# CHAPTER 28

# Interest Rate Derivatives: Models of the Short Rate

The models for pricing interest rate options that we have presented so far make the assumption that the probability distribution of an interest rate, a bond price, or some other variable at a future point in time is lognormal. They are widely used for valuing instruments such as caps, European bond options, and European swap options. However, they have limitations. They do not provide a description of how interest rates evolve through time. Consequently, they cannot be used for valuing interest rate derivatives such as American-style swap options, callable bonds, and structured notes.

This chapter and the next discuss alternative approaches for overcoming these limitations. These involve building what is known as a *term structure model*. This is a model describing the evolution of all zero-coupon interest rates.[1] This chapter focuses on term structure models constructed by specifying the behavior of the short-term interest rate, $r$.

## 28.1 BACKGROUND

The short rate, $r$, at time $t$ is the rate that applies to an infinitesimally short period of time at time $t$. It is sometimes referred to as the *instantaneous short rate*. Bond prices, option prices, and other derivative prices depend only on the process followed by $r$ in a risk-neutral world. The process for $r$ in the real world is irrelevant. The risk-neutral world we consider here will be the traditional risk-neutral world where, in a very short time period between $t$ and $t + \Delta t$, investors earn on average $r(t)\,\Delta t$. All processes for $r$ that we present will be processes in this risk-neutral world.

From equation (25.19), the value at time $t$ of an interest rate derivative that provides a payoff of $f_T$ at time $T$ is

$$\hat{E}[e^{-\bar{r}(T-t)} f_T] \tag{28.1}$$

where $\bar{r}$ is the average value of $r$ in the time interval between $t$ and $T$, and $\hat{E}$ denotes expected value in the traditional risk-neutral world.

[1] Note that when a term structure model is used we do not need to make the convexity, timing, and quanto adjustments discussed in the previous chapter.

As usual we define $P(t, T)$ as the price at time $t$ of a zero-coupon bond that pays off \$1 at time $T$. From equation (28.1),

$$P(t, T) = \hat{E}[e^{-\bar{r}(T-t)}] \tag{28.2}$$

If $R(t, T)$ is the continuously compounded interest rate at time $t$ for a term of $T - t$, then

$$P(t, T) = e^{-R(t,T)(T-t)} \tag{28.3}$$

so that

$$R(t, T) = -\frac{1}{T-t} \ln P(t, T) \tag{28.4}$$

and, from equation (28.2),

$$R(t, T) = -\frac{1}{T-t} \ln \hat{E}[e^{-\bar{r}(T-t)}] \tag{28.5}$$

This equation enables the term structure of interest rates at any given time to be obtained from the value of $r$ at that time and the risk-neutral process for $r$. It shows that once we have fully defined the process for $r$, we have fully defined everything about the initial zero curve and its evolution through time.

## 28.2 EQUILIBRIUM MODELS

Equilibrium models usually start with assumptions about economic variables and derive a process for the short rate, $r$. They then explore what the process for $r$ implies about bond prices and option prices.

In a one-factor equilibrium model, the process for $r$ involves only one source of uncertainty. Usually the risk-neutral process for the short rate is described by an Itô process of the form

$$dr = m(r)\,dt + s(r)\,dz$$

The instantaneous drift, $m$, and instantaneous standard deviation, $s$, are assumed to be functions of $r$, but are independent of time. The assumption of a single factor is not as restrictive as it might appear. A one-factor model implies that all rates move in the same direction over any short time interval, but not that they all move by the same amount. The shape of the zero curve can therefore change with the passage of time.

We now consider three one-factor equilibrium models:

$m(r) = \mu r$; $s(r) = \sigma r$ (Rendleman and Bartter model)

$m(r) = a(b - r)$; $s(r) = \sigma$ (Vasicek model)

$m(r) = a(b - r)$; $s(r) = \sigma\sqrt{r}$ (Cox, Ingersoll, and Ross model)

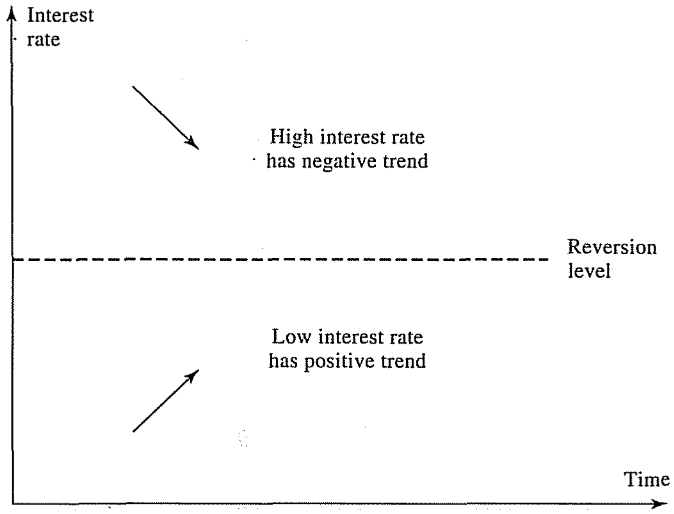
### The Rendleman and Bartter Model

In Rendleman and Bartter's model, the risk-neutral process for $r$ is[2]

$$dr = \mu r\,dt + \sigma r\,dz$$

[2] See R. Rendleman and B. Bartter, "The Pricing of Options on Debt Securities," *Journal of Financial and Quantitative Analysis*, 15 (March 1980): 11–24.

**Figure 28.1** Mean reversion.




where $\mu$ and $\sigma$ are constants. This means that $r$ follows geometric Brownian motion. The process for $r$ is of the same type as that assumed for a stock price in Chapter 13. It can be represented using a binomial tree similar to the one used for stocks in Chapter 11.[3]

The assumption that the short-term interest rate behaves like a stock price is a natural starting point but is less than ideal. One important difference between interest rates and stock prices is that interest rates appear to be pulled back to some long-run average level over time. This phenomenon is known as *mean reversion*. When $r$ is high, mean reversion tends to cause it to have a negative drift; when $r$ is low, mean reversion tends to cause it to have a positive drift. Mean reversion is illustrated in Figure 28.1. The Rendleman and Bartter model does not incorporate mean reversion.

There are compelling economic arguments in favor of mean reversion. When rates are high, the economy tends to slow down and there is low demand for funds from borrowers. As a result, rates decline. When rates are low, there tends to be a high demand for funds on the part of borrowers and rates tend to rise.

## The Vasicek Model

In Vasicek's model, the risk-neutral process for $r$ is

$$dr = a(b - r)\,dt + \sigma\,dz$$

where $a$, $b$, and $\sigma$ are constants.[4] This model incorporates mean reversion. The short rate is pulled to a level $b$ at rate $a$. Superimposed upon this "pull" is a normally distributed stochastic term $\sigma\,dz$.

[3] The way that the interest rate tree is used is explained later in the chapter.

[4] See O. A. Vasicek, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5 (1977): 177–88.

Vasicek shows that equation (28.2) can be used to obtain the following expression for the price at time $t$ of a zero-coupon bond that pays \$1 at time $T$:

$$P(t, T) = A(t, T)e^{-B(t,T)r(t)} \tag{28.6}$$

In this equation $r(t)$ is the value of $r$ at time $t$,

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a} \tag{28.7}$$

and

$$A(t, T) = \exp\left[\frac{(B(t, T) - T + t)(a^2 b - \sigma^2/2)}{a^2} - \frac{\sigma^2 B(t, T)^2}{4a}\right] \tag{28.8}$$

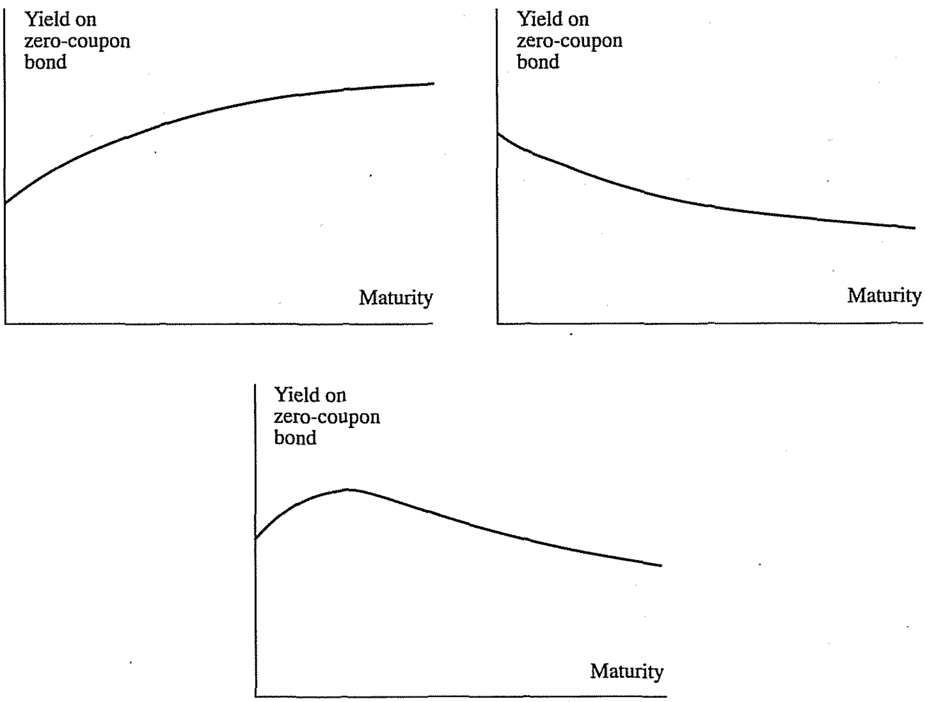
When $a = 0$, $B(t, T) = T - t$ and $A(t, T) = \exp[\sigma^2(T - t)^3/6]$.

Using equation (28.4), we get

$$R(t, T) = -\frac{1}{T-t}\ln A(t, T) + \frac{1}{T-t}B(t, T)r(t) \tag{28.9}$$

showing that the entire term structure can be determined as a function of $r(t)$ once $a$, $b$, and $\sigma$ are chosen. The shape can be upward-sloping, downward-sloping, or slightly "humped" (see Figure 28.2).

**Figure 28.2** Possible shapes of term structure when Vasicek's model is used.



## The Cox, Ingersoll, and Ross Model

In Vasicek's model the short-term interest rate, $r$, can become negative. Cox, Ingersoll, and Ross have proposed an alternative model where rates are always non-negative.[5] The risk-neutral process for $r$ in their model is

$$dr = a(b - r)\,dt + \sigma\sqrt{r}\,dz$$

This has the same mean-reverting drift as Vasicek, but the standard deviation of the change in the short rate in a short period of time is proportional to $\sqrt{r}$. This means that, as the short-term interest rate increases, its standard deviation increases.

Cox, Ingersoll, and Ross show that, in their model, bond prices have the same general form as those in Vasicek's model,

$$P(t, T) = A(t, T)e^{-B(t,T)r}$$

but the functions $B(t, T)$ and $A(t, T)$ are different,

$$B(t, T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}$$

and

$$A(t, T) = \left[\frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma}\right]^{2ab/\sigma^2}$$

with $\gamma = \sqrt{a^2 + 2\sigma^2}$. Upward-sloping, downward-sloping, and slightly humped yield curves are possible. As in the case of Vasicek's model, the long rate, $R(t, T)$, is linearly dependent on $r(t)$. This means that the value of $r(t)$ determines the level of the term structure at time $t$. The general shape of the term structure at time $t$ is independent of $r(t)$, but does depend on $t$.

## Two-Factor Equilibrium Models

A number of researchers have investigated the properties of two-factor equilibrium models. For example, Brennan and Schwartz have developed a model where the process for the short rate reverts to a long rate, which in turn follows a stochastic process.[6] The long rate is chosen as the yield on a perpetual bond that pays $1 per year. Because the yield on this bond is the reciprocal of its price, Itô's lemma can be used to calculate the process followed by the yield from the process followed by the price of the bond. The bond is a traded security. This simplifies the analysis because the expected return on the bond in a risk-neutral world must be the risk-free interest rate.

---

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

[6] See M.J. Brennan and E.S. Schwartz, "A Continuous Time Approach to Pricing Bonds," *Journal of Banking and Finance*, 3 (July 1979): 133–55; M.J. Brennan and E.S. Schwartz, "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *Journal of Financial and Quantitative Analysis*, 21, 3 (September 1982): 301–29.

Another two-factor model, proposed by Longstaff and Schwartz, starts with a general equilibrium model of the economy and derives a term structure model where there is stochastic volatility.[7] The model proves to be analytically quite tractable.

## 28.3 NO-ARBITRAGE MODELS

The disadvantage of the equilibrium models we have presented is that they do not automatically fit today's term structure of interest rates. By choosing the parameters judiciously, they can be made to provide an approximate fit to many of the term structures that are encountered in practice. But the fit is not usually an exact one and, in some cases, no reasonable fit can be found. Most traders find this unsatisfactory. Not unreasonably, they argue that they can have very little confidence in the price of a bond option when the model does not price the underlying bond correctly. A 1% error in the price of the underlying bond may lead to a 25% error in an option price.

A *no-arbitrage model* is a model designed to be exactly consistent with today's term structure of interest rates. The essential difference between an equilibrium and a no-arbitrage model is therefore as follows. In an equilibrium model, today's term structure of interest rates is an output. In a no-arbitrage model, today's term structure of interest rates is an input.

In an equilibrium model, the drift of the short rate (i.e., the coefficient of $dt$) is not usually a function of time. In a no-arbitrage model, the drift is, in general, dependent on time. This is because the shape of the initial zero curve governs the average path taken by the short rate in the future in a no-arbitrage model. If the zero curve is steeply upward-sloping for maturities between $t_1$ and $t_2$, then $r$ has a positive drift between these times; if it is steeply downward-sloping for these maturities, then $r$ has a negative drift between these times.

It turns out that some equilibrium models can be converted to no-arbitrage models by including a function of time in the drift of the short rate. We now consider the Ho–Lee, Hull–White (one- and two-factor), and Black–Karasinski models.

### The Ho–Lee Model

Ho and Lee proposed the first no-arbitrage model of the term structure in a paper in 1986.[8] They presented the model in the form of a binomial tree of bond prices with two parameters: the short-rate standard deviation and the market price of risk of the short rate. It has since been shown that the continuous-time limit of the model is

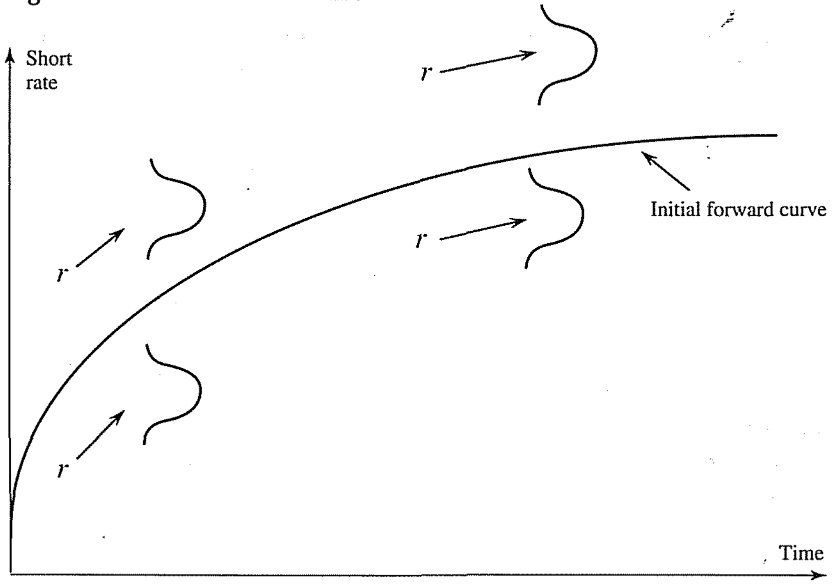
$$dr = \theta(t)\,dt + \sigma\,dz \tag{28.10}$$

where $\sigma$, the instantaneous standard deviation of the short rate, is constant and $\theta(t)$ is a function of time chosen to ensure that the model fits the initial term structure. The variable $\theta(t)$ defines the average direction that $r$ moves at time $t$. This is independent of the level of $r$. Interestingly, Ho and Lee's parameter that concerns the market price of

[7] See F. A. Longstaff and E. S. Schwartz, "Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model," *Journal of Finance*, 47, 4 (September 1992): 1259–82.

[8] See T. S. Y. Ho and S.-B. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, 41 (December 1986): 1011–29.

**Figure 28.3** The Ho–Lee model.




risk proves to be irrelevant when the model is used to price interest rate derivatives. This is analogous to risk preferences being irrelevant in the pricing of stock options.

The variable $\theta(t)$ can be calculated analytically (see Problem 28.13). It is

$$\theta(t) = F_t(0, t) + \sigma^2 t \tag{28.11}$$

where the $F(0, t)$ is the instantaneous forward rate for a maturity $t$ as seen at time zero and the subscript $t$ denotes a partial derivative with respect to $t$. As an approximation, $\theta(t)$ equals $F_t(0, t)$. This means that the average direction that the short rate will be moving in the future is approximately equal to the slope of the instantaneous forward curve. The Ho–Lee model is illustrated in Figure 28.3. The slope of the forward curve defines the average direction that the short rate is moving at any given time. Superimposed on this slope is the normally distributed random outcome.

In the Ho–Lee model, zero-coupon bonds and European options on zero-coupon bonds can be valued analytically. The expression for the price of a zero-coupon bond at time $t$ in terms of the short rate is

$$P(t, T) = A(t, T)e^{-r(t)(T-t)} \tag{28.12}$$

where

$$\ln A(t, T) = \ln\frac{P(0, T)}{P(0, t)} + (T - t)F(0, t) - \tfrac{1}{2}\sigma^2 t(T - t)^2$$

In these equations, time zero is today. Times $t$ and $T$ are general times in the future with $T \geqslant t$. The equations, therefore, define the price of a zero-coupon bond at a future time $t$ in terms of the short rate at time $t$ and the prices of bonds today. The latter can be calculated from today's term structure.

## The Hull–White (One-Factor) Model

In a paper published in 1990, Hull and White explored extensions of the Vasicek model that provide an exact fit to the initial term structure.[9] One version of the extended Vasicek model that they consider is

$$dr = [\theta(t) - ar]\,dt + \sigma\,dz \tag{28.13}$$

or

$$dr = a\left[\frac{\theta(t)}{a} - r\right]dt + \sigma\,dz$$

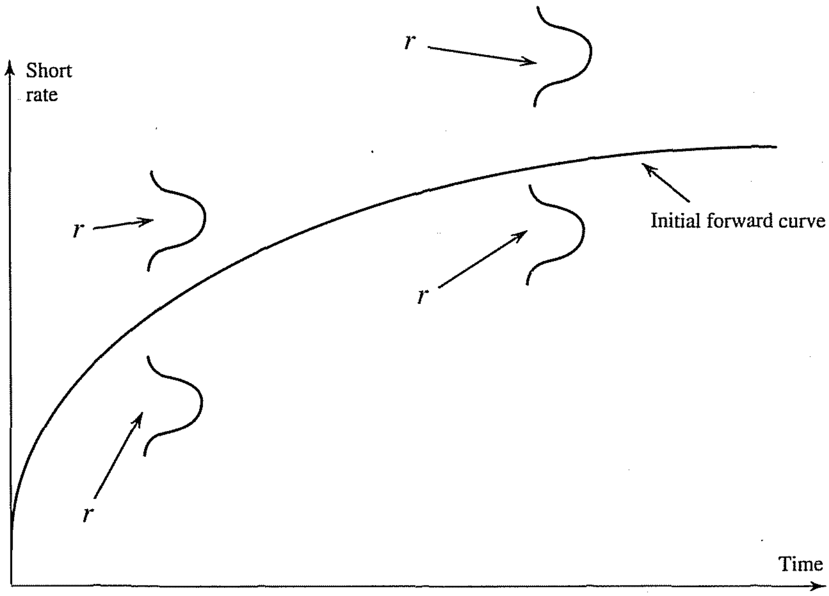
where $a$ and $\sigma$ are constants. This is known as the Hull–White model. It can be characterized as the Ho–Lee model with mean reversion at rate $a$. Alternatively, it can be characterized as the Vasicek model with a time-dependent reversion level. At time $t$, the short rate reverts to $\theta(t)/a$ at rate $a$. The Ho–Lee model is a particular case of the Hull–White model with $a = 0$.

The model has the same amount of analytic tractability as Ho–Lee. The $\theta(t)$ function can be calculated from the initial term structure (see Problem 28.14):

$$\theta(t) = F_t(0, t) + aF(0, t) + \frac{\sigma^2}{2a}(1 - e^{-2at}) \tag{28.14}$$

The last term in this equation is usually fairly small. If we ignore it, the equation implies that the drift of the process for $r$ at time $t$ is $F_t(0, t) + a[F(0, t) - r]$. This shows that, on

**Figure 28.4** The Hull–White model.




[9] See J. Hull and A. White, "Pricing Interest Rate Derivative Securities," *Review of Financial Studies*, 3, 4 (1990): 573–92.

average, $r$ follows the slope of the initial instantaneous forward rate curve. When it deviates from that curve, it reverts back to it at rate $a$. The model is illustrated in Figure 28.4.

Bond prices at time $t$ in the Hull–White model are given by

$$P(t, T) = A(t, T)e^{-B(t,T)r(t)} \tag{28.15}$$

where

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a} \tag{28.16}$$

and

$$\ln A(t, T) = \ln\frac{P(0, T)}{P(0, t)} + B(t, T)F(0, t) - \frac{1}{4a^3}\sigma^2(e^{-aT} - e^{-at})^2(e^{2at} - 1) \tag{28.17}$$

Equations (28.15), (28.16), and (28.17) define the price of a zero-coupon bond at a future time $t$ in terms of the short rate at time $t$ and the prices of bonds today. The latter can be calculated from today's term structure.

## The Black–Karasinski Model

The Ho–Lee and Hull–White models have the disadvantage that the short-term interest rate, $r$, can become negative. A model that allows only positive interest rates is a model propose by Black and Karasinski:[10]

$$d\ln r = [\theta(t) - a(t)\ln(r)]\,dt + \sigma(t)\,dz \tag{28.18}$$

The variable $\ln r$ follows the same process as $r$ in the Hull–White model. Whereas the value of $r$ at a future time is normal in the Ho–Lee and Hull–White models, it is lognormal in the Black–Karasinski model.

The Black–Karasinski model does not have as much analytic tractability as Ho–Lee or Hull–White. For example, it is not possible to produce formulas for valuing bonds in terms of $r$ using the model.

## The Hull–White Two-Factor Model

A no-arbitrage model that involves a similar idea to the two-factor equilibrium model suggested by Brennan and Schwartz is[11]

$$df(r) = [\theta(t) + u - af(r)]\,dt + \sigma_1\,dz_1 \tag{28.19}$$

where $u$ has an initial value of zero and follows the process

$$du = -bu\,dt + \sigma_2\,dz_2$$

As in the one-factor models just considered, the parameter $\theta(t)$ is chosen to make the model consistent with the initial term structure. The stochastic variable $u$ is a component of the reversion level of $r$ and itself reverts to a level of zero at rate $b$. The

[10] See F. Black and P. Karasinski, "Bond and Option Pricing When Short Rates Are Lognormal," *Financial Analysts Journal*, July/August (1991), 52–59.

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

parameters $a$, $b$, $\sigma_1$, and $\sigma_2$ are constants and $dz_1$ and $dz_2$ are Wiener processes with instantaneous correlation $\rho$.

This model provides a richer pattern of term structure movements and a richer pattern of volatilities than one-factor models of $r$. For more information on the model, see Technical Note 14 on the author's website.

## 28.4 OPTIONS ON BONDS

Some of the models we have presented allow options on zero-coupon bonds to be valued analytically. For the Vasicek, Ho–Lee, and Hull–White models, the price at time zero of a call option that matures at time $T$ on a zero-coupon bond maturing at time $s$ is

$$LP(0,s)N(h) - KP(0,T)N(h-\sigma_P) \tag{28.20}$$

where $L$ is the principal of the bond, $K$ is its strike price, and

$$h = \frac{1}{\sigma_P}\ln\frac{LP(0,s)}{P(0,T)K} + \frac{\sigma_P}{2}$$

The price of a put option on the bond is

$$KP(0,T)N(-h+\sigma_P) - LP(0,s)N(-h)$$

In the case of the Vasicek and Hull–White models,

$$\sigma_P = \frac{\sigma}{a}[1-e^{-a(s-T)}]\sqrt{\frac{1-e^{-2aT}}{2a}}$$

In the case of the Ho–Lee model,

$$\sigma_P = \sigma(s-T)\sqrt{T}$$

Equation (28.20) is essentially the same as Black's model for pricing bond options in Section 26.2. The bond price volatility is $\sigma_P/\sqrt{T}$ and the standard deviation of the logarithm of the bond price at time $T$ is $\sigma_P$. As explained in Section 26.3, an interest rate cap or floor can be expressed as a portfolio of options on zero-coupon bonds. It can, therefore, be valued analytically using the equations just presented.

There are also formulas for valuing options on zero-coupon bonds in the Cox, Ingersoll, and Ross model, which we presented in Section 28.2. These involve integrals of the noncentral chi-square distribution.

### Options on Coupon-Bearing Bonds

In a one-factor model of $r$, all zero-coupon bonds move up in price when $r$ decreases and all zero-coupon bonds move down in price when $r$ increases. As a result, a one-factor model allows us to express a European option on a coupon-bearing bond as the sum of European options on zero-coupon bonds. The procedure is as follows:

1. Calculate $r^*$, the critical value of $r$ for which the price of the coupon-bearing bond equals the strike price of the option on the bond at option maturity.

2. Calculate the prices of options on the zero-coupon bonds that comprise the coupon-bearing bond. Set the strike price of each option equal to the value the corresponding zero-coupon bond will have at time $T$ when $r = r^*$.
3. Set the price of the option on the coupon-bearing bond equal to the sum of the prices on the options on zero-coupon bonds calculated in step 2.

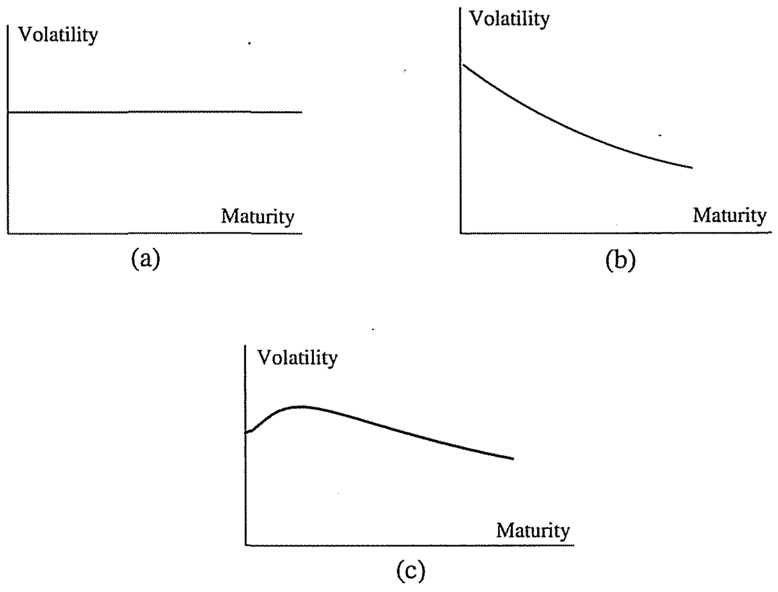
This allows options on coupon-bearing bonds to be valued for the Vasicek, Cox, Ingersoll, and Ross, Ho–Lee, and Hull–White models As explained in Business Snapshot 26.2, a European swap option can be viewed as an option on a coupon-bearing bond. It can, therefore, be valued using this procedure. For more details on the procedure, see Technical Note 15 on the author's website.

## 28.5 VOLATILITY STRUCTURES

The models we have looked at give rise to different volatility environments. In Figure 28.5 we show the volatility of the 3-month forward rate as a function of maturity for Ho–Lee, Hull–White one-factor and Hull–White two-factor models. The term structure of interest rates is assumed to be flat.

For Ho–Lee the volatility of the 3-month forward rate is the same for all maturities. In the one-factor Hull–White model the effect of mean reversion is to cause the volatility of the 3-month forward rate to be a declining function of maturity. In the Hull–White two-factor model when parameters are chosen appropriately, the volatility

**Figure 28.5** Volatility of 3-month forward rate as a function of maturity for (a) the Ho–Lee model, (b) the Hull–White one-factor model, and (c) the Hull–White two-factor model (when parameters are chosen appropriately).



of the 3-month forward rate has a "humped" look. The latter is consistent with empirical evidence and implied cap volatilities discussed in Section 26.3.

## 28.6 INTEREST RATE TREES

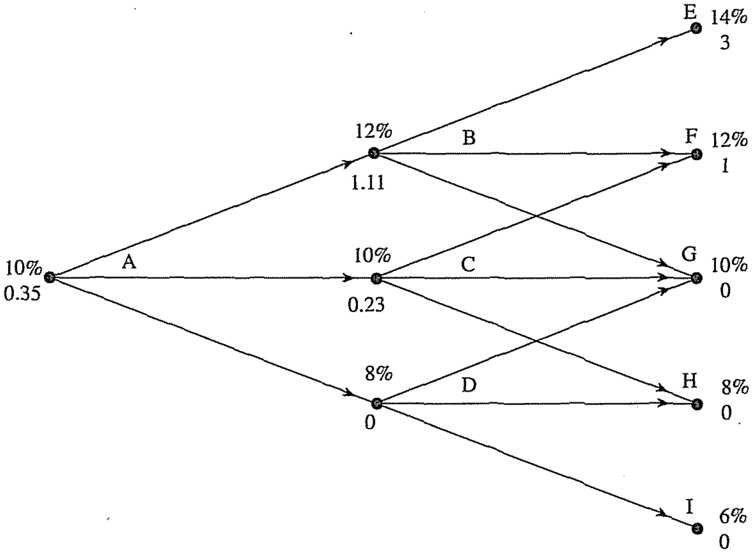
An interest rate tree is a discrete-time representation of the stochastic process for the short rate in much the same way as a stock price tree is a discrete-time representation of the process followed by a stock price. If the time step on the tree is $\Delta t$, the rates on the tree are the continuously compounded $\Delta t$-period rates. The usual assumption when a tree is constructed is that the $\Delta t$-period rate, $R$, follows the same stochastic process as the instantaneous rate, $r$, in the corresponding continuous-time model. The main difference between interest rate trees and stock price trees is in the way that discounting is done. In a stock price tree, the discount rate is usually assumed to be the same at each node (or a function of time). In an interest rate tree, the discount rate varies from node to node.

It often proves to be convenient to use a trinomial rather than a binomial tree for interest rates. The main advantage of a trinomial tree is that it provides an extra degree of freedom, making it easier for the tree to represent features of the interest rate process such as mean reversion. As mentioned in Section 17.8, using a trinomial tree is equivalent to using the explicit finite difference method.

### Illustration of Use of Trinomial Trees

To illustrate how trinomial interest rate trees are used to value derivatives, we consider the simple example shown in Figure 28.6. This is a two-step tree with each time step equal to 1 year in length so that $\Delta t = 1$ year. We assume that the up, middle, and down

**Figure 28.6** Example of the use of trinomial interest rate trees. Upper number at each node is rate; lower number is value of instrument.



probabilities are 0.25, 0.50, and 0.25, respectively, at each node. The assumed $\Delta t$-period rate is shown as the upper number at each node.[12]

The tree is used to value a derivative that provides a payoff at the end of the second time step of

$$\max[100(R - 0.11), 0]$$

where $R$ is the $\Delta t$-period rate. The calculated value of this derivative is the lower number at each node. At the final nodes, the value of the derivative equals the payoff. For example, at node E, the value is $100 \times (0.14 - 0.11) = 3$. At earlier nodes, the value of the derivative is calculated using the rollback procedure explained in Chapters 11 and 17. At node B, the 1-year interest rate is 12%. This is used for discounting to obtain the value of the derivative at node B from its values at nodes E, F, and G as

$$[0.25 \times 3 + 0.5 \times 1 + 0.25 \times 0]e^{-0.12 \times 1} = 1.11$$

At node C, the 1-year interest rate is 10%. This is used for discounting to obtain the value of the derivative at node C as

$$(0.25 \times 1 + 0.5 \times 0 + 0.25 \times 0)e^{-0.1 \times 1} = 0.23$$

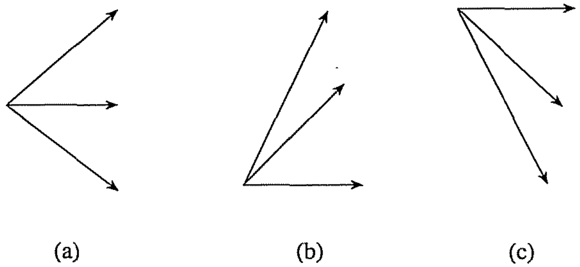
At the initial node, A, the interest rate is also 10% and the value of the derivative is

$$(0.25 \times 1.11 + 0.5 \times 0.23 + 0.25 \times 0)e^{-0.1 \times 1} = 0.35$$

## Nonstandard Branching

It sometimes proves convenient to modify the standard branching pattern, which is used at all nodes in Figure 28.6. Three alternative branching possibilities are shown in Figure 28.7. The usual branching is shown in Figure 28.7(a). It is "up one/straight along/down one". One alternative to this is "up two/up one/straight along", as shown in Figure 28.7(b). This proves useful for incorporating mean reversion when interest rates are very low. A third branching pattern shown in Figure 28.7(c) is "straight along/down one/down two". This is useful for incorporating mean reversion when interest rates are very high. We illustrate the use of different branching patterns in the following section.

**Figure 28.7** Alternative branching methods in a trinomial tree.




[12] We explain later how the probabilities and rates on an interest rate tree are determined.

## 28.7 A GENERAL TREE-BUILDING PROCEDURE

Hull and White have proposed a robust two-stage procedure for constructing trinomial trees to represent a wide range of one-factor models.[13] This section first explains how the procedure can be used for the Hull–White model in equation (28.13) and then shows how it can be extended to represent other models.

### First Stage

The Hull–White model for the instantaneous short rate $r$ is

$$dr = [\theta(t) - ar]\,dt + \sigma\,dz$$

We suppose that the time step on the tree is constant and equal to $\Delta t$.[14]

We assume that the $\Delta t$ rate, $R$, follows the same process as $r$.

$$dR = [\theta(t) - aR]\,dt + \sigma\,dz$$

Clearly, this is reasonable in the limit as $\Delta t$ tends to zero. The first stage in building a tree for this model is to construct a tree for a variable $R^*$ that is initially zero and follows the process

$$dR^* = -aR^*\,dt + \sigma\,dz$$

This process is symmetrical about $R^* = 0$. The variable $R^*(t + \Delta t) - R^*(t)$ is normally distributed. If terms of higher order than $\Delta t$ are ignored, the expected value of $R^*(t + \Delta t) - R^*(t)$ is $-aR^*(t)\Delta t$ and the variance of $R^*(t + \Delta t) - R^*(t)$ is $\sigma^2 \Delta t$.

We define $\Delta R$ as the spacing between interest rates on the tree and set

$$\Delta R = \sigma\sqrt{3\Delta t}$$

This proves to be a good choice of $\Delta R$ from the viewpoint of error minimization.

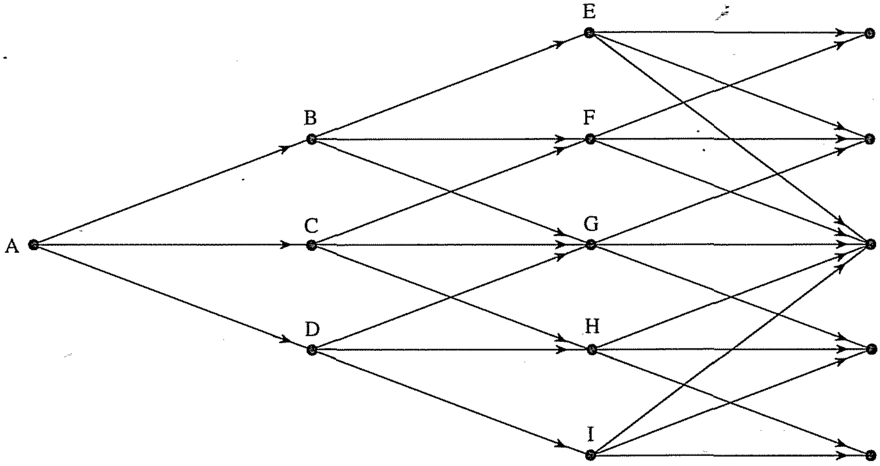
Our objective during the first stage of this procedure is to build a tree similar to that shown in Figure 28.8 for $R^*$. To do this, we must resolve which of the three branching methods shown in Figure 28.7 will apply at each node. This will determine the overall geometry of the tree. Once this is done, the branching probabilities must also be calculated.

Define $(i, j)$ as the node where $t = i\,\Delta t$ and $R^* = j\,\Delta R$. (The variable $i$ is a positive integer and $j$ is a positive or negative integer.) The branching method used at a node must lead to the probabilities on all three branches being positive. Most of the time, the branching shown in Figure 28.7(a) is appropriate. When $a > 0$, it is necessary to switch from the branching in Figure 28.7(a) to the branching in Figure 28.7(c) for a sufficiently large $j$. Similarly, it is necessary to switch from the branching in Figure 28.7(a) to the branching in Figure 28.7(b) when $j$ is sufficiently negative. Define $j_{max}$ as the value of $j$ where we switch from the Figure 28.7(a) branching to the Figure 28.7(c) branching and $j_{min}$ as the value of $j$ where we switch from the Figure 28.7(a) branching to the Figure 28.7(b) branching. Hull and White show that probabilities are always positive if

[13] See J. Hull and A. White, "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives*, 2, 1 (1994): 7–16; and J. Hull and A. White, "Using Hull–White Interest Rate Trees," *Journal of Derivatives*, (Spring 1996): 26–36.

[14] See Technical Note 16 on the author's website for a discussion of how nonconstant time steps can be used.

**Figure 28.8** Tree for $R^*$ in Hull–White model (first stage).




| *Node*: | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| $R$(%) | 0.000 | 1.732 | 0.000 | −1.732 | 3.464 | 1.732 | 0.000 | −1.732 | −3.464 |
| $p_u$ | 0.1667 | 0.1217 | 0.1667 | 0.2217 | 0.8867 | 0.1217 | 0.1667 | 0.2217 | 0.0867 |
| $p_m$ | 0.6666 | 0.6566 | 0.6666 | 0.6566 | 0.0266 | 0.6566 | 0.6666 | 0.6566 | 0.0266 |
| $p_d$ | 0.1667 | 0.2217 | 0.1667 | 0.1217 | 0.0867 | 0.2217 | 0.1667 | 0.1217 | 0.8867 |

we set $j_{max}$ equal to the smallest integer greater than $0.184/(a\,\Delta t)$ and $j_{min}$ equal to $-j_{max}$.[15] Define $p_u$, $p_m$, and $p_d$ as the probabilities of the highest, middle, and lowest branches emanating from the node. The probabilities are chosen to match the expected change and variance of the change in $R^*$ over the next time interval $\Delta t$. The probabilities must also sum to unity. This leads to three equations in the three probabilities.

As already mentioned, the mean change in $R^*$ in time $\Delta t$ is $-aR^*\Delta t$ and the variance of the change is $\sigma^2\Delta t$. At node $(i, j)$, $R^* = j\,\Delta r$. If the branching has the form shown in Figure 28.7(a), the $p_u$, $p_m$, and $p_d$ at node $(i, j)$ must satisfy the following three equations:

$$p_u\Delta R - p_d\Delta R = -aj\,\Delta R\,\Delta t$$

$$p_u\Delta R^2 + p_d\Delta R^2 = \sigma^2\Delta t + a^2j^2\Delta R^2\Delta t^2$$

$$p_u + p_m + p_d = 1$$

Using $\Delta R = \sigma\sqrt{3\Delta t}$, the solution to these equations is

$$p_u = \tfrac{1}{6} + \tfrac{1}{2}(a^2j^2\Delta t^2 - aj\,\Delta t)$$

$$p_m = \tfrac{2}{3} - a^2j^2\Delta t^2$$

$$p_d = \tfrac{1}{6} + \tfrac{1}{2}(a^2j^2\Delta t^2 + aj\,\Delta t)$$

[15] The probabilities are positive for any value of $j_{max}$ between $0.184/(a\,\Delta t)$ and $0.816/(a\,\Delta t)$ and for any value of $j_{min}$ between $-0.184/(a\,\Delta t)$ and $-0.816/(a\,\Delta t)$. Changing the branching at the first possible node proves to be computationally most efficient.

Similarly, if the branching has the form shown in Figure 28.7(b), the probabilities are

$$p_u = \tfrac{1}{6} + \tfrac{1}{2}(a^2 j^2 \Delta t^2 + aj\,\Delta t)$$
$$p_m = -\tfrac{1}{3} - a^2 j^2 \Delta t^2 - 2aj\,\Delta t$$
$$p_d = \tfrac{7}{6} + \tfrac{1}{2}(a^2 j^2 \Delta t^2 + 3aj\,\Delta t)$$

Finally, if the branching has the form shown in Figure 28.7(c), the probabilities are

$$p_u = \tfrac{7}{6} + \tfrac{1}{2}(a^2 j^2 \Delta t^2 - 3aj\,\Delta t)$$
$$p_m = -\tfrac{1}{3} - a^2 j^2 \Delta t^2 + 2aj\,\Delta t$$
$$p_d = \tfrac{1}{6} + \tfrac{1}{2}(a^2 j^2 \Delta t^2 - aj\,\Delta t)$$

To illustrate the first stage of the tree construction, suppose that $\sigma = 0.01$, $a = 0.1$, and $\Delta t = 1$ year. In this case, $\Delta R = 0.01\sqrt{3} = 0.0173$, $j_{\max}$ is set equal to the smallest integer greater than 0.184/0.1, and $j_{\min} = -j_{\max}$. This means that $j_{\max} = 2$ and $j_{\min} = -2$ and the tree is as shown in Figure 28.8. The probabilities on the branches emanating from each node are shown below the tree and are calculated using the equations above for $p_u$, $p_m$, and $p_d$.

Note that the probabilities at each node in Figure 28.8 depend only on $j$. For example, the probabilities at node B are the same as the probabilities at node F. Furthermore, the tree is symmetrical. The probabilities at node D are the mirror image of the probabilities at node B.

## Second Stage

The second stage in the tree construction is to convert the tree for $R^*$ into a tree for $R$. This is accomplished by displacing the nodes on the $R^*$-tree so that the initial term structure of interest rates is exactly matched. Define

$$\alpha(t) = R(t) - R^*(t)$$

We calculate the $\alpha$'s iteratively so that the initial term structure is matched exactly.[16] Define $\alpha_i$ as $\alpha(i\,\Delta t)$, the value of $R$ at time $i\,\Delta t$ on the $R$-tree minus the corresponding value of $R^*$ at time $i\,\Delta t$ on the $r^*$-tree. Define $Q_{i,j}$ as the present value of a security that pays off \$1 if node $(i, j)$ is reached and zero otherwise. The $\alpha_i$ and $Q_{i,j}$ can be calculated using forward induction in such a way that the initial term structure is matched exactly.

## Illustration of Second Stage

Suppose that the continuously compounded zero rates in the example in Figure 28.8 are as shown in Table 28.1. The value of $Q_{0,0}$ is 1.0. The value of $\alpha_0$ is chosen to give the

[16] It is possible to estimate $\alpha(t)$ analytically. Since

$$dR = [\theta(t) - aR]\,dt + \sigma\,dz \quad \text{and} \quad dR^* = -aR^*\,dt + \sigma\,dz$$

it follows that

$$d\alpha = [\theta(t) - a\alpha(t)]\,dt$$

If we ignore the distinction between $r$ and $R$, the solution to this is

$$\alpha(t) = F(0, t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2$$

However, these are instantaneous $\alpha$'s and do not lead to the tree calculations exactly matching the term structure of interest rates.

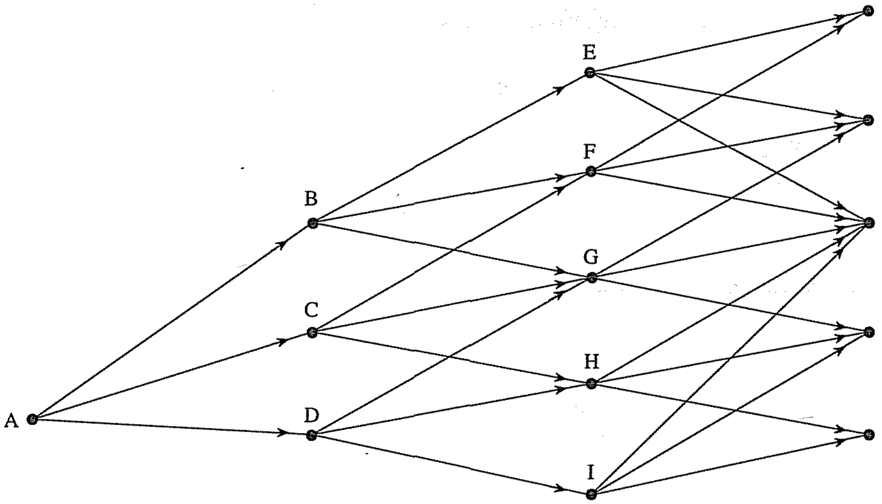
**Table 28.1** Zero rates for example in Figures 28.8 and 28.9.

| *Maturity* | *Rate (%)* |
|---|---|
| 0.5 | 3.430 |
| 1.0 | 3.824 |
| 1.5 | 4.183 |
| 2.0 | 4.512 |
| 2.5 | 4.812 |
| 3.0 | 5.086 |

right price for a zero-coupon bond maturing at time $\Delta t$. That is, $\alpha_0$ is set equal to the initial $\Delta t$-period interest rate. Because $\Delta t = 1$ in this example, $\alpha_0 = 0.03824$. This defines the position of the initial node on the $R$-tree in Figure 28.9. The next step is to calculate the values of $Q_{1,1}$, $Q_{1,0}$, and $Q_{1,-1}$. There is a probability of 0.1667 that the (1, 1) node is reached and the discount rate for the first time step is 3.82%. The value of $Q_{1,1}$ is therefore $0.1667e^{-0.0382} = 0.1604$. Similarly, $Q_{1,0} = 0.6417$ and $Q_{1,-1} = 0.1604$.

Once $Q_{1,1}$, $Q_{1,0}$, and $Q_{1,-1}$ have been calculated, we are in a position to determine $\alpha_1$. This is chosen to give the right price for a zero-coupon bond maturing at time $2\Delta t$. Because $\Delta R = 0.01732$ and $\Delta t = 1$, the price of this bond as seen at node B is $e^{-(\alpha_1+0.01732)}$. Similarly, the price as seen at node C is $e^{-\alpha_1}$ and the price as seen at

**Figure 28.9** Tree for $R$ in Hull–White model (the second stage).




| *Node*: | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| $R$(%) | 3.824 | 6.937 | 5.205 | 3.473 | 9.716 | 7.984 | 6.252 | 4.520 | 2.788 |
| $p_u$ | 0.1667 | 0.1217 | 0.1667 | 0.2217 | 0.8867 | 0.1217 | 0.1667 | 0.2217 | 0.0867 |
| $p_m$ | 0.6666 | 0.6566 | 0.6666 | 0.6566 | 0.0266 | 0.6566 | 0.6666 | 0.6566 | 0.0266 |
| $p_d$ | 0.1667 | 0.2217 | 0.1667 | 0.1217 | 0.0867 | 0.2217 | 0.1667 | 0.1217 | 0.8867 |

node D is $e^{-(\alpha_1-0.01732)}$. The price as seen at the initial node A is therefore

$$Q_{1,1}e^{-(\alpha_1+0.01732)} + Q_{1,0}e^{-\alpha_1} + Q_{1,-1}e^{-(\alpha_1-0.01732)} \tag{28.21}$$

From the initial term structure, this bond price should be $e^{-0.04512\times 2} = 0.9137$. Substituting for the $Q$'s in equation (28.21), we obtain

$$0.1604e^{-(\alpha_1+0.01732)} + 0.6417e^{-\alpha_1} + 0.1604e^{-(\alpha_1-0.01732)} = 0.9137$$

or

$$e^{\alpha_1}(0.1604e^{-0.01732} + 0.6417 + 0.1604e^{0.01732}) = 0.9137$$

or

$$\alpha_1 = \ln\left[\frac{0.1604e^{-0.01732} + 0.6417 + 0.1604e^{0.01732}}{0.9137}\right] = 0.05205$$

This means that the central node at time $\Delta t$ in the tree for $R$ corresponds to an interest rate of 5.205% (see Figure 28.9).

The next step is to calculate $Q_{2,2}$, $Q_{2,1}$, $Q_{2,0}$, $Q_{2,-1}$, and $Q_{2,-2}$. The calculations can be shortened by using previously determined $Q$ values. Consider $Q_{2,1}$ as an example. This is the value of a security that pays off \$1 if node F is reached and zero otherwise. Node F can be reached only from nodes B and C. The interest rates at these nodes are 6.937% and 5.205%, respectively. The probabilities associated with the B–F and C–F branches are 0.6566 and 0.1667. The value at node B of a security that pays \$1 at node F is therefore $0.6566e^{-0.06937}$. The value at node C is $0.1667e^{-0.05205}$. The variable $Q_{2,1}$ is $0.6566e^{-0.06937}$ times the present value of \$1 received at node B plus $0.1667e^{-0.05205}$ times the present value of \$1 received at node C; that is,

$$Q_{2,1} = 0.6566e^{-0.0693} \times 0.1604 + 0.1667e^{-0.05205} \times 0.6417 = 0.1998$$

Similarly, $Q_{2,2} = 0.0182$, $Q_{2,0} = 0.4736$, $Q_{2,-1} = 0.2033$, and $Q_{2,-2} = 0.0189$.

The next step in producing the $R$-tree in Figure 28.9 is to calculate $\alpha_2$. After that, the $Q_{3,j}$'s can then be computed. We can then calculate $\alpha_3$; and so on.

## Formulas for $\alpha$'s and $Q$'s

To express the approach more formally, we suppose that the $Q_{i,j}$ have been determined for $i \leqslant m$ ($m \geqslant 0$). The next step is to determine $\alpha_m$ so that the tree correctly prices a zero-coupon bond maturing at $(m+1)\,\Delta t$. The interest rate at node $(m, j)$ is $\alpha_m + j\,\Delta R$, so that the price of a zero-coupon bond maturing at time $(m+1)\Delta t$ is given by

$$P_{m+1} = \sum_{j=-n_m}^{n_m} Q_{m,j}\exp[-(\alpha_m + j\,\Delta R)\Delta t] \tag{28.22}$$

where $n_m$ is the number of nodes on each side of the central node at time $m\,\Delta t$. The solution to this equation is

$$\alpha_m = \frac{\ln\sum_{j=-n_m}^{n_m} Q_{m,j}e^{-j\Delta R\Delta t} - \ln P_{m+1}}{\Delta t}$$

Once $\alpha_m$ has been determined, the $Q_{i,j}$ for $i = m + 1$ can be calculated using

$$Q_{m+1,j} = \sum_k Q_{m,k} q(k, j) \exp[-(\alpha_m + k \, \Delta R) \, \Delta t]$$

where $q(k, j)$ is the probability of moving from node $(m, k)$ to node $(m + 1, j)$ and the summation is taken over all values of $k$ for which this is nonzero.

## Extension to Other Models

The procedure that has just been outlined can be extended to more general models of the form

$$df(r) = [\theta(t) - a f(r)] \, dt + \sigma \, dz \tag{28.23}$$

This family of models has the property that they can fit any term structure.[17]

As before, we assume that the $\Delta t$ period rate, $R$, follows the same process as $r$:

$$df(R) = [\theta(t) - a f(R)] \, dt + \sigma \, dz$$

We start by setting $x = f(R)$, so that

$$dx = [\theta(t) - ax] \, dt + \sigma \, dz$$

The first stage is to build a tree for a variable $x^*$ that follows the same process as $x$ except that $\theta(t) = 0$ and the initial value is zero. The procedure here is identical to the procedure already outlined for building a tree such as that in Figure 28.8.

As in Figure 28.9, we then displace the nodes at time $i \, \Delta t$ by an amount $\alpha_i$ to provide an exact fit to the initial term structure. The equations for determining $\alpha_i$ and $Q_{i,j}$ inductively are slightly different from those for the $f(R) = R$ case. The value of $Q$ at the first node, $Q_{0,0}$, is set equal to 1. Suppose that the $Q_{i,j}$ have been determined for $i \leqslant m$ $(m \geqslant 0)$. The next step is to determine $\alpha_m$ so that the tree correctly prices an $(m + 1)\Delta t$ zero-coupon bond. Define $g$ as the inverse function of $f$ so that the $\Delta t$-period interest rate at the $j$th node at time $m \, \Delta t$ is

$$g(\alpha_m + j \, \Delta x)$$

The price of a zero-coupon bond maturing at time $(m + 1)\Delta t$ is given by

$$P_{m+1} = \sum_{j=-n_m}^{n_m} Q_{m,j} \exp[-g(\alpha_m + j \, \Delta x)\Delta t] \tag{28.24}$$

This equation can be solved using a numerical procedure such as Newton–Raphson. The value $\alpha_0$ of $\alpha$ when $m = 0$, is $f\big(R(0)\big)$.

Once $\alpha_m$ has been determined, the $Q_{i,j}$ for $i = m + 1$ can be calculated using

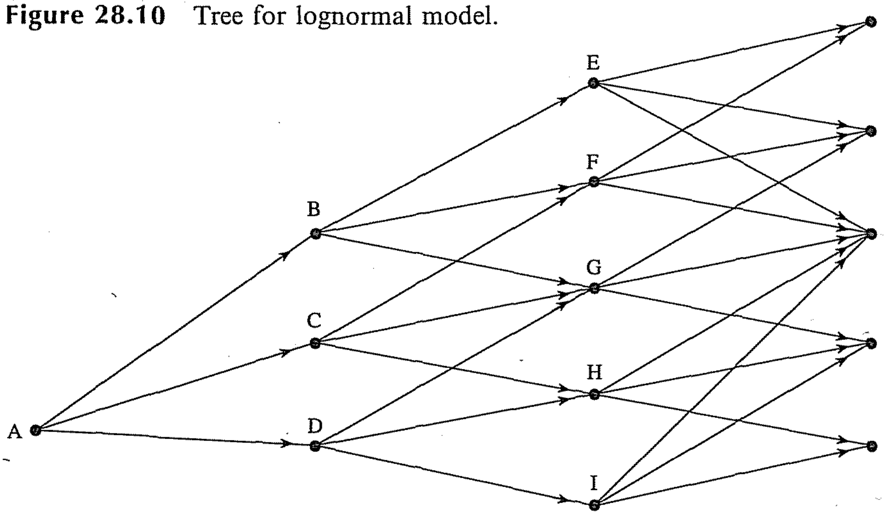
$$Q_{m+1,j} = \sum_k Q_{m,k} q(k, j) \exp[-g(\alpha_m + k \, \Delta x)\Delta t]$$

[17] Not all no-arbitrage models have this property. For example, the extended-CIR model, considered by Cox, Ingersoll, and Ross (1985) and Hull and White (1990), which has the form

$$dr = [\theta(t) - ar] \, dt + \sigma\sqrt{r} \, dz$$

cannot fit yield curves where the forward rate declines sharply. This is because the process is not well defined when $\theta(t)$ is negative.

**Figure 28.10** Tree for lognormal model.




| *Node:* | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| $x$ | −3.373 | −2.875 | −3.181 | −3.487 | −2.430 | −2.736 | −3.042 | −3.349 | −3.655 |
| $R(\%)$ | 3.430 | 5.642 | 4.154 | 3.058 | 8.803 | 6.481 | 4.772 | 3.513 | 2.587 |
| $p_u$ | 0.1667 | 0.1177 | 0.1667 | 0.2277 | 0.8609 | 0.1177 | 0.1667 | 0.2277 | 0.0809 |
| $p_m$ | 0.6666 | 0.6546 | 0.6666 | 0.6546 | 0.0582 | 0.6546 | 0.6666 | 0.6546 | 0.0582 |
| $p_d$ | 0.1667 | 0.2277 | 0.1667 | 0.1177 | 0.0809 | 0.2277 | 0.1667 | 0.1177 | 0.8609 |

where $q(k, j)$ is the probability of moving from node $(m, k)$ to node $(m+1, j)$ and the summation is taken over all values of $k$ where this is nonzero.

Figure 28.10 shows the results of applying the procedure to the model

$$d\ln(r) = [\theta(t) - a\ln(r)]\,dt + \sigma\,dz$$

when $a = 0.22$, $\sigma = 0.25$, $\Delta t = 0.5$, and the zero rates are as in Table 28.1.

## Choosing $f(r)$

When $f(r) = r$ we obtain the Hull–White model in equation (28.13); when $f(r) = \ln(r)$ we obtain the Black–Karasinksi model in equation (28.18). In most circumstances these two models appear to perform about equally well in fitting market data on actively traded instruments such as caps and European swap options. The main advantage of the $f(r) = r$ model is its analytic tractability. Its main disadvantage is that negative interest rates are possible. In most circumstances, the probability of negative interest rates occurring under the model is very small, but some analysts are reluctant to use a model where there is any chance at all of negative interest rates. The $f(r) = \ln r$ model has no analytic tractability, but has the advantage that interest rates are always positive. Another advantage is that traders naturally think in terms of $\sigma$'s arising from a lognormal model rather than $\sigma$'s arising from a normal model.

There is a problem in choosing a satisfactory model for countries with low interest rates. The normal model is unsatisfactory because, when the initial short rate is low, the probability of negative interest rates in the future is no longer negligible. The lognormal model is unsatisfactory because the volatility of rates (i.e., the $\sigma$ parameter in the lognormal model) is usually much greater when rates are low than when they are high. (For example, a volatility of 100% might be appropriate when the short rate is less than 1%, while 20% might be appropriate when it is 4% or more.) A model that appears to work well is one where $f(r)$ is chosen so that rates are lognormal for $r$ less than 1% and normal for $r$ greater than 1%.[18]

## Using Analytic Results in Conjunction with Trees

When a tree is constructed for the $f(r) = r$ version of the Hull–White model, the analytic results in Section 28.3 can be used to provide the complete term structure and European option prices at each node. It is important to recognize that the interest rate on the tree is the $\Delta t$-period rate $R$. It is not the instantaneous short rate $r$.

From equations (28.15), (28.16), and (28.17) it can be shown (see Problem 28.21) that

$$P(t, T) = \hat{A}(t, T)e^{-\hat{B}(t,T)R} \tag{28.25}$$

where

$$\ln \hat{A}(t, T) = \ln \frac{P(0, T)}{P(0, t)} - \frac{B(t, T)}{B(t, t + \Delta t)} \ln \frac{P(0, t + \Delta t)}{P(0, t)} - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t, T)[B(t, T) - B(t, t + \Delta t)] \tag{28.26}$$

and

$$\hat{B}(t, T) = \frac{B(t, T)}{B(t, t + \Delta t)}\Delta t \tag{28.27}$$

(In the case of the Ho–Lee model, we set $\hat{B}(t, T) = T - t$ in these equations.)

We should, therefore, calculate bond prices using equation (28.25), not equation (28.15).

***Example 28.1***

As an example, we use the zero rates in Table 28.2. The rates for maturities between those indicated are generated using linear interpolation.

We price a 3-year ($= 3 \times 365$ days) European put option on a zero-coupon bond that will expire in 9 years ($= 9 \times 365$ days). Interest rates are assumed to follow the Hull–White ($f(r) = r$) model. The strike price is 63, $a = 0.1$, and $\sigma = 0.01$. We construct a 3-year tree and calculate zero-coupon bond prices at the final nodes analytically as just described. As shown in Table 28.3, the results from the tree are consistent with the analytic price of the option.

This example provides a good test of the implementation of the model because the gradient of the zero curve changes sharply immediately after the expiration of the option. Small errors in the construction and use of the tree are liable to have a big effect on the option values obtained. (The example is used in Sample Application G of the DerivaGem Application Builder software.)

---

[18] See J. Hull and A. White "Taking Rates to the Limit," *Risk*, December (1997): 168–69.

**Table 28.2** Zero curve with all rates continuously compounded.

| *Maturity* | *Days* | *Rate* (%) |
|---|---|---|
| 3 days | 3 | 5.01772 |
| 1 month | 31 | 4.98284 |
| 2 months | 62 | 4.97234 |
| 3 months | 94 | 4.96157 |
| 6 months | 185 | 4.99058 |
| 1 year | 367 | 5.09389 |
| 2 years | 731 | 5.79733 |
| 3 years | 1,096 | 6.30595 |
| 4 years | 1,461 | 6.73464 |
| 5 years | 1,826 | 6.94816 |
| 6 years | 2,194 | 7.08807 |
| 7 years | 2,558 | 7.27527 |
| 8 years | 2,922 | 7.30852 |
| 9 years | 3,287 | 7.39790 |
| 10 years | 3,653 | 7.49015 |

## Tree for American Bond Options

The DerivaGem software accompanying this book implements the normal and the lognormal model, as well as Black's model, for valuing European bond options, caps/floors, and European swap options. In addition, American-style bond options can be handled. Figure 28.11 shows the tree produced by the software when it is used to value a 1.5-year American call option on a 10-year bond using four time steps and the lognormal model. The parameters used in the lognormal model are $a = 5\%$ and $\sigma = 20\%$. The underlying bond lasts 10 years, has a principal of 100, and pays a coupon of 5% per annum semiannually. The yield curve is flat at 5% per annum. The strike price is 105. As explained in Section 26.2 the strike price can be a cash strike price or a quoted strike price. In this case it is a quoted strike price. The bond price shown on the tree is the cash bond price. The accrued interest at each node is shown below the tree. The cash strike price is calculated as the quoted strike price plus accrued interest.

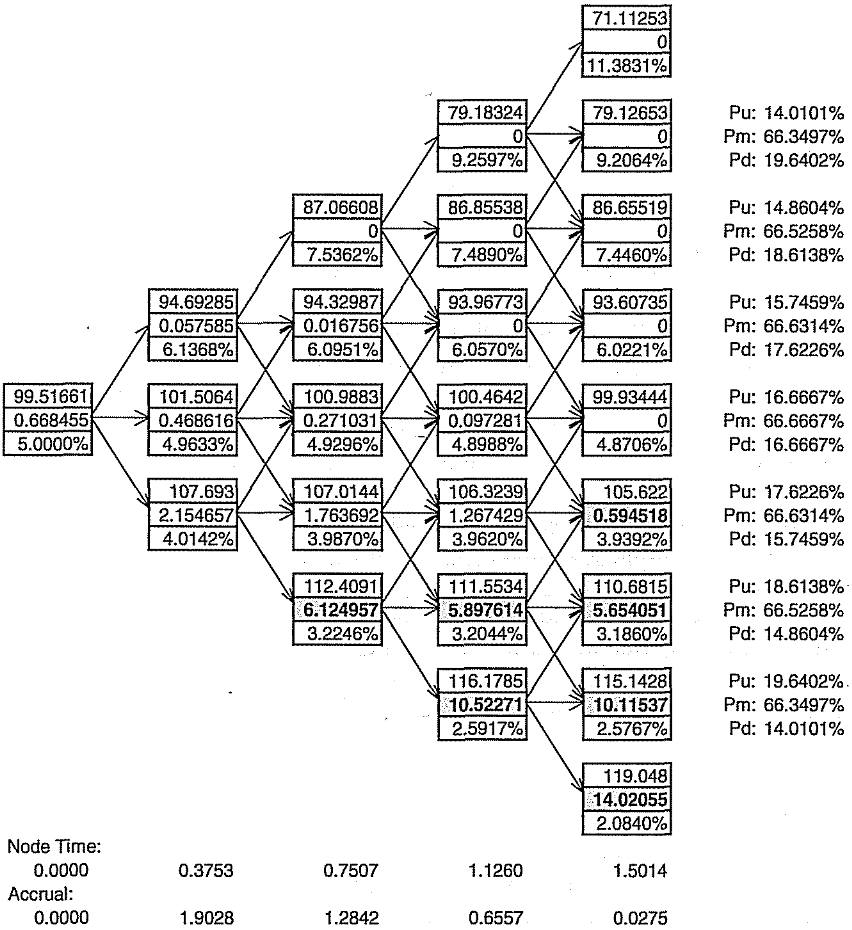
**Table 28.3** Value of a three-year put option on a nine-year zero-coupon bond with a strike price of 63: $a = 0.1$ and $\sigma = 0.01$; zero curve as in Table 28.2.

| *Steps* | *Tree* | *Analytic* |
|---|---|---|
| 10 | 1.8658 | 1.8093 |
| 30 | 1.8234 | 1.8093 |
| 50 | 1.8093 | 1.8093 |
| 100 | 1.8144 | 1.8093 |
| 200 | 1.8097 | 1.8093 |
| 500 | 1.8093 | 1.8093 |

**Figure 28.11** Tree, produced by DerivaGem, for valuing an American bond option.

At each node:
Upper value = Cash Bond Price
Middle value = Option Price
Lower value = dt-period Rate
Shaded values are as a result of early exercise

Strike price = 105
Time step, dt = 0.3753 years, 137.00 days




The quoted bond price is the cash bond price minus accrued interest. The payoff from the option is the cash bond price minus the cash strike price. Equivalently it is the quoted bond price minus the quoted strike price.

The tree gives the price of the option as 0.668. A much larger tree with 100 time steps gives the price of the option as 0.699. Two points should be noted about Figure 28.11:

1. The software measures the time to option maturity as a whole number of days. For example, when an option maturity of 1.5 years is input, the life of the option is assumed to be 1.5014 years (or 1 year and 183 days).

2. The price of the 10-year bond cannot be computed analytically when the lognormal model is assumed. It is computed numerically by rolling back through a much larger tree than that shown.

## 28.8 CALIBRATION

Up to now, we have assumed that the volatility parameters $a$ and $\sigma$ are known. We now discuss how they are determined. This is known as calibrating the model.

The volatility parameters are determined from market data on actively traded options (e.g., broker quotes on caps and swap options such as those in Tables 26.1 and 26.2). These will be referred to as the *calibrating instruments*. The first stage is to choose a "goodness-of-fit" measure. Suppose there are $n$ calibrating instruments. A popular goodness-of-fit measure is

$$\sum_{i=1}^{n}(U_i - V_i)^2$$

where $U_i$ is the market price of the $i$th calibrating instrument and $V_i$ is the price given by the model for this instrument. The objective of calibration is to choose the model parameters so that this goodness-of-fit measure is minimized.

If $a$ and $\sigma$ are constant, there are only two volatility parameters. The models can be extended so that $a$ or $\sigma$, or both, are functions of time. Step functions can be used. Suppose, for example, we elect to make $a$ constant and $\sigma$ a function of time. We might choose times $t_1, t_2, \dots, t_n$ and assume $\sigma(t) = \sigma_0$ for $t \leqslant t_1$, $\sigma(t) = \sigma_i$ for $t_i < t \leqslant t_{i+1}$ $(1 \leqslant i \leqslant n-1)$, and $\sigma(t) = \sigma_n$ for $t > t_n$. There would then be a total of $n+2$ volatility parameters: $a, \sigma_0, \sigma_1, \dots$, and $\sigma_n$. The number of volatility parameters should not be greater than the number of calibrating instruments.

The minimization of the goodness-of-fit measure can be accomplished using the Levenberg–Marquardt procedure.[19] When $a$ or $\sigma$, or both, are functions of time, a penalty function is often added to the goodness-of-fit measure to that the functions are "well behaved". In the example just mentioned, where $\sigma$ is a step function, we would choose the objective function as

$$\sum_{i=1}^{n}(U_i - V_i)^2 + \sum_{i=1}^{n} w_{1,i}(\sigma_i - \sigma_{i-1})^2 + \sum_{i=1}^{n-1} w_{2,i}(\sigma_{i-1} + \sigma_{i+1} - 2\sigma_i)^2$$

The second term provides a penalty for large changes in $\sigma$ between one step and the next. The third term provides a penalty for high curvature in $\sigma$. Appropriate values for $w_{1,i}$ and $w_{2,i}$ are based on experimentation and are chosen to provide an appropriate level of smoothness in the $\sigma$ function.

The calibrating instruments chosen should be as similar as possible to the instrument being valued. Suppose, for example, that we wish to value a Bermudan-style swap option that lasts 10 years and can be exercised on any payment date between year 5 and year 9 into a swap maturing 10 years from today. The most relevant calibrating instruments are $5 \times 5$, $6 \times 4$, $7 \times 3$, $8 \times 2$, and $9 \times 1$ European swap options. (An $n \times m$ European swap option is an $n$-year option to enter into a swap lasting for $m$ years beyond the maturity of the option.)

---

[19] For a good description of this procedure, see W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.

The advantage of making $a$ or $\sigma$, or both, functions of time is that the models can be fitted more precisely to the prices of instruments that trade actively in the market. The disadvantage is that the volatility structure becomes nonstationary. The volatility term structure given by the model in the future is liable to be quite different from that existing in the market today.[20]

A somewhat different approach to calibration is to use all available calibrating instruments to calculate "global-best-fit" $a$ and $\sigma$ parameters. The parameter $a$ is held fixed at its best-fit value. The model can then be used in the same way as Black–Scholes. There is a one-to-one relationship between options prices and the $\sigma$ parameter. The model can be used to convert tables such as Table 26.1 and 26.2 into tables of implied $\sigma$'s.[21] These tables can be used to assess the $\sigma$ most appropriate for pricing the instrument under consideration.

## 28.9 HEDGING USING A ONE-FACTOR MODEL

We outlined some general approaches to hedging a portfolio of interest rate derivatives in Section 26.6. They can be used with the term structure models discussed in this chapter. The calculation of deltas, gammas, and vegas involves making small changes to either the zero curve or the volatility environment and recomputing the value of the portfolio.

Note that, although we often assume there is one factor when pricing interest rate derivatives, we do not assume only one factor when hedging. For example, the deltas we calculate allow for many different movements in the yield curve, not just those that are possible under the model chosen. The practice of taking account of changes that cannot happen under the model considered, as well as those that can, is known as *outside model hedging* and is standard practice for traders.[22] The reality is that relatively simple one-factor models if used carefully usually give reasonable prices for instruments, but good hedging schemes must explicitly or implicitly assume many factors.

## SUMMARY

The traditional models of the term structure used in finance are known as equilibrium models. These are useful for understanding potential relationships between variables in the economy, but have the disadvantage that the initial term structure is an output from the model rather than an input to it. When valuing derivatives, it is important that the model used be consistent with the initial term structure observed in the market. No-arbitrage models are designed to have this property. They take the initial term structure as given and define how it can evolve.

---

[20] For a discussion of the implementation of a model where $a$ and $\sigma$ are functions of time, see Technical Note 16 on the author's website.

[21] Note that in a term structure model the implied $\sigma$'s are not the same as the implied volatilities calculated from Black's model in Tables 26.1 and 26.2. The procedure for computing implied $\sigma$'s is as follows. The Black volatilities are converted to prices using Black's model. An iterative procedure is then used to imply the $\sigma$ parameter in the term structure model from the price.

[22] A simple example of outside model hedging is in the way that the Black–Scholes model is used. The Black–Scholes model assumes that volatility is constant—but traders regularly calculate vega and hedge against volatility changes.

This chapter has provided a description of a number of one-factor no-arbitrage models of the short rate. These are very robust and can be used in conjunction with any set of initial zero rates. The simplest model is the Ho–Lee model. This has the advantage that it is analytically tractable. Its chief disadvantage is that it implies that all rates are equally variable at all times. The Hull–White model is a version of the Ho–Lee model that includes mean reversion. It allows a richer description of the volatility environment while preserving its analytic tractability. Lognormal one-factor models have the advantage that they avoid the possibility of negative interest rates but, unfortunately, they have no analytic tractability.

## FURTHER READING

### *Equilibrium Models*

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

Longstaff, F. A. and E. S. Schwartz, "Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model," *Journal of Finance*, 47, 4 (September 1992): 1259–82.

Vasicek, O. A.,"An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5 (1977): 177–88.

### *No-Arbitrage Models*

Black, F., and P. Karasinski, "Bond and Option Pricing When Short Rates Are Lognormal," *Financial Analysts Journal*, July/August (1991): 52–59.

Ho, T. S. Y., and S.-B. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, 41 (December 1986): 1011–29.

Hull, J., and A. White, "Bond Option Pricing Based on a Model for the Evolution of Bond Prices," *Advances in Futures and Options Research*, 6 (1993): 1–13.

Hull, J., and A. White, "Pricing Interest Rate Derivative Securities," *The Review of Financial Studies*, 3, 4 (1990): 573–92.

Hull, J., and A. White, "Using Hull–White Interest Rate Trees," *Journal of Derivatives*, Spring (1996): 26–36.

Kijima, M., and I. Nagayama, "Efficient Numerical Procedures for the Hull–White Extended Vasicek Model," *Journal of Financial Engineering*, 3 (September/December 1994): 275–92.

Kijima, M., and I. Nagayama, "A Numerical Procedure for the General One-Factor Interest rate Model" *Journal of Financial Engineering*, 5 (December 1996): 317–37.

Li, A., P. Ritchken, and L. Sankarasubramanian, "Lattice Models for Pricing American Interest Rate Claims," *Journal of Finance*, 50, 2 (June 1995): 719–37.

Rebonato, R., *Interest Rate Option Models*. Chichester: Wiley, 1998.

## Questions and Problems (Answers in Solutions Manual)

28.1. What is the difference between an equilibrium model and a no-arbitrage model?

28.2. Suppose that the short rate is currently 4% and its standard deviation is 1% per annum. What happens to the standard deviation when the short rate increases to 8% in (a) Vasicek's model; (b) Rendleman and Bartter's model; and (c) the Cox, Ingersoll, and Ross model?

28.3. If a stock price were mean reverting or followed a path-dependent process there would be market inefficiency. Why is there not a market inefficiency when the short-term interest rate does so?

28.4. Explain the difference between a one-factor and a two-factor interest rate model.

28.5. Can the approach described in Section 28.4 for decomposing an option on a coupon-bearing bond into a portfolio of options on zero-coupon bonds be used in conjunction with a two-factor model? Explain your answer.

28.6. Suppose that $a = 0.1$ and $b = 0.1$ in both the Vasicek and the Cox, Ingersoll, Ross model. In both models, the initial short rate is 10% and the initial standard deviation of the short-rate change in a short time $\Delta t$ is $0.02\sqrt{\Delta t}$. Compare the prices given by the models for a zero-coupon bond that matures in year 10.

28.7. Suppose that $a = 0.1$, $b = 0.08$, and $\sigma = 0.015$ in Vasicek's model, with the initial value of the short rate being 5%. Calculate the price of a 1-year European call option on a zero-coupon bond with a principal of \$100 that matures in 3 years when the strike price is \$87.

28.8. Repeat Problem 23.7 valuing a European put option with a strike of \$87. What is the put–call parity relationship between the prices of European call and put options? Show that the put and call option prices satisfy put–call parity in this case.

28.9. Suppose that $a = 0.05$, $b = 0.08$, and $\sigma = 0.015$ in Vasicek's model with the initial short-term interest rate being 6%. Calculate the price of a 2.1-year European call option on a bond that will mature in 3 years. Suppose that the bond pays a coupon of 5% semiannually. The principal of the bond is 100 and the strike price of the option is 99. The strike price is the cash price (not the quoted price) that will be paid for the bond.

28.10. Use the answer to Problem 23.9 and put–call parity arguments to calculate the price of a put option that has the same terms as the call option in Problem 23.9.

28.11. In the Hull–White model, $a = 0.08$ and $\sigma = 0.01$. Calculate the price of a 1-year European call option on a zero-coupon bond that will mature in 5 years when the term structure is flat at 10%, the principal of the bond is \$100, and the strike price is \$68.

28.12. Suppose that $a = 0.05$ and $\sigma = 0.015$ in the Hull–White model with the initial term structure being flat at 6% with semiannual compounding. Calculate the price of a 2.1-year European call option on a bond that will mature in 3 years. Suppose that the bond pays a coupon of 5% per annum semiannually. The principal of the bond is 100 and the strike price of the option is 99. The strike price is the cash price (not the quoted price) that will be paid for the bond.

28.13. Use a change of numeraire argument to show that the relationship between the futures rate and forward rate for the Ho–Lee model is as shown in Section 6.4. Use the relationship to verify the expression for $\theta(t)$ given for the Ho–Lee model in equation (28.11). (*Hint*: The futures price is a martingale when the market price of risk is zero. The forward price is a martingale when the market price of risk is a zero-coupon bond maturing at the same time as the forward contract.)

28.14. Use a similar approach to that in Problem 28.13 to derive the relationship between the futures rate and the forward rate for the Hull–White model. Use the relationship to verify the expression for $\theta(t)$ given for the Hull–White model in equation (28.14).

28.15. Suppose $a = 0.05$, $\sigma = 0.015$, and the term structure is flat at 10%. Construct a trinomial tree for the Hull–White model where there are two time steps, each 1 year in length.

28.16. Calculate the price of a 2-year zero-coupon bond from the tree in Figure 28.6.

28.17. Calculate the price of a 2-year zero-coupon bond from the tree in Figure 28.9 and verify that it agrees with the initial term structure.

28.18. Calculate the price of an 18-month zero-coupon bond from the tree in Figure 28.10 and verify that it agrees with the initial term structure.

28.19. What does the calibration of a one-factor term structure model involve?

28.20. Use the DerivaGem software to value $1 \times 4$, $2 \times 3$, $3 \times 2$, and $4 \times 1$ European swap options to receive fixed and pay floating. Assume that the 1-, 2-, 3-, 4-, and 5-year interest rates are 6%, 5.5%, 6%, 6.5%, and 7%, respectively. The payment frequency on the swap is semiannual and the fixed rate is 6% per annum with semiannual compounding. Use the Hull–White model with $a = 3\%$ and $\sigma = 1\%$. Calculate the volatility implied by Black's model for each option.

28.21. Prove equations (28.25), (28.26), and (28.27).

## Assignment Questions

28.22. Construct a trinomial tree for the Ho–Lee model where $\sigma = 0.02$. Suppose that the the initial zero-coupon interest rate for a maturities of 0.5, 1.0, and 1.5 years are 7.5%, 8%, and 8.5%. Use two time steps, each 6 months long. Calculate the value of a zero-coupon bond with a face value of \$100 and a remaining life of 6 months at the ends of the final nodes of the tree. Use the tree to value a 1-year European put option with a strike price of 95 on the bond. Compare the price given by your tree with the analytic price given by DerivaGem.

28.23. A trader wishes to compute the price of a 1-year American call option on a 5-year bond with a face value of 100. The bond pays a coupon of 6% semiannually and the (quoted) strike price of the option is \$100. The continuously compounded zero rates for maturities of 6 months, 1 year, 2 years, 3 years, 4 years, and 5 years are 4.5%, 5%, 5.5%, 5.8%, 6.1%, and 6.3%. The best-fit reversion rate for either the normal or the lognormal model has been estimated as 5%.

A 1-year European call option with a (quoted) strike price of 100 on the bond is actively traded. Its market price is \$0.50. The trader decides to use this option for calibration. Use the DerivaGem software with 10 time steps to answer the following questions:

(a) Assuming a normal model, imply the $\sigma$ parameter from the price of the European option.
(b) Use the $\sigma$ parameter to calculate the price of the option when it is American.
(c) Repeat (a) and (b) for the lognormal model. Show that the model used does not significantly affect the price obtained providing it is calibrated to the known European price.
(d) Display the tree for the normal model and calculate the probability of a negative interest rate occurring.
(e) Display the tree for the lognormal model and verify that the option price is correctly calculated at the node where, with the notation of Section 28.7, $i = 9$ and $j = -1$.

28.24. Use the DerivaGem software to value $1 \times 4$, $2 \times 3$, $3 \times 2$, and $4 \times 1$ European swap options to receive floating and pay fixed. Assume that the 1-, 2-, 3-, 3-, and 5-year interest rates are 3%, 3.5%, 3.8%, 4.0%, and 4.1%, respectively. The payment frequency on the swap is semiannual and the fixed rate is 4% per annum with semiannual compounding. Use the lognormal model with $a = 5\%$, $\sigma = 15\%$, and 50 time steps. Calculate the volatility implied by Black's model for each option.

28.25. Verify that the DerivaGem software gives Figure 28.11 for the example considered. Use the software to calculate the price of the American bond option for the lognormal and normal models when the strike price is 95, 100, and 105. In the case of the normal model, assume that $a = 5\%$ and $\sigma = 1\%$. Discuss the results in the context of the heaviness of the tails arguments of Chapter 16.

28.26. Modify Sample Application G in the DerivaGem Application Builder software to test the convergence of the price of the trinomial tree when it is used to price a 2-year call option on a 5-year bond with a face value of 100. Suppose that the strike price (quoted) is 100, the coupon rate is 7% with coupons being paid twice a year. Assume that the zero curve is as in Table 28.2. Compare results for the following cases:
(a) Option is European; normal model with $\sigma = 0.01$ and $a = 0.05$
(b) Option is European; lognormal model with $\sigma = 0.15$ and $a = 0.05$
(c) Option is American; normal model with $\sigma = 0.01$ and $a = 0.05$
(d) Option is American; lognormal model with $\sigma = 0.15$ and $a = 0.05$