14

A Brief Survey of Term Structure Models

This chapter presents very brief descriptions of several important models. The list of models is certainly not exhaustive, and our descriptions will be far from complete. As mentioned previously, there are many good references for this material in book form already, and our goal here is merely to provide a short introduction.

14.1 Ho-Lee

The Ho-Lee [36] model is a binomial version of the Vasicek model without mean reversion, in which the one-period interest rate is assumed to have a deterministic drift. This was the first widely-used model that enabled one to fit the current yield curve.

Consider discrete times $0 = t_0 < t_1 < \cdots t_N$ with $t_i - t_{i-1} = \Delta t$ for each i. We denote the one-period interest rate from date t_i to t_{i+1} by $r(t_i)$. We express the rate as an annualized continuously compounded rate, so the one-period discount factor from date t_i to t_{i+1} is $e^{-r(t_i)\Delta t}$. We could put this in a continuous-time framework by assuming the short rate is constant $(= r(t_i))$ during each time interval (t_i, t_{i+1}). As always, the risk-neutral measure is the probability measure corresponding to the numeraire

$$R(t) = \exp\left(\int_0^t r(s)\, ds\right) . \tag{14.1}$$

However, we will only be doing valuation at the discrete dates $t_0, \ldots, t_N$ and at date t_i the accumulation factor (14.1) is

$$R(t_i) = \exp\left(\sum_{j=0}^{i-1} r(t_j)\, \Delta t\right) .$$

Thus, the continuous-time framework is not necessary.

The model assumes that over each time period $t_i - t_{i-1}$ the change $\Delta r(t_i) = r(t_i) - r(t_{i-1})$ in the one-period rate is

$$\Delta r(t_i) = \theta(t_i)\,\Delta t \pm \sigma\sqrt{\Delta t}\,,$$

where the risk-neutral probability of "+" and "−" is $1/2$ each and θ is a non-random function. As in the discussion of the extended Vasicek model, it is convenient to define a random process $\hat{r}$ by $\hat{r}(0) = r(0)$ and

$$\Delta \hat{r}(t_i) = \pm\sigma\sqrt{\Delta t}$$

for each i. Also, define $\phi(t_i) = \sum_{j=1}^{i}\theta(t_i)$ for $i \geq 1$. Then we have, for $i > 0$,

$$r(t_i) = \phi(t_i) + \hat{r}(t_i)\,.$$

The following illustrates a three-period tree with initial one-period rate $r_0 = 5\%$, $\Delta t = 1$, $\theta(t_i) = 0$ for all i, and $\sigma = 1\%$.
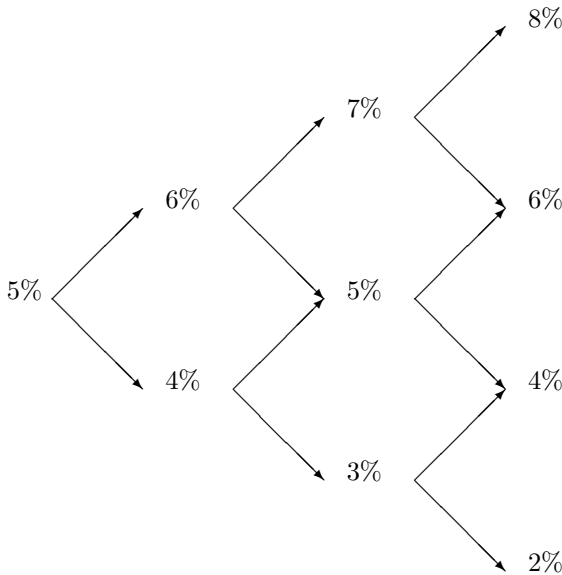

To value a fixed-income derivative, we discount the terminal value backwards through the tree as in Chaps. 5 and 9. The new feature is that the discount rate is changing over time. So, we now have three trees to consider: the tree for the underlying, the tree for the one-period interest rate, and the tree for the derivative value. However, the tree for the underlying can be created from the interest-rate tree, so the interest-rate tree is the basic input instead of the tree for the underlying.

To clarify this, we will start with the simplest example: valuing a discount bond. Consider the above interest-rate tree and a discount bond that matures

at date 3. The value of the bond is 1 at maturity. Because we are using continuous compounding, the discount factor is $e^{-.07} = 0.932$ at the top node at date 2. This is also the value of the discount bond at that node. Likewise, the value of the bond at the middle node at date 2 is $e^{-.05} = 0.951$. This implies that the value of the bond at the top node at date 1 is

$$e^{-.06}\left(\frac{1}{2} \times 0.932 + \frac{1}{2} \times 0.951\right) = 0.887 \, .$$

Continuing in this way, we derive the following tree for the discount bond, concluding that its price at date 0 should be \$0.861.
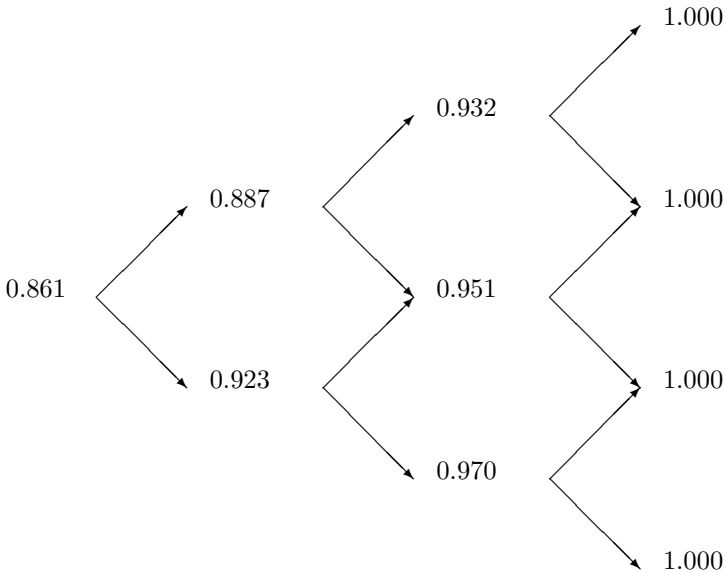
In general, the price at date 0 of a discount bond maturing at date t_n is

$$P(0, t_n) = E^R \left[\exp\left(-\sum_{i=0}^{n-1} r(t_i)\,\Delta t\right)\right]$$

$$= E^R \left[\exp\left(-r(t_0)\,\Delta t - \sum_{i=1}^{n-1} [\hat{r}(t_i) + \phi(t_i)]\,\Delta t\right)\right]$$

$$= \exp\left(-r(t_0)\,\Delta t - \sum_{i=1}^{n-1} \phi(t_i)\,\Delta t\right) E^R \left[\exp\left(-\sum_{i=1}^{n-1} \hat{r}(t_i)\,\Delta t\right)\right] \, .$$

$$(14.2)$$

Given the parameter σ, the expectation in the last line of the above can easily be computed.

The parameters $\phi(t_n)$ can be chosen to equate model prices of discount bonds maturing at $t_1, \ldots, t_N$ to market prices. This is done simply by choosing $\phi(t_n)$ to satisfy the following equation, which we will derive below:

$$e^{-\phi(t_n)\,\Delta t} = \frac{2e^{r(0)\,\Delta t}}{e^{n\sigma\sqrt{\Delta t}\,\Delta t} + e^{-n\sigma\sqrt{\Delta t}\,\Delta t}} \times \frac{P^{\mathrm{mkt}}(0, t_{n+1})}{P^{\mathrm{mkt}}(0, t_n)}. \tag{14.3}$$

Because the ratio of market prices $P^{\mathrm{mkt}}(0, t_{i+1})/P^{\mathrm{mkt}}(0, t_i)$ is the reciprocal of one plus the market forward rate at date 0 for loans from date t_i to t_{i+1}, this formula for the parameters $\phi(t_i)$ is equivalent to equating model forward rates to market forward rates, as was discussed for the Hull-White and continuous-time Ho-Lee models in Sect. 13.6. In fact, the fitting of the Ho-Lee model to market bond prices is often expressed by saying "the market forward rate curve is an input to the model." This idea was developed further by Heath, Jarrow and Morton [33]—see Sect. 14.6.

As for options on equities and currencies, the binomial model for interest rates is especially useful for valuing early exercise features. However, it should be noted that, even though this model is very easy to use, it has important limitations. The assumption of a constant volatility for the one-period rate and no mean reversion implies excessive uncertainty about the level of the one-period rate at long horizons, as discussed in Sect. 13.1. To offset this, one could use a smaller volatility when valuing long-maturity options. However, for options with early exercise features, this would imply too little uncertainty about the level of the one-period rate at short horizons and thereby undervalue the early exercise option.

We will conclude this section with a proof of formula (14.3). Let $\varepsilon(t_i)$ denote independent random variables that equal $\pm\sigma\sqrt{\Delta t}$ with probability $1/2$ each under the risk-neutral measure. Then we can write

$$\hat{r}(t_i) = r(0) + \sum_{j=1}^{i} \varepsilon(t_j) \,.$$

This implies that

$$\sum_{i=0}^{n-1} \hat{r}(t_i) = (n-1)r(0) + \sum_{i=1}^{n-1}\sum_{j=1}^{i} \varepsilon(t_j)$$

$$= (n-1)r(0) + (n-1)\varepsilon(t_1) + (n-2)\varepsilon(t_2) + \cdots + \varepsilon(t_{n-1})$$

$$= (n-1)r(0) + \sum_{i=1}^{n-1} (n-i)\varepsilon(t_i) \,.$$

Therefore, (14.2) gives us

$$P(0, t_n) = \exp\left(-nr(0)\,\Delta t - \sum_{i=1}^{n-1}\phi(t_i)\,\Delta t\right) E^R\left[\exp\left(-\sum_{i=1}^{n-1}(n-i)\varepsilon(t_i)\Delta t\right)\right] \,.$$

Moreover,

$$E^R \left[\exp \left(- \sum_{i=1}^{n-1} (n-i)\varepsilon(t_i) \, \Delta t \right) \right] = E^R \left[\prod_{i=1}^{n-1} \exp \left(-(n-i)\varepsilon(t_i) \, \Delta t \right) \right]$$

$$= \prod_{i=1}^{n-1} E^R \left[\exp \left(-(n-i)\varepsilon(t_i) \, \Delta t \right) \right]$$

$$= \prod_{i=1}^{n-1} \left(\frac{1}{2} \exp \left((n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-(n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) \right).$$

Thus,

$$P(0, t_n) = \exp \left(-nr(0) \, \Delta t - \sum_{i=1}^{n-1} \phi(t_i) \, \Delta t \right)$$

$$\times \prod_{i=1}^{n-1} \left(\frac{1}{2} \exp \left((n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-(n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) \right).$$

Likewise,

$$P(0, t_{n+1}) = \exp \left(-(n+1)r(0) \, \Delta t - \sum_{i=1}^{n} \phi(t_i) \, \Delta t \right)$$

$$\times \prod_{i=1}^{n} \left(\frac{1}{2} \exp \left((n+1-i)\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-(n+1-i)\sigma\sqrt{\Delta t} \, \Delta t \right) \right),$$

which we can write as

$$\exp \left(-(n+1)r(0) \, \Delta t - \sum_{i=1}^{n} \phi(t_i) \, \Delta t \right)$$

$$\times \prod_{i=0}^{n-1} \left(\frac{1}{2} \exp \left((n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-(n-i)\sigma\sqrt{\Delta t} \, \Delta t \right) \right)$$

$$= \exp \left(-r(0) \, \Delta t - \phi(t_n) \, \Delta t \right)$$

$$\times \left(\frac{1}{2} \exp \left(n\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-n\sigma\sqrt{\Delta t} \, \Delta t \right) \right) P(0, t_n).$$

Thus, the ratio of model prices is

$$\frac{P(0, t_{n+1})}{P(0, t_n)} = \exp \left(-r(0) \, \Delta t - \phi(t_n) \, \Delta t \right)$$

$$\times \left(\frac{1}{2} \exp \left(n\sigma\sqrt{\Delta t} \, \Delta t \right) + \frac{1}{2} \exp \left(-n\sigma\sqrt{\Delta t} \, \Delta t \right) \right).$$

Equating this to the ratio of market prices $P^{\mathrm{mkt}}(0, t_{n+1})/P^{\mathrm{mkt}}(0, t_n)$ gives the formula (14.3).

14.2 Black-Derman-Toy

The Black-Derman-Toy [4] model is, like the Ho-Lee model and the Black-Karasinski model discussed in the next section, a binomial model of the one-period interest rate. The model assumes that

$$\Delta \log r(t_i) = \eta(t_i)\,\Delta t \pm \sigma(t_i)\sqrt{\Delta t}, \tag{14.4}$$

where "+" and "−" have probability one-half each under the risk-neutral measure. The volatility $\sigma(t_i)$ has the interpretation of the *percentage* volatility of the one-period rate (rather than the absolute volatility, as in the Vasicek and Ho-Lee models). The model implies that the one-period rate will always be nonnegative.

A significant feature of the model is that the volatility is allowed to be time-varying. This would produce a non-recombining tree except that the drift $\eta(t_i)$ is allowed to vary across the date–t_{i-1} nodes, i.e., to be a random variable, depending on the level of the one-period rate at date t_{i-1}. To understand this, consider the following two-period tree. For convenience, we write a_i for $\eta(t_i)\,\Delta t$ and b_i for $\sigma(t_i)\sqrt{\Delta t}$.
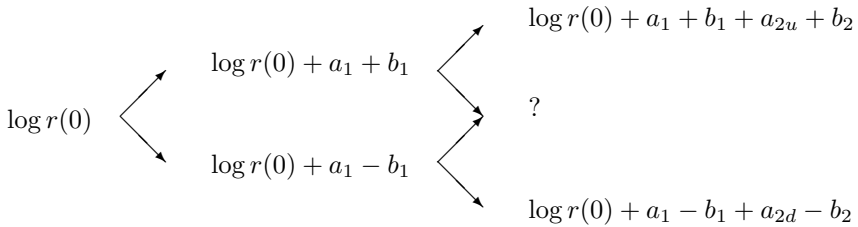
We have written a_{2u} and a_{2d} to demonstrate that the drift between date 1 and date 2 can vary, depending on whether we are at the top or bottom node at date 1. If we arrive at the node marked with a question mark via a down move from the top node at date 1, the value will be

$$\log r(0) + a_1 + b_1 + a_{2u} - b_2 \ .$$

On the other hand, if we arrive at it via an up move from the bottom node at date 1, the value will be

$$\log r(0) + a_1 - b_1 + a_{2d} + b_2 \ .$$

For the tree to be recombining, these values must be the same, which implies

$$a_{2d} = a_{2u} + 2b_1 - 2b_2 \ .$$

In general, at each date there are two free parameters: the volatility and the drift at one of the nodes, the drifts at the other nodes then being determined by the requirement that the tree be recombining.

The dependence of the drift on the node can be expressed as a linear dependence on the logarithm of the one-period rate. In other words, it is possible (and convenient) to write the Black-Derman-Toy model (14.4) as

$$\Delta \log r(t_i) = \kappa(t_i)[\theta(t_i) - \log r(t_{i-1})] \, \Delta t \pm \sigma(t_i)\sqrt{\Delta t}, \qquad (14.5)$$

where now the functions $\kappa(t_i)$, $\theta(t_i)$, and $\sigma(t_i)$ are deterministic—i.e., depending on time but constant across nodes at each date.

In the two-period example above, we have

$$a_{2u} = \kappa(t_2)[\theta(t_2) - \{\log r(t_0) + a_1 + b_1\}] \, \Delta t \,,$$

and

$$a_{2d} = \kappa(t_2)[\theta(t_2) - \{\log r(t_0) + a_1 - b_1\}] \, \Delta t \,,$$

so the relation $a_{2d} = a_{2u} + 2b_1 - 2b_2$ is equivalent to

$$\kappa(t_2) = \frac{b_1 - b_2}{b_1 \, \Delta t} = -\frac{1}{\sigma(t_1)} \times \frac{\sigma(t_2) - \sigma(t_1)}{\Delta t} = -\frac{1}{\sigma(t_1)} \times \frac{\Delta \sigma(t_2)}{\Delta t}. \qquad (14.6)$$

This same relationship holds at each node at each date (just consider the two-period example as two periods extending from any node in the tree); thus, in general, we have

$$\kappa(t_i) = -\frac{1}{\sigma(t_{i-1})} \times \frac{\Delta \sigma(t_i)}{\Delta t}. \qquad (14.7)$$

Equations (14.5) and (14.7) define the Black-Derman-Toy model. The free parameters at each date are $\sigma(t_i)$ and $\theta(t_i)$, and the parameter $\kappa(t_i)$ is defined by (14.7). Alternatively, one can view $\kappa(t_i)$ and $\theta(t_i)$ as free parameters and define $\sigma(t_i)$ from (14.7). Note that $\kappa(t_i)$ can be interpreted as a mean-reversion parameter for $\log r(t_i)$, at least when it is positive (i.e., when $\Delta \sigma(t_i) < 0$). Because there are two free parameters at each date rather than one (as in the Ho-Lee model) it is possible to match both market bond prices and market yield volatilities or cap prices.

In continuous time, we would write (14.7) as

$$\kappa(t) = -\frac{d \log \sigma(t)}{dt} \,.$$

Therefore, the continuous-time version of the Black-Derman-Toy model is

$$d \log r(t) = -\frac{d \log \sigma(t)}{dt}[\theta(t) - \log r(t)] \, dt + \sigma(t) \, dB(t) \,,$$

with B being a Brownian motion under the risk-neutral measure.

14.3 Black-Karasinski

The Black-Karasinski [5] model is similar to the Black-Derman-Toy model—it assumes (14.5) for the changes in the logarithm of the one-period rate—but it removes the linkage (14.7) between the mean-reversion parameter and the volatility. It does this by allowing the lengths of the time steps to vary. Denote the length of the time-step $t_i - t_{i-1}$ by τ_i. Consider again the two-period example of the previous section. As a necessary condition for the tree to be recombining, we deduced (in (14.6)) that

$$\kappa(t_2) = \frac{b_1 - b_2}{b_1 \tau_2} \ .$$

In this model, we have $b_i = \sigma(t_i)\sqrt{\tau_i}$. Making this substitution, we obtain

$$\kappa(t_2) = \frac{\sigma(t_1)\sqrt{\tau_1} - \sigma(t_2)\sqrt{\tau_2}}{\sigma(t_1)\sqrt{\tau_1}\tau_2} \ ,$$

which we can rewrite as

$$\kappa(t_2)\sigma(t_1)\sqrt{\tau_1}\tau_2 + \sigma(t_2)\sqrt{\tau_2} - \sigma(t_1)\sqrt{\tau_1} = 0 \ .$$

This is a quadratic equation in the unknown $\sqrt{\tau_2}$ with the unique positive solution (assuming $\kappa(t_2) > 0$)

$$\sqrt{\tau_2} = \frac{\sqrt{\sigma(t_2)^2 + 4\kappa(t_2)\sigma(t_1)^2\sqrt{\tau_1}} - \sigma(t_2)}{2\kappa(t_2)\sigma(t_1)\sqrt{\tau_1}} \ .$$

This relation must hold at each date. Thus, squaring both sides, we obtain the general formula

$$\tau_i = \frac{\left[\sqrt{\sigma(t_i)^2 + 4\kappa(t_i)\sigma(t_{i-1})^2\sqrt{\tau_{i-1}}} - \sigma(t_i)\right]^2}{4\kappa(t_i)^2\sigma(t_{i-1})^2\tau_{i-1}}. \tag{14.8}$$

To summarize, the Black-Karasinski model is given by (14.5) with three free parameters—$\kappa(t_i)$, $\theta(t_i)$, and $\sigma(t_i)$—at each date. It is implemented in a recombining tree by defining the length of each time step $\tau_i = t_i - t_{i-1}$ for $i \geq 2$ according to the formula (14.8). The length of the first time step τ_1 can be chosen arbitrarily.

14.4 Cox-Ingersoll-Ross

Cox, Ingersoll, and Ross [19] introduced a continuous-time model[1] in which the short rate satisfies

---

[1] Cox, Ingersoll and Ross (hereafter CIR) also discuss a variety of other continuous-time models, but this particular model is so well known that it is often simply called *the* CIR model.

$$dr(t) = \kappa[\theta - r(t)]\, dt + \sigma\sqrt{r(t)}\, dB(t), \tag{14.9}$$

where κ, θ, and σ are positive constants and B is a Brownian motion under the risk-neutral measure. Like the Vasicek model, this short rate process has a long-run mean of θ. The difference between the CIR model and the Vasicek model is that the volatility in the CIR model is proportional to the square root of the short rate rather than being constant. Because of this fact, the short rate can never be negative. Intuitively, the reason is that the volatility $\sigma\sqrt{r(t)}$ is very small if $r(t)$ is near zero, so the drift will dominate the change in $r(t)$, pushing it upwards towards θ. This interest rate model was mimicked by Heston [34] in his stochastic volatility model discussed in Chap. 4. We will briefly discuss three topics in connection with this model: discount bond prices, calibrating the model to the current market, and pricing fixed-income derivatives.

Discount Bond Prices in the CIR Model

Discount bond prices can be most easily computed in the CIR model by solving the fundamental partial differential equation (pde) discussed in Chap. 10. Let $P(t, u)$ denote the price at date t of a discount bond maturing at date u, having remaining time to maturity of $\tau = u - t$. The discount bond price will depend on the remaining time to maturity and the short rate at date t, because, as in the Vasicek model, the short rate is the only random factor in this model. Thus, there must be some deterministic function f such that $P(t, u) = f(r(t), \tau)$. As in Chap. 10, the fundamental pde is obtained by applying Itô's formula to f to compute df in terms of the partial derivatives of f and then using the fact that the expected return of the discount bond (hence the drift of df/f) must equal the short rate under the risk-neutral measure.

From Itô's formula and the definition (14.9) of dr, we have

$$\begin{aligned}
df &= \frac{\partial f}{\partial \tau}\, d\tau + \frac{\partial f}{\partial r}\, dr + \frac{1}{2}\frac{\partial^2 f}{\partial r^2}(dr)^2 \\
&= -\frac{\partial f}{\partial \tau}\, dt + \frac{\partial f}{\partial r}\left\{\kappa[\theta - r]\, dt + \sigma\sqrt{r}\, dB\right\} + \frac{1}{2}\frac{\partial^2 f}{\partial r^2}\sigma^2 r\, dt \\
&= \left(-\frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial r}\kappa[\theta - r] + \frac{1}{2}\frac{\partial^2 f}{\partial r^2}\sigma^2 r\right) dt + \frac{\partial f}{\partial r}\sigma\sqrt{r}\, dB \ . \tag{14.10}
\end{aligned}$$

Equating the drift to $rf\, dt$ gives us the fundamental pde:

$$-\frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial r}\kappa[\theta - r] + \frac{1}{2}\frac{\partial^2 f}{\partial r^2}\sigma^2 r = rf \ . \tag{14.11}$$

This equation should be solved for the function f subject to the boundary condition that the value of the discount bond is one at maturity; i.e., $f(r, 0) = 1$ for all r.

The solution can be obtained by "guessing" a solution of the same form as the Vasicek bond pricing formula (13.6a), namely[2]

$$f(r, \tau) = \exp\left(-a(\tau) - b(\tau)r\right) \tag{14.12}$$

for deterministic functions a and b. The boundary condition

$$f(r, 0) = \exp\left(-a(0) - b(0)r\right) = 1$$

for all r implies $a(0) = b(0) = 0$, and it can easily be checked that the fundamental pde is equivalent to

$$b'(\tau) = 1 - \kappa b(\tau) - \frac{1}{2}\sigma^2 b^2(\tau) , \tag{14.13a}$$

and

$$a'(\tau) = \kappa\theta b(\tau) , \tag{14.13b}$$

where the "primes" denote derivatives. By differentiating, one can verify that the solution of (14.13a) (which is called a "Riccati equation") with the boundary condition $b(0) = 0$ is

$$b(\tau) = \frac{2\left(e^{\gamma\tau} - 1\right)}{c(\tau)} , \tag{14.14a}$$

where

$$\gamma = \sqrt{\kappa^2 + 2\sigma^2} \qquad \text{and} \qquad c(\tau) = (\kappa + \gamma)\left(e^{\gamma\tau} - 1\right) + 2\gamma . \tag{14.14b}$$

Integrating (14.13b) then gives

$$a(\tau) = -\frac{2\kappa\theta}{\sigma^2}\left[\frac{(\kappa + \gamma)\tau}{2} + \log\frac{2\gamma}{c(\tau)}\right] . \tag{14.14c}$$

To summarize,

The price at date t of a discount bond maturing at $u > t$ in the CIR model is
$$P(t, u) = \exp\left(-a(\tau) - b(\tau)r(t)\right) , \tag{14.15}$$
where $\tau = u - t$ and $a(\tau)$ and $b(\tau)$ are defined in (14.14).

Note that (14.12) implies $\partial f/\partial r = -b(\tau)f$. Substituting this into (14.10) gives us the discount bond return as

$$\frac{\mathrm{d}P(t, u)}{P(t, u)} = \frac{\mathrm{d}f}{f} = r(t)\,\mathrm{d}t - b(\tau)\sigma\sqrt{r(t)}\,\mathrm{d}B(t) . \tag{14.16}$$

---

[2] This guess works because the CIR model, like the Vasicek model, is an "affine model." See Sect. 14.5.

This is again similar to the Vasicek model except for the appearance of the $\sqrt{r}$ factor in the volatility. Because of this factor, the volatility is random. Thus, the option pricing formulas of previous chapters cannot be directly applied to price discount bond options (and hence caps, floors, coupon bond options, and swaptions). Nevertheless, the ideas underlying those formulas can be applied to obtain similar valuation formulas.

Hedge ratios depend on relative volatilities, so they are determined by the function b, just as discussed in Sect. 13.4 for the Vasicek model.

Calibrating the CIR Model to the Yield Curve

The CIR model can be calibrated to current market conditions by taking one or more of the parameters κ, θ and σ to be time-varying, as in the extended Vasicek model. This was suggested by Cox, Ingersoll and Ross. However, the simplest way to calibrate the model to discount bond prices, which was also suggested by Cox, Ingersoll and Ross, is to take the short rate to be the sum of a non-random function of time and a square-root process as defined in (14.9). Specifically, let

$$r(t) = \phi(t) + \hat{r}(t) \ ,$$

where ϕ is a non-random function and $\hat{r}$ satisfies

$$\mathrm{d}\hat{r}(t) = \kappa[\theta - \hat{r}(t)]\,\mathrm{d}t + \sigma\sqrt{\hat{r}(t)}\,\mathrm{d}B(t) \ , \tag{14.17}$$

with $\hat{r}(0) = r(0)$. Then, as in the Hull-White model, discount bond prices are given by

$$P(t,u) = \exp\left(-\int_t^u \phi(s)\,\mathrm{d}s\right) E_t^R\left[\exp\left(-\int_t^u \hat{r}(s)\,\mathrm{d}s\right)\right] \ .$$

Moreover, the expectation in this equation is the discount bond pricing function calculated in the previous subsection, so we have

$$P(t,u) = \exp\left(-\int_t^u \phi(s)\,\mathrm{d}s\right)\exp\left(-a(\tau) - b(\tau)\hat{r}(t)\right) \ . \tag{14.18}$$

In particular, discount bond prices at date 0 are

$$P(0,u) = \exp\left(-\int_0^u \phi(s)\,\mathrm{d}s\right)\exp\left(-a(u) - b(u)r(0)\right) \ ,$$

so the model can be calibrated to market prices $P^{\mathrm{mkt}}(0,u)$ by setting

$$\exp\left(-\int_0^u \phi(s)\,\mathrm{d}s\right) = \exp\left(a(u) + b(u)r(0)\right)P^{\mathrm{mkt}}(0,u) \tag{14.19}$$

for each u. Note that the calibration does not affect discount bond returns: the expected return under the risk-neutral measure must still be the short

rate and the volatility is unaffected by a deterministic factor. Therefore, we have, as in (14.16),

$$\frac{\mathrm{d}P(t,u)}{P(t,u)} = r(t)\,\mathrm{d}t - b(\tau)\sigma\sqrt{\hat{r}(t)}\,\mathrm{d}B(t)\ . \tag{14.20}$$

This can be verified by applying Itô's formula to (14.18).

Pricing Fixed-Income Derivatives in the CIR Model

In the previous chapter, it was shown for the extended Vasicek model that pricing formulas for caps, floors, coupon bond options, and swaptions can be derived from a pricing formula for discount bond options. The same is true in the CIR model — caps and floors are of course portfolios of discount bond options and, in a single-factor model such as the Vasicek or CIR model, coupon bond options and swaptions can also be priced as portfolios of discount bond options. Here we will explain briefly how to price discount bond options in the CIR model.

Consider a call option maturing at date T with the underlying being a discount bond maturing at $u > T$. Let K denote the strike price. From our fundamental pricing formula (1.17), the value at date 0 of the option is

$$P(0,u) \times \mathrm{prob}^u\big(P(T,u) > K\big) - P(0,T) \times \mathrm{prob}^T\big(P(T,u) > K\big)\ , \quad (14.21)$$

where prob^u denotes the probability measure using the discount bond maturing at u as numeraire and prob^T denotes the probability measure using the discount bond maturing at T as the numeraire. Using the calibration of the previous subsection, the price of the underlying at date T will be, according to the model,

$$P(T,u) = \exp\left(-\int_T^u \phi(s)\,\mathrm{d}s - a(u-T) - b(u-T)\hat{r}(T)\right)\ .$$

Therefore, the option will finish in the money if and only if

$$\frac{-\int_T^u \phi(s)\,\mathrm{d}s - a(u-T) - \log K}{b(u-T)} > \hat{r}(T)\ . \tag{14.22}$$

Thus, to price discount bond options, we need to compute the probabilities that $\hat{r}(T)$ is less than a given number, using discount bonds as numeraires.

Consider the discount bond maturing at u. The calculation for the discount bond maturing at T can be done in the same way. We use the fact that

$$\frac{\mathrm{e}^{\int_0^t r(s)\,\mathrm{d}s}}{P(t,u)}$$

is a martingale, when $P(t,u)$ is used as the numeraire, for $t \le u$. Let $Z(t)$ denote this ratio and apply Itô's formula for ratios to derive

$$\frac{\mathrm{d}Z(t)}{Z(t)} = r(t)\,\mathrm{d}t - \frac{\mathrm{d}P(t,u)}{P(t,u)} + \left(\frac{\mathrm{d}P(t,u)}{P(t,u)}\right)^2 .$$

Substituting from (14.20) now gives us

$$\begin{aligned}
\frac{\mathrm{d}Z(t)}{Z(t)} &= b(\tau)\sigma\sqrt{\hat{r}(t)}\,\mathrm{d}B(t) + b(\tau)^2\sigma^2\hat{r}(t)\,\mathrm{d}t \\
&= b(\tau)\sigma\sqrt{\hat{r}(t)}\left[\mathrm{d}B(t) + b(\tau)\sigma\sqrt{\hat{r}(t)}\,\mathrm{d}t\right] .
\end{aligned}$$

Given that Z is a martingale, $\mathrm{d}Z/Z$ cannot have a drift, so it must be that B^* defined by $B^*(0) = 0$ and

$$\mathrm{d}B^*(t) = \mathrm{d}B(t) + b(\tau)\sigma\sqrt{\hat{r}(t)}\,\mathrm{d}t$$

is a martingale, and hence a Brownian motion, when $P(t,u)$ is used as the numeraire. Substituting this into the definition (14.17) of $\hat{r}$, we have

$$\begin{aligned}
\mathrm{d}\hat{r}(t) &= \kappa[\theta - \hat{r}(t)]\,\mathrm{d}t + \sigma\sqrt{\hat{r}(t)}\left[\mathrm{d}B^*(t) - b(\tau)\sigma\sqrt{\hat{r}(t)}\,\mathrm{d}t\right] \\
&= \kappa[\theta - \hat{r}(t)]\,\mathrm{d}t + \sigma\sqrt{\hat{r}(t)}\,\mathrm{d}B^*(t) - \sigma^2 b(\tau)\hat{r}(t)\,\mathrm{d}t \\
&= \kappa^*(t)[\theta^*(t) - \hat{r}(t)]\,\mathrm{d}t + \sigma\sqrt{\hat{r}(t)}\,\mathrm{d}B^*(t) , \qquad (14.23)
\end{aligned}$$

where we define

$$\kappa^*(t) = \kappa + \sigma^2 b(u - t) \qquad \text{and} \qquad \theta^*(t) = \frac{\kappa\theta}{\kappa^*(t)} .$$

Thus, using a discount bond as numeraire, the process $\hat{r}$ is still a square root process, but now with a time-dependent long-run mean and mean-reversion rate.

The random variable $\hat{r}(T)$ defined by $\hat{r}(0) = 0$ and equation (14.23) for $t \leq T$ is a transformation of a random variable having what is called a "non-central chi-square" distribution. See Appendix B.3 for further discussion and calculation of the probabilities $\mathrm{prob}^u\big(P(T,u) > K\big)$ and $\mathrm{prob}^T\big(P(T,u) > K\big)$.

14.5 Longstaff-Schwartz

Cox, Ingersoll and Ross suggest adding two independent square-root processes to obtain a two-factor model. This means that we would take $r(t) = x_1(t) + x_2(t)$, where

$$\mathrm{d}x_i(t) = \kappa_i[\theta_i - x_i(t)]\,\mathrm{d}t + \sigma_i\sqrt{x_i(t)}\,\mathrm{d}B_i(t) , \qquad (14.24)$$

where B_1 and B_2 are independent Brownian motions under the risk-neutral measure and κ_i, θ_i and σ_i are positive constants for $i = 1, 2$. Longstaff and

Schwartz [47] investigate this model further, including providing an "equilibrium" foundation, deriving discount bond option prices, and estimating the coefficients. The model is usually called the Longstaff-Schwartz model.

An important observation made by Longstaff and Schwartz is that the model can be rewritten so that the short rate and its volatility are the factors (rather than the unobservable x_1 and x_2).[3] Note that the instantaneous variance of $r = x_1 + x_2$ is

$$\left(\sigma_1\sqrt{x_1(t)}\,\mathrm{d}B_1(t) + \sigma_2\sqrt{x_2(t)}\,\mathrm{d}B_2(t)\right)^2 = \left(\sigma_1^2 x_1(t) + \sigma_2^2 x_2(t)\right)\,\mathrm{d}t\ .$$

Define $V(t) = \sigma_1^2 x_1(t) + \sigma_2^2 x_2(t)$. We can solve for x_1 and x_2 in terms of r and V as

$$x_1 = \frac{\sigma_2^2 r - V}{\sigma_2^2 - \sigma_1^2} \tag{14.25a}$$

$$x_2 = \frac{V - \sigma_1^2 r}{\sigma_2^2 - \sigma_1^2}\ , \tag{14.25b}$$

provided $\sigma_1 \neq \sigma_2$. Making these substitutions for x_1 and x_2 on the right-hand side of (14.24) and noting that $\mathrm{d}r = \mathrm{d}x_1 + \mathrm{d}x_2$ and $\mathrm{d}V = \sigma_1^2\,\mathrm{d}x_1 + \sigma_2^2\,\mathrm{d}x_2$, we obtain the following equations presented by Longstaff and Schwartz:

$$\mathrm{d}r = \left(\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha}r - \frac{\xi - \delta}{\beta - \alpha}V\right)\mathrm{d}t$$

$$+ \alpha\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}\,\mathrm{d}B_1 + \beta\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}\,\mathrm{d}B_2\ , \tag{14.26a}$$

$$\mathrm{d}V = \left(\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha}r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha}V\right)\mathrm{d}t$$

$$+ \alpha^2\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}\,\mathrm{d}B_1 + \beta^2\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}\,\mathrm{d}B_2\ , \tag{14.26b}$$

where $\delta = \kappa_1$, $\xi = \kappa_2$, $\alpha = \sigma_1^2$, $\beta = \sigma_2^2$, $\gamma = \kappa_1\theta_1/\sigma_1^2$, and $\eta = \kappa_2\theta_2/\sigma_2^2$. Thus, this can be regarded as a two-factor model in which the factors are the short rate and its instantaneous variance, with the six parameters δ, ξ, α, β, γ, and η.

The simplest way to compute discount bond prices in this model is to return to the definition $r = x_1 + x_2$. Discount bond prices are

---

[3] Of course, the volatility is also not directly observable. Longstaff and Schwartz use a GARCH model to estimate it and then use the time series of estimated volatilities and the time series of short rates to estimate the parameters of the model.

$$P(t, u) = E_t \left[\exp \left(- \int_t^u r(s) \, ds \right) \right]$$

$$= E_t \left[\exp \left(- \int_t^u x_1(s) \, ds \right) \exp \left(- \int_t^u x_2(s) \, ds \right) \right]$$

$$= E_t \left[\exp \left(- \int_t^u x_1(s) \, ds \right) \right] E_t \left[\exp \left(- \int_t^u x_2(s) \, ds \right) \right] ,$$

due to the independence of x_1 and x_2. Moreover, these expectations have the same form as discount bond prices in the CIR model, namely

$$E_t \left[\exp \left(- \int_t^u x_i(s) \, ds \right) \right] = \exp \left(-a_i(\tau) - b_i(\tau) x_i(t) \right) , \qquad (14.27)$$

where the functions a_i and b_i are defined in (14.14), using the parameters κ_i, θ_i and σ_i. The expectations (14.27) can be written in terms of $r(t)$ and $V(t)$ by substituting from (14.25).

The Vasicek, CIR and Longstaff-Schwartz models are examples of "affine models." An affine model is defined by a set of factors $x_1, \ldots, x_n$, where

- The short rate is an affine function of the factors;[4] i.e., $r(t) = \alpha_0 + \sum_{i=1}^n \alpha_i x_i(t)$ for constants α_i,
- The drift of each factor is an affine function of the factors.
- The instantaneous variance of each factor is an affine function of the factors.
- The instantaneous covariance of each pair of factors is an affine function of the factors.

In any affine model, discount bond prices are of the form

$$P(t, u) = \exp \left(-a(\tau) - \sum_{i=1}^n b_i(\tau) x_i(t) \right) \qquad (14.28)$$

for deterministic functions a and b_i for $i = 1, \ldots, n$, as we have seen is true for the Vasicek, CIR, and Longstaff-Schwartz models. Most, but certainly not all, of the continuous-time models studied in the finance literature are affine.

In any single-factor affine model, the short rate can be used as the factor. Thus, the general affine one-factor model is of the form

$$dr = \kappa(\theta - r) \, dt + \sqrt{\alpha + \beta r} \, dB ,$$

for constants κ, θ, α and β, where B is a Brownian motion under the risk-neutral measure. The Vasicek model is the special case $\beta = 0$ (α being the same as the parameter σ^2). The CIR model is the special case $\alpha = 0$ (β being the same as the parameter σ^2).

---

[4] An affine function of a real variable x is a function $f(x) = a + bx$ for constants a and b. This is often called a linear function, but technically a linear function is of the form $f(x) = bx$. Thus, an affine function is a constant plus a linear function.

Because yields of discount bonds are affine functions of the factors in an affine model, as (14.28) shows, the short rate and yields at $n-1$ fixed times-to-maturity τ_i can be chosen to be the n factors (except in the rare case that the linear transformation from factors to the short rate and yields fails to be invertible). The transformation from factors to yields is analogous to the transformation from (x_1, x_2) to (r, V) in the Longstaff-Schwartz model. Important papers on affine models include Duffie and Kan [24] and Dai and Singleton [22].

14.6 Heath-Jarrow-Morton

Heath, Jarrow and Morton [33] propose an alternative framework for modelling. Rather than modelling the evolution of the short rate (and possibly other factors such as the volatility of the short rate or other yields), Heath, Jarrow and Morton (hereafter, HJM) propose modelling the evolution of instantaneous forward rates. They derive a formula for the drifts of instantaneous forward rates under the risk-neutral measure, in terms of the volatilities of the forward rates. A model is therefore completely defined by specifying the volatilities of forward rates. A model of this type is easily fit to market discount bond prices by simply using the initial term structure of forward rates as an input. By calibrating the volatility structure, the model can also be fit to other market prices. Any of the continuous-time models we have discussed can be written in the HJM form. The virtue of the HJM approach is that it facilitates the construction of new models. The disadvantage of the HJM approach is that models of this form will generally be path-dependent—bond prices and the prices of other fixed-income instruments at any point in time depend on the entire history of the forward rate processes, rather than depending only on the values of a small set of factors. This makes computation quite difficult, just as computation with non-recombining binomial trees is much more difficult than with recombining trees.

As discussed in Sect. 13.6, the forward rate at date t for an instantaneous loan at date $u \geq t$ is $f(t, u)$ defined by

$$f(t, u) = -\frac{\mathrm{d} \log P(t, u)}{\mathrm{d}u} \; . \tag{14.29}$$

The short rate at time t is the forward rate for maturity date t; i.e., $r(t) = f(t, t)$. By integrating (14.29), one can see that discount bond prices are written in terms of forward rates as:

$$P(t, u) = \exp\left(-\int_t^u f(t, s)\, \mathrm{d}s\right) \; . \tag{14.29'}$$

Heath, Jarrow and Morton use the definition (14.29') and the fact that the expected return of a discount bond under the risk-neutral measure must be

the short rate to derive a formula for the drifts of forward rates under the risk-neutral measure.

Assume, for the sake of simplicity, that there is only a single source of uncertainty (i.e., a single Brownian motion) driving the yield curve. Then, for each fixed u, the forward rate $f(t, u)$ at date $t < u$ will evolve as

$$\mathrm{d}f(t, u) = \mu(t, u)\,\mathrm{d}t + \sigma(t, u)\,\mathrm{d}B(t)\,, \tag{14.30}$$

for some μ and σ, where B is a Brownian motion under the risk-neutral measure. In general $\mu(t, u)$ and $\sigma(t, u)$ could depend on the entire history of the Brownian motion through date t. Heath, Jarrow and Morton show that

$$\mu(t, u) = \sigma(t, u) \int_t^u \sigma(t, s)\,\mathrm{d}s\,. \tag{14.31}$$

One can show that $\int_t^u \sigma(t, s)\,\mathrm{d}s$ is the volatility of $P(t, u)$. Therefore, (14.31) states that the drift of the forward rate is the product of the volatilities of the forward rate and the discount bond return. A model is fully specified by specifying initial forward rates—i.e., $f(0, u)$ for all u—and the volatility processes $\sigma(t, u)$ for each u and $t \le u$. The generalization to multiple Brownian motions is straightforward and allows for forward rates that are not instantaneously perfectly correlated. An important application of the HJM modelling framework is work by Brace, Gatarek and Musiela [9] and Miltersen, Sandmann, and Sondermann [52], who derive conditions on the volatility processes $\sigma(t, u)$ that guarantee forward LIBOR rates of a fixed maturity (e.g., quarterly or semi-annual rates) have deterministic volatilities, thus justifying the use of Black's formula in Sect. 12.4 for valuing caps and floors.

To see how some of the models we have discussed can be written in the HJM form, let us re-examine the Hull-White model. From the Hull-White bond price formula (13.14), the instantaneous forward rate in the Hull-White model is

$$
\begin{aligned}
f(t, u) &= \phi(u) + \frac{\partial}{\partial u} a(u - t) + \hat{r}(t) \frac{\partial}{\partial u} b(u - t) \\
&= \phi(u) - \frac{\sigma^2}{2\kappa^2}\left(1 + \mathrm{e}^{-2\kappa(u-t)} - 2\mathrm{e}^{-\kappa(u-t)}\right) + \mathrm{e}^{-\kappa(u-t)}\hat{r}(t)\,.
\end{aligned}
$$

Applying Itô's formula yields

$$\mathrm{d}f(t, u) = -\frac{\sigma^2}{\kappa}\left(\mathrm{e}^{-2\kappa(u-t)} - \mathrm{e}^{-\kappa(u-t)}\right)\mathrm{d}t + \sigma\mathrm{e}^{-\kappa(u-t)}\,\mathrm{d}B\,.$$

Therefore, in the HJM notation,

$$\mu(t, u) = -\frac{\sigma^2}{\kappa}\left(\mathrm{e}^{-2\kappa(u-t)} - \mathrm{e}^{-\kappa(u-t)}\right) \qquad \text{and} \qquad \sigma(t, u) = \sigma\mathrm{e}^{-\kappa(u-t)}\,.$$

A direct calculation shows that these functions μ and σ satisfy the HJM equation (14.31), as we knew they must, given that the HJM equation is

based only on the assumption that expected returns of discount bonds equal the short rate under the risk-neutral measure. The initial forward rate curve in the Hull-White model is $f(0, u) = \phi(u)$ which is chosen to fit the market forward rate curve, as discussed in Sect. 13.6. Thus, rather than defining the Hull-White model as we did in Chap. 13, it could be defined alternatively as an HJM model in which the volatility process is the deterministic function $\sigma(t, u) = \sigma e^{-\kappa(u-t)}$ for positive constants σ and κ. The normal distribution of the short rate in the Hull-White model is a consequence of the volatility $\sigma(t, u)$ being non-random.

Likewise, the Cox-Ingersoll-Ross model fit to market bond prices as we discussed in Sect. 14.4 could be described as an HJM model. Calculations of the sort we have just done show that the volatility process in the CIR model is $\sigma(t, u) = \sigma b'(\tau)\sqrt{\hat{r}(t)}$, where $\tau = u - t$ and the function b is defined in (14.14). In this case, the volatility is random, but it depends only on the short rate at date t. Similarly, in any factor model, such as the Longstaff-Schwartz model, the volatilities $\sigma(t, u)$ of the forward rates will depend only on the factors at each date t. As mentioned at the beginning of this section, such a factor structure simplifies calculations considerably.

HJM models are sometimes written in a slightly different fashion than we have done here. If we define

$$\Sigma(t, u) = \int_t^u \sigma(t, s)\, \mathrm{d}s$$

then we have $\mathrm{d}\Sigma(t, u)/\mathrm{d}u = \sigma(t, u)$, and the HJM equation (14.31) can be written as

$$\mu(t, u) = \Sigma(t, u)\frac{\mathrm{d}\Sigma(t, u)}{\mathrm{d}u}\ ,$$

so the evolution of forward rates can be written as

$$\mathrm{d}f(t, u) = \Sigma(t, u)\frac{\mathrm{d}\Sigma(t, u)}{\mathrm{d}u}\, \mathrm{d}t + \frac{\mathrm{d}\Sigma(t, u)}{\mathrm{d}u}\, \mathrm{d}B(t)\ .$$

For example, in the Hull-White model we have $\Sigma(t, u) = \sigma b(\tau)$ where b is defined in (13.14), and in the CIR model, we have $\Sigma(t, u) = \sigma b(\tau)\sqrt{\hat{r}(t)}$, where b is defined in (14.14).

14.7 Market Models Again

Many fixed-income derivatives (e.g., caps, floors, and swaps) have cash flows that depend on simple interest rates (e.g., LIBOR). In this chapter and the preceding chapter, we discussed valuation formulas for fixed-income derivatives based on (i) models of the short rate or one-period rate and possibly other factors, or (ii) models of instantaneous forward rates. However, it is possible, and simpler for many purposes, to model simple interest rates directly. For example, we observed in Chap. 12 that Black's formula can be

applied to value caps and floors when the underlying simple interest rates have nonrandom volatilities. Models of this type are called "market models" or "LIBOR models" or, sometimes, "BGM models," the last name referring to the paper of Brace, Gatarek and Musiela [9].[5] This class of models has become quite popular in recent years. A thorough and very readable account is given by Rebonato [56].

As in Chap. 12, we will use "LIBOR" as a generic name for simple interest rates. One important fact about forward LIBOR rates that we have already essentially derived is that they are martingales under the corresponding forward measures. As was discussed in Sect. 7.7, a probability measure corresponding to a discount bond being the numeraire is called a "forward measure," because the forward price of any contract maturing at the same time as the discount bond is a martingale under that measure. We showed in Sect. 12.3 that forward LIBOR rates are forward prices of portfolios that pay spot rates. Specifically, considering a LIBOR rate of term (also called "tenor") Δt and the forward LIBOR rate corresponding to loans over a period u to $u + \Delta t$, the forward LIBOR rate is a martingale under the measure corresponding to the discount bond maturing at $u + \Delta t$.

To price derivatives other than caps and floors (e.g., swaptions), it is important to know the dynamics of forward LIBOR rates under other probability measures as well—for example, it is useful to know the dynamics under the forward measures corresponding to discount bonds maturing at dates $T \neq u + \Delta t$, or under the risk-neutral measure, or under the measure that uses as numeraire the portfolio consisting of rolling over an investment at spot LIBOR rates.[6] We will derive here the dynamics under different forward measures.

Consider dates $t_1 < t_2 < \cdots < t_N$ with $t_i - t_{i-1} = \Delta t$ for each i. At dates $t \leq t_i$, we denote the forward LIBOR rate for the time period (t_i, t_{i+1}) by $\mathcal{R}_i(t)$. The forward LIBOR rate satisfies equation (12.2), which we repeat here:

$$\frac{P(t, t_i)}{P(t, t_{i+1})} = 1 + \mathcal{R}_i(t) \, \Delta t \ . \tag{14.32}$$

Fix a date t_n. We will compute the drift of each rate $\mathcal{R}_i(t)$ when we use the discount bond maturing at t_n as the numeraire.

Let $\sigma_i(t)$ denote the volatility of rate $\mathcal{R}_i(t)$ at date t. This means that

$$\frac{\mathrm{d}\mathcal{R}_i(t)}{\mathcal{R}_i(t)} = \mu_i(t) \, \mathrm{d}t + \sigma_i(t) \, \mathrm{d}B_i(t) \ , \tag{14.33}$$

---

[5] Other important work on this topic includes Miltersen, Sandmann, and Sondermann [52] and Jamshidian [43].

[6] Actually, for this theory, it is not even necessary that the short rate exist, so the risk-neutral measure may not even be defined. The risk-neutral measure uses as numeraire the portfolio that consists of continuously rolling over an investment at the instantaneously risk-free rate, and the more natural object in a market model is the portfolio that consists of rolling over an investment at spot LIBOR rates.

for some μ_i, where the B_i are Brownian motions when $P(t, t_n)$ is used as the numeraire. The different rates should be correlated, so the Brownian motions will be correlated. Let ρ_{ij} denote the correlation of B_i and B_j. We will show that the drifts μ_i are determined by the volatilities and correlations, in analogy to the HJM result for instantaneous forward rates. Specifying the volatilities and correlations of forward rates, and inputting initial forward rates, must therefore determine the value of any security whose cash flows depend on the LIBOR rates of term Δt at the dates $t_1, \ldots, t_N$. Of course, this does not mean that there are simple formulas. Obviously, the easiest case is when the volatilities and correlations are nonrandom. In this case, we can use Black's formula to price caps and floors as in Chap. 12. However, even when the σ_i are nonrandom, forward swap rates will have random volatilities, as mentioned in Sect. 12.6.

If $i = n - 1$, then (14.32) implies

$$\mathcal{R}_i(t) = \frac{P(t, t_{n-1}) - P(t, t_n)}{P(t, t_n)\, \Delta t} \, .$$

Hence, it is the ratio of a non-dividend-paying asset (portfolio) price to the price of the numeraire asset. Consequently, it is a martingale, and we have $\mu_{n-1} = 0$. This is the case discussed in the second paragraph of this section. Consider now $i \neq n - 1$.

Define

$$Y(t) = \frac{P(t, t_{i+1})}{P(t, t_n)} \, , \qquad (14.34)$$

and

$$Z(t) = \mathcal{R}_i(t) Y(t) \, . \qquad (14.35)$$

Note that Y is the ratio of a non-dividend-paying asset price to the price of the numeraire asset and hence is a martingale. Furthermore, (14.32) gives us

$$Z(t) = \frac{P(t, t_i) - P(t, t_{i+1})}{P(t, t_n)\, \Delta t} \, ,$$

and hence Z is also the ratio of a non-dividend-paying asset (portfolio) price to the price of the numeraire asset and consequently a martingale. Itô's formula applied to (14.35) yields

$$\frac{\mathrm{d}Z}{Z} = \frac{\mathrm{d}\mathcal{R}_i}{\mathcal{R}_i} + \frac{\mathrm{d}Y}{Y} + \left(\frac{\mathrm{d}\mathcal{R}_i}{\mathcal{R}_i} \right) \left(\frac{\mathrm{d}Y}{Y} \right) \, .$$

Because both Z and Y are martingales, the drift of $\mathrm{d}\mathcal{R}_i / \mathcal{R}_i$ must cancel the product (covariance) term in this equation, implying

$$\mu_i \, \mathrm{d}t = - \left(\frac{\mathrm{d}\mathcal{R}_i}{\mathcal{R}_i} \right) \left(\frac{\mathrm{d}Y}{Y} \right) \, . \qquad (14.36)$$

To compute the covariance, it is helpful to define

$$X_j(t) = 1 + \mathcal{R}_j(t)\,\Delta t \qquad (14.37)$$

for $j = 1, \ldots, N$. Then we have

$$\frac{dX_j}{X_j} = \left(\frac{\mathcal{R}_j\,\Delta t}{1 + \mathcal{R}_j\,\Delta t}\right)\left(\mu_j\,dt + \sigma_j\,dB_j\right). \qquad (14.38)$$

We distinguish two cases. If $i < n - 1$, then the definitions (14.32), (14.34) and (14.37) imply

$$Y(t) = \frac{P(t, t_{i+1})}{P(t, t_{i+2})} \times \frac{P(t, t_{i+2})}{P(t, t_{i+3})} \cdots \times \frac{P(t, t_{n-1})}{P(t, t_n)}$$

$$= X_{i+1}(t) \times X_{i+2}(t) \times \cdots \times X_{n-1}(t).$$

In this case, (14.33) and (14.38) yield

$$\left(\frac{d\mathcal{R}_i}{\mathcal{R}_i}\right)\left(\frac{dY}{Y}\right) = \sum_{j=i+1}^{n-1} \left(\frac{d\mathcal{R}_i}{\mathcal{R}_i}\right)\left(\frac{dX_j}{X_j}\right)$$

$$= \sum_{j=i+1}^{n-1} \left(\frac{\mathcal{R}_j\,\Delta t}{1 + \mathcal{R}_j\,\Delta t}\right)\sigma_i\sigma_j\rho_{ij}\,dt.$$

On the other hand, if $i > n - 1$, then the definitions (14.32), (14.34) and (14.37) imply

$$\frac{1}{Y(t)} = \frac{P(t, t_n)}{P(t, t_{n+1})} \times \frac{P(t, t_{n+1})}{P(t, t_{n+2})} \times \cdots \times \frac{P(t, t_i)}{P(t, t_{i+1})}$$

$$= X_n(t) \times X_{n+1}(t) \times \cdots \times X_i(t).$$

In this case, (14.33) and (14.38) yield

$$\left(\frac{d\mathcal{R}_i}{\mathcal{R}_i}\right)\left(\frac{dY}{Y}\right) = -\sum_{j=n}^{i} \left(\frac{d\mathcal{R}_i}{\mathcal{R}_i}\right)\left(\frac{dX_j}{X_j}\right)$$

$$= -\sum_{j=n}^{i} \left(\frac{\mathcal{R}_j\,\Delta t}{1 + \mathcal{R}_j\,\Delta t}\right)\sigma_i\sigma_j\rho_{ij}\,dt.$$

We conclude:

When we use the discount bond maturing at t_n as the numeraire, the drift of (expected percentage change in) the forward rate $\mathcal{R}_i$ is

$$\mu_i(t) = \begin{cases} \sum_{j=i+1}^{n-1} \left(\frac{\mathcal{R}_j(t)\,\Delta t}{1 + \mathcal{R}_j(t)\,\Delta t}\right)\sigma_i\sigma_j\rho_{ij} & \text{if } i < n - 1, \\ 0 & \text{if } i = n - 1, \\ -\sum_{j=n}^{i} \left(\frac{\mathcal{R}_j(t)\,\Delta t}{1 + \mathcal{R}_j(t)\,\Delta t}\right)\sigma_i\sigma_j\rho_{ij} & \text{if } i > n - 1. \end{cases} \qquad (14.39)$$

Problems

14.1. Create an Excel worksheet demonstrating a four-period Ho-Lee model. Allow the user to input σ, Δt, and $P^{\text{mkt}}(t_n)$ for $n = 1, \ldots, 5$. Compute $\phi(t_n)$ from (14.3) for $n = 1, \ldots, 4$. Create the one-period interest rate tree (starting from $r(0) = -\log P^{\text{mkt}}(t_1)/\Delta t$) and the valuation tree for a discount bond maturing at t_5. Verify that the tree gives the price $P^{\text{mkt}}(t_5)$. Note: to create a binomial tree in a spreadsheet, it is probably easiest to put the topmost (or bottommost) nodes along one row and the other nodes in a triangle below (or above).

14.2. Modify the preceding exercise to include the valuation tree for a caplet with t_4 as its reset date and t_5 as its payment date. Note that the payoff of the caplet at date t_5 is

$$\max\left(0, \mathcal{R}(t_4) - \bar{\mathcal{R}}\right) \Delta t,$$

where

$$\mathcal{R}(t_4) = \frac{1}{P(t_4, t_5)} - 1.$$

Allow the user to input $\bar{\mathcal{R}}$.

14.3. Create a function `Ho_Lee_Caplet` that values a caplet in the Ho-Lee model. Look up market discount bond prices from a function such as `DiscountBondPrice` in Prob. 12.1 and calibrate to the market from (14.3). The inputs to the function should be $\bar{\mathcal{R}}$, σ, T_1, T_2, N_1 and N_2, where T_1 is the reset date for the caplet, T_2 is the payment date for the caplet, N_1 is the number of periods between date 0 and T_1, and N_2 is the number of periods between T_1 and T_2. Note that the payoff of the caplet at date T_2 is

$$\max\left(0, \mathcal{R}(T_1) - \bar{\mathcal{R}}\right) \times (T_2 - T_1),$$

where

$$\mathcal{R}(T_1) = \frac{1}{P(T_1, T_2)} - 1.$$

14.4. Create an Excel worksheet demonstrating a four-period Black-Derman-Toy model. Allow the user to input Δt, $r(0)$, and $\sigma(t_i)$ and $\theta(t_i)$ for $i = 1, \ldots, 4$. Create the one-period interest rate and the valuation tree for a discount bond maturing at t_5.

14.5. Modify the preceding exercise by including the valuation tree for a caplet with reset date t_4 and payment date t_5, as in Prob. 14.2.

14.6. Repeat Probs. 14.4 and 14.5 for the Black-Karasinski model, allowing the user to also input $\kappa(t_i)$ for $i = 1, \ldots, 4$.

14.7. Create a VBA function `CIR_Caplet_MC` that values a caplet in the CIR model using Monte Carlo, without calibrating the model to the current yield curve. Simulate the CIR process as described in Sect. 4.5 for the Heston model. The inputs should be $\mathcal{R}$, $r(0)$, κ, θ, σ, T_1, T_2, N and M, where T_1 is the reset date for the caplet, T_2 is the payment date for the caplet, N is the number of periods between 0 and T_1, and M is the number of simulations. The payoff of the caplet is as in Prob. 14.3, where $P(T_1, T_2)$ is the function of $r(T_1)$ and $T_2 - T_1$ given in (14.18) with $\phi = 0$.

14.8. Modify the function in the preceding exercise to create a function `CIR_Calibrated_Caplet_MC` that values a caplet in the CIR model using Monte Carlo, with the model calibrated to the market. Look up market discount bond prices from a function such as `DiscountBondPrice` in Prob. 12.1. To compute $\phi(t_i)$ for $i = 1, \ldots, N$ from market bond prices, use (14.19) for dates t_i and t_{i+1} as in Sect. 13.8.