CHAPTER
TWENTY-FIVE

Equilibrium Term Structure Models

8. What's your problem? Any moron can understand bond pricing models.

Top Ten Lies Finance Professors Tell Their Students[1]

Many interest rate models have been proposed in the literature and used in practice. This chapter surveys equilibrium models, and the next chapter covers no-arbitrage models. Because the spot rates satisfy

$$r(t, T) = -\frac{\ln P(t, T)}{T - t},$$

the discount function $P(t, T)$ suffices to establish the spot rate curve. Most models to follow are short rate models, in which the short rate is the sole source of uncertainty. Unless stated otherwise, the market price of risk λ is assumed to be zero; the processes are hence risk-neutral to start with.

25.1 The Vasicek Model

Vasicek proposed the model in which the short rate follows [855]

$$dr = \beta(\mu - r)\, dt + \sigma\, dW.$$

The short rate is thus pulled to the long-term mean level μ at rate β. Superimposed on this "pull" is a normally distributed stochastic term $\sigma\, dW$. The idea of mean reversion for interest rates dates back to Keynes [232]. This model seems relevant to interest rates in Germany and the United Kingdom [248].

Because the process is an Ornstein–Uhlenbeck process,

$$E[\, r(T) \mid r(t) = r\,] = \mu + (r - \mu)\, e^{-\beta(T-t)}$$

from Eq. (14.13). The term structure equation under the Vasicek model is

$$-\frac{\partial P}{\partial T} + \beta(\mu - r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} = rP.$$

The price of a zero-coupon bond paying one dollar at maturity can be shown to be

$$P(t, T) = A(t, T)\, e^{-B(t,T)\, r(t)}, \tag{25.1}$$

https://doi.org/10.1017/CBO9780511546839.026 Published online by Cambridge University Press

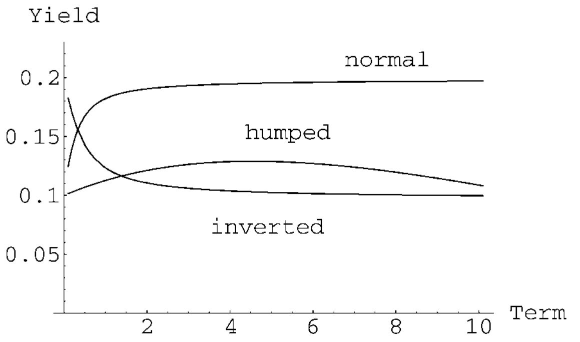

Figure 25.1: Term structure shapes. The parameters (β, μ, σ, r) are $(5.9, 0.2, 0.3, 0.1)$, $(3.9, 0.1, 0.3, 0.2)$, and $(0.1, 0.4, 0.11, 0.1)$ for normal, inverted, and humped term structures, respectively.

where

$$A(t,T) = \begin{cases} \exp\left[\frac{\{B(t,T)-T+t\}(\beta^2\mu-\sigma^2/2)}{\beta^2} - \frac{\sigma^2 B(t,T)^2}{4\beta}\right], & \text{if } \beta \neq 0 \\ \exp\left[\frac{\sigma^2(T-t)^3}{6}\right], & \text{if } \beta = 0 \end{cases},$$

$$B(t,T) = \begin{cases} \frac{1-e^{-\beta(T-t)}}{\beta}, & \text{if } \beta \neq 0 \\ T-t, & \text{if } \beta = 0 \end{cases}.$$

This model has some unpleasant properties; for example, if $\beta = 0$, then P goes to infinity as $T \to \infty$, like the Merton model. However, sensibly, P goes to zero as $T \to \infty$ if $\beta \neq 0$. Even so, P may exceed one for a finite T. See Fig. 25.1 for the shapes of the spot rate curve. The spot rate volatility structure is the curve $[\,\partial r(t,T)/\partial r\,]\sigma = \sigma B(t,T)/(T-t)$. When $\beta > 0$, the curve tends to decline with maturity. The speed of mean reversion, β, controls the shape of the curve; indeed, higher β leads to greater attenuation of volatility with maturity. It is not hard to verify that duration $-\frac{\partial P(t,T)/\partial r}{P(t,T)}$ equals $B(t,T)$. Duration decreases toward $1/\beta$ as the term lengthens if there is mean reversion ($\beta > 0$). On the other hand, duration equals the term to maturity $T-t$ if there is no mean reversion ($\beta = 0$), much like the static world. Interestingly, duration is independent of the interest rate volatility σ.

➤ **Exercise 25.1.1** Connect the Vasicek model with the AR(1) process.

➤ **Exercise 25.1.2** (1) Show that the long rate is $\mu - \sigma^2/(2\beta^2)$, independent of the current short rate. (2) Derive the liquidity premium for the $\beta \neq 0$ case.

➤ **Exercise 25.1.3** Show that Eq. (25.1) satisfies the term structure equation.

➤ **Exercise 25.1.4** Verify that $dP/P = r\,dt - B(t,T)\,\sigma\,dW$ is the bond price process for the Vasicek model.

➤ **Exercise 25.1.5** Show that the Ito process for the instantaneous forward rate $f(t,T)$ under the Vasicek model with $\beta \neq 0$ is

$$df = \frac{\sigma^2}{\beta}\, e^{-\beta(T-t)}\left[1 - e^{-\beta(T-t)}\right]dt + \sigma e^{-\beta(T-t)}\, dW.$$

(Hint: Section 24.5 and Exercise 25.1.4.)

25.1.1 Options on Zero-Coupon Bonds

Consider a European call with strike price X expiring at time T on a zero-coupon bond with par value \$1 and maturing at time $s > T$. Its price is given by the following Black–Scholes-like formula [506]:

$$P(t, s)\, N(x) - XP(t, T)\, N(x - \sigma_v),$$

where

$$\begin{aligned} x &\equiv \frac{1}{\sigma_v} \ln\left(\frac{P(t, s)}{P(t, T)\, X}\right) + \frac{\sigma_v}{2}, \\ \sigma_v &\equiv v(t, T)\, B(T, s), \\ v(t, T)^2 &\equiv \begin{cases} \frac{\sigma^2\left[1-e^{-2\beta(T-t)}\right]}{2\beta}, & \text{if } \beta \neq 0 \\ \sigma^2(T-t), & \text{if } \beta = 0 \end{cases}. \end{aligned}$$

Note that $v(t, T)^2$ is the variance of $r(t, T)$ by Eq. (14.14). The put–call parity says that

$$\text{call} = \text{put} + P(t, s) - P(t, T)\, X.$$

The price of a European put is thus $XP(t, T)\, N(-x + \sigma_v) - P(t, s)\, N(-x)$.

➤ **Exercise 25.1.6** Verify that the variance of $\ln P(t, T)$ is σ_v^2.

25.1.2 Binomial Approximation

We consider a binomial model for the short rate in the time interval $[0, T]$ divided into n identical pieces. Let $\Delta t \equiv T/n$ and

$$p(r) \equiv \frac{1}{2} + \frac{\beta(\mu - r)\sqrt{\Delta t}}{2\sigma}.$$

The following binomial model converges in distribution to the Vasicek model [696]:

$$r(k+1) = r(k) + \sigma\sqrt{\Delta t}\; \xi(k), \quad 0 \leq k < n,$$

where $\xi(k) = \pm 1$, with

$$\text{Prob}[\,\xi(k) = 1\,] = \begin{cases} p[\,r(k)\,], & \text{if } 0 \leq p(r(k)) \leq 1 \\ 0, & \text{if } p(r(k)) < 0 \\ 1, & \text{if } 1 < p(r(k)) \end{cases}.$$

Observe that the probability of an up move, p, is a decreasing function of the interest rate r. This is consistent with mean reversion.

The rate is the same whether it is the result of an up move followed by a down move or a down move followed by an up move; in other words, the binomial tree combines. The key feature of the model that makes it happen is its *constant* volatility, σ. For a general process Y with nonconstant volatility, the resulting binomial tree may not combine. Fortunately, if Y can be transformed into one with constant volatility, say X, then we can first construct a combining tree for X and then apply the inverse transformation on each node to obtain a combining tree for Y. This idea will be explored in Subsection 25.2.2.

➤ **Exercise 25.1.7** Prove that

$$\frac{E[\,r(k+1)-r(k)\,]}{\Delta t}=\begin{cases}\beta[\,\mu-r(k)\,], & \text{if } 0\le p(r(k))\le 1\\ \sigma/\sqrt{\Delta t}, & \text{if } p(r(k))<0\\ -\sigma/\sqrt{\Delta t}, & \text{if } 1<p(r(k))\end{cases}$$

and $\text{Var}[\,r(k+1)-r(k)\,]\to\sigma^2\Delta t$.

➤ **Exercise 25.1.8** Show that discretizing the Vasicek model directly by (14.7) does not result in a combining binomial tree.

➢ **Programming Assignment 25.1.8** Write a program to implement the binomial tree for the Vasicek model. Price zero-coupon bonds and compare the results against Eq. (25.1).

25.2 The Cox-Ingersoll-Ross Model

Cox, Ingersoll, and Ross (CIR) proposed the following square-root short-rate model [234]:

$$dr=\beta(\mu-r)\,dt+\sigma\sqrt{r}\,dW. \tag{25.2}$$

Although the randomly moving interest rate is elastically pulled toward the long-term value μ, as in the Vasicek model, the diffusion differs by a multiplicative factor $\sqrt{r}$. The parameter β determines the speed of adjustment. The short rate can reach zero only if $2\beta\mu<\sigma^2$.

The price of a zero-coupon bond paying \$1 at maturity is [470]

$$P(t,T)=A(t,T)\,e^{-B(t,T)\,r(t)}, \tag{25.3}$$

where

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

$$B(t,T)=\frac{2\left[\,e^{\gamma(T-t)}-1\,\right]}{(\beta+\gamma)\left[\,e^{\gamma(T-t)}-1\,\right]+2\gamma},$$

$$\gamma=\sqrt{\beta^2+2\sigma^2}.$$

A formula for consols is also available [266]. Figure 25.2 illustrates the shapes of the spot rate curve. In general, the curve is normal if the current short rate $r(t)$ is below the long rate $r(t,\infty)$, becomes inverted if $r(t)>\mu$, and is slightly humped if $r(t,\infty)<r(t)<\mu$ [493]. Figure 25.3 illustrates the long rate and duration of zero-coupon bonds. To incorporate the market price of risk into bond prices, replace each occurrence of β in $A(t,T)$ (except the exponent $2\beta\mu/\sigma^2$), $B(t,T)$, and γ with $\beta+\lambda$. Consult Subsection 14.3.2 for additional properties of the square-root process.

Two implications of the CIR model are at odds with empirical evidence: constant long rate and perfect correlation in yield changes along the term structure. The CIR model has been subject to many empirical studies (e.g., [11, 135, 138, 173, 257, 384, 835]). The model seemed to fit the term structure of *real* interest rates in the United Kingdom well until the end of 1992 [137]. (Real rates have been generally less volatile than nominal rates.) It also seems relevant to Denmark and Sweden [248].

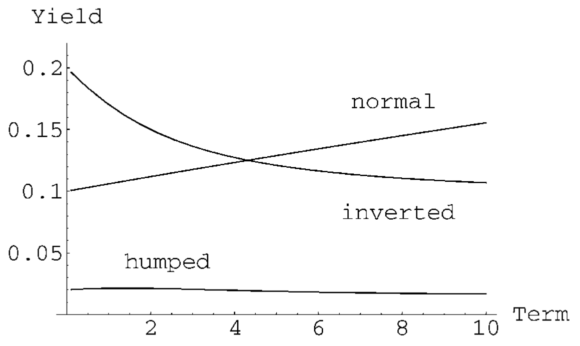

Figure 25.2: Term structure shapes. The values for the parameters (β, μ, σ, r) are (0.02, 0.7, 0.02, 0.1), (0.7, 0.1, 0.3, 0.2), and (0.06, 0.09, 0.5, 0.02) for normal, inverted, and humped term structures, respectively. The long rates are 0.512436, 0.0921941, and 0.0140324, respectively.

➤ **Exercise 25.2.1** Show that the long rate is $2\beta\mu/(\beta+\gamma)$, independent of the short rate.

➤ **Exercise 25.2.2** Show that Eq. (25.3) satisfies the term structure equation.

➤ **Exercise 25.2.3** Verify that $dP/P = r\,dt - B(t, T)\,\sigma\sqrt{r}\,dW$ is the bond price process for the CIR model.

➤ **Exercise 25.2.4 (Affine Models)** For any short rate model $dr = \mu(r, t)\,dt + \sigma(r, t)\,dW$ that produces zero-coupon bond prices of the form $P(t, T) = A(t, T)\,e^{-B(t,T)\,r(t)}$, show that the spot rate volatility structure is the curve $\sigma(r, t)\,B(t, T)/(T-t)$.

➤ **Exercise 25.2.5** (1) Write the bond price formula in terms of $\phi_1 \equiv \gamma$, $\phi_2 \equiv (\beta+\gamma)/2$, and $\phi_3 \equiv 2\beta\mu/\sigma^2$. (2) How do we estimate σ, given estimates for ϕ_1, ϕ_2, and ϕ_3?

➤ **Exercise 25.2.6** Consider a yield curve option with payoff $\max(0, r(T, T_1) - r(T, T_2))$ at expiration T, where $T < T_1$ and $T < T_2$. The security is based on the yield spread of two different maturities, $T_1 - T$ and $T_2 - T$. Assume either the Vasicek or the CIR model. Show that this option is equivalent to a portfolio of caplets on the $(T_2 - T)$-year spot rate.

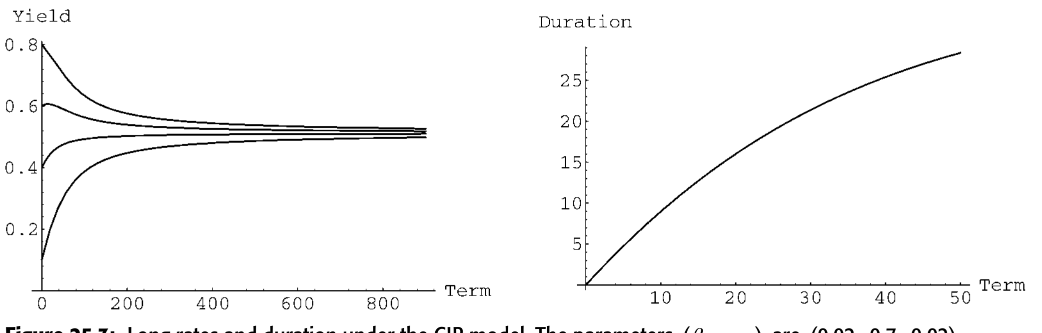

Figure 25.3: Long rates and duration under the CIR model. The parameters (β, μ, σ) are (0.02, 0.7, 0.02). The long-rate plot uses 0.8, 0.6, 0.4, and 0.1 as the initial rates. The duration of zero-coupon bonds uses $r = 0.1$,

$$
\begin{array}{cccccc}
 & & & & & \nearrow \\
 & & & & x+2\sqrt{\Delta t} & \\
 & & & \nearrow & & \searrow \\
 & & x+\sqrt{\Delta t} & & & \\
 & \nearrow & & \searrow & & \nearrow \\
x & & & & x & \\
 & \searrow & & \nearrow & & \searrow \\
 & & x-\sqrt{\Delta t} & & & \\
 & & & \searrow & & \nearrow \\
 & & & & x-2\sqrt{\Delta t} & \\
 & & & & & \searrow
\end{array}
\qquad
\begin{array}{cccccc}
 & & & & & \nearrow \\
 & & & & f(x+2\sqrt{\Delta t}) & \\
 & & & \nearrow & & \searrow \\
 & & f(x+\sqrt{\Delta t}) & & & \\
 & \nearrow & & \searrow & & \nearrow \\
f(x) & & & & f(x) & \\
 & \searrow & & \nearrow & & \searrow \\
 & & f(x-\sqrt{\Delta t}) & & & \\
 & & & \searrow & & \nearrow \\
 & & & & f(x-2\sqrt{\Delta t}) & \\
 & & & & & \searrow
\end{array}
$$

Figure 25.4: Binomial tree for the CIR model.

➢ **Programming Assignment 25.2.7** Implement the implicit method in Exercise 24.4.3 for zero-coupon bonds under the CIR model.

25.2.1 Binomial Approximation

Suppose we want to approximate the short rate process in the time interval $[0, T]$ divided into n periods of duration $\Delta t \equiv T/n$. Assume that $\mu, \beta \geq 0$. A direct discretization of the process is problematic because the resulting binomial tree will *not* combine (see Exercise 25.2.14, part (1)). Instead, consider the transformed process $x(r) \equiv 2\sqrt{r}/\sigma$. It follows

$$dx = m(x)\,dt + dW,$$

where $m(x) \equiv 2\beta\mu/(\sigma^2 x) - (\beta x/2) - 1/(2x)$. Because this new process has a constant volatility, its associated binomial tree combines.

The combining tree for r can be constructed as follows. First, construct a tree for x. Then transform each node of the tree into one for r by means of the inverse transformation $r = f(x) \equiv x^2\sigma^2/4$ (see Fig. 25.4). The probability of an up move at each node r is

$$p(r) \equiv \frac{\beta(\mu - r)\,\Delta t + r - r^-}{r^+ - r^-}, \tag{25.4}$$

where $r^+ \equiv f(x + \sqrt{\Delta t})$ denotes the result of an up move from r and $r^- \equiv f(x - \sqrt{\Delta t})$ the result of a down move [268, 696, 746]. Finally, set the probability $p(r)$ to one as r goes to zero to make the probability stay between zero and one. See Fig. 25.6 for the algorithm.

For a concrete example, consider the process

$$0.2\,(0.04 - r)\,dt + 0.1\sqrt{r}\,dW$$

for the time interval $[0, 1]$ given the initial rate $r(0) = 0.04$. We use $\Delta t = 0.2$ (year) for the binomial approximation. Figure 25.5(a) shows the resulting binomial short rate tree with the up-move probabilities in parentheses. To give an idea how these numbers come into being, consider the node that is the result of an up move from the root. Because the root has $x = 2\sqrt{r(0)}/\sigma = 4$, this particular node's x value equals $4 + \sqrt{\Delta t} = 4.4472135955$. Now use the inverse transformation to obtain the short rate $x^2 \times (0.1)^2/4 \approx 0.0494442719102$. Other short rates can be similarly obtained.

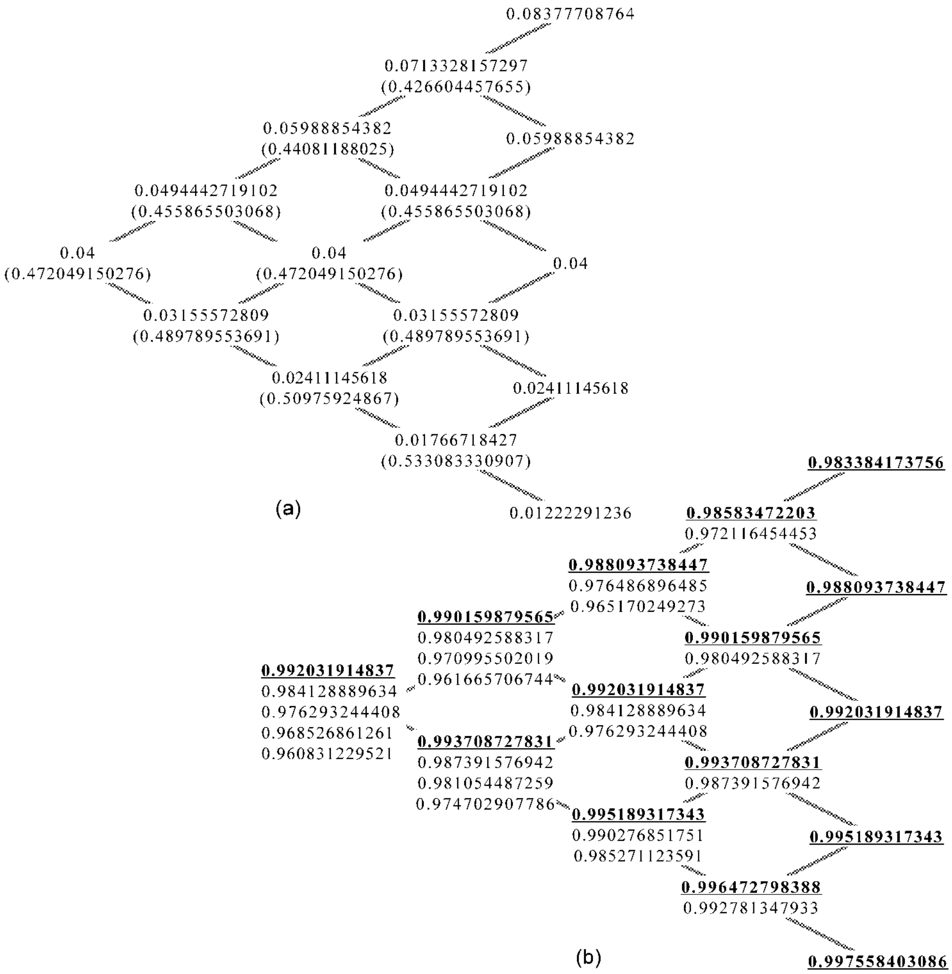

Figure 25.5: Short rate and bond price trees for the CIR model.

Once the short rates are in place, computing the probabilities is easy. Note that the up-move probability decreases as interest rates increase and decreases as interest rates decline. This phenomenon agrees with mean reversion.

➤ **Exercise 25.2.8** Derive $E[\, r(k+1) - r(k)\,]$ and $\text{Var}[\, r(k+1) - r(k)\,]$.

➤ **Exercise 25.2.9** Show that $p(r) = (1/2) + (1/2)\, m(x(r))\sqrt{\Delta t}$.

➢ **Programming Assignment 25.2.10** Write a program to implement the binomial short rate tree and the bond price tree for the CIR model. Compare the results against Eq. (25.3).

Term Structure Dynamics

The tree of short rates can be used to calculate the one-period bond prices (the underlined numbers in Fig. 25.5(b)). For example, the rate after two up moves, 0.05988854382, gives rise to

$$e^{-0.05988854382 \times 0.2} = 0.988093738447.$$

Binomial CIR model for zero-coupon bonds:

```
input:    τ, r, β, μ, σ, n;
real      r⁺, r⁻, r', x, x', p, Δt, Δx, P[n+1];
integer   i, j;
Δt := τ/n;
x := 2√r/σ;
Δx := √Δt;
for (i = 0 to n) P[i] := 1;
for (j = n−1 down to 0)
          for (i = 0 to j) {
                x' := x + (j − 2i) Δx;
                r' := x'²σ²/4;
                r⁺ := (x' + Δx)²σ²/4;
                r⁻ := (x' − Δx)²σ²/4;
                if [r' = 0] p := 1;
                else p := (β(μ − r') Δt + r' − r⁻)/(r⁺ − r⁻);
                if [p < 0] p := 0;
                if [p > 0] p := 1;
                P[i] := (p × P[i] + (1 − p) × P[i+1])/e^(r'Δt);
          }
return P[0];
```

Figure 25.6: Binomial CIR model for zero-coupon bonds.

The one-period bond prices and the local expectations theory then completely determine the evolution of the term structure in Fig. 25.5(b) as follows. Suppose we are interested in the m-period zero-coupon bond price at a node A given that $(m-1)$-period zero-coupon bond prices have been available in the two successive nodes P_{u} (up move) and P_{d} (down move). Let the probability of an up move be p. Then the desired price equals

$$P_{\mathrm{A}}(p P_{\mathrm{u}} + (1-p)\, P_{\mathrm{d}}),$$

where P_{A} is the one-period bond price at A. For instance, the five-period zero-coupon bond price at time zero, 0.960831229521, can be derived with $P_{\mathrm{A}} = 0.992031914837$, $p = 0.472049150276$, $P_{\mathrm{u}} = 0.961665706744$, and $P_{\mathrm{d}} = 0.974702907786$. Once the discount factors are in place, they can be used to obtain the spot rates. For instance, the spot rates at time zero are $(0.04, 0.039996, 0.0399871, 0.0399738, 0.0399565)$ based on Fig. 25.5(b).

➤ **Exercise 25.2.11** Suppose we want to calculate the price of some interest-rate-sensitive security by using the binomial tree for the CIR or the Vasicek model. Assume further that we opt for the Monte Carlo method with antithetic variates. One difficulty with the standard paradigm covered in Subsection 18.2.3 is that, here, the probability at each node varies, and all paths are hence not equally probable. How do we handle this difficulty?

➤ **Exercise 25.2.12** (1) The binomial short rate tree as described requires $\Theta(n^2)$ memory space. How do we perform backward induction on the tree with only $O(n)$ space? (2) Describe a scheme that needs only $O(n^2)$ space for the bond price tree.

Convergence

The binomial approximation converges fast. For example, analytical formula (25.3) gives the following market discount factors at times zero and Δt after an up move.

	Discount factor	
Year	**(Now)**	**(Time Δt; up state)**
0.2	0.992032	0.990197
0.4	0.984131	0.980566
0.6	0.976299	0.971102
0.8	0.968536	0.961804
1.0	0.960845	

The numbers derived by the binomial model in Fig. 25.5(b) can be seen to stand up quite well even at $\Delta t = 0.2$. In fact, we could have started from the short rate tree and generated zero-coupon bond prices at the terminal nodes by Eq. (25.3). This nice feature, made possible by the availability of closed-form formulas, can speed up derivatives pricing. We need to build the tree up to only the maturity of the derivative instead of that of its underlying bond, which could be much longer (see Comment 23.2.1 for more on this point).

25.2.2 A General Method for Constructing Binomial Models

The binomial approximations for the Vasicek and the CIR models follow this general guideline. Given a continuous-time process $dy = \alpha(y,t)\,dt + \sigma(y,t)\,dW$, we first make sure the binomial model's drift and diffusion converge to those of the continuous-time process by setting the probability of an up move to

$$\frac{\alpha(y,t)\,\Delta t + y - y_{\mathrm{u}}}{y_{\mathrm{u}} - y_{\mathrm{d}}},$$

where $y_{\mathrm{u}} + \equiv y + \sigma(y,t)\sqrt{\Delta t}$ and $y_{\mathrm{d}} \equiv y - \sigma(y,t)\sqrt{\Delta t}$ represent the two rates that follow the current rate y. Note that the displacements are identical at $\sigma(y,t)\sqrt{\Delta t}$.

As it stands, the binomial tree may not combine: An up move followed by a down move may not reach the same value as a down move followed by an up move as in general

$$\sigma(y,t)\sqrt{\Delta t} - \sigma(y_{\mathrm{u}},t)\sqrt{\Delta t} \neq -\sigma(y,t)\sqrt{\Delta t} + \sigma(y_{\mathrm{d}},t)\sqrt{\Delta t}\,.$$

When $\sigma(y,t)$ is a constant independent of y, equality holds and the tree combines. To achieve this, define the transformation $x(y,t) \equiv \int^{y} \sigma(z,t)^{-1}\,dz$. Then x follows $dx = m(y,t)\,dt + dW$ for some function $m(y,t)$ (see Exercise 25.2.13). The key is that the diffusion term is now a constant, and the binomial tree for x combines. The probability of an up move remains

$$\frac{\alpha(y(x,t),t)\,\Delta t + y(x,t) - y_{\mathrm{d}}(x,t)}{y_{\mathrm{u}}(x,t) - y_{\mathrm{d}}(x,t)},$$

where $y(x,t)$ is the inverse transformation of $x(y,t)$ from x back to y. Note that $y_{\mathrm{u}}(x,t) \equiv y(x+\sqrt{\Delta t}, t+\Delta t)$ and $y_{\mathrm{d}}(x,t) \equiv y(x-\sqrt{\Delta t}, t+\Delta t)$ [696].

For example, the transformation is $\int^{r} (\sigma\sqrt{z})^{-1}\,dz = 2\sqrt{r}/\sigma$ for the CIR model. As another example, the transformation is $\int^{S} (\sigma z)^{-1}\,dz = (1/\sigma)\ln S$ for the Black–Scholes model. The familiar BOPM in fact discretizes $\ln S$, not S.

➤ **Exercise 25.2.13** Verify that the transformation $x(y, t) \equiv \int^{y} \sigma(z, t)^{-1}\, dz$ turns the process $dy = \alpha(y, t)\, dt + \sigma(y, t)\, dW$ into one for x whose diffusion term is one.

➤ **Exercise 25.2.14** (1) Show that the binomial tree for the untransformed CIR model does not combine. (2) Show that the binomial tree for the geometric Brownian motion $dr = r\mu\, dt + r\sigma\, dW$ does combine even though its volatility is *not* a constant.

25.2.3 Multifactor CIR Models

One-factor models such as the Vasicek and the CIR models are driven by a single source of uncertainty such as the short rate. To address the weaknesses of these models, reviewed in Section 25.5, multifactor models have been proposed. In one two-factor CIR model, the short rate is the sum of two factors r_1 and r_2: $r \equiv r_1 + r_2$ [474]. The risk-neutral processes for r_1 and r_2 are

$$dr_1 = \beta_1(\mu_1 - r_1)\, dt + \sigma_1 \sqrt{r_1}\, dW_1,$$
$$dr_2 = \beta_2(\mu_2 - r_2)\, dt + \sigma_2 \sqrt{r_2}\, dW_2,$$

and ρ is the correlation between dW_1 and dW_2. The partial differential equation for the zero-coupon bond is

$$-\frac{\partial P}{\partial T} + \beta_1(\mu_1 - r_1)\frac{\partial P}{\partial r_1} + \beta_2(\mu_2 - r_2)\frac{\partial P}{\partial r_2} + \frac{\sigma_1^2 r_1}{2}\frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2 r_2}{2}\frac{\partial^2 P}{\partial r_2^2} + \rho\sigma_1\sigma_2\sqrt{r_1 r_2}\,\frac{\partial^2 P}{\partial r_1 \partial r_2} = rP.$$

Because both factors have an impact on yields from the very short end of the term structure, this model behaves like a one-factor model [149].

25.3 Miscellaneous Models

Ogden proposed the following short rate process:

$$dr = \beta(\mu - r)\, dt + \sigma r\, dW,$$

where $\beta \geq 0$ denotes the speed of adjustment and μ is the steady-state interest rate [702]. The predictable part of the change in rates, $\beta(\mu - r)\, dt$, incorporates the mean reversion toward the long-term mean. Clearly the size of the change in rates is greater the further the current rate deviates from its mean. The unpredictable part says that the interest rate is more volatile, in absolute terms, when it is high than when it is low. Dothan's model is lognormal [282]:

$$\frac{dr}{r} = \alpha\, dt + \sigma\, dW.$$

Because interest rates do not grow without bounds, the $dr = \sigma r\, dW$ version may be preferred.

Constantinides developed a family of models to address some of the shortcomings of the CIR model while maintaining positive interest rates and closed-form formulas for various prices of interest rate derivatives [223]. The simplest of the models is

$$dr = 2a\left(1 - \frac{\sigma^2}{a}\right)(\sigma^2 - 2axy)\, dt + 4a\left(1 - \frac{\sigma^2}{a}\right) y\sigma\, dW_1,$$

where a, σ, and α are constants satisfying certain inequalities, $y \equiv x - \alpha + \frac{\alpha}{2(1-\sigma^2/a)}$, and x is the Ornstein–Uhlenbeck process $dx = -ax\,dt + \sigma\,dW_2$. The two Wiener processes W_1 and W_2 are uncorrelated. This model is able to produce inverted-humped yield curves, which are not possible with the CIR model.

Chan, Karolyi, Longstaff, and Sanders (CKLS) proposed the following model [173]:

$$dr = (\alpha + \beta r)\,dt + \sigma r^{\gamma}\,dW.$$

It subsumes the Vasicek model, the CIR model, the Ogden model, and the Dothan model, as well as many others. Using 1-month T-bills, they found that $\gamma \geq 1$ captures short rate dynamics better than $\gamma < 1$. They also reported positive relations between interest rate volatility and the level of interest rate. Their finding of weak evidence of mean reversion is not shared by the data from several European countries, however [248]. Other researchers suggest that $\gamma \geq 1$ overestimates the importance of rate levels on interest rate volatility [93, 126, 563].

Brennan and Schwartz proposed the following two-factor model:

$$d\ln r = \beta(\ln \ell - \ln r)\,dt + \sigma_1\,dW_1,$$
$$\frac{d\ell}{\ell} = a(r, \ell, b_2)\,dt + \sigma_2\,dW_2,$$

where ρ is the correlation between dW_1 and dW_2 [121, 123]. Unlike the two-factor CIR model, the two factors here are at the two ends of the yield curve, i.e., short and long rates. The short rate has mean reversion to the long rate and follows a lognormal process, whereas the long rate follows another lognormal process. This model seems popular [38, 653]. See [462] for problems with this model.

Fong and Vasicek proposed the following model whose two stochastic factors are the short rate and its instantaneous variance v:

$$dr = \beta(\mu_r - r)\,dt + \sqrt{v}\,dW_1,$$
$$dv = \gamma(\mu_v - v)\,dt + \xi\sqrt{v}\,dW_2,$$

where μ_r is the long-term mean of the short rate and μ_v is the long-term mean of the variance of the short rate [362]. See [387, 793] for additional information. A related three-factor model makes the long-term mean μ_r stochastic [46].

➤ **Exercise 25.3.1** To construct a combining binomial tree for the CKLS model, what function of r should be modeled?

25.4 Model Calibration

Two standard approaches to calibrating models are the time-series approach and the cross-sectional approach. In the **time-series** approach, the time series of short rates is used to estimate the parameters of the process. Although it may help in validating the proposed interest rate process, this approach alone cannot be used to estimate the risk premium parameter λ. The model prices based on the estimated parameters may also deviate a lot from those in the market.

The **cross-sectional** approach uses a cross section of observed bond prices. The parameters are to be such that the model prices closely match those in the market. After this procedure, the calibrated model can be used to price interest rate derivatives. Unlike the time-series approach, the cross-sectional approach is unable to separate out the interest rate risk premium from the model parameters. Furthermore, empirical evidence indicates that these estimates may not be stable over time [77, 746]. The common practice of repeated recalibration, albeit pragmatic, is not theoretically sound. A joint cross-section/time-series estimation is also possible [257].

If the model contains only a finite number of parameters, which is true of the Vasicek and the CIR models, a complete match with the market data must be the result of pure luck. This consideration calls for models that have an infinite parameter vector. One way to achieve this is to let some parameters in a finite-dimensional model be deterministic functions of time [234]. Many no-arbitrage models take this route. It must be emphasized that making parameters time dependent does *not* render a model multifactor. Each factor in a multifactor model must represent a distinct source of uncertainty, which a time-dependent parameter does not do, even though it does provide the model with greater flexibility [38].

Calibration cannot correct model specification error. The price of a derivative is the cost of carrying out a self-financing replicating strategy based on its delta, we recall. Delta hedges that fail to replicate the derivative will provide incorrect prices. Hence a misspecified model does not price or hedge correctly even if it has been calibrated [42, 149]. For instance, if the drift of the short rate is not linear, as some evidence suggests [11], then all models that postulate a linear drift err. Of course, it is possible for a wrong model to be useful as an interpolator of prices within a set of claims similar to the ones used in calibration. There is no support, however, for using such a model to price claims very different from the ones in the calibrating set.

➤ **Exercise 25.4.1** Two methods were mentioned for calibrating the Black–Scholes option pricing model: historical volatility and implied volatility. Which corresponds to the time-series approach and which to the cross-sectional approach?

25.5 One-Factor Short Rate Models

One-model short rate models have several shortcomings. To begin with, they throw away much information. By using only the short rate, they ignore other rates on the yield curve. Such models also restrict the volatility to be a function of interest rate *levels* only [126].

When changes in the term structure are driven by a single factor, the prices of all bonds move in the same direction at the same time even though their magnitudes may differ. The returns on all bonds thus become highly correlated. In reality, there seems to be a certain amount of independence between short- and long-term rates [38, 304].[2] One-factor models therefore cannot accommodate nondegenerate correlation structures across maturities. Not surprisingly, derivatives whose values depend on the correlation structure across distinct sectors of the yield curve, such as yield curve options, are mispriced by one-factor models [149].

In one-factor models, the shape of the term structure is typically limited to being monotonically increasing, monotonically decreasing, and slightly humped. The calibrated models also may not generate term structures as concave as the data suggest [41]. The term structure empirically changes in slope and curvature as well as makes parallel moves (review Subsection 19.2.5). This is inconsistent with the restriction that the movements of all segments of the term structure be perfectly correlated. One-factor models are therefore incomplete [607].

Generally speaking, one-factor models generate hedging errors for complex securities [91], and their hedging accuracy is poor [42, 149]. They may nevertheless be acceptable for applications such as managing portfolios of similar-maturity bonds or valuation of securities with cash flows determined predominantly by the overall level of interest rates [198].

Models in which bond prices depend on two or more sources of uncertainty lead to families of yield curves that can take a greater variety of shapes and can better represent reality [46, 793]. Multifactor models include the Brennan–Schwartz model, the Richard model [741], the Langetieg model, the Longstaff–Schwartz model, and the Chen–Scott model [183]. However, a multifactor model is much harder to think about and work with. It also takes much more computer time – the curse of dimensionality raises its head again. These practical concerns limit the use of multifactor models to two-factor ones [38]. Working with different one-factor models before moving on to multifactor ones may be a wise recommendation [84, 482].

The price of a European option on a coupon bond can be calculated from those on zero-coupon bonds as follows. Consider a European call expiring at time T on a bond with par value \$1. Let X denote the strike price. The bond has cash flows $c_1, c_2, \ldots, c_n$ at times $t_1, t_2, \ldots, t_n$, where $t_i > T$ for all i. The payoff for the option is clearly

$$\max\left(\sum_{i=1}^{n} c_i P(r(T), T, t_i) - X, 0\right).$$

At time T, there is a unique value r^* for $r(T)$ that renders the coupon bond's price equal to the strike price X. We can obtain this r^* by solving $X = \sum_i c_i P(r, T, t_i)$ numerically for r, which is straightforward if analytic formulas are known for zero-coupon bond prices. The solution is also unique for one-factor models as the bond price is a monotonically decreasing function of r. Let $X_i \equiv P(r^*, T, t_i)$, the value at time T of a zero-coupon bond with par value \$1 and maturing at time t_i if $r(T) = r^*$. Note that $P(r(T), T, t_i) >= X_i$ if and only if $r(T) <= r^*$. As $X = \sum_i X_i$, the option's payoff equals

$$\sum_{i=1}^{n} c_i \times \max(P(r(T), T, t_i) - X_i, 0).$$

Thus the call is a package of n options on the underlying zero-coupon bond [506].

➤ **Exercise 25.5.1** Suppose that the spot rate curve $r(r, a, b, t, T) \equiv r + a(T - t) + b(T - t)^2$ is implied by a three-factor model. Which of the factors, r, a, and b, affects slope, curvature, and parallel moves, respectively?

➤ **Exercise 25.5.2** Repeat the preceding argument for European puts on coupon bonds and show that the payoff equals $\sum_{i=1}^{n} c_i \times \max(X_i - P(r(T), T, t_i), 0)$.

Concluding Remarks and Additional Reading

When a financial series is described by a stochastic differential equation like

$$dX_t = \mu(X_t)\, dt + \sigma(X_t)\, dW_t,$$

the specific parametric forms chosen for μ and σ may be based more on analytic or computational tractability than economic considerations. This arbitrariness presents a potential problem for every **parametric model**: specification error from picking the wrong functional form. In fact, one study claims that none of the existing parametric interest rate models fit historical data well [11] (this finding is contested in [728]). **Nonparametric models** in contrast make no parametric assumptions about the functional forms of the drift μ and/or the diffusion σ [819]. Instead, one or both functions are to be estimated nonparametrically from the discretely observed data. The requirement is that approximations to the true drift and diffusion converge pointwise to μ and σ at a rate $(\Delta t)^k$, where Δt is the time between successive observations and $k > 0$. As a result, the approximation errors should be small as long as observations are made frequently enough. See [517, 611, 613] for the estimation of Ito processes.

This chapter surveyed equilibrium models and pointed out some of their weaknesses. One way to address them is the adoption of no-arbitrage models, to which we will turn in the next chapter. Another approach is the use of additional factors. Nonparametric models are yet another option. Unlike equity derivatives, no single dominant model emerges.

For the pricing of interest rate caps, consult [616] (the CIR case) and [617] (the Vasicek case). See [184, 185, 186, 257, 630, 645, 803] for more information on multifactor CIR models and parameter estimation techniques. Refer to [291, 477] for more discussions on one-factor models. Finally, see [301, 302] for discussions on long rates.

NOTES

1. www.cob.ohio-state.edu/˜fin/journal/lies.htm.
2. Real rates seem to be more correlated [135].