Chapter 17

Models of the Term Structure of Interest Rates

The objective of this chapter is to provide an introduction to the main approaches for modeling the term structure of interest rates and for valuing fixed-income derivatives. It is not meant to be a comprehensive review of this subject, as the literature on term structure models is voluminous, and many surveys on this topic, including (Dai and Singleton 2004), (Dai and Singleton 2003), (Maes 2003), (Piazzesi 2003), (Rebonato 2004), and (Yan 2001) have been written in recent years. This chapter discusses valuation models of bonds and bond derivatives that are assumed to be default-free, such as Treasury bills, notes, bonds, and their derivatives. The following chapter analyzes the effects of default risk on bond prices.

The chapter is broken into two main sections. In the first we discuss models that derive the equilibrium bond prices of different maturities in terms of particular state variables. One way to think about these models is that the state variables are the models’ "input" while the bond values of different maturities
CHAPTER 17. TERM STRUCTURE MODELS

are the models' "output." The second section covers models that value fixed-income derivatives, such as interest rate caps and swaptions, in terms a given set of bond prices. In contrast, these models take the term structure of observed bond prices as the input and have derivative values as the models’ output.

17.1 Equilibrium Term Structure Models

Equilibrium term structure models describe the prices (or, equivalently, the yields) of different maturity bonds as functions of one or more state variables or "factors." The Vasicek (Vasicek 1977) model, introduced in Chapter 9 (see equation 9.28), and the Cox, Ingersoll, and Ross (Cox, Ingersoll, and Ross 1985b) model, presented in Chapter 13 (see equation 13.51) were examples of single-factor models. The single factor in the Vasicek model was the instantaneous maturity interest rate, denoted \( r(t) \), which was assumed to follow the Ornstein-Uhlenbeck process (9.17). In Cox, Ingersoll, and Ross’s one-factor model, the factor determined the expected returns of the economy’s production processes, and, in equilibrium, the instantaneous maturity interest rate was proportional to this factor and inherited its dynamics. This interest rate followed the square root process in equation (13.49).

Let us generalize the pricing relationships for default-free zero-coupon bonds by considering a situation where multiple factors determine their prices. The derivation of these bond values will utilize our results from earlier chapters. We start by assuming that there are \( n \) state variables, \( x_i, i = 1, ..., n \), that follow the multi-variate diffusion process

\[
dx = a(t, x) \, dt + b(t, x) \, dz
\]

(17.1)

where \( x = (x_1, ..., x_n)' \), \( a(t, x) \) is a \( nx1 \) vector, \( b(t, x) \) is a \( nxn \) matrix, and
$dz = (dz_1...dz_n)'$ is an $n \times 1$ vector of independent Brownian motion processes so that $dz_i dz_j = 0$ for $i \neq j$. This specification allows for any general correlation structure for the state variables when $b(t,x)$ has full rank.

Define $P(t,T,x)$ as the date $t$ price of a default-free, zero-coupon bond that pays 1 at date $T$. Itô’s lemma gives the process followed by this bond’s price.

$$dP(t,T,x)/P(t,T,x) = \mu_p(t,T,x) dt + \sigma_p(t,T,x)' dz \quad (17.2)$$

where

$$\mu_p(t,T,x) = (a(t,x)' P_x + P_t + 1/2 \text{Trace} [b(t,x) b(t,x)' P_{xx}]) / P(t,T,x) \quad (17.3)$$

and $\sigma_p(t,T,x)$ is an $n \times 1$ vector equal to

$$\sigma_p(t,T,x) = b(t,x)' P_x / P(t,T,r) \quad (17.4)$$

and where $P_x$ is an $n \times 1$ vector whose $i^{th}$ element equals the partial derivative $P_{ix}$, $P_{xx}$ is an $n \times n$ matrix whose $i,j^{th}$ element is second order partial derivative $P_{ixj}$, and Trace$[A]$ is the sum of the diagonal elements of a square matrix $A$.

Similar to the Black-Scholes hedging argument discussed in Chapter 9 and applied to derive the Vasicek model, we can form a hedge portfolio of $n+1$ bonds having distinctly different maturities. By appropriately choosing the portfolio weights for these $n+1$ bonds, the $n$ sources of risk can be hedged so that the portfolio generates a riskless return. In the absence of arbitrage, this portfolio’s return must equal the riskless rate, $r(t,x)$. Making this no-arbitrage restriction produces the implication that each bond’s expected rate of return must satisfy

---

1 As discussed in Chapter 10, the independence assumption is not important. If there are correlated sources of risk (Brownian motions), they can be re-defined by a linear transformation to be represented by $n$ orthogonal risk sources.
\[ \mu_p(t, T, x) = r(t, x) + \Theta(t, x)' \sigma_p(t, T, x) \quad (17.5) \]

where \( \Theta(t, x) = (\theta_1 \ldots \theta_n)' \) is the \( n \times 1 \) vector of market prices of risks associated with each of the Brownian motions in \( dz = (dz_1 \ldots dz_n)' \). By equating (17.5) to the process for \( \mu_p(t, T, x) \) given by Itô’s lemma in (17.3), we obtain the equilibrium partial differential equation

\[ \frac{1}{2} \text{Trace} [b(t, x) b(t, x)' P_{xx}] + [a(t, x) - b(t, x) \Theta]' P_x - rP + P_t = 0 \quad (17.6) \]

Given functional forms for \( a(t, x) \), \( b(t, x) \), \( \Theta(t, x) \), \( r(t, x) \), this PDE can be solved subject to the boundary condition \( P(T, T, x) = 1 \). Note that equation (17.6) depends on the expected changes in the factors under the risk-neutral measure \( Q \), \( a(t, x) - b(t, x) \Theta \), rather than the factors’ expected changes under the physical measure \( P \), \( a(t, x) \). Hence, to price bonds, one could simply specify only the factors’ risk-neutral processes.\(^2\) This insight is not surprising because we saw in Chapter 10 that the Feynman-Kac solution to this PDE is the risk-neutral pricing equation (10.58).

\[ P(t, T, x) = \hat{E}_t \left[ e^{-\int_t^T r(s, x)ds} \right] \quad (17.7) \]

In addition to the pricing relations (17.6) and (17.7), we saw that a third pricing approach can be based on the pricing kernel that follows the process

\[ \frac{dM}{M} = -r(t, x) dt - \Theta(t, x)' dz \quad (17.8) \]

In this case, pricing can be accomplished under the physical measure based on

\(^2\) However, if the factors are observative variables for which data are available, it may be necessary to specify their physical processes if empirical implementations of the model require estimates for \( a(x, t) \) and \( b(x, t) \).
the formula
\[ P(t, T, x) = E_t \left[ \frac{M(T)}{M(t)} \right] \] (17.9)

Thus far, we have placed few restrictions on the factors and their relationship to the short rate, \( r(t, x) \), other than to assume that the factors follow the Markov diffusion processes (17.1). Let us next consider some popular parametric forms.

### 17.1.1 Affine Models

We start with models in which the yields of zero coupon bonds are linear or "affine" functions of state variables. This class of models includes those of Vasicek (Vasicek 1977) and Cox, Ingersoll, and Ross (Cox, Ingersoll, and Ross 1985b). Affine models are attractive because they lead to bond price formulae that are relatively easy to compute and because the parameters of the state variable processes can often be estimated using relatively straightforward econometric techniques.

Recall that a zero-coupon bond’s continuously compounded yield, \( Y(t, T, x) \), is defined from its price by the relation
\[ P(t, T, x) = e^{-Y(t, T, x)(T-t)} \] (17.10)

One popular class of models assumes that zero-coupon bonds’ continuously-compounded yields are affine functions of the factors. Defining the time until maturity as \( \tau = T - t \), this assumption can be written as
\[ Y(t, T, x) \tau = A(\tau) + B(\tau)'x \] (17.11)

where \( A(\tau) \) is a scalar function and \( B(\tau) \) is an \( nx1 \) vector of functions that do not depend on the factors, \( x \). Because at maturity \( P(T, T, x) = 1 \), (17.11) implies that \( A(0) = 0 \) and \( B(0) \) is an \( nx1 \) vector of zeros. Another implication
of (17.11) is that the short rate is also affine in the factors since

\[ r(t, x) = \lim_{T \to t} Y(t, T, x) = \lim_{\tau \downarrow 0} \frac{A(\tau) + B(\tau)'x}{\tau} \tag{17.12} \]

so that we can write \( r(t, x) = \alpha + \beta'x \), where \( \alpha = \partial A(0)/\partial \tau \) is a scalar and \( \beta = \partial B(0)/\partial \tau \) is an \( nx1 \) vector of constants.

Under what conditions regarding the factors’ dynamics would the no-arbitrage, equilibrium bond yields be affine in the state variables? To answer this, let us substitute the affine yield assumption of (17.10) and (17.11) into the general no-arbitrage PDE of (17.6). Doing so, one obtains

\[
\frac{1}{2}B(\tau)'b(t, x)b(t, x)'B(\tau) - [a(t, x) - b(t, x)\Theta]'B(\tau) \\
+ \frac{\partial A(\tau)}{\partial \tau} + \frac{\partial B(\tau)'}{\partial \tau}x = \alpha + \beta'x
\tag{17.13}
\]

Darrell Duffie and Rui Kan (Duffie and Kan 1996) characterize sufficient conditions for a solution to equation (17.13). Specifically, two of the conditions are that the factors’ risk-neutral instantaneous expected changes and variances are affine in \( x \). In other words, if the state variables’ risk-neutral drifts and variances are affine in the state variables, so will the equilibrium bond price yields. These conditions can be written as

\[
a(t, x) - b(t, x)\Theta = \kappa(x - \bar{x}) \tag{17.14}
\]

\[
b(t, x) = \Sigma \sqrt{s(x)} \tag{17.15}
\]

where \( \bar{x} \) is an \( nx1 \) vector of constants, \( \kappa \) and \( \Sigma \) are \( nxn \) matrices of constants, and \( s(x) \) is an \( nxn \) diagonal matrix with the \( i^{th} \) diagonal term

\[
s_i(x) = s_{ii} + s_i'x \tag{17.16}
\]
where \( s_{oi} \) is a scalar constant and \( s_{1i} \) is an \( nx1 \) vector of constants. Now, because the state variables’ covariance matrix equals \( b(t, x) b(t, x)' = \Sigma s(x) \Sigma' \), additional conditions are needed to ensure that this covariance matrix remains positive definite for all possible realizations of the state variable, \( x \). Qiang Dai and Kenneth Singleton (Dai and Singleton 2000) and Darrell Duffie, Damir Filipovic, and Walter Schachermayer (Duffie, Filipovic, and Schachermayer 2002) derive these conditions.\(^3\)

Given (17.14), (17.15), and (17.16), the partial differential equation in (17.13) can be re-written as

\[
\frac{1}{2} B(\tau)' \Sigma s(x) \Sigma' B(\tau) - [\kappa(x-x)'] B(\tau) + \frac{\partial A(\tau)}{\partial \tau} + \frac{\partial B(\tau)'}{\partial \tau} = \alpha + \beta' x
\] (17.17)

Note that this equation is linear in the state variables, \( x \). For the equation to hold for all values of \( x \), the constant terms in the equation must sum to zero and the terms multiplying each element of \( x \) must also sum to zero. These conditions imply

\[
\frac{\partial A(\tau)}{\partial \tau} = \alpha + (\kappa x)' B(\tau) - \frac{1}{2} \sum_{i=1}^{n} [\Sigma' B(\tau)]_{i}^{2} s_{0i}
\] (17.18)

\[
\frac{\partial B(\tau)'}{\partial \tau} = \beta - \kappa' B(\tau) - \frac{1}{2} \sum_{i=1}^{n} [\Sigma' B(\tau)]_{i}^{2} s_{1i}
\] (17.19)

where \( [\Sigma' B(\tau)]_{i} \) is the \( i^{th} \) element of the \( nx1 \) vector \( \Sigma' B(\tau) \). Equations (17.18) and (17.19) are a system of first order ordinary differential equations that can

\(^3\)These conditions can have important consequences regarding the correlation between the state variables. For example, if the state variables follow a multivariate Ornstein Uhlenbeck process, so that the model is a multifactor version of the Vasicek model given in (9.28), (9.29), and (9.30), then any general correlation structure between the state variables is permitted. However, if the state variables follow a multivariate square-root process, so that the model is a multifactor version of the Cox, Ingersoll and Ross model given in (13.51), (13.52), and (13.53), then the correlation between the state variables must be non-negative.
be solved subject to the boundary conditions \( A(0) = 0 \) and \( B(0) = 0 \). In some cases, such as a multiple state variable version of the Vasicek model (where \( s_{1i} = 0 \) \( \forall i \)), there exist closed-form solutions. In other cases, fast and accurate numerical solutions to these ordinary differential equations can be computed using techniques such as a Runge Kutta algorithm.

While affine term structure models require that the state variables’ risk-neutral expected changes be affine in the state variables, there is more flexibility regarding the state variables’ drifts under the physical measure. Note that the state variables’ expected change under the physical measure is

\[
a(t, x) = \kappa (\bar{x} - x) + \Sigma \sqrt{s(x)} \Theta
\]  

(17.20)

so that specification of the market prices of risk, \( \Theta \), is required to determine the physical drifts of the state variables. Qiang Dai and Kenneth Singleton (Dai and Singleton 2000) study the "completely affine" case where both the physical and risk-neutral drifts are affine, while Gregory Duffee (Duffee 2002) and Jefferson Duarte (Duarte 2004) consider extensions of the physical drifts that permit nonlinearities. Because the means, volatilities, and risk premia of bond prices estimated from time series data depend on the physical moments of the state variables, the flexibility in choosing the parametric form for \( \Theta \) can allow the model to better fit historical bond price data.

---

4Dai and Singleton analyze \( \Theta = \sqrt{s(x)} \lambda_1 \) where \( \lambda_1 \) is an \( nx1 \) vector of constants. Duffee considers the "essentially affine" modeling of the market price of risk of the form \( \Theta = \sqrt{s(x)} \lambda_1 + \sqrt{s(x)} \lambda_2 x \), where \( s(x) \) is an \( nxn \) diagonal matrix whose \( i^{th} \) element equals \( (s_{0i} + s'_{1i} x)^{-1} \) if \( (s_{0i} + s'_{1i} x) > 0 \), zero otherwise, and \( \lambda_2 \) is an \( nxn \) matrix of constants. This specification allows time variation in the market prices of risk for Gaussian state variables (such as state variables that follow Ornstein-Uhlenbeck processes), allowing their signs to switch over time. Duarte extend Duffee’s modeling to add a square root term. This "semi affine square-root" model takes the form \( \Theta = \Sigma^{-1} \lambda_0 + \sqrt{s(x)} \lambda_1 + \sqrt{s(x)} \lambda_2 x \) where \( \lambda_2 \) is an \( nx1 \) vector of constants. See, also, work by Patrick Cheridito, Damir Filipovic, and Robert Kimmel (Cheridito, Filipovic, and Kimmel 2003) for extensions in modeling the market price of risk for affine models.
Another class of models assumes that the yields of zero coupon bonds are quadratic functions of state variables. This assumption can be expressed as

\[ Y(t, T, x) = A(\tau) + B(\tau)^\prime x + x^\prime C(\tau) x \quad (17.21) \]

where \( C(\tau) \) is an \( n \times n \) matrix, and, with no loss of generality, can be assumed to be symmetric. Similar to our analysis of affine models, since \( P(T, T, x) = 1 \), we must have \( A(0) = 0 \), \( B(0) \) equal to an \( n \times 1 \) vector of zeros, and \( C(0) \) equal to an \( n \times n \) matrix of zeros. In addition, the yield on a bond of instantaneous maturity must be of the form

\[ r(t, x) = \alpha + \beta^\prime x + x^\prime \gamma x \]

where \( \alpha = \partial A(0) / \partial \tau \), \( \beta = \partial B(0) / \partial \tau \), and \( \gamma = \partial C(0) / \partial \tau \) is an \( n \times n \) symmetric matrix of constants. Note that if \( \gamma \) is a positive semi-definite matrix and \( \alpha - \frac{1}{4} \beta^\prime \gamma^{-1} \beta \geq 0 \), then the interest rate can be restricted from becoming negative.\(^5\) Substituting \( P(t, T, x) = \exp \left( -A(\tau) - B(\tau)^\prime x - x^\prime C(\tau) x \right) \) into the general partial differential equation (17.6), we obtain

\[
\frac{1}{2} \left[ [B(\tau) + 2C(\tau) x]^\prime \left( b(t, x) - b(t, x) \Theta \right) + b(t, x)^\prime [B(\tau) + 2C(\tau) x] \right] - \text{Trace} \left[ b(t, x)^\prime C(\tau) b(t, x) \right] - |a(t, x) - b(t, x) \Theta| \left[ B(\tau) + 2C(\tau) x \right] + \frac{\partial A(\tau)}{\partial \tau} x + x^\prime \frac{\partial C(\tau)}{\partial \tau} x
\]

\[
= \alpha + \beta^\prime x + x^\prime \gamma x \quad (17.22)
\]

In addition to yields being quadratic in the state variables, quadratic-gaussian models then assume that the state variable, \( x \), has a multivariate normal (Gaussian) distribution.

\(^5\) The lower bound for \( r(t) \) is \( \alpha - \frac{1}{4} \beta^\prime \gamma^{-1} \beta \), which occurs when \( x = -\frac{1}{2} \gamma^{-1} \beta \).
distribution. Specifically, it is assumed that $\mathbf{x}$ follows a multivariate Ornstein-Uhlenbeck process:

$$a(t, \mathbf{x}) - b(t, \mathbf{x}) \Theta = \kappa (\mathbf{x} - \mathbf{x}) \quad (17.23)$$

$$b(t, \mathbf{x}) = \Sigma \quad (17.24)$$

Substituting these assumptions into the partial differential equation (17.22), one obtains

$$\frac{1}{2} \left[ [B(\tau) + 2C(\tau) \mathbf{x}]' \Sigma \Sigma' [B(\tau) + 2C(\tau) \mathbf{x}] \right]$$

$$- \text{Trace} \left[ \Sigma' C(\tau) \Sigma \right] - [\kappa (\mathbf{x} - \mathbf{x})]' [B(\tau) + 2C(\tau) \mathbf{x}]$$

$$+ \frac{\partial A(\tau)}{\partial \tau} + \frac{\partial B(\tau)}{\partial \tau} \mathbf{x} + \mathbf{x}' \frac{\partial C(\tau)}{\partial \tau} \mathbf{x}$$

$$= \alpha + \beta' \mathbf{x} + \gamma \mathbf{x} \quad (17.25)$$

For this equation to hold for all values of $\mathbf{x}$, it must be the case that the sums of the equation’s constant terms, the terms proportional to the elements of $\mathbf{x}$, and the terms that are products of the elements of $\mathbf{x}$ must each equal zero. This leads to the system of first order ordinary differential equations

$$\frac{\partial A(\tau)}{\partial \tau} = \alpha + (\kappa \mathbf{x})' B(\tau) - \frac{1}{2} B(\tau)' \Sigma \Sigma' B(\tau) + \text{Trace} \left[ \Sigma' C(\tau) \Sigma \right] \quad (17.26)$$

$$\frac{\partial B(\tau)}{\partial \tau} = \beta - \kappa' B(\tau) - 2C(\tau)' \Sigma \Sigma B(\tau) + 2C(\tau)' \kappa \mathbf{x} \quad (17.27)$$

$$\frac{\partial C(\tau)}{\partial \tau} = \gamma - 2\kappa' C(\tau) - 2C(\tau)' \Sigma \Sigma' C(\tau) \quad (17.28)$$

which are solved subject to the aforementioned boundary conditions, $A(0) = 0$, $B(0) = 0$, and $C(0) = 0$. Dong-Hyun Ahn, Robert Dittmar, and Ronald
17.1. EQUILIBRIUM TERM STRUCTURE MODELS

Gallant (Ahn, Dittmar, and Gallant 2002) show that the models of Francis Longstaff (Longstaff 1989), David Beaglehole and Mark Tenney (Beaglehole and Tenney 1992), and George Constantinides (Constantinides 1992) are special cases of quadratic-gaussian models. They also demonstrate that since quadratic-gaussian models allow a non-linear relationship between yields and state variables, these models can outperform affine models in explaining historical bond yield data.

However, quadratic-gaussian models are more difficult to estimate from historical data because, unlike affine models, there is not a one-to-one mapping between bond yields and the elements of the vector of state variables. For example, suppose that at a given point in time we observed bond yields of \( n \) different maturities, say \( Y(t, T_i, x), i = 1, \ldots, n \). Denoting \( \tau_i = T_i - t \), if yields are affine functions of the state variables then

\[
Y(t, T_i, x) \tau_i = A(\tau_i) + B(\tau_i)'x,
\]

\( i = 1, \ldots, n \), represent a set of \( n \) linear equations in the \( n \) elements of the state variable \( x \). Solving these equations for the state variables \( x_1, x_2, \ldots, x_n \) effectively allows one to observe the individual state variables from the observed yields. By observing a time-series of these state variables, the parameters of their physical process could be estimated.

This approach cannot be used when yields are quadratic functions of the state variables since with

\[
Y(t, T_i, x) \tau_i = A(\tau_i) + B(\tau_i)'x + x'C(\tau_i)x
\]

there is not a one-to-one mapping between yields and state variables \( x_1, x_2, \ldots, x_n \). There are multiple values of the state variable vector, \( x \), consistent with the set of yields.\(^6\) This difficulty requires an different approach to inferring the most likely state variable vector. Ahn, Dittmar, and Gallant use an efficient method of moments technique that simulates the state variable, \( x \), to estimate the state variable vector that best fits the data.

\(^6\)For example, if \( n = 1 \), there are two state variable roots of the quadratic yield equation.
17.1.3 Other Equilibrium Models

Term structure models have been modified to allow state variable processes to differ from strict diffusions. Such models can no longer rely on the Black-Scholes hedging argument to identify market prices of risk and a risk-neutral pricing measure. Because fixed-income markets may not be dynamically complete, these models need to make additional assumptions regarding the market prices of risks that cannot be hedged.

A number of researchers, including Chang-Mo Ahn and Howard Thompson (Ahn and Thompson 1988), Sanjiv Das and Silverio Foresi (Das and Foresi 1996), Darrell Duffie, Jun Pan, and Kenneth Singleton (Duffie, Pan, and Singleton 2000), Sanjiv Das (Das 2002), and George Chacko and Sanjiv Das (Chacko and Das 2002), have extended equilibrium models to allow state variables to follow jump-diffusion processes. An interesting application of a model with jumps in a short-term interest rate is presented by Monika Piazzesi (Piazzesi 2005) who studies changes in the federal funds rate made by the Federal Reserve.

Other affine equilibrium models have been set in discrete-time, where the assumed existence of a discrete-time pricing kernel allows one to find solutions for equilibrium bond prices that have a recursive structure. Examples of models of this type include work by Tong-Sheng Sun (Sun 1992), David Backus and Stanley Zin (Backus and Zin 1994), and V. Cvsa and Peter Ritchken (Cvsa and Ritchken 2001). Term structure models also have been generalized to include discrete regime shifts in the processes followed by state variables. See work by Vasant Naik and Moon Hoe Lee (Naik and Lee 1997) and Ravi Bansal and Hao Zhou (Bansal and Zhou 2002) for models of this type.

Let us now turn to term structure models whose primary purpose is not to determine the term structure of zero coupon bond prices as a function of
state variables. Rather, their objective is to determine the value of fixed-income derivatives as a function of a given term structure of bond prices.

17.2 Valuation Models for Interest Rate Derivatives

Models for valuing bonds and bond derivatives have different uses. The equilibrium models of the previous section can provide insights as to the nature of term structure movements. They allow us to predict how factor dynamics will affect the prices of bonds of different maturities. Equilibrium models may also be of practical use to bond traders who wish to identify bonds of particular maturities whose market valuations appear to be over- or under-priced based on their predicted model prices. Such information could suggest profitable bond trading strategies.

However, bond prices are modeled for other objectives, such as the pricing of derivatives whose payoffs depend on the future prices of bonds or yields. Equilibrium models may be less than satisfactory for this purpose because it is bond derivatives, not the underlying bond prices themselves, that one wishes to value. In this context, one would like to use observed market prices for bonds as an input into the valuation formulas for derivatives, not model the value of the underlying bonds themselves. For such a derivative-pricing exercise, one would like the model to "fit," or be consistent with, the market prices of the underlying bonds.

In a discrete-time, binomial model setting, Thomas Ho and Sang Bin Lee (Ho and Lee 1986) first introduced the concept of pricing fixed-income derivatives by taking the initial term structure of bond prices as given and then making assumptions regarding the risk-neutral distribution of future interest
CHAPTER 17. TERM STRUCTURE MODELS

rates. This binomial approach was modified for different risk-neutral interest rate distributions by Fischer Black, Emanuel Derman, and William Toy (Black, Derman, and Toy 1990) and Fischer Black and Piotr Karasinski (Black and Karasinski 1991). While these binomial models are not covered in this chapter, we will discuss a similar approach that is set in continuous time. It shares the characteristics of fitting the currently observed market prices of bonds and of making particular assumptions regarding the risk-neutral distribution of future interest rates.

17.2.1 Heath-Jarrow-Morton Models

The approach by David Heath, Robert Jarrow, and Andrew Morton ((Heath, Jarrow, and Morton 1992)), hereafter referred to as HJM, differs from the previous equilibrium term structure models because it does not begin by specifying a set of state variables, $x$, that determines the current term structure of bond prices. Rather, their approach takes the initial term structure of bond prices as given (observed) and then specifies how this term structure evolves in the future in order to value derivatives whose payoffs depend on future term structures. Because models of this type do not derive the term structure from more basic state variables, they cannot provide insights regarding how economic fundamentals determine the maturity structure of zero-coupon bond prices. Instead, HJM models are used to value fixed-income derivative securities: securities such as bond and interest rate options whose payoffs depend on future bond prices or yields.

An analogy to the HJM approach can be drawn from the risk-neutral valuation of equity options. Recall that in Chapter 10 equation (10.47), we assumed that the risk-neutral process for the price of a stock, $S(t)$, was geometric Brownian motion, making this price lognormally distributed under the risk-neutral
measure. From this assumption, and given the initial price of the stock, $S(t)$, the Black-Scholes formula for the value of a call option written on this stock was derived in equations (10.51) and (10.52). Note that we did not attempt to determine the initial value of the stock in terms of some fundamental state variables, say $S(t, x)$. Rather, the initial stock price, $S(t)$, was taken as given and an assumption about this stock price’s volatility, namely that it was constant over time, was made.

The HJM approach to valuing fixed-income derivatives is similar but slightly more complex because it takes as given the entire initial term structure of bond prices, $P(t, T) \forall T \geq 1$, not just a single asset price. It then assumes risk-neutral processes for how this initial set of bond prices change over time and does not attempt to derive these initial prices in terms of state variables, say $P(t, T, x)$. However, the way that HJM specify the processes followed by bond prices is somewhat indirect. They begin by specifying processes for bond forward rates. A fundamental result of the HJM analysis is to show that, in the absence of arbitrage, there must be a particular relationship between the drift and volatility parameters of forward rate processes and that only an assumption regarding the form of forward rate volatilities is needed for pricing derivatives.

Let us start by defining forward rates. Recall from Chapter 7 that a forward contract is an agreement between two parties where the long (short) party agrees to purchase (deliver) an underlying asset in return for paying (receiving) the forward price. Consider a forward contract agreed to at date $t$, where the contract matures at date $T \geq t$ and the underlying asset is a zero-coupon bond that matures at date $T + \tau$ where $\tau \geq 0$. Let $F(t, T, \tau)$ be the equilibrium forward price agreed to by the parties. Then this contract requires the long party to pay $F(t, T, \tau)$ at date $T$ in return for receiving a cashflow of $\$1$ (the zero-coupon bond’s maturity value) at date $T + \tau$. In the absence of arbitrage,
CHAPTER 17. TERM STRUCTURE MODELS

the value of these two cashflows at date \( t \) must sum to zero, implying

\[
-F(t, T, \tau) P(t, T) + P(t, T + \tau) = 0
\]  

(17.29)

so that the equilibrium forward price equals the ratio of the bond prices maturing at dates \( T + \tau \) and \( T \), \( F(t, T, \tau) = \frac{P(t, T + \tau)}{P(t, T)} \). From this forward price a continuously-compounded forward rate, \( f(t, T, \tau) \), is defined as

\[
e^{-f(t, T, \tau)\tau} \equiv F(t, T, \tau) = \frac{P(t, T + \tau)}{P(t, T)}
\]  

(17.30)

\( f(t, T, \tau) = -\frac{\ln[P(t, T + \tau)/P(t, T)]}{\tau} \) is the implicit per-period rate of return (interest rate) that the long party earns by investing \( $F(t, T, T + \tau) \) at date \( T \) and by receiving \( $1 \) at date \( T + \tau \). Now consider the case of such a forward contract where the underlying bond matures very shortly (e.g., the next day or instant) after the maturity of the forward contract. This permits us to define an instantaneous forward rate as

\[
f(t, T) \equiv \lim_{\tau \downarrow 0} f(t, T, \tau) = \lim_{\tau \downarrow 0} -\frac{\ln[P(t, T + \tau)] - \ln[P(t, T)]}{\tau} = -\frac{\partial \ln[P(t, T)]}{\partial T}
\]  

(17.31)

Equation (17.31) is a simple differential equation that can be solved to obtain

\[
P(t, T) = e^{-\int_t^T f(t, s) \, ds}
\]  

(17.32)

Since this bond’s continuously-compounded yield to maturity is defined from the relation \( P(t, T) = e^{-Y(t, T)(T-t)} \), we can write \( Y(t, T) = \frac{1}{T-t} \int_t^T f(t, s) \, ds \). Thus, a bond’s yield equals the average of the instantaneous forward rates for horizons out to the bond’s maturity. In particular, the yield on an instantaneous maturity bond is given by \( r(t) = f(t, t) \).
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

Because the term structure of instantaneous forward rates, \( f(t, T) \forall T \geq t \), can be determined from the term structure of bond prices, \( P(t, T) \forall T \geq t \), or yields, \( Y(t, T) \forall T \geq t \), specifying the evolution of forward rates over time is equivalent to specifying the dynamics of bond prices. HJM assume that forward rates for all horizons are driven by a finite-dimensional Brownian motion:

\[
df(t, T) = \alpha(t, T) dt + \sigma(t, T) dz
\]

(17.33)

where \( \sigma(t, T) \) is an \( nx1 \) vector of volatility functions and \( dz \) is an \( nx1 \) vector of independent Brownian motions. Note that since there are an infinite number of instantaneous forward rates, one for each future horizon, equation (17.33) represents infinitely many processes that are driven by the same \( n \) Brownian motions.

Importantly, the absence of arbitrage places restrictions on \( \alpha(t, T) \) and \( \sigma(t, T) \). To show this, let us start by deriving the process followed by bond prices, \( P(t, T) \), implied by the forward rate processes. Note that since \( \ln[P(t, T)] = - \int_t^T f(t, s) ds \), if we differentiate with respect to date \( t \), we find that the process followed by the log bond price is

\[
d\ln[P(t, T)] = f(t, t) dt - \int_t^T df(t, s) ds
\]

(17.34)

\[
= r(t) dt - \int_t^T \left[ \alpha(t, s) dt + \sigma(t, s)' dz(t) \right] ds
\]

Fubini’s Theorem allows us to switch the order of integration:

\[
d\ln[P(t, T)] = r(t) dt - \int_t^T \alpha(t, s) ds dt - \int_t^T \sigma(t, s)' dz(t)
\]

(17.35)

\[
= r(t) dt - \alpha_I(t, T) dt - \sigma_I(t, T)' dz(t)
\]

where we have used the shorthand notation \( \alpha_I(t, T) \equiv \int_t^T \alpha(t, s) ds \) and \( \sigma_I(t, T) \equiv \)
\[ \int_t^T \sigma(t, s) \, ds \] to designate these integrals which are known functions as of date \( t \). Using Itô’s lemma we can derive the bond’s rate of return process from the log process in (17.35).

\[
\frac{dP(t, T)}{P(t, T)} = \left[ r(t) - \alpha_I(t, T) + \frac{1}{2} \sigma_I(t, T)' \sigma_I(t, T) \right] \, dt - \sigma_I(t, T)' \, dz. \tag{17.36}
\]

Now recall from (17.5) that the absence of arbitrage requires that the bond’s expected rate of return equal the instantaneous risk-free return plus the product of the bond’s volatilities and the market prices of risk. This is written as

\[
r(t) - \alpha_I(t, T) + \frac{1}{2} \sigma_I(t, T)' \sigma_I(t, T) = r(t) - \Theta(t)' \sigma_I(t, T) \tag{17.37}
\]

or

\[
\alpha_I(t, T) = \frac{1}{2} \sigma_I(t, T)' \sigma_I(t, T) + \Theta(t)' \sigma_I(t, T) \tag{17.38}
\]

Equation (17.36) and (17.38) shows that the bond price process depends only on the instantaneous risk-free rate, the volatilities of the forward rates, and the market prices of risk. This no-arbitrage condition also has implications for the risk-neutral process followed by forward rates. If we substitute \( dz = d\tilde{z} - \Theta(t) \, dt \) in (17.33), we obtain

\[
df(t, T) = [\alpha(t, T) - \sigma(t, T)' \Theta(t)] \, dt + \sigma(t, T)' \, d\tilde{z}
\]

\[
= \hat{\alpha}(t, T) \, dt + \sigma(t, T)' \, d\tilde{z} \tag{17.39}
\]

where \( \hat{\alpha}(t, T) \equiv \alpha(t, T) - \sigma(t, T)' \Theta(t) \) is the risk-neutral drift observed at date \( t \) for the forward rate at date \( T \). Define \( \hat{\alpha}_I(t, T) \equiv \int_t^T \hat{\alpha}(t, s) \, ds \) as the integral over the drifts of all forward rates from date \( t \) to date \( T \). Then using
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

(17.38) we have

$$\hat{\alpha}_I (t, T) = \int_t^T \hat{\alpha} (s, s) \, ds = \int_t^T \alpha (s, s) \, ds - \int_t^T \sigma (s, s) \, d\Theta (t)$$

$$= \alpha_I (t, T) - \Theta (t) \sigma_I (t, T)$$

$$= \frac{1}{2} \sigma_I (t, T) \sigma_I (t, T) + \Theta (t) \sigma_I (t, T) - \Theta (t) \sigma_I (t, T)$$

$$= \frac{1}{2} \sigma_I (t, T) \sigma_I (t, T)$$

(17.40)

or $\int_t^T \hat{\alpha} (s, s) \, ds = \frac{1}{2} \left( \int_t^T \sigma (s, s) \, ds \right) \left( \int_t^T \sigma (t, s) \, ds \right).$ This shows that in the absence of arbitrage, the risk-neutral drifts of forward rates are completely determined by their volatilities. Indeed, if we differentiate $\hat{\alpha}_I (t, T)$ with respect to $T$ to recover $\hat{\alpha} (t, T)$ we obtain

$$df (t, T) = \sigma (t, T) \sigma_I (t, T) \, dt + \sigma (t, T) d\tilde{w}$$

(17.41)

Equation (17.41) has an important implication, namely, that if we want to model the risk-neutral dynamics of forward rates in order to price fixed-income derivatives, we only need specify the form of the forward rates’ volatility functions.\footnote{In general, these volatility functions may be stochastic, as they could be specified to depend on current levels of the forward rates, that is, $\sigma (t, T, f (t, T))$.} One can also use (17.41) to derive the risk-neutral dynamics of the instantaneous maturity interest rate, $r (t) = f (t, t)$, which is required for discounting risk-neutral payoffs. Suppose dates are ordered such that $0 \leq t \leq T$.

In integrated form (17.41) becomes

$$f (t, T) = f (0, T) + \int_0^t \sigma (u, T) \sigma_I (u, T) \, du + \int_0^t \sigma (u, T) d\tilde{w} (u)$$

(17.42)
and for \( r(t) = f(t, t) \) this becomes

\[
  r(t) = f(0, t) + \int_0^t \sigma(u, t)' \sigma_I(u, t) du + \int_0^t \sigma(u, t)' d\tilde{z}(u) \tag{17.43}
\]

Differentiating with respect to \( t \) leads to\(^8\)

\[
  dr(t) = \frac{\partial f(0, t)}{\partial t} dt + \sigma(t, t)' \sigma_I(t, t) dt + \int_0^t \frac{\partial \sigma(u, t)'}{\partial t} \sigma_I(u, t) du \, dt + \sigma(t, t)' d\tilde{z}(u) \, dt + \int_0^t \sigma(u, t)' d\tilde{z}(u) \, dt + \sigma(t, t)' d\tilde{z} \tag{17.44}
\]

where we have used the fact that \( \sigma_I(t, t) = 0 \) and \( \partial \sigma_I(u, t) / \partial t = \sigma(u, t) \).

With these results, one can now value fixed-income derivatives. As an example, define \( C(t) \) as the current date \( t \) price of a European-type contingent claim that has a payoff at date \( T \). This payoff is assumed to depend on the forward rate curve (equivalently, the term structure of bond prices or yields) at date \( T \), which we write as \( C(T, f(T, T + \delta)) \) where \( \delta \geq 0 \). The contingent claim’s risk-neutral valuation equation is

\[
  C(t, f(t, t + \delta)) = \mathbb{E}_t \left[ e^{-\int_t^T r(s) ds} C(T, f(T, T + \delta)) \mid f(t, t + \delta), \forall \delta \geq 0 \right] \tag{17.45}
\]

where the expectation is conditioned on information of the current date \( t \) forward rate curve, \( f(t, t + \delta) \) \( \forall \delta \geq 0 \). Equation (17.45) is the risk-neutral expectation of the claim’s discounted payoff, conditional on information of all currently-observed forward rates. In this manner, the contingent claim’s formula can be

\(^8\)Note that the dynamics of \( dr \) are more complicated than simply setting \( T = t \) in equation (17.41) because both arguments of \( f(t, t) = r(t) \) are varying simultaneously. Equation (17.44) is equivalent to \( dr = df(t, t) + \frac{\partial f(u, \omega)}{\partial u} |_{u=t} dt \).
assured of fitting the current term structure of interest rates, since the forward rate curve, $f(t, t + \delta)$, is an input. Only for special cases regarding the type of contingent claim and the assumed forward rate volatilities can the expectation in (17.45) be computed analytically. In general, it can be computed by a Monte Carlo simulation of a discrete-time analogue to the continuous-time risk-neutral forward rate and instantaneous interest rate processes in (17.41) and (17.44).9

Valuing American-type contingent claims using the HJM approach can be more complicated because, in general, one needs to discretize forward rates to produce a lattice (e.g., binomial tree) and check the nodes of the lattice to see if early exercise is optimal.10 However, HJM forward rates will not necessarily follow Markov processes. From (17.41) and (17.44), one see that if the forward rate volatility functions are specified to depend on the level of forward rates themselves, $\sigma(t, s, f(t, s))$, then the evolution of $f(t, T)$ and $r(t)$ depends on the entire history of forward rates between two dates such as 0 and $t$. It will not be possible to express forward rates as $f(0, T, x(0))$ and $f(t, T, x(t))$ where $x(t)$ is a set of finite state variables.11 Non-Markov processes lead to lattice structures where the nodes do not re-combine. This can make computation extremely time consuming because the number of nodes grows exponentially (rather than linearly in the case of recombining nodes) with the number of time steps. Hence, to value American contingent claims using the HJM framework, it is highly desirable to pick volatility structures that lead to forward rate processes that are Markov.12 The following is a simple example of an HJM model where

---

9 An example is presented by Kaushik Amin and Andrew Morton (Amin and Morton 1994). They value Eurodollar futures and options assuming different one-factor ($n = 1$) specifications for forward rate volatilities. Their models are nested in the functional form $\sigma(t, T) = [\sigma_0 + \sigma_1(T - t)]e^{-\alpha(T - t)}f(t, T)\gamma$.

10 Recall that this method was used in Chapter 7 to value an American option.

11 The reason why one may want to assume that forward rate volatilities depend on their own level is to preclude negative forward rates, a necessary condition if currency is not to dominate bonds in a nominal term structure model. For example, similar to the square-root model of Cox, Ingersoll, and Ross, one could specify $\sigma(t, T) = \sigma f(t, T)^{\delta}$.

12 Note, also, that non-Markov short-rate and forward rate process imply that contingent claims cannot be valued by solving an equilibrium partial differential equation, such as was
forward rate volatilities are deterministic and lead to a Markov process for the instantaneous interest rate.

**Example: Extended Vasicek**

Consider the value of a European option maturing at date $T$ where the underlying asset is a zero coupon bond maturing at date $T + \tau$. It is assumed that $n = 1$ and the forward rates’ volatilities decline exponentially with their time horizons:

$$\sigma(t, T) = \sigma_r e^{-\alpha(T-t)}$$  \hspace{1cm} (17.46)

where $\sigma_r$ and $\alpha$ are positive constants. From (17.36), this implies that the rate of return volatility of a zero-coupon bond equals

$$\sigma_I(t, T) = \int_t^T \sigma_I(t, s) ds = \int_t^T \sigma_r e^{-\alpha(s-t)} ds = \frac{\sigma_r}{\alpha} \left( 1 - e^{-\alpha(T-t)} \right) \hspace{1cm} (17.47)$$

Note that this volatility function is the same as the Vasicek model of the term structure given in (9.31). Hence, the bond price’s risk-neutral process is $dP(t, T) / P(t, T) = r(t) dt - \frac{\sigma_r}{\alpha} \left( 1 - e^{-\alpha(T-t)} \right) d\zeta$. For option valuation, it remains to derive the instantaneous-maturity interest rate and its dynamics. From (17.43) and (17.44) we have:

$$r(t) = f(0, t) + \int_0^t \frac{\sigma_r^2}{\alpha} \left( e^{-\alpha(t-u)} - e^{-2\alpha(t-u)} \right) du + \int_0^t \sigma_r e^{-\alpha(t-u)} d\zeta(u) \hspace{1cm} (17.48)$$

done in Chapter 9 in equation (9.27). For conditions on forward rate volatilities that lead to Markov structures, see the work of Andrew Carverhill (Carverhill 1994), Peter Ritchken and L. Sankarasubramanian (Ritchken and Sankarasubramanian 1995), and Koji Inui and Masaaki Kijima (Inui and Kijima 1998).
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

\[ dr = \frac{\partial f(0,t)}{\partial t} dt + \int_0^t \left[ \sigma_r^2 e^{-2\alpha(t-u)} - \alpha^2 e^{-\alpha(t-u)} \right] du + \sigma_r d\bar{\varepsilon} \]

Substituting (17.48) into (17.49) and simplifying leads to

\[ dr = \frac{\partial f(0,t)}{\partial t} dt + \int_0^t \left[ \sigma_r^2 e^{-2\alpha(t-u)} \right] du + \alpha \left[ f(0,t) - r(t) \right] dt + \sigma_r d\bar{\varepsilon} \]

where \( \tau(t) = \frac{1}{\alpha} \frac{\partial f(0,t)}{\partial t} + f(0,t) + \sigma_r^2 (1 - e^{-2\alpha t}) / \alpha^2 \) is the risk-neutral central tendency of the short-rate process that is a deterministic function of time. The process in (17.50) is Markov in that the only stochastic variable affecting its future distribution is the current level of \( r(t) \). However, it differs from the standard Vasicek model which assumes that the risk-neutral process for \( r(t) \) has a long run mean that is constant.\(^{13}\) By making the central tendency, \( \tau(t) \), to be a deterministic function of the currently observed forward rate curve, \( f(0,t) \) \( \forall t \geq 0 \), the model's implied date 0 price of a zero coupon bond, \( P(0,T) \), will coincide exactly with observed prices.\(^{14}\) This model was proposed by John Hull and Alan White ((Hull and White 1990), (Hull and White 1993)), and is referred to as the "Extended Vasicek" model.\(^{15}\)

Since, as with the standard Vasicek model, the extended Vasicek model has

---

\(^{13}\)Recall from equation (10.63) that the unconditional mean of the risk-neutral interest rate is \( \tau + qr \alpha / \alpha \), where \( \tau \) is the mean of the physical process and \( q \) is the market price of interest rate risk.

\(^{14}\)It is left as an exercise to verify that when \( \tau(t) = \frac{1}{\alpha} \frac{\partial f(0,t)}{\partial t} + f(0,t) + \sigma_r^2 (1 - e^{-2\alpha t}) / \alpha^2 \), then \( P(0,T) = \bar{E} \left[ \exp \left\{ -\int_0^T r(s) ds \right\} \right] = \exp \left( -\int_0^T f(0,s) ds \right) \).

\(^{15}\)Hull and White show that, besides \( \tau(t) \), the parameters \( \alpha(t) \) and \( \sigma_r(t) \) can also be extended to be deterministic functions of time. With these extensions, \( r(t) \) remains normally distributed and analytic solutions to options on discount bonds can be obtained. Making \( \alpha(t) \) and \( \sigma_r(t) \) to be time varying allows one to fit other aspects of the term structure, such as observed volatilities of forward rates.
bond return volatilities being a deterministic function of time, the expectation in (17.45) for the case of a European option has an analytic solution. Alternatively, the results of Merton (Merton 1973b) given in (9.45) to (9.47) on the pricing of options when interest rates are random can be applied to derive the solution. However, instead of Chapter 9’s assumption of the underlying asset being an equity that follows geometric Brownian motion, the underlying asset is a bond that matures at date $T + \tau$. For a call option with exercise price $X$, the boundary condition is $c(T) = \max[P(T, T + \tau) - X, 0]$. This leads to the solution

$$
c(t) = P(t, T + \tau) N(d_1) - P(t, T) X N(d_2) = e^{-\int_t^T f(t, s) ds} P(t, T + \tau) N(d_1) - e^{-\int_t^T f(t, s) ds} P(t, T) X N(d_2)$$

where $d_1 = \left[ \ln \left( \frac{P(t, T + \tau)}{(P(t, T) X)} \right) + \frac{1}{2} v(t, T)^2 \right] / v(t, T)$, $d_2 = d_1 - v(t, T)$, and where\(^{16}\)

$$
v(t, T)^2 = \int_t^T \left[ \sigma_1^2(t, u + \tau) + \sigma_1^2(t, u) - 2\rho \sigma_1(t, u + \tau) \sigma_1(t, u) \right] du = \frac{\sigma_\tau^2}{2\alpha^3} \left( 1 - e^{-2\alpha(T-t)} \right) \left( 1 - e^{-\sigma \tau} \right)^2$$

This solution illustrates a general principle of the HJM approach, namely, that formulas can be derived whose inputs match the initial term structure of bond prices ($P(t, T)$ and $P(t, T + \tau)$) or, equivalently, the initial forward rate curve ($f(t, s) \forall s \geq t$).

17.2.2 Market Models

\(^{16}\)Note that when applying Merton’s derivation to the case of the underlying asset being a bond, then $\rho$, the return correlation between bonds maturing at dates $T$ and $T + \tau$, equals 1. This is because there is a single Brownian motion determining the stochastic component of returns.
As shown in the previous section, HJM models begin with a particular specification for instantaneous-maturity, continuously-compounded, forward rates, and then derivative values are calculated based on these initial forward rates. However, instantaneous-maturity forward rates are not directly observable, and in many applications they must be approximated from data on bond yields or discrete-maturity forward or futures rates that are unavailable at every maturity. A class of models that is a variation on the HJM approach can sometimes avoid this approximation error and may lead to more simple, analytic solutions for particular types of derivatives. These models are known as "Market Models" and are designed to price derivatives whose payoffs are a function of a discrete maturity, rather than instantaneous maturity, forward interest rate. Examples of such derivatives include interest rate caps and floors and swaptions. Let us illustrate the market model approach by way of these examples.

Example: An Interest Rate Cap

Consider valuing a European option written on a discrete forward rate, such as one based on the London Interbank Offer Rate (LIBOR). Define \( L(t, T, \tau) \) as the date \( t \) annualized, \( \tau \)-period compounded, forward interest rate for borrowing or lending over the period from future date \( T \) to \( T + \tau \).

In terms of current date \( t \) discount bond prices \( (P(t, t + \delta)) \), forward price \( (F(t, T, \tau)) \), and continuously-compounded forward rate \( (f(t, T, \tau)) \), this discrete forward rate is defined by the relation

\[
\frac{P(t, T + \tau)}{P(t, T)} = F(t, T, \tau) = e^{-f(t, T, \tau)\tau} = \frac{1}{1 + \tau L(t, T, \tau)}
\]

\[ (17.53) \]

The convention for LIBOR is to set the compounding interval equal to the underlying instrument’s maturity. For example, if \( \tau = \frac{1}{4} \) years, then three-month LIBOR is quarterly-compounded. If \( \tau = \frac{1}{2} \) years, then six-month LIBOR is semi-annually compounded.
CHAPTER 17. TERM STRUCTURE MODELS

Note that when \( T = t \), \( P(t, t + \tau) = 1 / [1 + \tau L(t, t, \tau)] \) defines \( L(t, t, \tau) \) as the current "spot" \( \tau \)-period LIBOR.\(^{18} \) An example of an option written on LIBOR is a caplet that matures at date \( T + \tau \) and is based on the realized spot rate \( L(T, T, \tau) \). Assuming this caplet has an exercise cap rate \( X \), its date \( T + \tau \) payoff is

\[
c(T + \tau) = \tau \max[L(T, T, \tau) - X, 0]
\]

(17.54)

that is, the option payoff at date \( T + \tau \) depends on the \( \tau \)-period spot LIBOR at date \( T \).\(^{19} \) Because uncertainty regarding the LIBOR rate is resolved at date \( T \), which is \( \tau \) period’s prior to the caplet’s settlement (payment) date, we can also write

\[
c(T) = P(T, T + \tau) \max[\tau L(T, T, \tau) - \tau X, 0]
\]

(17.55)

\[
= P(T, T + \tau) \max\left[\frac{1}{P(T, T + \tau)} - 1 - \tau X, 0\right]
\]

\[
= \max\left[1 - \frac{1 + \tau X}{P(T, T + \tau)}, 0\right] = \max\left[1 - \frac{1 + \tau X}{1 + \tau L(T, T, \tau)}, 0\right]
\]

which illustrates that a caplet maturing at date \( T + \tau \) is equivalent to a put option that matures at date \( T \), has an exercise price of 1, and is written on a zero coupon bond that has a payoff of \( 1 + \tau X \) at its maturity date of \( T + \tau \). Similarly, a floorlet, whose date \( T + \tau \) payoff equals \( \tau \max[X - L(T, T, \tau), 0] \), can be shown to be equivalent to a call option on a zero coupon bond.\(^{20} \)

To value a caplet using a market model approach, let us first analyze the

\[^{18}\text{This modeling assumes that LIBOR is the yield on a default-free discount bond. However, LIBOR is not a fully default-free interest rate, such as a Treasury security rate. It represents the borrowing rate of a large, generally high credit quality, bank. Typically, the relatively small amount of default risk is ignored when applying market models to derivatives based on LIBOR.}\]

\[^{19}\text{Caplets are based on a notional principal amount, which here is assumed to be $1. The value of a caplet having a notional principal of $N is simply $N times the value of a caplet with a notional principal of $1, that is, is payoff is $N \max[L(T, T, \tau) - X, 0].}\]

\[^{20}\text{Therefore, the HJM - extended Vasicek solution in (17.51) to (17.52) is a one method for valuing a floorlet. A straightforward modification of this formula could also value a caplet.}\]
dynamics of $L(t,T,\tau)$. Re-arranging (17.53) gives

$$\tau L(t,T,\tau) = \frac{P(t,T)}{P(t,T+\tau)} - 1 \quad (17.56)$$

We can derive the stochastic process followed by this forward rate in terms of the bond prices’ risk neutral process. Note that from (?? along with $dz = dBz - \Theta(t)\,dt$, we have $dP(t,T)/P(t,T) = r(t)\,dt - \sigma_I(t,T)\,dBz$. Applying Ito’s lemma to (17.56), we obtain

$$dL(t,T,\tau) = \left[\sigma_I(t,T+\tau)\left[\sigma_I(t,T+\tau) - \sigma_I(t,T)\right]\right] dt + \left[\sigma_I(t,T+\tau) - \sigma_I(t,T)\right] dBz \quad (17.57)$$

In principle, now we could value a contingent claim written on $L(t,T,\tau)$ by calculating the claim’s discounted expected terminal payoff assuming $L(t,T,\tau)$ follows the process in (17.57). However, as will become clear, there is an alternative probability measure to the one generated by $dBz$ that can be used to calculate a contingent claim’s expected payoff, and this alternative measure is analytically more convenient for this particular forward rate application.

To see this, consider the new transformation $dBz = dBz + \sigma_I(t,T+\tau)\,dt = dBz + \left[\Theta(t) + \sigma_I(t,T+\tau)\right] dt$. Substituting into (17.57) results in

$$dL(t,T,\tau) \over L(t,T,\tau) = \left[\sigma_I(t,T+\tau) - \sigma_I(t,T)\right] dBz \quad (17.58)$$

so that under the probability measure generated by $dBz$, the process followed by $L(t,T,\tau)$ is a martingale. This probability measure is referred to as the forward rate measure at date $T+\tau$. Note that since $L(t,T,\tau)$ is linear in the bond price

---

21Specifically, if $c(t,L(t,T,\tau))$ is the contingent claim’s value, it could be calculated as $\hat{E}_t\left[ e^{-\int_t^T r(s)\,ds} c(T,L(t,T,\tau)) \right]$ where $r(t)$ and $L(t,T,\tau)$ are assumed to follow risk-neutral processes.
CHAPTER 17. TERM STRUCTURE MODELS

\( P(t,T) \) deflated by \( P(t,T + \tau) \), the forward rate measure at date \( T + \tau \) works by deflating all security prices by the price of the discount bond that matures at date \( T + \tau \). This contrasts with the risk-neutral measure where security prices are deflated by the value of the money market account which follows the process \( dB(t) = r(t)B(t)dt \).

Not only does \( L(t,T,\tau) \) follow a martingale under the forward measure, but so does the value of all other securities. To see this, let the date \( t \) price of a contingent claim be given by \( c(t) \). In the absence of arbitrage, its price process is of the form

\[
\frac{dc}{c} = [r(t) + \Theta(t)\sigma_c(t)]dt + \sigma_c(t)d\tilde{z}
\]

(17.59)

Now define the deflated contingent claim’s price as \( C(t) = c(t)/P(t,T + \tau) \).

Applying Itô’s lemma gives

\[
\frac{dC}{C} = [\Theta(t) + \sigma_I(t,T + \tau)]'\sigma_c(t) C(t) dt + [\sigma_c(t) + \sigma_I(t,T + \tau)]'d\tilde{z}
\]

(17.60)

and making the forward measure transformation \( d\tilde{z} = dz + [\Theta(t) + \sigma_I(t,T + \tau)]dt \), (17.60) becomes the martingale process

\[
\frac{dC}{C} = [\sigma_c(t) + \sigma_I(t,T + \tau)]'d\tilde{z}
\]

(17.61)

so that \( C(t) = \tilde{E}_t[C(t + \delta)] \forall \delta \geq 0 \), where \( \tilde{E}_t[\cdot] \) is the date \( t \) expectation under the forward measure. Now, to show why this transformation can be convenient, suppose that this contingent claim is the caplet described earlier. This deflated
caplet’s value is given by

$$C(t) = \tilde{E}_t[C(T + \tau)]$$

$$= \tilde{E}_t\left[\tau_{\max}[L(T, T, \tau) - X, 0]\right]$$

(17.62)

Noting that $C(t) = c(t) / P(t, T + \tau)$ and realizing that $P(T + \tau, T + \tau) = 1$, we can re-write this as

$$c(t) = P(t, T + \tau) \tilde{E}_t[\tau_{\max}[L(T, T, \tau) - X, 0]]$$

(17.63)

A common practice is to assume that $L(T, T, \tau)$ is lognormally distributed under the date $T + \tau$ forward measure. This means that $[\sigma_I(t, T + \tau) - \sigma_I(t, T)]$ in 17.58 must be a vector of non-stochastic functions of time that can be calibrated to match observed bond or forward rate volatilities. Noting that $L(t, T, \tau)$ also has a zero drift, leads to a similar formula first proposed by Fischer Black (Black 1976) for valuing options on commodity futures:

$$c(t) = \tau P(t, T + \tau) [L(t, T, \tau) N(d_1) - X N(d_2)]$$

(17.64)

where $d_1 = \left[\ln(L(t, T, \tau)/X) + \frac{1}{2}v(t, T)^2\right]/v(t, T)$, $d_2 = d_1 - v(t, T)$, and

$$v(t, T)^2 = \int_t^T |\sigma_I(s, T + \tau) - \sigma_I(s, T)|^2 ds$$

(17.65)

Equation (17.64) is similar to equation (10.57) derived in Chapter 10 for the

22 Assuming a lognormal distribution for $L(t, T, \tau)$ is attractive because it prevents this discrete forward rate from becoming negative, thereby also restricting yields on discount bonds to be non-negative. Note that if instantaneous-maturity forward rates are assumed to be lognormally distributed, HJM show that they will be expected to become infinite in finite time. This is inconsistent with an arbitrage-free bond prices. Fortunately, such an explosion of rates does not occur when forward rates are discrete (Brace, Gatarek, and Musiela 1997).

23 Note that since $\sigma_I(t, t + \delta)$ is an integral of instantaneous forward rate volatilities, the lognormality of $\sigma_I(t, t + \tau) - \sigma_I(t, T)$ puts restrictions on instantaneous forward rates under an HJM modeling approach. However, we need not focus on this issue for pricing applications involving a discrete forward rate.
CHAPTER 17. TERM STRUCTURE MODELS

case of a call option on a forward or futures price where the underlying is lognormally-distributed and interest rates are non-stochastic.

An interest rate cap is a portfolio of caplets written on the same \( \tau \)-period LIBOR but maturing at different dates \( T = T_1, T_2, ..., T_n \) where typically \( T_{j+1} = T_j + \tau \). Standard practice is to value each individual caplet in the portfolio in the manner described above, where the caplet maturing at date \( T_j \) is priced using the date \( T_j + \tau \) forward measure. Often, caps are purchased by issuers of floating-rate bonds whose bond payments coincide with the caplet maturity dates. Doing so insures the bond issuer against having to make a floating coupon rate greater than \( X \) (plus a credit spread). Since a floating rate bond’s coupon rate payable at date \( T + \tau \) is most commonly tied to the \( \tau \)-period LIBOR at date \( T \), caplet payoffs follow this same structure. Analogous to a cap, an interest rate floor is a portfolio of floorlets, and can be valued using the same technique described in this section.

**Example: A Swaption**

Frequently, a market model approach is applied to value another common interest rate derivative, a swaption. A swaption is an option to become a party in an interest rate swap at a given future maturity date and at a pre-specified swap rate. Let us, then, define the interest rate swap underlying this swaption. A standard (plain vanilla) swap is an agreement between two parties to exchange fixed interest rate coupon payments for floating interest rate coupon payments at dates \( T_1, T_2, ..., T_{n+1} \) where \( T_{j+1} = T_j + \tau \) and \( \tau \) is the maturity of the LIBOR of the floating rate coupon payments. Thus, if \( K \) is the swap’s fixed annualized coupon rate, then at date \( T_{j+1} \) the fixed-rate payer’s net payment is \( \tau [K - L(T_j, T_j, \tau)] \) while that of the floating rate payer is exactly the
Note that the swap’s series of floating rate payments plus an additional $1 at date $T_{n+1}$ can be replicated by starting with $1$ at time $T_0 = T_1 - \tau$ and repeatedly investing this $1$ in $\tau$-maturity LIBOR deposits. These are the same cashflows that one would obtain by investing $1$ in a floating rate bond at date $T_0$. Similarly, the swap’s series of fixed rate payments plus an additional $1$ at date $T_{n+1}$ can be replicated by buying a fixed-coupon bond that pays coupons of $\tau K$ at each swap date and pays a principal of $1$ at its maturity date of $T_{n+1}$. Based on this insight, one can see that the value of a swap to the floating rate payer is the difference between a fixed coupon bond having coupon rate $K$, and a floating coupon bond having coupons tied to $\tau$-period LIBOR. Thus, if $t \leq T_0 = T_1 - \tau$, then the date $t$ value of the swap to the floating rate payer is

$$\tau K \sum_{j=1}^{n+1} P(t, T_j) + P(t, T_{n+1}) - P(t, T_0)$$

(17.66)

When a standard swap agreement is initiated at time $T_0$, the fixed rate $K$ is set such that the value of the swap in (17.66) is zero. This concept of setting $K$ to make the agreement fair (similar to forward contracts) can be extended to dates prior to $T_0$. One can define $s_{0,n}(t)$ as the forward swap rate that makes the date $t$ value of the swap (starting at date $T_0$ and making $n$ subsequent exchanges) equal to zero. Setting $K = s_{0,n}(t)$ and equating (17.66) to zero,

---

24 Recall that $L(T_j, T_j, \tau)$ is the spot $\tau$-period LIBOR at date $T_j$. Also, as discussed in the preceding footnote, this exchange is based on a notional principal of $1$. For a notional principal of $N$, all payments are multiplied by $N$.

25 Thus, $1$ invested at time $T_0$ produces a return of $1 + \tau L(T_0, T_0, \tau)$ at $T_1$. Keeping the cashflow of $\tau L(T_0, T_0, \tau)$ and re-investing the $1$ will then produce a return of $1 + \tau L(T_1, T_1, \tau)$ at $T_2$. Keeping the cashflow of $\tau L(T_1, T_1, \tau)$ and re-investing the $1$ will then produce a return of $1 + \tau L(T_2, T_2, \tau)$ at $T_3$. This process is repeated until at time $T_{n+1}$ a final return of $1 + \tau L(T_n, T_n, \tau)$ is obtained.

26 Notice that $P(t, T_0)$ is the date $t$ value of the floating rate bond while the remaining terms are the value of the fixed rate bond.
one obtains

\[
s_{0,n}(t) = \frac{P(t,T_0) - P(t,T_{n+1})}{\tau \sum_{j=1}^{n+1} P(t,T_j)} = \frac{P(t,T_0) - P(t,T_{n+1})}{B_{1,n}(t)}
\]

where \( B_{1,n}(t) \equiv \tau \sum_{j=1}^{n+1} P(t,T_j) \) is a portfolio of zero coupon bonds that each pay \( \tau \) at the times of the swap’s exchanges.

Now a standard swaption is an option to become either a fixed rate payer or floating rate payer at a fixed swap rate \( X \) at a specified future date. Thus, if the maturity of the swaption is date \( T_0 \), at which time the holder of the swaption has the right, but not the obligation, to become a fixed rate payer (floating rate receiver), this option’s payoff equals\(^{27}\)

\[
c(T_0) = \max \left[ B_{1,n}(T_0) [s_{0,n}(T_0) - X], 0 \right]
\]

Note from the first line of (17.68) that when the option is in the money, then \( B_{1,n}(T_0) [s_{0,n}(T_0) - X] \) is the date \( T_0 \) value of the fixed rate payer’s savings from having the swaption relative to entering into a swap at the fair spot rate \( s_{0,n}(T_0) \).

In the second line of (17.68), we have substituted from (17.67) \( s_{0,n}(T_0) B_{1,n}(T_0) = P(T_0,T_0) - P(T_0,T_{n+1}) = 1 - P(T_0,T_{n+1}) \). This illustrates that a swaption is equivalent to an option on a coupon bond with coupon rate \( X \) and exercise price of 1.

To value this swaption at date \( t \leq T_0 \), a convenient approach is to recognize from (17.68) that the swaption’s payoff is proportional to \( B_{1,n}(T_0) \equiv \tau \sum_{j=1}^{n+1} P(T_0,T_j) \). This suggests that \( B_{1,n}(t) \) is a convenient deflator for valu-
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

By normalizing all security prices by \( B_{1,n}(t) \), we will value the swaption using the so-called "forward swap measure."

Similar to valuation under the risk-neutral or forward measure of the previous section, let us define \( C(t) = c(t) / B_{1,n}(t) \). Also define \( dz = dz + [\Theta(t) + \sigma_{B_{1,n}}(t)] dt \) where \( \sigma_{B_{1,n}}(t) \) is the date \( t \) vector of instantaneous volatilities of the zero coupon bond portfolio’s value, \( B_{1,n}(t) \). Similar to the derivation in equations (17.59) to (17.61), we have

\[
\frac{dC}{C} = \left[ \sigma_c(t) + \sigma_{B_{1,n}}(t, T + \tau) \right] dz
\]

(17.69)

so that all deflated asset prices under the forward swap measure follow martingale processes. Thus,

\[
C(t) = E_t [C(T_0)]
\]

(17.70)

\[
= E_t \left[ \max \left[ \frac{B_{1,n}(T_0) [s_{0,n}(T_0) - X]}{B_{1,n}(T_0)} \right] \right]
\]

\[
= E_t [\max [s_{0,n}(T_0) - X, 0]]
\]

Re-written in terms of the undeflated swaption’s current value, \( c(t) = C(t) B_{1,n}(t) \), (17.70) becomes

\[
c(t) = B_{1,n}(t) E_t [\max [s_{0,n}(T_0) - X, 0]]
\]

(17.71)

so that the expected payoff under the forward swap measure is discounted by the current value of a portfolio of zero coupon bonds that mature at the times of the swap’s exchanges.

Importantly, note that \( s_{0,n}(t) = [P(t, T_0) - P(t, T_{n+1})] / B_{1,n}(T_0) \) is the ratio of the difference between two security prices deflated by \( B_{1,n}(t) \). In the absence of arbitrage, it must also follow a martingale process under the forward
CHAPTER 17. TERM STRUCTURE MODELS

A convenient and commonly made assumption is that this forward swap rate is lognormally distributed under the forward swap measure:

\[ \frac{dS_{0,n}(t)}{S_{0,n}(t)} = \sigma_{s_{0,n}}(t)'d\bar{Z} \]  

(17.72)

so that \( \sigma_{s_{0,n}}(t) \) is a vector of deterministic functions of time that can be calibrated to match observed forward swap volatilities or zero coupon bond volatilities.\(^{28}\) This assumption results in (17.71) taking a Black-Scholes-type form

\[ c(t) = B_{1,n}(t) [s_{0,n}(t) N(d_1) - X N(d_2)] \]  

(17.73)

where

\[ d_1 = \left[ \ln \left( \frac{s_{0,n}(t)}{X} \right) + \frac{1}{2} v(t,T_0)^2 \right] /v(t,T_0), \quad d_2 = d_1 - v(t,T_0) , \]

\[ v^2(t,T_0) = \int_t^{T_0} \sigma_{s_{0,n}}(u)'\sigma_{s_{0,n}}(u) du. \]  

(17.74)

17.2.3 Random Field Models

The term structure models that we have studied thus far have specified a finite number of Brownian motion processes as the source of uncertainty determining the evolution of bond prices or forward rates. For example, the bond price processes in equilibrium models (equation 17.2) and HJM models (equation 17.36) were driven by an \( n \times 1 \) vector of Brownian motions, \( dz \). One implication of this is that a Black-Scholes hedge portfolio of \( n \) different maturity bonds can be used to perfectly replicate the risk of any other maturity bond. As shown in Chapter 10, in the absence of arbitrage, the fact that any bond’s risk can be hedged with other bonds places restrictions on bonds’ expected excess rates of return and results in a unique vector of market prices of risk, \( \Theta(t) \), associated

\(^{28}\)Applying Itô’s lemma to (17.67) allows one to derive the volatility of \( s_{0,n}(t) \) in terms of zero coupon bond volatilities.
with $\text{d}z$. This implied a bond price process of the form

$$
\frac{dP(t, T)}{P(t, T)} = [r(t) + \Theta(t)\sigma_p(t, T)] \, dt + \sigma_p(t, T) \, \text{d}z
$$

(17.75)

Moreover, the Black-Scholes hedge, by making the market dynamically complete and by identifying a unique $\Theta(t)$ associated with $\text{d}z$, allows us to perform risk-neutral valuation by the transformation $\text{d}z = \text{d}z + \Theta(t) \, dt$ or valuation using the pricing kernel $dM/M = -r(t) \, dt - \Theta(t) \, \text{d}z$.

However, the elegance of these models comes with an empirical downside. The fact that all bond prices depend on the same $n \times 1$ vector $\text{d}z$ places restrictions on the covariance of bonds’ rates of return. For example, when $n = 1$, the rates of return on all bonds are instantaneously perfectly correlated. While in these models, the correlation can be made less perfect by increasing $n$, doing so introduces more parameters that require estimation.

A related empirical implication of (17.75) or (17.33) is that it restricts the possible future term structures of bond prices or forward rates. In other words, starting from the current date $t$ set of bond prices $P(t, T) \, \forall T > t$, an arbitrary future term structure, $P(t + dt, T) \, \forall T > t + dt$, cannot always be achieved by any realization of $\text{d}z$. This is because a given future term structure has an infinite number of bond prices (each of a different maturity), but the finiteness of $\text{d}z$ allows matching this future term structure at only a finite number of maturity horizons. Hence, models based on a finite $\text{d}z$ are almost certainly inconsistent with future observed bond prices and forward rates. Because of this, empiricists must assume that data on bond prices (or yields) are observed with "noise" or that, in the case of HJM-type models, parameters (that the model assumes to be constant) must be re-calibrated at each observation date.

\footnote{For example, consider $n = 1$. In this case, all bond prices must either rise or fall with a given realization of $\text{d}z$. This model would not permit a situation where short maturity bond prices fell but long maturity bond prices rose. The model could produce a realization of $\text{d}z$ that matches long maturity bond prices or short maturity bond prices, but not both.}
to match the new term structure of forward rates.

Random field models are an attempt to avoid these empirical deficiencies. Research in this area includes that of David Kennedy (Kennedy 1994), (Kennedy 1997), Robert Goldstein (Goldstein 2000), Pedro Santa-Clara and Didier Sornette (Santa-Clara and Sornette 2001), and Robert Kimmel (Kimmel 2004). These models specify that each zero coupon bond price, $P(t,T)$, or each instantaneous forward rate, $f(t,T)$, is driven by a Brownian motion process that is unique to the bond’s or rate’s maturity, $T$. For example, a model of this type might assume that a bond’s risk-neutral process satisfies

$$dP(t,T)/P(t,T) = r(t)dt + \sigma_p(t,T)d\tilde{z}_T \quad \forall T > t \quad (17.76)$$

where $d\tilde{z}_T(t)$ is a single Brownian motion process (under the risk-neutral measure) that is unique to the bond that matures at date $T$. The set of Brownian motions for all zero coupon bonds $\{\tilde{z}_T(t)\}_{T>t}$ comprise a Brownian "field" or "sheet." This continuum of Brownian motions has two dimensions: calendar time, $t$, and time to maturity, $T$. The elements affecting different bonds are linked by an assumed correlation structure:

$$d\tilde{z}_{T_1}(t)d\tilde{z}_{T_2}(t) = \rho(t,T_1,T_2)dt \quad (17.77)$$

where $\rho(t,T_1,T_2) > 0$ is specified to be a particular continuous, differentiable function with $\rho(t,T,T) = 1$ and $\frac{\partial \rho(t,T_1,T_2)}{\partial T_1}|_{T_1=T_2} = 0$. For example, one simple specification involving only a single parameter is $\rho(t,T_1,T_2) = e^{-\varrho |T_1 - T_2|}$ where $\varrho$ is a positive constant.

One can also model the physical process for bond prices corresponding to

---

30 An alternative way of specifying a random field model is to assume that the risk-neutral processes for instantaneous forward rates are of the form $df(t,T) = [\sigma(t,T)\int_t^T \sigma(t,s)c(t,s)ds] dt + \sigma(t,T)d\tilde{z}_T$ where $d\tilde{z}_{T_1}d\tilde{z}_{T_2} = c(t,T_1,T_2)dt$. This specification extends the HJM equation (17.41) to a random field driving forward rates.
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

(17.76). If \( \theta_T(t) \) is the market price of risk associated with \( dz_T(t) \), then making the transformation \( dz_T = \tilde{z}_T + \theta_T(t) \, dt \), one obtains

\[
dP(t,T)/P(t,T) = [r(t) + \theta_T(t) \sigma_p(t, T)] \, r(t) \, dt + \sigma_p(t, T) \, dz_T \quad \forall T > t
\]

(17.78)

with \( dz_T(t), T > t \) satisfying the same correlation function as in (17.77). Analogous to the finite-factor pricing kernel process in (17.8), a pricing kernel for this random field model would be

\[
dM/M = -r(t) \, dt - \int_t^\infty [\theta_T(t) \, dz_T(t)] \, dT
\]

(17.79)

so that an integral of the products of market prices of risk and Brownian motions replaces the usual sum of these products that occur for the finite factor case.\(^{31}\)

The benefit of a model like (17.76) and (17.77) is that a realization of the Brownian field can generate any future term structure of bond prices or forward rates and, hence, be consistent with empirical observation and not require model re-calibration. Moreover, with only a few additional parameters, random field models can provide a flexible covariance structure among different maturity bonds. Specifically, unlike finite-dimensional equilibrium models or HJM models, the covariance matrix of different maturity bond returns or forward rates will always be non-singular no matter how many bonds are included. This could be important when valuing particular fixed-income derivatives where the underlying is a portfolio of zero coupon bonds, and the correlation between these bonds affects the overall portfolio volatility.

\(^{31}\)Note, however, that a random field model is not the same as a standard finite factor model extended to an infinite number of factors. As shown in (17.78), a random field model has a single Brownian motion driving each bond price or forward rate. A factor model, such as (17.2) or (17.36), extended to infinite factors would have the same infinite set of Brownian motions driving each bond price.
CHAPTER 17. TERM STRUCTURE MODELS

However, this rich covariance structure requires stronger theoretical assumptions for valuing derivatives compared to finite-dimensional diffusion models. A given bond’s return can no longer be perfectly replicated by a portfolio of other bonds, and, thus, a Black-Scholes hedging argument cannot be used to identify a unique market price of risk associated with each $d_z(t)$.\footnote{Robert Goldstein (Goldstein 2000) characterizes random field models of the term structure as being analogous to the APT model ((Ross 1976)). As discussed in Chapter 3, the APT assumes that a given asset’s return depends on the risk from a finite number of factors along with the asset’s own idiosyncratic risk. Thus, the asset is imperfectly correlated with any portfolio containing a finite number of other assets. Similarly, in a random field model, a given bond’s return is imperfectly correlated with any portfolio containing a finite number of other bonds. Taking the analogy a step further, perhaps market prices of risk in a random field model can be characterized using the notion of asymptotic arbitrage, rather than exact arbitrage.} The market for fixed-income securities is no longer dynamically complete. Hence, one must assume, perhaps due to an underlying preference-based general equilibrium model, that there exists particular $\theta_T(t)$ associated with each $d_z(t)$, or, equivalently, that a risk-neutral pricing exists.

Random field models can be parameterized by assuming particular functions for bond price or forward rate volatilities. For example, Pierre Collin-Dufresne and Robert Goldstein (Collin-Dufresne and Goldstein 2003) propose a stochastic volatility model where, in equation (17.76), $\sigma_p(t,T) = \sigma(t,T) \sqrt{\Sigma(t)}$, where $\sigma(t,T)$ is a deterministic function and where $\Sigma(t)$ is a volatility factor, common to all bonds, that follows the square-root process

$$d\Sigma(t) = \kappa (\Sigma - \Sigma(t)) dt + \theta \sqrt{\Sigma(t)} d\tilde{z}_\Sigma$$ \hspace{1cm} (17.80)$$

where $d\tilde{z}_\Sigma$ is a Brownian motion (under the risk-neutral measure) that is assumed to be independent of the Brownian field $\{d\tilde{z}_T\} \forall T > t$. Based on this parameterization, which is similar to a one factor affine model, they derive solutions for various interest rate derivatives.\footnote{Robert Kimmel ((Kimmel 2004)) also derives models with stochastic volatility driven by multiple factors.}
17.2. VALUATION MODELS FOR INTEREST RATE DERIVATIVES

If, similar to David Kennedy (Kennedy 1994), one makes the more simple assumption that \( \sigma_p(t, T) \) and (17.76) and \( \rho(t, T_1, T_2) \) in (17.77) are deterministic functions, then options on bonds, such as caplets and florets, have a Black-Scholes-type valuation formula. For example, suppose as in the HJM-extended Vasicek case of (17.51) to (17.52) that we value a European call option that matures at date \( T \), is written on a zero coupon bond that matures at date \( T + \tau \), and has an exercise price of \( X \). Similar to (17.63), we can value this option using the date \( T \) forward rate measure:

\[
    c(t) = P(t, T) \, \tilde{E}_t \left[ \max \left[ p(T, T + \tau) - X, 0 \right] \right] \tag{17.81}
\]

where \( p(t, T + \tau) \equiv P(t, T + \tau)/P(t, T) \) is the deflated price of the bond that matures at date \( T + \tau \). Applying Itô’s lemma to the risk-neutral process for bond prices in (17.76), we obtain:

\[
    \frac{dp(t, T + \tau)}{p(t, T + \tau)} = \sigma_p(t, T) \left[ \sigma_p(t, T) - \rho(t, T, T + \tau) \sigma_p(t, T + \tau) \right] dt \\
    + \sigma_p(t, T + \tau) \, d\tilde{z}_{T+\tau} - \sigma_p(t, T) \, d\tilde{z}_T \tag{17.82}
\]

We can re-write \( d\tilde{z}_{T+\tau} = \rho(t, T, T + \tau) d\tilde{z}_T + \sqrt{1 - \rho(t, T, T + \tau)^2} \, d\tilde{z}_{U,T} \) where \( d\tilde{z}_{U,T} \) is a Brownian motion uncorrelated with \( d\tilde{z}_T \), so that the stochastic component in (17.82) can be written as \( \sigma_p(t, T + \tau) \sqrt{1 - \rho(t, T, T + \tau)^2} \, d\tilde{z}_{U,T} \)

\[
    + [\sigma_p(t, T + \tau) \rho(t, T, T + \tau) - \sigma_p(t, T)] \, d\tilde{z}_T. \tag{17.83}
\]

Then making the transformation to the date \( T \) forward measure, \( d\tilde{z}_T = d\tilde{z}_T + \sigma_p(t, T) \), the process for

---

34This re-writing puts the risk-neutral process for \( p(t, T + \tau) \) in the form of our prior analysis in which the vector of Brownian motions, \( d\tilde{z} \), was assumed to have independent elements. This allows us to make the transformation to the forward measure in the same manner as was done earlier.
\[ dp(t, T + \tau) \]

\[ \frac{dp(t, T + \tau)}{p(t, T + \tau)} = \sigma_p(t, T + \tau) \sqrt{1 - \rho(t, T, T + \tau)^2} d\tilde{\omega}_{U,T} \]

\[ + [\sigma_p(t, T + \tau) \rho(t, T, T + \tau) - \sigma_p(t, T)] d\tilde{\omega}_T \]

\[ = \sigma(t, T, \tau) d\tilde{\omega} \quad (17.83) \]

where

\[ \sigma(t, T, \tau)^2 \equiv \sigma_p(t, T + \tau)^2 + \sigma_p(t, T)^2 - 2\rho(t, T, T + \tau)\sigma_p(t, T + \tau)\sigma_p(t, T) \quad (17.84) \]

Thus, \( p(t, T + \tau) \) is lognormally distributed under the forward rate measure, so that \( (17.81) \) has the Black-Scholes-Merton-type solution

\[ c(t) = P(t, T) [p(t, T + \tau) N(d_1) - X N(d_2)] \quad (17.85) \]

\[ = P(t, T + \tau) N(d_1) - P(t, T) N(d_2) \]

where \( d_1 = \left[ \ln \left( \frac{p(t, T + \tau)}{X} \right) + \frac{1}{2} v(t, T)^2 \right] / v(t, T) \), \( d_2 = d_1 - v(t, T) \), and

\[ v(t, T)^2 = \int_t^T \sigma(u, T, \tau)^2 du \quad (17.86) \]

and \( \sigma(u, T, \tau) \) is defined in \( (17.84) \). While this formula is similar to the Vasicek-based ones in \( (9.45) \) and \( (17.51) \), the volatility function in \( (17.84) \) may permit a relatively more flexible form for matching observed data.

17.3 Summary

This chapter has briefly surveyed some of the important theoretical developments in modeling bond yield curves and valuing fixed income securities. The
chapter's presentation has been in the context of continuous time models and, to keep its length manageable, many similar models set in discrete time have been left out. Moreover, questions regarding numerical implementation and parameter estimation for specific models could not be answered in the short presentations given here.

There is a continuing search for improved ways of describing the term structure of bond prices and of valuing fixed income derivatives. Researchers in this field have different objectives, and the models that we presented reflect this diversity. Much academic research focuses on analyzing equilibrium models in hopes of better understanding the underlying macro-economic factors that shape the term structure bond yields. In contrast, practitioner research tends to focus on finding better ways of valuing and hedging bond-related derivatives. Researchers' efforts in these fields remains strong, and interest in the related subject of valuing defaultable fixed income securities may be growing more quickly. The next chapter takes up this topic of default-risky bond pricing.

17.4 Exercises

1. Consider the following example of a two-factor term structure model (Jegadeesh and Pennacchi 1996) (Balduzzi, Das, and Foresi 1998). The instantaneous-maturity interest rate is assumed to follow the physical process

\[ dr(t) = \alpha [\gamma (t) - r(t)] dt + \sigma_r dz_r \]

\(^{35}\) Treatments of models set in discrete-time include books by Robert Jarrow (Jarrow 2002), Bruce Tuckman (Tuckman 2002) and Thomas Ho and Sang Bin Lee (Ho and Lee 2004).
and the physical process for the interest rate’s stochastic "central tendency," \( \gamma(t) \) satisfies

\[
\frac{d\gamma(t)}{dt} = \delta \left[ \bar{\gamma} - \gamma(t) \right] dt + \sigma_{\gamma} \frac{d\gamma}{dz}
\]

where \( dz_r dz_\gamma = \rho dt \) and \( \alpha > 0, \sigma_r, \delta > 0, \bar{\gamma}, \sigma_\gamma, \) and \( \rho \) are constants.

In addition, define the constant market prices of risk associated with \( dz_r \) and \( dz_\gamma \) to be \( \theta_r \) and \( \theta_\gamma \). Re-write this model using the affine model notation used in this chapter and solve for the equilibrium price of a zero coupon bond, \( P(t, T) \).

2. Consider the following one-factor quadratic-gaussian model. The single state variable, \( x(t) \), follows the risk-neutral process

\[
dx(t) = \kappa (x - \bar{x}) dt + \sigma_x \frac{dz}{\sqrt{\rho}}
\]

and the instantaneous-maturity interest rate is given by \( r(t, x) = \alpha + \beta x(t) + \gamma x(t)^2 \). Assume \( \kappa, \bar{x}, \alpha, \) and \( \gamma \) are positive constants, and that \( \alpha - \frac{1}{2} \beta^2 / \gamma \geq 0 \) where \( \beta \) is also a constant. Solve for the equilibrium price of a zero coupon bond, \( P(t, T) \).

3. Show that for the extended Vasicek model when \( \bar{\gamma}(t) \equiv \frac{1}{\alpha} \frac{d}{dt} f(0, t) + f(0, t) + \frac{\sigma^2}{2} (1 - e^{-2\alpha t}) / \alpha^2 \), then \( P(0, T) = \mathbb{E} \left[ \exp \left( - \int_0^T r(s) ds \right) \right] = \exp \left( - \int_0^T f(0, s) ds \right) \).

4. Determine the value of an \( n \) payment interest rate floor using the LIBOR market model.