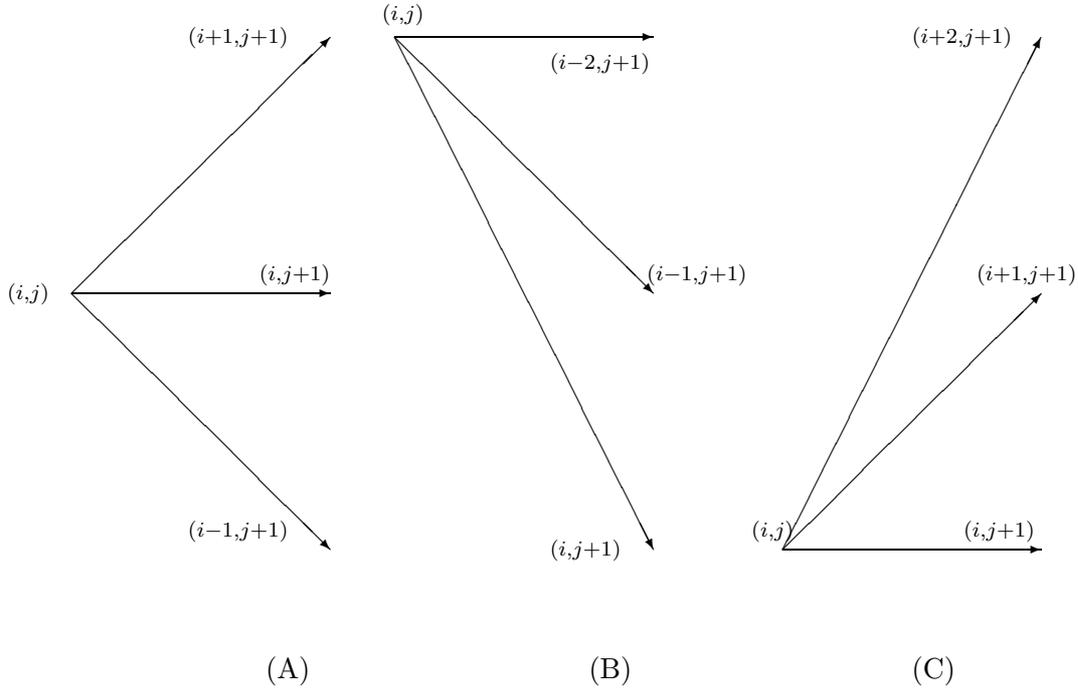


Figure 2. Branching Processes. In order to control the state spanned by the tree, the common branching processes (A) is altered to either (B) for high interest rates, or (C) for low interest rates. With the latter branching processes, negative interest rates within the tree can be avoided.

Panel (B) would avoid the possibility that the trinomial tree spans interest rates, which are unreasonably high. Depending on the branching process used, the system of equations for the tree probabilities in (6) has to be adjusted accordingly. However, we do not necessarily have to rely on the assumption that the tree evolves to the three neighboring states. We can be more general in two directions:

- The tree evolves from state (i, j) to the three states $\{(i + k_1, j + 1), (i + k_2, j + 1), (i + k_3, j + 1)\}$, requiring $k_1 \neq k_2 \neq k_3$.
- The tree directly evolves from state (i, j) to a threesome of states at time $j + h$, $h \geq 1$.

With these two generalizations, the equations to match the first two moments are given by

$$\begin{aligned} M_1^h &= X_{i,j} + \delta \left(\pi_{i,j}^{k_1,h} k_1 + \pi_{i,j}^{k_2,h} k_2 + \pi_{i,j}^{k_3,h} k_3 \right), \\ M_2^h &= \pi_{i,j}^{k_1,h} (X_{i,j} + k_1 \delta)^2 + \pi_{i,j}^{k_2,h} (X_{i,j} + k_2 \delta)^2 + \pi_{i,j}^{k_3,h} (X_{i,j} + k_3 \delta)^2. \end{aligned}$$

In this general setup, we then obtain

$$\begin{aligned} \pi_{i,j}^{k_1,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_2 + k_3)) M_1^h + (X_{i,j} + \delta k_2)(X_{i,j} + \delta k_3)}{\delta^2(k_1 - k_2)(k_1 - k_3)}, \\ \pi_{i,j}^{k_2,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_1 + k_3)) M_1^h + (X_{i,j} + \delta k_1)(X_{i,j} + \delta k_3)}{\delta^2(k_2 - k_1)(k_2 - k_3)}, \\ \pi_{i,j}^{k_3,h} &= \frac{M_2^h - (2X_{i,j} + \delta(k_1 + k_2)) M_1^h + (X_{i,j} + \delta k_1)(X_{i,j} + \delta k_2)}{\delta^2(k_3 - k_1)(k_3 - k_2)}. \end{aligned}$$

With these formulas for the probabilities at hand, we can construct a large structure of possible tree geometries. One possible structure of a generalized trinomial tree is plotted in Figure 3. Obviously, there are many degrees of freedom for building a trinomial tree. In general however, the more degrees of freedom, the less stable will the tree be when it comes to pricing derivatives. Therefore, one has to be careful, as in most cases flexibility comes at the price of less stability.

3. Calibrating the Tree

In practice, term structure models are implemented by calibrating them to the prices and volatilities of some subset of traded instruments. These instruments include e.g. US T-bonds, interest rate swaps and interest rate options like caps and swaptions. Typically, the drift of the short rate process is matched to the current term structure of bond prices. In Section 3.1

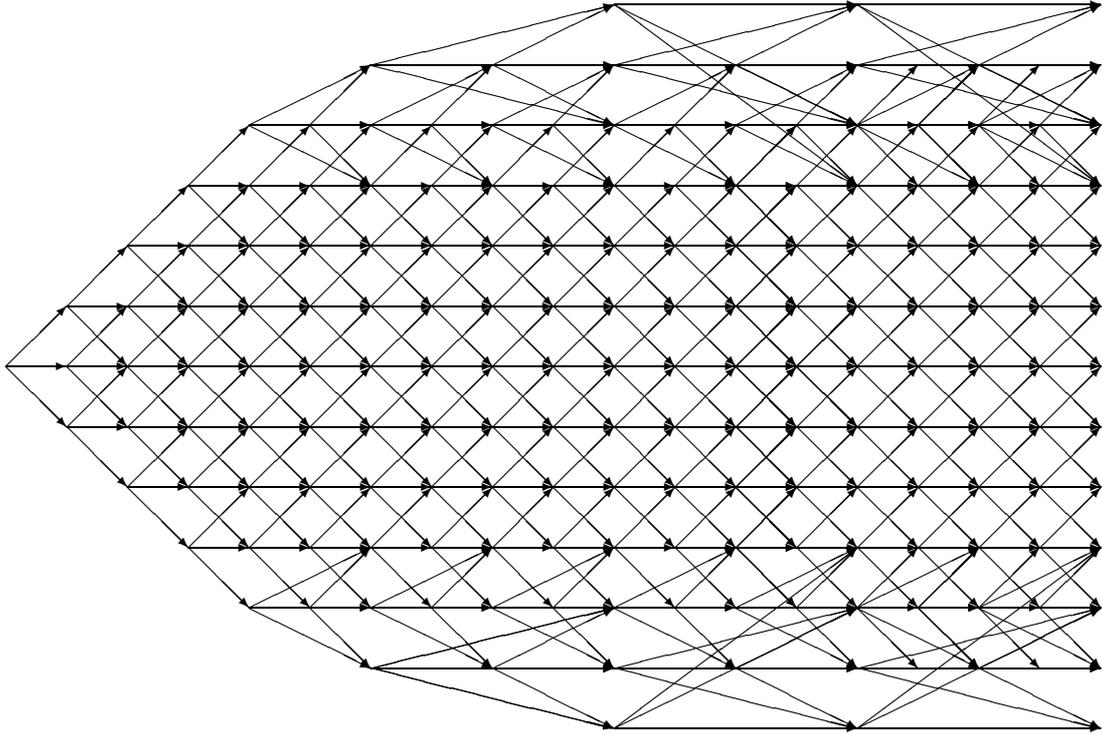


Figure 3. Possible tree structure. For high and for low values of X we slow down the tree. Furthermore, in the last two branches we use the alternative branching processes as illustrated in Figure 2.

we present a novel method to achieve this goal. In a second step, the volatility function of the short rate may then be chosen to match the term structure of volatilities of the yield curve, or the term structure of implied volatilities of at-the-money interest rate options. The latter is of particular importance when it comes to pricing of exotic interest rate options. We elaborate on this method in Section 3.2.

3.1. Matching Bond Prices Using Forward Measure

The common procedure for matching the tree to the term structure of bond prices is based on forward induction as used by Hull and White (1994) and first proposed by Jamshidian (1991).

This procedure obviously becomes computationally demanding for more involved functions $g(X, t)$. For completeness, we review this method in the Appendix.

In this section, however, we use a novel approach that builds on the use of the forward measure. Matching can be done analytically for a wide range of short rate models. Thus, compared to the analytical implementation of the Hull-White model recently proposed by Grant and Vora (2001), our procedure is not restricted to the extended Vasicek model. An additional advantage of our method is that we can directly match the tree to the discretely discounted forward rates, which is a more natural approach than matching to the instantaneous short rate. But most importantly, since we entirely circumvent the forward induction procedure, our method is significantly much more efficient. To this end, we exploit two properties of the forward measure:

Property 1 *Under the forward measure the forward rate is an unbiased estimator of the future interest rate.*

The first property allows us to match the tree in a straightforward manner and, moreover, the level shift is given in closed-form for a wide range of term structure models. Consider, for instance, the lognormal model of Black and Karasinski (1991). In the Hull-White framework, the tree matching procedure under lognormal short rates requires the use of a root search algorithm, which needs to be applied in every time slice of the tree. In our approach however, the level shift can be determined analytically for a large class of term structure models.

Property 2 *In discrete time the one-period forward measure equals the risk-neutral measure.*

From the second property it follows that the backward induction to determine derivative prices remains the same as in the standard trinomial tree. Therefore, no further adjustments have to be made and prices can be calculated the same way as with a tree spanned under the risk-neutral measure.

Changing the probability measure to the forward measure is a common tool in pricing derivative instruments, as it considerably facilitates the calculation of the corresponding expectation.⁵ Suppose that $\int_0^T g(X, s) ds < \infty$. Then, the T -forward measure \mathbb{P}^T is defined by

$$\frac{d\mathbb{P}^T}{d\mathbb{P} | \mathcal{F}_T} = \frac{\exp\left(-\int_0^T g(X, s) ds\right)}{\mathbb{E}\left[\exp\left(-\int_0^T g(X, s) ds\right)\right]} = P(t, T)^{-1} \exp\left(-\int_0^T g(X, s) ds\right).$$

Then, as stated in Property 1, we have

$$\mathbb{E}^T(r_T | \mathcal{F}_t) = f(t, T),$$

where \mathbb{E}^T is the expectation operator under the T -forward measure. As it is more natural to work with simple compounded interest rates (e.g. LIBOR rates), we denote by $f_\Delta^*(T) = (P(T)/P(T + \Delta) - 1)/\Delta$ the annualized, discrete time forward rate prevailing at time $[T, T + \Delta]$ as observed from the initial term structure. In order to match the initial term structure, we have to insure that

$$\mathbb{E}^T(r_T | \mathcal{F}_t) = \mathbb{E}^T(g(X, T) | \mathcal{F}_t) = f_\Delta^*(T),$$

holds for each time step in the trinomial tree. How can this be achieved? Up to now, we have constructed the tree for X , such that for every time step the conditional expectation of ΔX is matched under the risk-neutral measure. Therefore, the probabilities $\pi_{i,j}^{k,h}$ are \mathbb{P} -probabilities. Note that, in discrete time, the one-period forward measure equals the risk-neutral measure (see Property 2). So far, however, we did not yet have specified the measures of path probabilities for more than one period. This additional degree of freedom will be used to efficiently match the tree to the initial term structure. How this can be achieved will be discussed next.

Given the appropriate technical conditions, the extended generator of X_t under \mathbb{P}^T following the SDE in (1) can be written as

$$\mathcal{L}^T g = \mathcal{L}g + \Gamma(\log P(t, T), g),$$

where Γ is the “carré du champ operator” corresponding to \mathcal{L} (see Davis (1998)), defined as

$$\Gamma(f, g) = \mathcal{L}(fg) - g\mathcal{L}f - f\mathcal{L}g, \quad f, g \in \mathcal{D}(\mathcal{L}).$$

This means that in continuous time, the drift of the process dX_t is changed from $\mu(X_t)$ to $\mu(X_t) - \sigma(X_t)\frac{\partial}{\partial X} \log P(t, T)$ under the forward measure \mathbb{P}^T :

$$\begin{aligned} dX_t &= \left(\mu(X_t) - \sigma(X_t)\frac{\partial}{\partial X} \log P(t, T) \right) dt + \sigma_X(X_t)dW_t^T \\ &= (\mu(X_t) - v(X, t, T)) dt + \sigma_X(X_t)dW_t^T, \end{aligned} \tag{7}$$

where $v(X, t, T)$ is the instantaneous volatility of the bond price process $dP(t, T)/P(t, T)$, and W^T is the standard Brownian motion under the measure \mathbb{P}^T .

Now, the level shift of the original tree can be determined as follows. Using continuous time notation, we need to find a level shift, such that the forward rate is an unbiased estimate of the future interest rate, i.e.

$$\mathbb{E}^T (g(X, T) | \mathcal{F}_t) = f(t, T),$$

where

$$\begin{aligned} \frac{d\mathbb{E}^T [g(X, T) | \mathcal{F}_t]}{dt} &= \mathbb{E}^T [\mathcal{L}^T g(X, T) | \mathcal{F}_t] \\ &= \mathbb{E} [\mathcal{L}g(X, T) + \Gamma(\log P(t, T), g(X, T)) | \mathcal{F}_t]. \end{aligned}$$

Thus,

$$\begin{aligned} f(t, T) &= f(t, t) + \int_t^T \mathbb{E} [\mathcal{L}g(X, s) + \Gamma(\log P(s, T), g(X, s)) | \mathcal{F}_t] ds \\ &= \mathbb{E} [g(X, T) | \mathcal{F}_t] + \int_t^T \mathbb{E} [\Gamma(\log P(s, T), g(X, s)) | \mathcal{F}_t] ds \\ &= \mathbb{E}^T [r_T | \mathcal{F}_t], \end{aligned} \tag{8}$$

subject to $g(X, t) = f(t, t) = r_t$. In order to match the tree, we therefore alter the original tree for the Markov process X in the following way:

- At each time slice, we change the level of the tree for $g(X, t)$ by a function $\eta(t, T, X)$ defined by

$$\eta(t, T, X) = \int_t^T \mathbb{E} [\Gamma(\log P(s, T), g(X, s)) | \mathcal{F}_t] ds. \quad (9)$$

- The tree for r_t is now spanned under forward probability measures. At time step j , we have $\mathbb{E}^{j\Delta} [r_{j\Delta} | \mathcal{F}_0] = f(0, j\Delta)$ under measure $\mathbb{P}^{j\Delta}$. The one period forward measure equals the risk-neutral probability measure.

To further clarify our point, we next discuss a three examples: the extended-Vasicek model, the lognormal model, and the quadratic model. We start by discussing the extended Vasicek model. Assuming

$$dX_t = -\kappa X_t dt + \sigma dW_t, \quad X_0 = 0, \quad (10)$$

we fix the initial date to 0. We define the function $g(X, t) = g(X_t) = X_t$. Since in the extended Vasicek model, the volatility of the bond price process is a function of time only, we can write the forward rate in (8) as

$$f(t, T) = \mathbb{E} [g(X_T) | \mathcal{F}_t] + \eta(t, T).$$

Then, according to equation (9), the level shift at time $t = 0$ becomes

$$\eta(0, \Delta) = f_{\Delta}^*(\Delta) - \mathbb{E} (X_0 | \mathcal{F}_0) = f_{\Delta}^*(\Delta).$$

The first level shift is just zero, since $f_{\Delta}^*(\Delta) = r_0$. For the next time-step, $[\Delta, 2\Delta]$, we simply obtain

$$\eta(0, 2\Delta) = f_{\Delta}^*(2\Delta).$$

Obviously, we do not have to make any tedious calculations at all. The level shift at time step $(j - 1)\Delta$ is simply given by

$$\eta(0, j\Delta) = f_{\Delta}^*(j\Delta). \quad (11)$$

Finally, we can determine the measure change $d\mathbb{P}^T/d\mathbb{P}$ as

$$\frac{d\mathbb{P}^T}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \exp \left(\int_0^T \frac{\partial \eta(u, T)}{\partial T} dW_u - \frac{1}{2} \int_0^T \left(\frac{\partial \eta(u, T)}{\partial T} \right)^2 du \right).$$

Since the derivation of the bond price is a straightforward task, the \mathbb{P}_T -dynamics of X_t in equation (7) are explicitly obtained as

$$dX_t = \left(\frac{e^{-\kappa(T-t)} - 1}{\kappa} \sigma^2 - \kappa X_t \right) dt + \sigma dW_t^T, \quad X_0 = 0, \quad (12)$$

Indeed, whenever we have an affine term structure model, where bond prices allow the representation

$$P(t, T) = \exp(A(t, T) + B(t, T)r_t),$$

the change of measure is

$$\frac{d\mathbb{P}^T}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \exp \left(\int_0^T B(u, T) \frac{\partial g(X, u)}{\partial X} \sigma(X) dW_u - \frac{1}{2} \int_0^T \left(B(u, T) \frac{\partial g(X, u)}{\partial X} \sigma(X) \right)^2 du \right).$$

As was already pointed out in Kijima and Nagayama (1994) and Pelsser (1994), the level shift can be calculated analytically for the Vasicek model. They argue that the level shift equals the expected value of the future interest rate. Thus, there is no forward induction necessary. Hull and White (1996) object that this procedure does not provide an exact fit to the initial

term structure, because the tree is a discrete time representation of an underlying continuous process. Hence, the tree is only fitted exactly using the forward induction procedure, which would justify the additional computational costs. Here, however, we calculate the expected value under the *forward measure* to determine the level shift for the one-period forward rate, which equals the discretized short rate. Hence, the tree is matched *exactly* to the initial term structure, while saving considerable amount of computation time.

We next turn to the lognormal models. Here, the advantage of our matching procedure becomes most evident, since we can match the initial term structure by using a closed-form expression. Using forward induction would require a root search algorithm in every time slice (see Hull and White (1994)). Again, we assume dX_t as given in equation (10), but now with $r_t = ae^{X_t}$. The short rate becomes lognormally distributed with expectation

$$\mathbb{E}[r_T | \mathcal{F}_t] = a \exp\left(X_t e^{-\kappa(T-t)}\right) + a \exp\left(\frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa(T-t)}\right)\right).$$

Therefore, for the tree centered at $X_0 = 0$, and with fixed κ and σ ,

$$r_0 = f_{\Delta}^*(\Delta) = a + a \exp\left(\frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa\Delta}\right)\right)$$

determines the constant a . Then, to match the whole term structure, we have to introduce a level shift given by

$$\eta(0, j\Delta) = f_{\Delta}^*(j\Delta) - a - a \exp\left(\frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa j\Delta}\right)\right)$$

to determine the level shift at the $(j - 1)\Delta$ -th time step. Again, we recall that with the standard Hull and White (1994) procedure one would have to perform a root search algorithm for each time step in order to determine the level shift for matching the tree with lognormal interest rates.

As a final example emphasizing the advantage of our procedure, we consider the general quadratic model as in Ahn, Dittmar, and Gallant (2002) and Leippold and Wu (2002). With the parameters a, b, c such that $r_t = a + bX_t + cX_t^2 > 0$, we can span an interest rate tree with strictly positive interest rates. Again, we restrict ourselves to choosing X_t as a simple Ornstein-Uhlenbeck process given in equation (10). Then, calculating the expected short rate, we first have to fix the constants a and c such that

$$r_0 = f_{\Delta}^*(\Delta) = a + ce^{-2\kappa\Delta} \frac{\sigma^2}{2\kappa} (e^{\kappa\Delta} - 1)$$

Without loss of generality, we set $c = 1$. Then, to match the entire initial term structure, the level shifts for $j > 1$ can readily be calculated as

$$\eta(0, j\Delta) = f^*(j\Delta) - a - e^{-2\kappa j\Delta} \frac{\sigma^2}{2\kappa} (e^{2\kappa j\Delta} - 1). \quad (13)$$

3.2. Calibrating to the Term Structure of Volatilities

In this section, we show how the trinomial tree can be fitted in an efficient way to the initial term structure of volatilities. Calibrating the volatility structure is often subject to some criticism. Hull and White (1996) suggest that in a Markov model there should be only one time-

dependent parameter. Whenever the volatility is modelled as time-dependent, the resulting non-stationarity in the volatility curve may have many unexpected effects. In particular, any instrument whose price depends on future volatilities is liable to be mispriced.

Before we start discussing, how the trinomial tree can be manipulated to match the volatility structure, we next explore the theoretical underpinning of calibrating the volatility. First, note that in the previous section we introduced a measure change to determine the level shift for the interest rate tree. An absolutely continuous change of measure only affects the drift of the process, but the quadratic variation will not be affected. Hence, for the tree construction, it is more appropriate to first match the volatility and then match the forward rate curve.

To alter the diffusion coefficient, we have to introduce a time change.⁶ As we are only considering a deterministic time change, we can use the following result.

Lemma 1 *Consider the continuous time-dependent functions $c(t) > 0$ and $\tau(t) = \int_0^t c(s)^{-1} ds \geq$*

0. Define

$$\tilde{W}_t = \int_0^{\tau(t)} \sqrt{c(s)} dW_s.$$

Then \tilde{W}_t is a $\mathcal{F}_{\tau(t)}$ -Brownian motion. Further,

$$\int_0^{\tau(t)} dW_s = \int_0^t \sqrt{\partial_s \tau(s)} d\tilde{W}_s. \quad (14)$$

As can be seen from equation (14), the time-changed Brownian motion alters the volatility of the original Brownian motion. Hence, this technique offers a convenient tool to match the trinomial tree to the term structure of volatilities.

In the previous section, where we matched the tree to the forward rate curve only, we implicitly assumed $\tau(t) = t$ and hence, $c(t) = 1$. By introducing the time change, we now construct a tree for the new process

$$dX_t = \mu_X(X)dt + \tilde{\sigma}_X(X, t)d\tilde{W}_t, \quad (15)$$

where $\tilde{\sigma}_X(X, t) = \sigma_X(X_t)\sqrt{\partial_t\tau(t)}$. Since we use now the process (15), care has to be taken when matching to the initial forward rate curve. The results in the previous section have to be adjusted accordingly. We will show below, how this can be achieved in an efficient manner. The introduction of a deterministic clock for volatility matching was already introduced in Schmidt (1997). Contrary to Schmidt (1997), we do not change the length of the interval between subsequent time-steps to fit this concept into our trinomial framework. Instead, we adjust the jump size in each time step in such a way that the tree probabilities remain unchanged. Hence, the jump size δ is becoming a function of time and we denote it as δ_j .

Assume now that we want to calibrate our interest rate tree to the term structure of forward rate volatilities.⁷ We further assume that we are given a set of one-period forward volatilities denoted by $V^*(j\Delta)$. To determine the tree, we have to find a $\sigma_X(X, t)$, such that

$$V^*(j\Delta) = \text{var}(g(j\Delta, j\Delta) | \mathcal{F}_t) = \text{var}(r_{j\Delta} | \mathcal{F}_t)$$

holds. This is achieved by adjusting the system of equations for the tree probabilities. Note that the probabilities now become time dependent. Restricting ourselves to $h = 1$ and $k = \{-1, 0, 1\}$, we obtain

$$\begin{aligned}
1 &= \pi_{i,j}^{-1,1} + \pi_{i,j}^{0,1} + \pi_{i,j}^{1,1}, \\
M_1 &= \delta_j(\pi_{i,j}^{1,1} - \pi_{i,j}^{-1,1}) + X_{i,j}, \\
M_2 &= \pi_{i-1,j}^{-1,1} (X_{i,j} - \delta_j)^2 + \pi_{i,j}^{0,1} X_{i,j}^2 + \pi_{i,j}^{-1,1} (X_{i,j} + \delta_j)^2, \\
V^*(j\Delta) &= \text{var}(g(X_{i,j}, j\Delta)|X_{i,j}).
\end{aligned}$$

Whenever X_t is Gaussian, the tree will still be centered at zero. Hence, the level shift to be applied at each time-step is determined as in the previous section.

3.3. Numerical Examples

As a first example we consider an affine and a quadratic term structure model based on a Gaussian state variable X_t .⁸ We assume that X_t follows the process

$$dX_t = -0.2X_t dt + 0.1dW_t.$$

We choose time steps as $\Delta = 1$. Furthermore, we set

$$\delta = \sqrt{3}\sigma \sqrt{\frac{1 - e^{-2\kappa\Delta}}{2\kappa}},$$

for numerical reasons (see Hull and White (1994)). The initial term structure is assumed to be given as

$$P^*(\Delta) = e^{-(0.08-0.05e^{-0.18\Delta})\Delta}. \quad (16)$$

Finally, we set

$$g(X) = \frac{1}{2}X_t,$$

for the affine Gaussian model and

$$g(X) = \frac{1}{2}X_t + X_t^2,$$

for the quadratic Gaussian model. We match the tree using the procedure outlined in Section 3.1. Then, for the short rate, the matching procedure is given by equation (11). Hence, the middle node for each time slice is shifted upwards according to the prevailing forward rate. For the quadratic term structure model, the level shift is given by equation (13). The upward shift by the forward rate has to be adjusted by a correction term. This correction term can be calculated in closed form for the quadratic model. As already pointed out, it can be calculated in closed form for any model which allows a closed-form expression for $\mathbb{E}^T [g(X_t) | \mathcal{F}_t]$. If such a closed-form expression is not available however, the calculation of this correction term can be done numerically using the path-probabilities in the tree, since for $T = j\Delta$,

$$\mathbb{E}^T [g(X_t) | \mathcal{F}_t] \approx \sum_i \Pi(i, j\Delta) X_{i,j},$$

where by $\Pi(i, j\Delta)$ we denote the sum of all path probabilities leading to state (i, j) . Certainly,

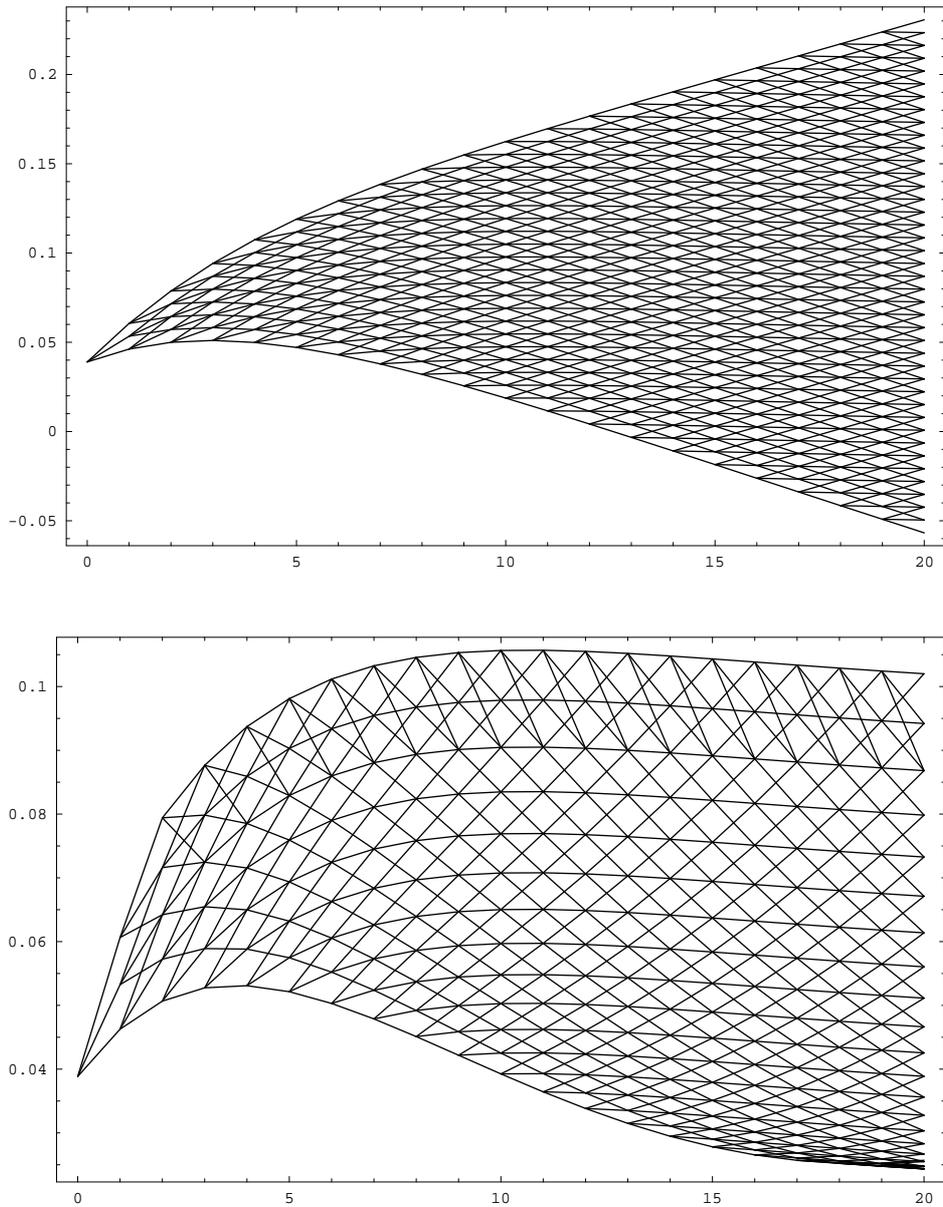


Figure 4. Trinomial trees for an affine (upper panel) and a quadratic (lower panel) Gaussian term structure model. The affine model leads to negative interest rates, when no alternative branching processes are introduced. The Gaussian model has only positive interest rates. To avoid to high interest rates for this model, we introduced the alternative branching process (B) from Figure 2.

using such a numerical procedure to determine the level shift comes at an additional computational cost. Nevertheless, we think that this cost is minor compared to that of using a forward induction method involving a root-search algorithm to match the tree to the term structure. The corresponding trees are plotted in Figure 4 for the affine model (upper panel) and the quadratic model (lower panel). As one can observe, the matching procedure gives rise to time dependent shifts of the original symmetric tree spanned for X . Furthermore, the disadvantage of using the affine tree is that negative interest rates are produced. One could now alter the geometry of the tree as outlined in Section 2. For more on alternative branching processes and its implementation in Mathematica we refer e.g. to Leippold and Wiener (1999). Another possibility to avoid negative interest rates is altering the function $g(X_t)$, e.g. by assuming this function to be quadratic in X_t . The trinomial tree for such a choice is plotted in the lower panel of Figure 4. Clearly, the short rate is not becoming negative. Note that the matching of this tree comes at almost no computational cost, as the shift is available in closed-form (see equation (13)). Also recall that, so far, we have spanned the tree under the forward measure and not under the risk-neutral measure.

Pricing European claims is a straightforward task. Under the forward measure, the time- t price V_t of a European claim with payoff V_T at expiration time T , we have

$$V_t = P^*(T) \mathbb{E}^T [V_T | \mathcal{F}_t].$$

Thus, when using the tree for valuation, we only need to discount the payoffs at T with the appropriate discount factor, i.e. the appropriate (and observable) bond prices. The same is true for all other path-independent options, which can be decomposed into a portfolio of European

Tree Depth	Gaussian			Lognormal	
	Analytical Value	Tree Value		Tree Value	
		absolute	relative	absolute	relative*
10	4.14431	4.14346	-0.0206 %	4.15055	-0.0109 %
50	4.14431	4.14411	-0.0048 %	4.15018	-0.0019 %
100	4.14431	4.14422	-0.0021 %	4.15014	-0.0010 %
150	4.14431	4.14425	-0.0015 %	4.15012	-0.0006 %
200	4.14431	4.14427	-0.0011 %	4.15011	-0.0004 %
250	4.14431	4.14428	-0.0008 %	4.15011	-0.0003 %
300	4.14431	4.14429	-0.0007 %	4.15011	-0.0003 %
500	4.14431	4.14429	-0.0005 %	4.15010	-0.0001 %
1000	4.14431	4.14430	-0.0002 %	4.15010	-0.0000 %

Table 1

One-year Put option on a three-year zero bond with face value 100. Strike price is 93.5. We set $g(X_t) = X_t$ and $dX_t = -0.2X_t dt + 0.01dW_t$. The tree was matched to the initial term structure using the procedure from Section 3.1. In case of the Gaussian model the relative approximation error was calculated relative to the closed-form solution. For the lognormal model, the deviations are relative to the tree value with depth 1000.

options. Even the calculation of American options is straightforward and the usual recursion method can be applied. Since in the current discrete time setting, the one-period forward measure equals the risk-neutral probability measure, we can determine American option prices by recursively working through the tree.

We end this section by comparing our numerical procedure with an analytical formula for bond options. In Table 1 we checked convergence of a European put option on a zero bond with face value 100, when interest rates follow either a Gaussian process or a lognormal process, i.e. $g(X_t) = X_t$ and $g(X_t) = e^{X_t}$ respectively. We assume that the put option matures in one year and is written on a zero maturing in three years. We further assume $g(X_t) = X_t$ and $dX_t = -0.2X_t dt + 0.01dW_t$. The initial term structure is given by (16). The strike price is set equal to 93.5. The trinomial tree is spanned under the forward measure. For the Gaussian model, we can compare the tree prices with the analytical formula. For the lognormal model

we calculated the relative deviation with respect to the model value obtained by a trinomial tree of depth 1000. It turns out that the trinomial tree spanned under the forward measure converges to the true value as expected. Its convergence is comparable with the standard trinomial trees, but again, we emphasize that our matching procedure is not only much more efficient, but also more flexible.

4. Conclusion

In this paper, we elaborated on some extensions and generalizations of the traditional trinomial tree models for interest rates. We furthermore showed how the tree matching procedure can be reformulated in a much more efficient way. Our approach is based on the forward measure methodology. For a large group of term structure models, this allows us to determine the level shifts in closed-form. Hence, our approach simplifies and considerably improves current practice. Furthermore, it is robust in the sense that it can still be applied when the tree is also required to match the term structure of volatilities.

Appendix

Matching by Forward Induction

Consider a trinomial tree with starting point $(0, 0)$. All quantities are expressed as annualized quantities. One year is partitioned into subperiods of length Δ . To simplify the subsequent analysis, we will consider standard branching processes only and set $h = 1$ and $k \in \{-1, 0, 1\}$.

Today's one-period bond price $P^*(\Delta)$ is assumed to be known, i.e. extracted from the market quotes by some estimation procedure. Then,

$$P^*(\Delta) = e^{-r_{0,0}\Delta},$$

with $r_{i,j}$ the annualized, continuously compounded short rate in state (i, j) prevailing over the time period $[j\Delta, (j+1)\Delta]$. Before considering the second time step, we introduce the concept of a state-price. The state-price is denoted by $Q_{i,j}$. In the following, the state-price $Q_{i,j}$ can be thought of today's price of a security that pays exactly \$1 if state (i, j) occurs, and \$0 in every other state. Then, $Q_{0,0} = 1$ and in the standard trinomial tree we have

$$P^*(j\Delta) = \sum_{i=-j}^j Q_{i,j}. \tag{A.1}$$

Now, moving from time Δ to 2Δ , we observe the following⁹:

$$P^*(2\Delta) = e^{-r_{0,0}\Delta} (\pi_1 e^{-r_{1,1}\Delta} + \pi_0 e^{-r_{0,1}\Delta} + \pi_{-1} e^{-r_{-1,1}\Delta}).$$

For large j , it would be rather cumbersome to write this up. Using the state-price formulation, we can considerably simplify the above procedure by writing the bond price as

$$P^*(2\Delta) = \sum_{i=-2}^2 Q_{i,2} = \sum_{i=-1}^1 Q_{i,1} e^{-r_{i,1}\Delta}.$$

The $Q_{i,\Delta}$ have still to be determined. This is achieved by forward induction. We know $Q_{0,0} = 1$ and $r_{0,0}$. So, for the next time-step

$$Q_{i,1} = \pi_i e^{-r_{0,0}\Delta}.$$

Generalizing the above procedure, the bond price $P^*((j+1)\Delta)$ can be written as

$$P^*((j+1)\Delta) = \sum_{i=-j}^j Q_{i,j} e^{-r_{i,j}\Delta}.$$

This form is much more amenable for determining the level shift needed to match the term structure. Once the interest rate at time-slice j is determined by matching the tree, the state-prices for the subsequent time step can be calculated as

$$Q_{i,j+1} = \sum_m Q_{m,j} \pi_m e^{-r_{m,j}\Delta},$$

where m is determined by the paths leading to node $(i, j+1)$. For now, we are still lacking a piece. Recall $r_t = g(X_t, t)$ and consider now, e.g. the function $r_t = g(\alpha_t + X_t)$ with α_t a deterministic function of time. Then,

$$P^*((j+1)\Delta) = \sum_{i=-j}^j Q_{i,j} e^{-g(\alpha_j + x_{i,j})\Delta}.$$

To determine the level shift α_j , we have to invert the above relation. For illustration, consider the affine function $r_t = \alpha_t + \beta X_t$. Hence, we obtain

$$\alpha_j = \frac{1}{\Delta} \log \left(\sum_{i=-j}^j Q_{i,j} e^{-\beta x_{i,j} \Delta} \right) - \frac{\log P^*((j+1)\Delta)}{\Delta} \quad (\text{A.2})$$

Note, the above procedure implicitly assumes the π 's to be probabilities under the risk-neutral measure. Therefore, when pricing claims using the matched tree, one should recall that the tree is spanned under the risk-neutral measure.

Notes

¹The extended generator of the process X is defined as $\mathcal{L} = \frac{\partial}{\partial t} + \frac{\partial}{\partial X}\mu_X(X) + \frac{1}{2}\frac{\partial}{\partial X^2}\sigma_X^2(X)$, see e.g. Arnold (1974), p. 180.

²General quadratic term structure models have been discussed by Ahn, Dittmar, and Gallant (2002), Leippold and Wu (2002).

³Such dynamics for the short interest rate would lead to infinite prices for Eurodollar futures as shown by Hogan and Weintraub (1993). Nevertheless, the Black and Karasinski (1991) model is one of the most popular models used in practice.

⁴We thank Ton Vorst for this remark.

⁵See e.g. Jamshidian (1989), Geman, El Karoui, and Rochet (1995), Benninga, Björk, and Wiener (2002)

⁶See e.g. Durrett (1996)

⁷Typically, they are either estimated historically or determined by at-the-money interest rate derivatives such as caps or swaptions.

⁸We implemented the models in *Mathematica*. The files can be obtained from the authors on request.

⁹Since in this simple setup probabilities are not time dependent, we just simplify the notation to π_i .

References

- Ahn, D.-H., R. F. Dittmar, and A. R. Gallant, 2002, “Quadratic Term Structure Models: Theory and Evidence,” *Review of Financial Studies*, forthcoming.
- Arnold, L., 1974, *Stochastic Differential Equations: Theory and Application*, Krieger.
- Beaglehole, D. R., and M. Tenney, 1992, “A Nonlinear Equilibrium Model of Term Structures of Interest Rates: Corrections and Additions,” *Journal of Financial Economics*, 32, 345–454.
- Benninga, S., T. Björk, and Z. Wiener, 2002, “On the Use of Numeraires in Option Pricing,” *Journal of Derivatives*, 10, 43–58.
- Black, F., E. Derman, and W. Toy, 1990, “A One-factor Model of Interest Rates and Its Application to Treasury Bond Options,” *Financial Analysts Journal*, 46, 33–39.
- Black, F., and P. Karasinski, 1991, “Bond and Option Pricing When Short Rates Are Lognormal,” *Financial Analysts Journal*, 47, 52–59.
- Cox, J. C., J. E. Ingersoll, and S. R. Ross, 1985, “A Theory of the Term Structure of Interest Rates,” *Econometrica*, 53, 385–408.
- Davis, M., 1998, “A note on the forward measure,” *Finance and Stochastics*, 2, 19–28.
- Durrett, R., 1996, *Stochastic Calculus*. Probability and Stochastics Series, CRC Press.
- Geman, H., N. El Karoui, and J.-C. Rochet, 1995, “Changes of Numéraire, Changes of Probability Measure and Option Pricing,” *Journal of Applied Probability*, 32, 77–105.
- Grant, D., and G. Vora, 2001, “An Analytical Implementation of the Hull and White Model,” *Journal of Derivatives*, 4, 54–60.
- Hogan, M., and K. Weintraub, 1993, “The log-normal interest rate model and Eurodollar futures,” Working paper, Citybank, New York.

- Hull, J., and A. White, 1994, "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives*, 2, 7–16.
- Hull, J., and A. White, 1996, "Using Hull-White Interest Rate Trees," *Journal of Derivatives*, 3, 26–36.
- Jamshidian, F., 1989, "An Exact Bond Option Formula," *Journal of Finance*, 44, 205–209.
- Jamshidian, F., 1991, "Forward Induction and the Construction of Yield Curve Diffusion Models," *Journal of Fixed Income*, 1, 62–74.
- Kijima, M., and I. Nagayama, 1994, "Efficient Numerical Procedures for the Hull-White Extended Vasicek Model," *Journal of Financial Engineering*, 5, 275–292.
- Leippold, M., and Z. Wiener, 1999, "Algorithms behind Term Structure Models II: Hull-White Model," Working paper, Hebrew University, Jerusalem.
- Leippold, M., and L. Wu, 2002, "Asset Pricing under the Quadratic Class," *Journal of Financial and Quantitative Analysis*, 37.
- Longstaff, F. A., 1989, "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 23, 195–224.
- Pelsser, A., 1994, "An Efficient Algorithm for Calculating Prices in the Hull-White Model," Working paper, Derivative Product Research and Development, ABN Amro Bank.
- Schmidt, W., 1997, "On a general class of one-factor models for the term structure of interest rates," *Finance and Stochastics*, 1, 3–24.
- Vasicek, O. A., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.