CHAPTER
TWENTY-SIX

No-Arbitrage Term Structure Models

The fox often ran to the hole by which they had come in, to find out if his body was still thin enough to slip through it.

The Complete Grimm's Fairy Tales

This chapter samples no-arbitrage models pioneered by Ho and Lee. Some of the salient features of such models were already covered, if implicit at that, in Chap. 23.

26.1 Introduction

Some of the difficulties facing equilibrium models were mentioned in Section 25.4. For instance, they usually require the estimation of the market price of risk and cannot fit the market term structure. However, consistency with the market is often mandatory in practice [457]. No-arbitrage models, in contrast, utilize the full information of the term structure. They accept the observed term structure as consistent with an unobserved and unspecified equilibrium. From there, arbitrage-free movements of interest rates or bond prices over time are modeled. By definition, the market price of risk must be reflected in the current term structure; hence the resulting interest rate process is risk-neutral.

No-arbitrage models can specify the dynamics of zero-coupon bond prices, forward rates, or the short rate [477, 482]. Bond price and forward rate models are usually non-Markovian (path dependent), whereas short rate models are generally constructed to be explicitly Markovian (path independent). Markovian models are easier to handle computationally than non-Markovian ones.

➤ **Exercise 26.1.1** Is the equilibrium or no-arbitrage model more appropriate in deciding which government bonds are overpriced?

26.2 The Ho–Lee Model

This path-breaking one-factor model enjoys popularity among practitioners [72]. Figure 26.1 captures the model's short rate process. The short rates at any given time are evenly spaced. Let p denote the risk-neutral probability that the short rate makes an up move. We shall adopt continuous compounding.

https://doi.org/10.1017/CBO9780511546839.027 Published online by Cambridge University Press

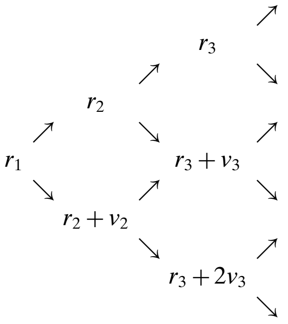

Figure 26.1: The Ho–Lee binomial interest rate tree. The distribution at any time converges to the normal distribution; the Ho–Lee model is a normal process in the limit. The v_is are related to short rate volatilities by Eq. (26.2).

The model starts with zero-coupon bond prices $P(t, t+1)$, $P(t, t+2), \ldots$, at time t identified with the root of the tree. Let the discount factors in the next period be

$$\begin{array}{ll} P_{\mathrm{d}}(t+1, t+2),\ P_{\mathrm{d}}(t+1, t+3), \ldots, & \text{if the short rate makes a down move,} \\ P_{\mathrm{u}}(t+1, t+2),\ P_{\mathrm{u}}(t+1, t+3), \ldots, & \text{if the short rate makes an up move.} \end{array}$$

By backward induction, it is not hard to see that, for $n \geq 2$,

$$P_{\mathrm{u}}(t+1, t+n) = P_{\mathrm{d}}(t+1, t+n)\, e^{-(v_2+\cdots+v_n)} \tag{26.1}$$

(see Exercise 26.2.1) and the n-period zero-coupon bond has yields

$$y_{\mathrm{d}}(n) \equiv -\frac{\ln P_{\mathrm{d}}(t+1, t+n)}{n-1},$$

$$y_{\mathrm{u}}(n) \equiv -\frac{\ln P_{\mathrm{u}}(t+1, t+n)}{n-1} = y_{\mathrm{d}}(n) + \frac{v_2 + \cdots + v_n}{n-1},$$

respectively. The volatility of the yield to maturity for this bond is therefore

$$\begin{aligned} \kappa_n &\equiv \sqrt{p y_{\mathrm{u}}(n)^2 + (1-p)\, y_{\mathrm{d}}(n)^2 - [\, p y_{\mathrm{u}}(n) + (1-p)\, y_{\mathrm{d}}(n)\,]^2} \\ &= \sqrt{p(1-p)}\ [\, y_{\mathrm{u}}(n) - y_{\mathrm{d}}(n)\,] \\ &= \sqrt{p(1-p)}\ \frac{v_2 + \cdots + v_n}{n-1}. \end{aligned}$$

In particular, we determine the short rate volatility by taking $n = 2$:

$$\sigma = \sqrt{p(1-p)}\ v_2. \tag{26.2}$$

The variance of the short rate therefore equals $p(1-p)(r_{\mathrm{u}} - r_{\mathrm{d}})^2$, where r_{u} and r_{d} are the two successor rates.[1]

The volatility term structure is composed of $\kappa_2, \kappa_3, \ldots$, independent of the r_is. It is easy to compute the v_is from the volatility structure (see Exercise 26.2.2), and vice versa. The r_is can be computed by forward induction. The volatility structure in the original Ho–Lee model is flat because it assumes that v_i are all equal to some constant. For the general Ho–Lee model that incorporates a term structure of volatilities, the volatility structure is supplied by the market.

➤ **Exercise 26.2.1** Verify Eq. (26.1).

➤ **Exercise 26.2.2** Show that $v_i = [\,(i-1)\,\kappa_i - (i-2)\,\kappa_{i-1})\,]/\sqrt{p(1-p)}\,$.

26.2.1 Bond Price Process

In a risk-neutral economy, the initial discount factors satisfy

$$P(t, t+n) = [\, p P_{\mathrm{u}}(t+1, t+n) + (1-p)\, P_{\mathrm{d}}(t+1, t+n)\,]\, P(t, t+1).$$

Combine the preceding equation with Eq. (26.1) and assume that $p = 1/2$ to obtain[2]

$$P_{\mathrm{d}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{2 \times \exp[\, v_2 + \cdots + v_n\,]}{1 + \exp[\, v_2 + \cdots + v_n\,]}, \tag{26.3}$$

$$P_{\mathrm{u}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{2}{1 + \exp[\, v_2 + \cdots + v_n\,]}. \tag{26.3$'$}$$

This defines the bond price process. The above system of equations establishes the price relations that must hold to prevent riskless arbitrages [304, 504]. The bond price tree combines (see Exercise 26.2.3).

In the original Ho–Lee model, v_i all equal some constant v. Then

$$P_{\mathrm{d}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{2\delta^{n-1}}{1 + \delta^{n-1}},$$

$$P_{\mathrm{u}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{2}{1 + \delta^{n-1}},$$

where $\delta \equiv e^{v} > 0$. The short rate volatility σ equals $v/2$ by Eq. (26.2). To annualize the numbers, simply apply $\sigma(\text{period}) = \sigma(\text{annual}) \times \sqrt{\Delta t}$ and $v(\text{period}) = v(\text{annual}) \times \Delta t$. As a consequence,

$$\delta(\text{annual}) = e^{2\sigma(\text{annual})(\Delta t)^{3/2}}. \tag{26.4}$$

The Ho–Lee model demonstrates clearly that no-arbitrage models price securities in a way consistent with the initial term structure. Furthermore, these models postulate dynamics that disallows intertemporal arbitrage opportunities. Derivatives are priced by taking expectations under the risk-neutral probability [359].

➤ **Exercise 26.2.3** Show that a rate rise followed by a rate decline produces the same term structure as that of a rate decline followed by a rate rise.

➤ **Exercise 26.2.4** Prove that Eqs. (26.3) and (26.3′) become

$$P_{\mathrm{d}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{\exp[\, v_2 + \cdots + v_n\,]}{p + (1-p) \times \exp[\, v_2 + \cdots + v_n\,]},$$

$$P_{\mathrm{u}}(t+1, t+n) = \frac{P(t, t+n)}{P(t, t+1)} \frac{1}{p + (1-p) \times \exp[\, v_2 + \cdots + v_n\,]}$$

for general risk-neutral probability p.

➤ **Exercise 26.2.5** Consider a portfolio of one zero-coupon bond with maturity T_1 and β zero-coupon bonds with maturity T_2. Find the β that makes the portfolio instantaneously riskless under the Ho–Lee model.

➢ **Programming Assignment 26.2.6** Write a linear-time program to calibrate the original Ho–Lee model. The inputs are Δt, the current market discount factors, and the short rate volatility σ, all annualized.

Yield Volatilities and Their Covariances

The one-period rate of return of an n-period zero-coupon bond is

$$r(t, t+n) \equiv \ln\left(\frac{P(t+1, t+n)}{P(t, t+n)}\right).$$

Because its value is either

$$\ln \frac{P_{\mathrm{d}}(t+1, t+n)}{P(t, t+n)}$$

or

$$\ln \frac{P_{\mathrm{u}}(t+1, t+n)}{P(t, t+n)},$$

the variance of the return is

$$\mathrm{Var}[\, r(t, t+n)\,] = p(1-p)[\,(n-1)\, v\,]^2 = (n-1)^2\sigma^2.$$

The covariance between $r(t, t+n)$ and $r(t, t+m)$ is $(n-1)(m-1)\,\sigma^2$ (see Exercise 26.2.7). As a result, the correlation between any two one-period returns is unity. Strong correlation between rates is inherent in all one-factor Markovian models.

➤ **Exercise 26.2.7** Prove that under a general p, the variance of the one-period return of n-period zero-coupon bonds equals $(n-1)^2\,\sigma^2$ and the covariance between the one-period returns of n- and m-period zero-coupon bonds equals $(n-1)(m-1)\,\sigma^2$.

26.2.2 Forward Rate Process

The forward rate at time t for money borrowed or lent from time $t+n$ to $t+n+1$ is

$$f(t, t+n) = -\ln\left(\frac{P(t, t+n+1)}{P(t, t+n)}\right)$$

from Eq. (24.3). The current state considered as the result of a downward rate move from time $t-1$ leads to

$$\begin{aligned} f(t, t+n) &= -\ln\left(\frac{\frac{P(t-1, t+n+1)}{P(t-1, t)}\ \frac{2\delta^{n+1}}{1+\delta^{n+1}}}{\frac{P(t-1, t+n)}{P(t-1, t)}\ \frac{2\delta^{n}}{1+\delta^{n}}}\right) \\ &= f(t-1, t+n) - \ln\left(\frac{1+\delta^n}{1+\delta^{n+1}}\right) - \ln\delta, \end{aligned}$$

and the current state considered as the result of an upward rate move leads to

$$f(t, t+n) = -\ln\left(\frac{\frac{P(t-1, t+n+1)}{P(t-1, t)}\ \frac{2}{1+\delta^{n+1}}}{\frac{P(t-1, t+n)}{P(t-1, t)}\ \frac{2}{1+\delta^{n}}}\right) = f(t-1, t+n) - \ln\left(\frac{1+\delta^n}{1+\delta^{n+1}}\right).$$

The preceding two equations can be combined to yield this forward rate process:

$$f(t, t+n) = f(t-1, t+n) - \ln\left(\frac{1+\delta^n}{1+\delta^{n+1}}\right) - \frac{1}{2}\,\ln\delta + \xi_{t-1}, \tag{26.5}$$

where ξ_s $(s \geq 0)$ is the following zero-mean random variable:

$$\xi_s = \begin{cases} -(1/2)\ \ln\delta, & \text{if down move occurs at time } s \\ (1/2)\ \ln\delta, & \text{if up move occurs at time } s \end{cases}.$$

Because $\text{Var}[\, f(t, t+n) - f(t-1, t+n)\,] = \sigma^2$ (see Exercise 26.2.8), the volatility σ can be estimated from historical data without the need to estimate the risk-neutral probability [436]. Equation (26.5) can be applied iteratively to obtain

$$f(t, t+n) = f(0, t+n) - \ln\left(\frac{1+\delta^n}{1+\delta^{n+t}}\right) - \frac{t}{2}\ \ln\delta + \sum_{s=1}^{t} \xi_{s-1}.$$

➤ **Exercise 26.2.8** Verify that $\text{Var}[\,\xi_s\,] = \sigma^2$.

➤ **Exercise 26.2.9** Prove that

$$-\ln\left(\frac{1+\delta^n}{1+\delta^{n+1}}\right) - \frac{1}{2}\ \ln\delta \to \sigma^2(T-\boldsymbol{t})(\Delta t)^2$$

if we substitute $\boldsymbol{t}/\Delta t$ for t and $(T-\boldsymbol{t})/\Delta t$ for n in Eq. (26.5) before applying Eq. (26.4). T, $\boldsymbol{t}$, Δt, and σ above are annualized. (The forward rate process hence converges to $df(\boldsymbol{t}, T) = \sigma^2(T-\boldsymbol{t})\, d\boldsymbol{t} + \sigma\, dW$.)

26.2.3 Short Rate Process

Because the short rate $r(t)$ equals $f(t, t)$,

$$r(t) = f(0, t) - \ln\left(\frac{2}{1+\delta^t}\right) - \frac{t}{2}\ \ln\delta + \sum_{s=1}^{t} \xi_{s-1}.$$

This implies the following difference equation:

$$r(t) = r(t-1) + f(0, t) - f(0, t-1) - \ln\left(\frac{1+\delta^{t-1}}{1+\delta^t}\right) - \frac{1}{2}\ \ln\delta + \xi_{t-1}. \tag{26.6}$$

The continuous-time limit of the Ho–Lee model is $dr = \theta(t)\, dt + \sigma\, dW$. This is essentially Vasicek's model with the mean-reverting drift replaced with a deterministic, time-dependent drift. A nonflat term structure of volatilities can be achieved if the short rate volatility is also made time varying, i.e., $dr = \theta(t)\, dt + \sigma(t)\, dW$ [508]. This corresponds to the discrete-time model in which v_i are not all identical.

➤ **Exercise 26.2.10** Prove that

$$-\ln\left(\frac{1+\delta^{t-1}}{1+\delta^t}\right) - \frac{1}{2}\ \ln\delta \to \sigma^2\boldsymbol{t}\,(\Delta t)^2$$

if we substitute $\boldsymbol{t}/\Delta t$ for t and apply Eq. (26.4). The $\boldsymbol{t}$, Δt, and σ above are annualized. (Short rate process (26.6) thus converges to $dr = \{\,[\,\partial f(0, \boldsymbol{t})/\partial \boldsymbol{t}\,] + \sigma^2 \boldsymbol{t}\,\}\, d\boldsymbol{t} + \sigma\, dW$.)

26.2.4 Problems with the Ho–Lee Model

Nominal interest rates must be nonnegative because we can hold cash. However, negative future interest rates are possible with the Ho–Lee model. This may not be a major concern for realistic volatilities and certain ranges of bond maturities [72]. More questionable is the fact that the short rate volatility is independent of the rate level [83, 173, 359, 645]. Given that the Ho–Lee model subsumes the Merton model and shares many of its unreasonable properties, how can it generate reasonable initial term structures? The answer lies in the model's unreasonable short rate dynamics.[3]

➤ **Exercise 26.2.11** Assess the claim that the problem of negative interest rates can be eliminated by making the short rate volatility time dependent.

Problems with No-Arbitrage Models in General

Interest rate movements should reflect shifts in the model's state variables (factors), not its parameters. This means that model parameters, such as the drift $\theta(t)$ in the continuous-time Ho–Lee model, should be stable over time. However, in practice, no-arbitrage models capture yield curve shifts through the recalibration of parameters. A new model is thus born everyday. This in effect says that the model estimated at some time does not describe the term structure of interest rates and their volatilities at other times. Consequently, a model's intertemporal behavior is suspect, and using it for hedging and risk management may be unreliable.

26.3 The Black–Derman–Toy Model

Black, Derman, and Toy (BDT) proposed their model in 1990 [84]. This model is extensively used by practitioners [72, 149, 215, 600, 731]. The BDT short rate process is the lognormal binomial interest rate process described in Chap. 23 and repeated in Fig. 26.2. The volatility structure is given by the market from which the short rate volatilities (thus v_i) are determined together with r_i. Our earlier binomial interest rate tree, in comparison, assumes that v_i are given a priori, and a related model of Salomon Brothers takes v_i to be constants [848]. Lognormal models preclude negative short rates.

The volatility structure defines the yield volatilities of zero-coupon bonds of various maturities. Let the yield volatility of the i-period zero-coupon bond be denoted by κ_i. Assume that P_{u} (P_{d}) is the price of the i-period zero-coupon bond one period from now if the short rate makes an up (down, respectively) move. Corresponding

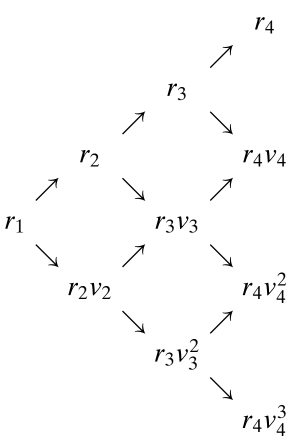

Figure 26.2: The BDT binomial interest rate tree. The distribution at any time converges to the lognormal distribution.

to these two prices are these yields to maturity:

$$y_{\mathrm{u}} \equiv P_{\mathrm{u}}^{-1/(i-1)} - 1, \quad y_{\mathrm{d}} \equiv P_{\mathrm{d}}^{-1/(i-1)} - 1.$$

The yield volatility is defined as $\kappa_i \equiv (1/2) \ln(y_{\mathrm{u}}/y_{\mathrm{d}})$.

26.3.1 Calibration

The inputs to the BDT model are riskless zero-coupon bond yields and their volatilities. For economy of expression, all numbers are period based. Suppose inductively that we have calculated $r_1, v_1, r_2, v_2, \ldots, r_{i-1}, v_{i-1}$, which define the binomial tree up to period $i-1$. We now proceed to calculate r_i and v_i to extend the tree to period i. Assume that the price of the i-period zero can move to P_{u} or P_{d} one period from now. Let y denote the current i-period spot rate, which is known. In a risk-neutral economy,

$$\frac{P_{\mathrm{u}} + P_{\mathrm{d}}}{2(1+r_1)} = \frac{1}{(1+y)^i}. \tag{26.7}$$

Obviously, P_{u} and P_{d} are functions of the unknown r_i and v_i. Viewed from now, the future $(i-1)$-period spot rate at time one is uncertain. Let y_{u} and y_{d} represent the spot rates at the up node and the down node, respectively, with κ^2 denoting the variance, or

$$\kappa_i = \frac{1}{2} \ln \left(\frac{P_{\mathrm{u}}{}^{-1/(i-1)} - 1}{P_{\mathrm{d}}{}^{-1/(i-1)} - 1} \right). \tag{26.8}$$

We use forward induction to derive a quadratic-time calibration algorithm [190, 625]. Recall that forward induction inductively figures out, by moving forward in time, how much \$1 at a node contributes to the price (review Fig. 23.7(a)). This number is called the state price and is the price of the claim that pays \$1 at that node and zero elsewhere.

Let the baseline rate for period i be $r_i = r$, let the multiplicative ratio be $v_i = v$, and let the state prices at time $i-1$ be $P_1, P_2, \ldots, P_i$, corresponding to rates $r, rv, \ldots, rv^{i-1}$, respectively. One dollar at time i has a PV of

$$f(r, v) \equiv \frac{P_1}{1+r} + \frac{P_2}{1+rv} + \frac{P_3}{1+rv^2} + \cdots + \frac{P_i}{1+rv^{i-1}},$$

and the yield volatility is

$$g(r, v) \equiv \frac{1}{2} \ln \left(\frac{\left(\frac{P_{\mathrm{u},1}}{1+rv} + \frac{P_{\mathrm{u},2}}{1+rv^2} + \cdots + \frac{P_{\mathrm{u},i-1}}{1+rv^{i-1}} \right)^{-1/(i-1)} - 1}{\left(\frac{P_{\mathrm{d},1}}{1+r} + \frac{P_{\mathrm{d},2}}{1+rv} + \cdots + \frac{P_{\mathrm{d},i-1}}{1+rv^{i-2}} \right)^{-1/(i-1)} - 1} \right).$$

In the preceding equation, $P_{\mathrm{u},1}, P_{\mathrm{u},2}, \ldots$ denote the state prices at time $i-1$ of the subtree rooted at the up node (like $r_2 v_2$ in Fig. 26.2), and $P_{\mathrm{d},1}, P_{\mathrm{d},2}, \ldots$ denote the state prices at time $i-1$ of the subtree rooted at the down node (like r_2 in Fig. 26.2). Now solve

$$f(r, v) = \frac{1}{(1+y)^i}, \quad g(r, v) = \kappa_i$$

Algorithm for calibrating the BDT model:

input: $n, S[1..n], \kappa[1..n]$;
real $P[0..n], P_u[0..n-1], P_d[0..n-1], r[1..n], v[1..n], r, v$;
integer i, j;
$P[0] := 0; P_u[0] := 0; P_d[0] := 0$; // Dummies; remain zero throughout.
$P[1] := 1; P_u[1] := 1; P_d[1] := 1$;
$r[1] := S[1]; v[1] := 0$;
for $(i = 2$ to $n)$ {
  $P[i] := 0$;
  for $(j = i$ down to $1)$ // State prices at time $i-1$.
    $P[j] := \frac{P[j-1]}{2\times(1+r[i-1]\times v[i-1]^{j-2})} + \frac{P[j]}{2\times(1+r[i-1]\times v[i-1]^{j-1})}$;
  Solve for r and v from
    $\sum_{j=1}^{i} \frac{P[j]}{(1+r\times v^{j-1})} = (1+S[i])^{-i}$ and
    $(\sum_{j=1}^{i-1} \frac{P_u[j]}{(1+r\times v^{j})})^{-1/(i-1)} - 1 = e^{2\times\kappa[i]} \times ((\sum_{j=1}^{i-1} \frac{P_d[j]}{(1+r\times v^{j-1})})^{-1/(i-1)} - 1)$;
  $r[i] := r; v[i] := v$;
  if $[i < n]$ {
    $P_u[i] := 0; P_d[i] := 0$;
    for $(j = i$ down to $1)$ { // State prices at time i.
      $P_u[j] := \frac{P_u[j-1]}{2\times(1+r[i]\times v[i]^{j-1})} + \frac{P_u[j]}{2\times(1+r[i]\times v[i]^{j})}$;
      $P_d[j] := \frac{P_d[j-1]}{2\times(1+r[i]\times v[i]^{j-2})} + \frac{P_d[j]}{2\times(1+r[i]\times v[i]^{j-1})}$;
    }
  }
}
return $r[\,]$ and $v[\,]$;

Figure 26.3: Algorithm for calibrating the BDT model. $S[i]$ is the i-period spot rate, $\kappa[i]$ is the yield volatility for period i, and n is the number of periods. All numbers are measured by the period. The period-i baseline rate and multiplicative ratio are stored in $r[i]$ and $v[i]$, respectively. The two-dimensional Newton–Raphson method of Eqs. (3.16) should be used to solve for r and v as the partial derivatives are straightforward to calculate.

for $r = r_i$ and $v = v_i$. This $O(n^2)$-time algorithm is given in Fig. 26.3. The continuous-time limit of the BDT model is $d\ln r = \{\theta(t) + [\sigma'(t)/\sigma(t)]\ln r\}\,dt + \sigma(t)\,dW$ [76, 149, 508]. Obviously the short rate volatility should be a declining function of time for the model to display mean reversion; in particular, constant volatility will not attain mean reversion.

➤ **Exercise 26.3.1** Describe the differential tree method with backward induction to calibrate the BDT model.

➢ **Programming Assignment 26.3.2** Implement the algorithm in Fig. 26.3.

➢ **Programming Assignment 26.3.3** Calibrate the BDT model with the secant method and evaluate its performance against the differential tree method.

26.3.2 The Black–Karasinski Model

The related Black–Karasinski model stipulates that the short rate follows

$$d\ln r = \kappa(t)[\theta(t) - \ln r]\,dt + \sigma(t)\,dW.$$

$$
\begin{array}{ccccc}
 & & & \nearrow & \\
 & \ln r_{\mathrm{d}}(t_2) & & & \\
\nearrow & & & \searrow & \\
\ln r(t_1) & & & & \ln r_{\mathrm{du}}(t_3) = \ln r_{\mathrm{ud}}(t_3) \\
\searrow & & & \nearrow & \\
 & \ln r_{\mathrm{u}}(t_2) & & & \\
 & & & \searrow &
\end{array}
$$

Figure 26.4: The Black–Karasinski tree. Time periods may not be equal in length.

This explicitly mean-reverting model depends on time through $\kappa(\cdot), \theta(\cdot)$, and $\sigma(\cdot)$. The Black–Karasinski model hence has one more degree of freedom than the BDT model. The speed of mean reversion $\kappa(t)$ and the short rate volatility $\sigma(t)$ are independent [85].

The discrete-time version of the Black–Karasinski model has the same representation as the BDT model. To maintain a combining binomial tree, however, requires some manipulations. These ideas are illustrated by Fig. 26.4 in which $t_2 \equiv t_1 + \Delta t_1$ and $t_3 \equiv t_2 + \Delta t_2$. Note that

$$
\begin{aligned}
\ln r_{\mathrm{d}}(t_2) &= \ln r(t_1) + \kappa(t_1)[\,\theta(t_1) - \ln r(t_1)\,]\,\Delta t_1 - \sigma(t_1)\sqrt{\Delta t_1},\\
\ln r_{\mathrm{u}}(t_2) &= \ln r(t_1) + \kappa(t_1)[\,\theta(t_1) - \ln r(t_1)\,]\,\Delta t_1 + \sigma(t_1)\sqrt{\Delta t_1}.
\end{aligned}
$$

To ensure that an up move followed by a down move coincides with a down move followed by an up move, we impose

$$
\begin{aligned}
&\ln r_{\mathrm{d}}(t_2) + \kappa(t_2)[\,\theta(t_2) - \ln r_{\mathrm{d}}(t_2)\,]\,\Delta t_2 + \sigma(t_2)\sqrt{\Delta t_2}\\
&\quad = \ln r_{\mathrm{u}}(t_2) + \kappa(t_2)[\,\theta(t_2) - \ln r_{\mathrm{u}}(t_2)\,]\,\Delta t_2 - \sigma(t_2)\sqrt{\Delta t_2}\,,
\end{aligned}
$$

which implies that

$$
\kappa(t_2) = \frac{1 - [\,\sigma(t_2)/\sigma(t_1)\,]\sqrt{\Delta t_2/\Delta t_1}}{\Delta t_2}.
$$

So from Δt_1, we can calculate the Δt_2 that satisfies the combining condition and then iterate.

➤ **Exercise 26.3.4** Show that the variance of r after Δt is approximately $[\,r(t)\,\sigma(t)\,]^2 \Delta t$.

➢ **Programming Assignment 26.3.5** Implement a forward-induction algorithm to calibrate the Black–Karasinski model given a constant κ.

26.3.3 Problems with Lognormal Models

Lognormal models such as the BDT, Black–Karasinski, and Dothan models share the problem that $E^{\pi}[\,M(t)\,] = \infty$ for any finite t if it is the continuously compounded rate that is modeled (review Subsection 24.2.1) [76]. Hence periodic compounding should be used. Another issue is computational. Lognormal models usually do not give analytical solutions to even basic fixed-income securities. As a result, to price short-dated derivatives on longterm bonds, the tree has to be built over the life of the underlying asset – which can be, say, 30 years – instead of the life of the claim – possibly only 1–2 years (review Comment 23.2.1). This problem can be somewhat mitigated if different time steps are adopted: Use a fine time step up to the maturity of the short-dated derivative and a coarse time step beyond the maturity [477]. A down side of this procedure is that it has to be carried out for each derivative.

26.4 The Models According to Hull and White

Hull and White proposed models that extend the Vasicek model and the CIR model [474]. They are called the extended Vasicek model and the extended CIR model. The extended Vasicek model adds time dependence to the original Vasicek model:

$$dr = (\theta(t) - a(t)\, r)\, dt + \sigma(t)\, dW.$$

Like the Ho–Lee model, this is a normal model, and the inclusion of $\theta(\cdot)$ allows for an exact fit to the current spot rate curve. As for the other two functions, $\sigma(t)$ defines the short-rate volatility and $a(t)$ determines the shape of the volatility structure. Under this model, many European-style securities can be evaluated analytically, and efficient numerical procedures can be developed for American-style securities. The Hull–White model is the following special case:

$$dr = (\theta(t) - ar)\, dt + \sigma\, dW.$$

When the current term structure is matched,

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + af(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at})$$

[477]. In the extended CIR model the short rate follows

$$dr = [\,\theta(t) - a(t)\, r\,]\, dt + \sigma(t)\sqrt{r}\, dW.$$

The functions $\theta(\cdot)$, $a(\cdot)$, and $\sigma(\cdot)$ are implied from market observables. With constant parameters, there exist analytical solutions to a small set of interest-rate-sensitive securities such as coupon bonds and European options on bonds.

For the BDT and the Ho–Lee models, once the initial volatility structure is specified, the future short rate volatility is completely determined. Conversely, if the future short rate volatility is specified, then the initial volatility structure is fully determined. However, we may want to specify the volatility structure and the short rate volatility separately because the future short rate volatility may have little impact on the yield volatility, which is about the uncertainty over the spot rate's value in the next period [476]. The extended Vasicek model has enough degrees of freedom to accommodate this request.

➤ **Exercise 26.4.1** Between a normal and a lognormal model, which overprices out-of-the-money calls on bonds *and* underprices out-of-the-money puts on bonds?

26.4.1 Calibration of the Hull–White Model with Trinomial Trees

Now a trinomial forward-induction scheme is described to calibrate the Hull–White model given a and σ [477]. As with the Ho–Lee model, in this model the set of achievable short rates is evenly spaced. Let r_0 be the annualized, continuously compounded short rate at time zero. Every short rate on the tree takes on a value $r_0 + j\Delta r$ for some integer j. Time increments on the tree are also equally spaced at Δt apart. (Binomial trees should not be used to model mean-reverting interest rates when Δt is a constant [475].) Hence nodes are located at times $i\Delta t$ for $i = 0, 1, 2, \ldots$. We

https://doi.org/10.1017/CBO9780511546839.027 Published online by Cambridge University Press

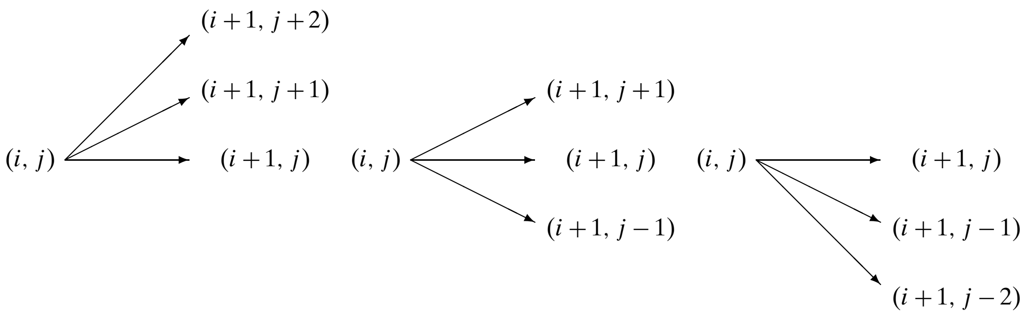

Figure 26.5: Three trinomial branching schemes in the Hull–White model. The choice is determined by the expected short rate at time t_{i+1} as seen from time t_i.

refer to the node on the tree with $t_i \equiv i\,\Delta t$ and $r_j \equiv r_0 + j\,\Delta r$ as the **(i, j) node**. The short rate at node (i, j), which equals r_j, is effective for the time period $[\,t_i, t_{i+1}\,]$. Use

$$\mu_{i,j} \equiv \theta(t_i) - ar_j \tag{26.9}$$

to denote the drift rate, or the expected change, of the short rate as seen from node (i, j). The three distinct possibilities for node (i, j) with three branches incident from it are shown in Fig. 26.5. The interest rate movement described by the middle branch may be an increase of Δr, no change, or a decrease of Δr. The upper and the lower branches bracket the middle branch. Define

$p_1(i, j) \equiv$ the probability of following the upper branch from node (i, j),
$p_2(i, j) \equiv$ the probability of following the middle branch from node (i, j),
$p_3(i, j) \equiv$ the probability of following the lower branch from node (i, j).

The root of the tree is set to the current short rate r_0. Inductively, the drift $\mu_{i,j}$ at node (i, j) is a function of $\theta(t_i)$. Once $\theta(t_i)$ is available, $\mu_{i,j}$ can be derived by means of Eq. (26.9). This in turn determines the branching scheme at every node (i, j) for each j, as we will see shortly. The value of $\theta(t_i)$ must thus be made consistent with the spot rate $r(0, t_{i+2})$.

The branches emanating from node (i, j) with their accompanying probabilities, $p_1(i, j)$, $p_2(i, j)$, and $p_3(i, j)$, must be chosen to be consistent with $\mu_{i,j}$ and σ. This is accomplished by letting the middle node be as close as possible to the current value of the short rate plus the drift. Let k be the number among $\{\, j-1, j, j+1\,\}$ that makes the short rate reached by the middle branch, r_k, closest to $r_j + \mu_{i,j}\Delta t$. Then the three nodes following node (i, j) are nodes $(i+1, k+1)$, $(i+1, k)$, and $(i+1, k-1)$. The resulting tree may have the geometry depicted in Fig. 26.6. The resulting tree combines.

The probabilities for moving along these branches are functions of $\mu_{i,j}$, σ, j, and k:

$$p_1(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} + \frac{\eta}{2\Delta r}, \tag{26.10}$$

$$p_2(i, j) = 1 - \frac{\sigma^2 \Delta t + \eta^2}{(\Delta r)^2}, \tag{26.10$'$}$$

$$p_3(i, j) = \frac{\sigma^2 \Delta t + \eta^2}{2(\Delta r)^2} - \frac{\eta}{2\Delta r}, \tag{26.10$''$}$$

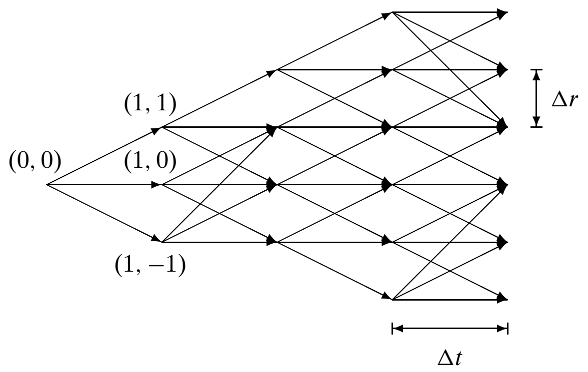

Figure 26.6: Trinomial tree for the Hull–White model. All the short rates at the nodes are known before *any* computation begins. They are simply $r_0 + j\,\Delta r$ for $j = 0, \pm 1, \pm 2, \ldots$. It is the branching schemes connecting the nodes that determine the term structure. Only the four nodes at times t_0 and t_1 are labeled here. The remaining nodes can be labeled similarly. The tree may not fully grow as some nodes are not reachable from the root $(0, 0)$ because of mean reversion.

where $\eta \equiv \mu_{i,j}\,\Delta t + (j-k)\,\Delta r$. As trinomial tree algorithms are but explicit methods in disguise (see Subsection 18.1.1), certain relations must hold for Δr and Δt to guarantee stability. It can be shown that their values must satisfy

$$\frac{\sigma\sqrt{3\Delta t}}{2} \leq \Delta r \leq 2\sigma\sqrt{\Delta t}$$

for the probabilities to lie between zero and one; for example, Δr can be set to $\sigma\sqrt{3\Delta t}$ [473]. It remains only to determine $\theta(t_i)$, to which we now turn.

At this point at time $t_i, r(0, t_1), r(0, t_2), \ldots, r(0, t_{i+1})$ have already been matched. By construction, the state prices $Q(i, k)$ for all k are known by now, where $Q(i, k)$ denotes the value of the state contingent claim that pays \$1 at node (i, k) and zero otherwise. The value at time zero of a zero-coupon bond maturing at time t_{i+2} is then

$$e^{-r(0,t_{i+2})(i+2)\,\Delta t} = \sum_j Q(i, j)\, e^{-r_j\,\Delta t}\, E^{\pi}\left[\, e^{-\widehat{r}(i+1)\,\Delta t} \,\middle|\, \widehat{r}(i) = r_j \,\right], \tag{26.11}$$

where $\widehat{r}(i)$ refers to the short-rate value at time t_i. The right-hand side represents the value of \$1 obtained by holding a zero-coupon bond until time t_{i+1} and then reinvesting the proceeds at that time at the prevailing short rate $\widehat{r}(i+1)$, which is stochastic. The expectation above can be approximated by

$$E^{\pi}\left[\, e^{-\widehat{r}(i+1)\,\Delta t} \,\middle|\, \widehat{r}(i) = r_j \right] \approx e^{-r_j\,\Delta t}\left[1 - \mu_{i,j}(\Delta t)^2 + \frac{\sigma^2(\Delta t)^3}{2}\right]. \tag{26.12}$$

Substitute approximation (26.12) into Eq. (26.11) and replace $\mu_{i,j}$ with $\theta(t_i) - a r_j$ to obtain

$$\theta(t_i) \approx \frac{\sum_j Q(i, j)\, e^{-2r_j\,\Delta t}[\,1 + a r_j(\Delta t)^2 + \sigma^2(\Delta t)^3/2\,] - e^{-r(0,t_{i+2})(i+2)\,\Delta t}}{(\Delta t)^2 \sum_j Q(i, j)\, e^{-2r_j\,\Delta t}}.$$

For the Hull–White model, the expectation in approximation (26.12) is actually known analytically:

$$E^{\pi}\left[\, e^{-\widehat{r}(i+1)\,\Delta t} \;\middle|\; \widehat{r}(i) = r_j \,\right] = e^{-r_j \Delta t + [-\theta(t_i) + a r_j + \sigma^2 \Delta t/2](\Delta t)^2}.$$

Therefore, alternatively,

$$\theta(t_i) = \frac{r(0, t_{i+2})(i+2)}{\Delta t} + \frac{\sigma^2 \Delta t}{2} + \frac{\ln \sum_j Q(i, j)\, e^{-2 r_j \Delta t + a r_j (\Delta t)^2}}{(\Delta t)^2}.$$

With $\theta(t_i)$ in hand, we can compute $\mu_{i,j}$, the probabilities, and finally the new state prices:

$$Q(i+1, j) = \sum_{(i,\, j^*) \text{ is connected to } (i+1,\, j) \text{ with probability } p_{j^*}} p_{j^*} e^{-r_{j^*} \Delta t}\, Q(i, j^*).$$

The total running time is quadratic. See Fig. 26.7 for an algorithm.

When using the Hull–White model, one can try different values of a and σ for each option or have an a value common to all options but use a different σ value for each option. Either approach can match all the option prices exactly. If the demand is for a single set of parameters that replicate all option prices, the Hull–White model can be calibrated to all the observed option prices by choosing a and σ that minimize the mean-square pricing error [482].

Algorithm for calibrating the Hull–White model:

```
input:    σ, a, Δt, n, S[1..n];
real      branch[n−1][−n..n], Q[−n..n], q[−n..n], θ, μ, r₀, Δr, p₁, p₂, p₃;
integer   i, j, k;
Q[0] := 1; r₀ := S[1];
Δr = σ√(3Δt);
branch[i][j] := ∞ for 0 ≤ i < n−1 and −n ≤ j ≤ n;   // Initial values ∉ {−1, 0, 1}.
for (i = 0 to n−2) {
        θ := S[i+2]×(i+2)/Δt + σ²Δt/2 + ln(Σ_{j=−i}^{i} Q[i][j]×exp[−2(r₀+jΔr)Δt+a(r₀+jΔr)(Δt)²])/(Δt)²; // θ(t_i).
        for (j = −i to i) { q[j] := 0; }
        for (j = −i to i)    { // Work on node (i, j)'s branching scheme.
            μ := θ − a(r₀ + jΔr);   // μ_{i,j}.
            Let k ∈ {−1, 0, 1} minimize |(r₀ + (j + k) Δr) − (r₀ + jΔr + μΔt)|;
            branch[i][j] := k;
            Use Eqs. (26.10) to calculate p₁, p₂, and p₃ with k = j + branch[i][j];
            q[k+1] := p₁ × Q[j] × e^{−(r₀+jΔr)Δt} + q[k+1]; // Add contribution to Q.
            q[k] := p₂ × Q[j] × e^{−(r₀+jΔr)Δt} + q[k];
            q[k−1] := p₃ × Q[j] × e^{−(r₀+jΔr)Δt} + q[k−1];
        }
        for (j = −i to i) { Q[j] := q[j]; }   // Update Q.
}
return branch[ ][ ];
```

Figure 26.7: Algorithm for calibrating the Hull–White model. $S[i]$ is the annualized i-period spot rate, $Q[\,]$ stores the state prices with initial values of zero, branch$[i][j]$ maintains the branching scheme for node (i, j), and n is the number of periods. Only the branching schemes are returned as they suffice to derive the short rates.

The algorithmic idea here is quite general and can be modified to apply to cases in which the diffusion term has the form σr^{β} [215]. A highly efficient algorithm exists that fully exploits the fact that the Hull–White model has a constant diffusion term [479].

➤ **Exercise 26.4.2** Verify approximation (26.12).

➢ **Programming Assignment 26.4.3** Implement the algorithm in Fig. 26.7.

➢ **Programming Assignment 26.4.4** Calibration takes the spot rate curve to reverse engineer the Hull–White model's parameters. However, just because the curve is matched by no means implies that the true model parameters and the estimated parameters are matched. Call a model **stable** if the model parameters can be approximated well by the estimated parameters. Verify that both the Hull–White model and the BDT model are stable [192, 885].

➢ **Programming Assignment 26.4.5** Implement a trinomial tree model for the Black–Karasinski model $d\ln r = (\theta(t) - a\ln r)\,dt + \sigma\,dW$.

26.4.2 Problems with the Models

When $\sigma(t)$ and $a(t)$ vary with time, the volatility structure will be nonstationary. Choosing $\sigma(t)$ and $a(t)$ to exactly fit the initial volatility structure then causes the volatility structure to evolve in unpredictable ways and makes option prices questionable. This observation holds for all Markovian models [76, 164]. Because it is in general dangerous to use time-varying parameters to match the initial volatility curve exactly, it has been argued that there should be no more than one time-dependent parameter in Markovian models and that it should be used to fit the initial spot rate curve only [482]. This line of reasoning favors the Hull–White model. Another way to maintain the volatility structure over time is to use the Heath–Jarrow–Morton (HJM) model.

26.5 The Heath–Jarrow–Morton Model

We have seen several Markovian short rate models. The Markovian approach, albeit computationally efficient, has the disadvantage that it is difficult to model the behavior of yields and bond prices of different maturities. The alternative **yield curve approach** regards the whole term structure as the state of a process and directly specifies how it evolves [725].

The influential model proposed by Heath, Jarrow, and Morton is a forward rate model [437, 511]. It is also a popular model [17, 198]. The HJM model specifies the initial forward rate curve and the forward rate volatility structure, which describes the volatility of each forward rate for a given maturity date. Like the Black–Scholes option pricing model, neither risk preference assumptions nor the drifts of forward rates are needed [515].

26.5.1 Forward-Rate Process

Within a finite time horizon $[0, U]$, we take as given the time-zero forward rate curve $f(0, t)$ for $t \in [0, U]$. Because this curve is used as the boundary value at

$t = 0$, perfect fit to the observed term structure is automatic. The forward rates are driven by n stochastic factors. Specifically the forward rate movements are governed by the stochastic process

$$df(t, T) = \mu(t, T)\, dt + \sum_{i=1}^{n} \sigma_i(t, T)\, dW_i,$$

where μ and σ_i may depend on the past history of the Wiener processes.

Take the one-factor model

$$df(t, T) = \mu(t, T)\, dt + \sigma(t, T)\, dW_t. \tag{26.13}$$

This is an infinite-dimensional system because there is an equation for each T. One-factor models seem to perform better than multifactor models empirically [19]. When is the bond market induced by forward rate model (26.13) arbitrage free in that there exists an equivalent martingale measure? For this to happen, there must exist a process $\lambda(t)$ such that for all $0 \leq t \leq T$, the drift equals

$$\mu(t, T) = \sigma(t, T) \int_t^T \sigma(t, s)\, ds + \sigma(t, T)\, \lambda(t). \tag{26.14}$$

The process $\lambda(t)$, which may depend on the past history of the Wiener process, is the market price of risk. Substitute this condition into Eq. (26.13) to yield the following arbitrage-free forward rate dynamics:

$$df(t, T) = \left[\sigma(t, T) \int_t^T \sigma(t, s)\, ds + \sigma(t, T)\, \lambda(t) \right] dt + \sigma(t, T)\, dW_t. \tag{26.15}$$

The market price of risk enters only into the drift. The short rate follows

$$dr(t) = \left[\sigma(t, t)\, \lambda(t) + \left. \frac{\partial f(t, T)}{\partial T} \right|_{T=t} \right] dt + \sigma(t, t)\, dW_t. \tag{26.16}$$

A unique equivalent martingale measure can be established under which the prices of interest rate derivatives do not depend on the market prices of risk. This fundamental result is summarized below.

THEOREM 26.5.1 *Assume that π is a martingale measure for the bond market and that the forward rate dynamics under π is given by $df(t, T) = \mu(t, T)\, dt + \sum_{i=1}^{n} \sigma_i(t, T)\, dW_i$. The volatility functions $\sigma_i(t, T)$ may depend on $f(t, T)$. (1) For all $0 < t \leq T$,*

$$\mu(t, T) = \sum_{i=1}^{n} \sigma_i(t, T) \int_t^T \sigma_i(t, u)\, du \tag{26.17}$$

holds under π almost surely. (2) The bond price dynamics under π is given by

$$\frac{dP(t, T)}{P(t, T)} = r(t)\, dt + \sum_{i=1}^{n} \sigma_{\mathrm{p},i}(t, T)\, dW_i,$$

where $\sigma_{\mathrm{p},i}(t, T) \equiv -\int_t^T \sigma_i(t, u)\, du$. (Choosing the volatility function $\sigma(t, T)$ of the forward rate dynamics under π uniquely determines the drift parameters under π and the prices of all claims.)

To use the HJM model, we first pick $\sigma(t, T)$. This is the modeling part. The drift parameters are then determined by Eq. (26.17). Now fetch today's forward rate curve $\{ f(0, T), T \geq 0 \}$ and integrate it to obtain the forward rates:

$$f(t, T) = f(0, T) + \int_0^t \mu(s, T)\, ds + \int_0^t \sigma(s, T)\, dW_s. \tag{26.18}$$

Compute the future bond prices by $P(t, T) = e^{-\int_t^T f(t,s)\, ds}$ if necessary. European-style derivatives can be priced by simulating many paths and taking the average.

From Eqs. (26.17) and (26.18),

$$\begin{aligned} r(t) &= f(t, t) = f(0, t) + \int_0^t df(s, t) \\ &= f(0, t) + \int_0^t \sigma_{\mathrm{p}}(s, t)\, \sigma(s, t)\, ds + \int_0^t \sigma(s, t)\, dW_s, \end{aligned}$$

where $\sigma_{\mathrm{p}}(s, t) \equiv \int_s^t \sigma(s, u)\, du$. Differentiate with respect to t and note that $\sigma_{\mathrm{p}}(t, t) = 0$ to obtain

$$\begin{aligned} dr(t) = {} & \frac{\partial f(0, t)}{\partial t}\, dt + \left\{ \int_0^t \left[\sigma_{\mathrm{p}}(s, t)\, \frac{\partial \sigma(s, t)}{\partial t} + \sigma(s, t)^2 \right] ds \right\} dt \\ & + \left[\int_0^t \frac{\partial \sigma(s, t)}{\partial t}\, dW_s \right] dt + \sigma(t, t)\, dW_t. \end{aligned} \tag{26.19}$$

Because the second and the third terms on the right-hand side depend on the history of σ_{p} and/or dW, they can make r non-Markovian. In the special case in which $\sigma_{\mathrm{p}}(t, T) = \sigma(T - t)$ for a constant σ, the short rate process r becomes Markovian and Eq. (26.19) reduces to

$$dr = \left[\frac{\partial f(0, t)}{\partial t} + \sigma^2 t \right] dt + \sigma\, dW.$$

Note that this is the continuous-time Ho–Lee model (review Exercise 26.2.10).

➤ **Exercise 26.5.1** What would $\mu(t, T)$ be if $\sigma(t, T) = \sigma e^{-\kappa(T-t)}$?

➤ **Exercise 26.5.2** Prove Eq. (26.16). (Hint: Use Eq. (26.15).)

➤ **Exercise 26.5.3** Consider the forward rate dynamics in Eq. (26.13) and define $r(t, \tau) \equiv f(t, t + \tau)$. Verify that

$$dr(r, \tau) = \left[\frac{\partial r(t, \tau)}{\partial \tau} + \mu(t, t + \tau) \right] dt + \sigma(t, t + \tau)\, dW_t.$$

Note that τ denotes the time to maturity.

Fixed-Income Option Pricing

For one-factor HJM models under the risk-neutral probability, the bond price process is

$$\frac{dP(t, T)}{P(t, T)} = r(t)\, dt + \sigma_{\mathrm{p}}(t, T)\, dW_t. \tag{26.20}$$

For a European option to buy at time s and at strike price X a zero-coupon bond maturing at time $T \geq s$, its value at time t is

$$C_t = P(t, T)\, N(d_t) - XP(t, s)\, N(d_t - \sigma_t), \tag{26.21}$$

where

$$d_t \equiv \sigma_t^{-1} \ln\left(\frac{P(t, T)}{XP(t, s)}\right) + \frac{\sigma_t}{2}$$

and $\sigma_t^2 \equiv \int_t^s [\sigma\sigma_{\mathrm{p}}(\tau, T) - \sigma_{\mathrm{p}}(\tau, s)]^2\, d\tau$ [164].

26.5.2 Markovian Short Rate Models

Markovian short rate models often simplify numerical procedures. First, the term structure at any time t is determined by t, the maturity, and the short rate at t. Second, the short rate dynamics can often be modeled by a combining tree. The HJM model's short rate process is usually non-Markovian. Under certain restrictions on volatility, however, the short rate contains all information relevant for pricing.

EXAMPLE 26.5.2 Suppose the volatility function is $\sigma(t, T) = \sigma$. Then the drift under π is $\mu(t, T) = \sigma \int_t^T \sigma\, ds = \sigma^2(T - t)$ by Theorem 26.5.1. The forward rate process is $df(t, T) = \sigma^2(T - t)\, dt + \sigma\, dW$, which is the Ho–Lee model (see Exercise 26.2.9). Integrate the above for each T to yield

$$\begin{aligned} f(t, T) &= f(0, T) + \int_0^t \sigma^2(T - s)\, ds + \int_0^t \sigma\, dW_s \\ &= f(0, T) + \sigma^2 t \left(T - \frac{t}{2}\right) + \sigma W(t). \end{aligned}$$

In particular, the short rate $r(t) \equiv f(t, t)$ is $r(t) = f(0, t) + (\sigma^2 t^2/2) + \sigma W(t)$. Because

$$\int_t^T f(t, s)\, ds = \int_t^T f(0, s)\, ds + \frac{\sigma^2}{2}\, tT(T - t) + \sigma(T - t)\, W(t),$$

the bond price is

$$P(t, T) = \frac{P(0, T)}{P(0, t)}\, e^{-(\sigma^2/2)\, tT(T-t) - \sigma(T-t)\, W(t)}.$$

Combining this with the short rate process above, we have

$$P(t, T) = \frac{P(0, T)}{P(0, t)}\, e^{f(0,t)(T-t) - (\sigma^2/2)\, t(T-t)^2 - r(t)(T-t)},$$

which expresses the bond price in terms of the short rate.

EXAMPLE 26.5.3 For the Hull–White model $dr = (\theta(t) - ar)\, dt + \sigma\, dW$, where $a, \sigma > 0$, the HJM formulation is $df(t, T) = \mu(t, T)\, dt + \sigma e^{-a(T-t)}\, dW$. The volatility structure of the forward rate, $\sigma(t, T) = \sigma e^{-a(T-t)}$, is exponentially decaying. The entire set of forward rates at any time t can be recovered from the short rate as follows:

$$f(t, T) = f(0, T) + e^{-a(T-t)}[\, r(t) - f(0, t) + \beta(t, T)\, \phi(t)\,],$$

where

$$\beta(t, T) \equiv \frac{1}{a}\left[1 - e^{-a(T-t)}\right], \qquad \phi(t) \equiv \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

The bond prices at time t and the state variable $r(t)$ are also related:

$$P(t, T) = \frac{P(0, T)}{P(0, t)}\, e^{-\beta(t,T)[\, r(t) - f(0,t)\,] - [\, \beta(t,T)^2 \phi(t)/2\,]}.$$

The duration of a zero-coupon bond at time t that matures at time T is $\beta(t, T)$. Finally,

$$dr(t) = \left[-a[\, r(t) - f(0, t)\,] + \frac{\partial f(0, t)}{\partial t} + \phi(t) \right] dt + \sigma\, dW_t$$

establishes the dynamics of the short rate.

EXAMPLE 26.5.4 When the σ_{p} in Eq. (26.20) is nonstochastic, $r(t)$ is Markovian if and only if $\sigma_{\mathrm{p}}(t, T)$ has the functional form $x(t)[\, y(T) - y(t)\,]$. The process for r then has the general form of the extended Vasicek model $dr = [\,\theta(t) - a(t)\, r\,]\, dt + \sigma(t)\, dW$ [477].

➤ **Exercise 26.5.4** (1) Price calls on a zero-coupon bond under the Ho–Lee model. (2) Price calls on a zero-coupon bond under the Hull–White model.

➤ **Exercise 26.5.5** Derive the Hull–White model's volatility structure.

➤ **Exercise 26.5.6** Verify that the model in Example 26.5.3 converges to the Ho–Lee model as $a \to 0$.

26.5.3 Binomial Approximation

Let Δt denote the duration of a period. The initial forward rate curve is $\{\, f(0, T), T = 0, \Delta t, 2\Delta t, \dots\}$, and $f(t, T)$ is the forward rate implied at time t for the time period $[\, T, T + \Delta t\,]$. During the next time period Δt, new information may arrive and cause the term structure to move. Let $f(t + \Delta t, T)$ be the new forward rate. The change $f(t + \Delta t, T) - f(t, T)$ depends on the forward rate, its maturity date T, and a host of other factors. See Fig. 26.8 for illustration.

Consider the binomial process in the actual economy:

$$f(t + \Delta t, T) = \begin{cases} f(t, T) + \mu(t, T)\, \Delta t + \sigma(t, T)\sqrt{\Delta t} & \text{with probability } q \\ f(t, T) + \mu(t, T)\, \Delta t - \sigma(t, T)\sqrt{\Delta t} & \text{with probability } 1 - q \end{cases},$$

where $q = 1/2 + o(\Delta t)$. The mean and the variance of the change in forward rates are $\mu(t, T)\, \Delta t + o(\Delta t)$ and $\sigma(t, T)^2 \Delta t + o(\Delta t)$, respectively. Convergence is guaranteed under mild conditions [434]. To make the process arbitrage free, a probability measure p must exist under which all claims can be priced as if the local expectations theory holds. We may assume that $p = 1/2$ for ease of computation. It can be shown that

$$\mu(t, T)\, \Delta t = \tanh(x)\, \sigma(t, T)\sqrt{\Delta t}\,, \tag{26.22}$$

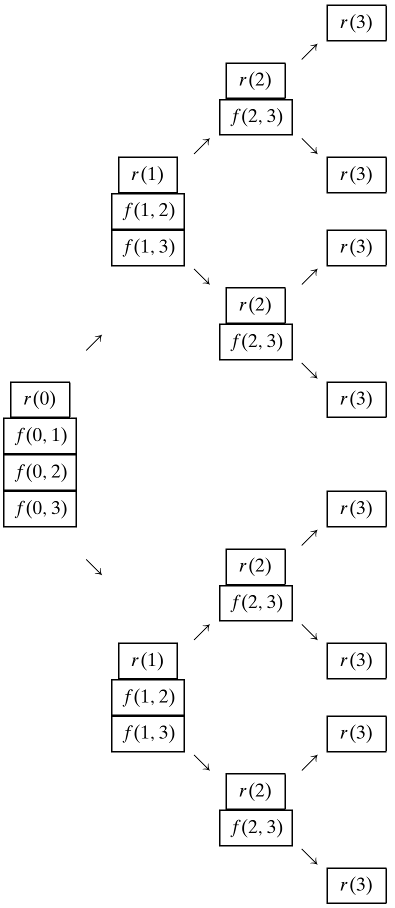

Figure 26.8: Binomial HJM model. For brevity, we use $f(i, j)$ to denote the forward rate for the period $[\, j\,\Delta t, (j+1)\,\Delta t\,]$ as seen from time $i\,\Delta t$. As always, $r(i) \equiv f(i, i)$ is the short rate. The binomial tree as a rule does not combine.

where

$$x \equiv \sqrt{\Delta t} \int_{t+\Delta t}^{T} \sigma(t, u)\, du, \quad \tanh(x) \equiv \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

This makes the price of any interest rate claim arbitrage free [746].

EXAMPLE 26.5.5 Consider $\sigma(t, T) = \sigma r(t)^\gamma e^{-\kappa(T-t)}$ for the forward rate volatility. Then

$$\mu(t, T)\,\Delta t = \tanh\left(\frac{\sigma r(t)^\gamma}{\kappa}\left[\, e^{-\kappa\Delta t} - e^{-\kappa(T-t)}\,\right]\sqrt{\Delta t}\right) \sigma r(t)^\gamma e^{-\kappa(T-t)} \sqrt{\Delta t}\,.$$

In particular, $\sigma(t, t)$ is the short rate volatility. This volatility is constant as in the Vasicek model when $\gamma = 0$. For $\gamma = 0.5$, the volatility is like that of the CIR model, whereas for $\gamma = 1$, the volatility is proportional to the short rate's level as in the BDT model. Take a three-period model with $\Delta t = 1$, $\gamma = 1$, $\kappa = 0.01$, and $\sigma = 0.3$. Then

$$\mu(t, T)\,\Delta t = \tanh\left(30 \times r(t)\left[\, e^{-0.01} - e^{-0.01\times(T-t)}\,\right]\right) \times 0.3 \times r(t)\, e^{-0.01\times(T-t)}.$$

Given a flat initial term structure at 5%, selected forward rates are shown in Fig. 26.9. The term structure at each node can be determined from the set of forward rates there.

➤ **Exercise 26.5.7** When pricing a derivative on bonds under the HJM model, does the tree have to be built over the life of the longer-term underlying bond or just over the life of the derivative?

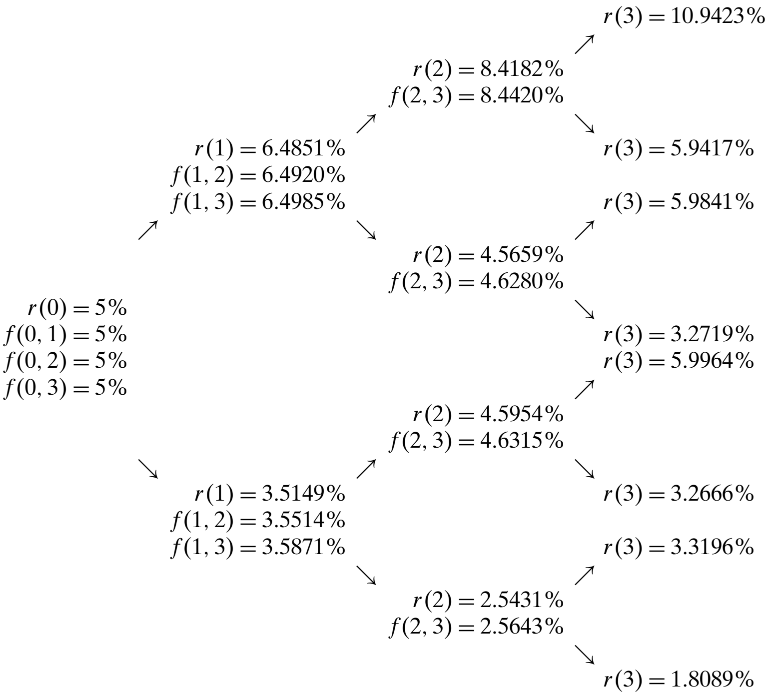

Figure 26.9: Sample binomial HJM tree. See Example 26.5.5 and Fig. 26.9.

A More Formal Setting

Assume that trades occur at times $\Delta t, 2\Delta t, 3\Delta t, \ldots$. As before, $f(t, T)$ denotes the forward interest rate contracted at time t for one period (of duration Δt) of borrowing or lending at time T. At each trading time t, the forward rates follow the stochastic difference equation:

$$f(t + \Delta t, T) - f(t, T) = \mu(t, T)\,\Delta t + \sigma(t, T)\,\xi(t + \Delta t)\sqrt{\Delta t}\,, \tag{26.23}$$

where $\xi(\cdot)$ are independent random variables with zero mean and unit variance. The value of $\xi(t)$ is realized before the trading at time t but after time $t - \Delta t$. The drift μ and the diffusion σ can be functions of the current or past values of forward rates. There is a single source of uncertainty (or factor), represented by ξ, that influences forward rates of all maturities. To permit forward rates of different maturities to vary independently, simply add additional sources of uncertainty to Eq. (26.23).

The price at time t of a discount bond of maturity at T is given by

$$P(t, T) = \exp\left[-\sum_{i=t/\Delta t}^{(T/\Delta t)-1} f(t, i\,\Delta t)\,\Delta t\right], \quad t = 0, \Delta t, 2\Delta t, \ldots$$

Substitute Eq. (26.23) into the preceding equation to yield

$$P(t, T) = \exp\Bigg[-\sum_{i=t/\Delta t}^{(T/\Delta t)-1}\Bigg(f(0, i\,\Delta t) + \sum_{j=0}^{i-1}\{\mu(j\,\Delta t, i\,\Delta t)\,\Delta t$$
$$+\,\sigma(j\,\Delta t, i\,\Delta t)\,\xi[(j+1)\,\Delta t]\sqrt{\Delta t}\,\}\Bigg)\Delta t\Bigg].$$

Now use the money market account $M(t)$ as numeraire and assume that it trades. Thus

$$M(0) = 1,$$
$$M(t) = \exp\left[\sum_{i=0}^{(t/\Delta t)-1} r(i\,\Delta t)\,\Delta t\right] = \exp\left[\sum_{i=0}^{(t/\Delta t)-1} f(i\,\Delta t, i\,\Delta t)\,\Delta t\right].$$

To avoid arbitrage, there must exist a probability measure π under which $P(t, T)/M(t)$ is a martingale; in particular,

$$P(t, T) = E_t^{\pi}[\, P(t+\Delta t, T)\, P(t, t+\Delta t)\,]. \tag{26.24}$$

The continuous-time limit under π yields $\mu(t, T) = \sigma(t, T) \int_t^T \sigma(t, u)\, du$.

We can take $\xi(\,\cdot\,) = \pm 1$, each with a probability of one-half, and then adjust the numbers on the nodes accordingly. This is what we did with the binomial interest rate tree, the Ho–Lee model, and the BDT model by fixing the risk-neutral probability to be 0.5. Because the martingale condition holds only as $\Delta t \to 0$, greater numerical accuracy may result by requiring it to hold under the binomial framework [19]. This leads to

$$\Delta t \int_{t+\Delta t}^{T} \mu(t, u)\, du = \ln \frac{e^x + e^{-x}}{2}, \tag{26.25}$$

where $x \equiv \sqrt{\Delta t} \int_{t+\Delta t}^{T} \sigma(t, u)\, du$. In Eq. (26.25) the convention is that $\mu(t, u)$ is constant for u between two adjacent trading times. The drifts $\mu(\,\cdot\,,\,\cdot\,)$ can be computed from Eq. (26.25) by solving them iteratively for $T = t + i\,\Delta t$, $i = 1, 2, \ldots$.

➤ **Exercise 26.5.8** Verify Eq. (26.25).

➤ **Exercise 26.5.9** Show that Eqs. (26.22) and (26.25) are equivalent in the limit.

26.6 The Ritchken–Sankarasubramanian Model

For the Ritchken–Sankarasubramanian (RS) model proposed by Ritchken and Sankarasubramanian [747], the forward rate volatility $\sigma(t, T)$ is related to the short rate volatility $\sigma(t, t)$ through an exogenously provided deterministic function $\kappa(x)$ by

$$\sigma(t, T) = \sigma(t, t)\, e^{-\int_t^T \kappa(x)\, dx}.$$

No particular restrictions are imposed on the short rate volatility $\sigma(t, t)$. This model precludes certain volatility structures. For instance, it does not permit volatilities of different forward rates to fluctuate according to different spot rates [600].

The term structure dynamics can be made Markovian with respect to two state variables. Bond prices – hence forward rates as well – at time t can be expressed in terms of the price information at time zero, the short rate $r(t)$, and a path-dependent statistic that represents the accumulated variance for the forward rate up to time t:

$$\phi(t) \equiv \int_0^t \sigma(u, t)^2\, du.$$

Note that $\phi(t)$ depends on the path the rate takes from time zero to t. In fact,

$$P(t, T) = \frac{P(0, T)}{P(0, t)}\, e^{-\beta(t,T)[r(t) - f(0,t)] - \beta(t,T)^2 \phi(t)/2}, \tag{26.26}$$

where $\beta(t, T) \equiv \int_t^T e^{-\int_t^u \kappa(x)\,dx}\,du$. The risk-neutral process follows

$$dr(t) = \mu(r, \phi, t)\,dt + \sigma(t, t)\,dW,$$
$$d\phi(t) = [\,\sigma(t, t)^2 - 2\kappa(t)\,\phi(t)\,]dt, \tag{26.27}$$

where

$$\mu(r, \phi, t) \equiv \kappa(t)[\,f(0, t) - r(t)\,] + \frac{df(0, t)}{dt} + \phi(t). \tag{26.28}$$

Because the short rate volatility could depend on both state variables $r(t)$ and $\phi(t)$, it may be expressed by $\sigma(r(t), \phi(t), t)$. Calibration is achieved by the appropriate choice of $\kappa(\cdot)$.

26.6.1 Binomial Approximation

Throughout this subsection, the forward rate volatility structure is given by

$$\sigma(t, T) = \sigma r(t)^{\gamma} e^{-\kappa(T-t)}.$$

(Note that the short rate volatility $\sigma(t, t)$ equals $\sigma r(t)^{\gamma}$.) Distant forward rates are therefore less volatile than near forward rates. For instance, with $\gamma = 0$, the extended Vasicek model results, and bond pricing formula (26.26) reduces to that for the extended Vasicek model (see Exercise 26.6.1). The CIR model emerges with $\gamma = 0.5$.

The forward rate at time t is given by

$$f(t, T) = f(0, T) + e^{-\kappa(T-t)}\,[\,r(t) - f(0, t) + \beta(t, T)\,\phi(t)\,],$$

where $\beta(t, T) \equiv [\,1 - e^{-\kappa(T-t)}\,]/\kappa$. When $\gamma \neq 0$, the $\phi(t)$ variable is determined by the path of rate values over the period $[\,0, t\,]$. Although the knowledge of $r(t)$ alone is insufficient to characterize the term structure at time t, that of $r(t)$ *and* $\phi(t)$ suffices. At time t, the duration of a zero-coupon bond maturing at time T is $\beta(t, T)$ [746].

The binomial tree model can be developed much as we did with the CIR model. First, take the following transformation to a form that has constant volatility:

$$Y(t) = \int \frac{1}{\sigma(r(t), \phi(t), t)}\,dr(t).$$

Let $r(t) = h(Y(t))$ be the inverse function. Then

$$dY(t) = m(Y, \phi, t)\,dt + dW_t,$$
$$d\phi(t) = [\,\sigma(r(t), \phi(t), t)^2 - 2\kappa(t)\,\phi(t)\,]\,dt,$$

where

$$m(Y, \phi, t) \equiv \frac{\partial Y(t)}{\partial t} + \mu(r, \phi, t)\,\frac{\partial Y(t)}{\partial r(t)} + \frac{1}{2}\,\sigma(r(t), \phi(t), t)^2\,\frac{\partial^2 Y(t)}{\partial r(t)^2}.$$

For example, for the **proportional model** with $\gamma = 1$, $Y(t) = (1/\sigma)\ln r(t)$ and

$$\sigma(r(t), \phi(t), t)^2 = \sigma^2 e^{2\sigma Y(t)}$$
$$m(Y, \phi, t) = \frac{1}{\sigma}\left\{\left[\kappa\big(f(0, t) - e^{\sigma Y(t)}\big) + \phi(t) + \frac{df(0, t)}{dt}\right] e^{-\sigma Y(t)} - \frac{\sigma^2}{2}\right\}.$$

Now, partition the interval $[0, T]$ into n periods each of length $\Delta t \equiv T/n$ and set up a combining binomial tree for Y. Initially, $Y_0 = Y(0)$ and $\phi_0 = 0$. During each time increment, the approximating Y_i can move to one of two values,

$$Y_i \begin{cases} Y_{i+1}^{+} = Y_i + \sqrt{\Delta t} \\ Y_{i+1}^{-} = Y_i - \sqrt{\Delta t} \end{cases}.$$

Let $r_i \equiv h(Y_i)$. Each node on the tree has two state variables. The tree evolves thus:

$$r_i, \phi_i \begin{cases} r_{i+1}^{+} = h(Y_{i+1}^{+}), \phi_{i+1} \\ r_{i+1}^{-} = h(Y_{i+1}^{-}), \phi_{i+1} \end{cases},$$

where $\phi_{i+1} = \phi_i + (\sigma^2 r_i^{2\gamma} - 2\kappa\phi_i)\,\Delta t$ by Eq. (26.27). The ϕ values are the same in both the up and the down nodes and have to be derived by forward induction. The probability of an up move to $(r_{i+1}^{+}, \phi_{i+1})$ given (r_i, ϕ_i) is

$$p(r_i, \phi_i, i) \equiv \frac{\mu(r_i, \phi_i, i\,\Delta t)\,\Delta t + r_i - r_{i+1}^{-}}{r_{i+1}^{+} - r_{i+1}^{-}},$$

where $\mu(\cdot, \cdot, \cdot)$ is the drift term in Eq. (26.28) with $\kappa(x) = \kappa$. Pricing can be carried out with backward induction, and the term structure at each node can be computed by Eq. (26.26).

EXAMPLE 26.6.1 For the proportional model, the transform is $r(t) = h(Y(t)) = e^{\sigma Y(t)}$. Let $r_i \equiv e^{\sigma Y_i}$. The binomial tree evolves according to

$$r_i, \phi_i \begin{cases} r_i e^{\sigma\sqrt{\Delta t}}, \phi_{i+1} \\ r_i e^{-\sigma\sqrt{\Delta t}}, \phi_{i+1} \end{cases},$$

where $\phi_{i+1} = \phi_i + (\sigma^2 r_i^2 - 2\kappa\phi_i)\,\Delta t$.

The size of the tree as described grows exponentially in n. In fact, the number of ϕ values at a node equals the number of distinct paths leading to it from the root because the ϕ value is path dependent. This observation should be clear from the binomial process. One remedy is to keep only the maximum and the minimum ϕ values at each node, interpolating m intermediate values linearly on demand (see Subsection 11.7.1 for the same idea). The model is expected to converge for sufficiently large n and m.

➤ **Exercise 26.6.1** Show that $\phi(t) = \sigma^2(1 - e^{-2\kappa t})/(2\kappa)$ for the extended Vasicek model.

Additional Reading

See [858] for a survey of interest rate models, [754] for a comparison of models, [32, 78, 378, 670] for theoretical analysis, [47, 139, 677] for empirical studies of the short rate and the HJM models, and [276, 507, 779, 780, 850] for more information on derivatives pricing. References [40, 183, 470, 510, 681, 731] also cover interest rate models. Consult [359] for an empirical study of the Ho–Lee model. See [74, 625] for calibration of the BDT model and [479, 483, 735] for that of the Hull–White model. A two-factor extended Vasicek model is proposed in [480]. For the HJM model, see

[58, 116, 165, 166, 435, 437] for theoretical analysis, [163, 197, 477, 499, 515, 747] for the conditions that make the short rate Markovian, [682] for risk management issues, [510] for estimating volatilities by use of principal components, [436, 510] for discrete-time models, and [160] for the Monte Carlo approach to pricing American-style fixed-income options. For the RS model, see [600, 746, 747] for additional numerical ideas and [92] for an empirical study. Reference [77] asks whether an interest rate model generates term structures within the functionals (say, polynomials) used for fitting the term structure; in particular, it shows that the Ho–Lee model and the Hull–White model are inconsistent with the Nelson–Siegel scheme. Volatility structures are investigated in [136, 393, 722]. A universal trinomial tree algorithm for any Ito process is proposed in [187].

NOTES

1. Contrast this with the lognormal model in Exercise 23.2.3.
2. The value of p can be chosen rather arbitrarily because, in the limit, only the volatility matters [436].
3. Under the premise that the forward rate tends to a constant for large t, it follows from Eq. (26.6) that $r(t) - r(t-1)$ is either essentially zero or $|\ln \delta| > 0$ [300].