The Behavior of Bonds and Interest Rates

Before discussing how a bond market-maker would delta-hedge, we first need to specify how bonds behave. Suppose we try to model a zero-coupon bond the same way we model a stock, by assuming that the bond price, $P(t, T)$ follows an Itô process:

$$\frac{dP}{P} = \alpha(r, t)dt - q(r, t)dZ \tag{24.1}$$

In this equation, the coefficients α and q cannot be constants and in fact must be modeled rather carefully to ensure that the bond satisfies its boundary conditions. For example, the bond must be worth $1 at maturity. Also, the volatility of the bond price should decrease as the bond approaches maturity—a given change in interest rates affects the price of a long-lived bond more than the price of a short-lived bond. Neither of these restrictions is automatically reflected in equation (24.1). In order to accommodate such behavior α and q must be carefully specified functions of the interest rate and time.

An alternative to beginning with equation (24.1) is to model the behavior of the interest rate and *solve* for the bond price. If we follow this approach, the bond price will *automatically* behave in an appropriate way, as long as the interest rate process is reasonable.

Suppose we assume that the short-term interest rate follows the Itô process

$$dr = a(r)dt + \sigma(r)dZ \tag{24.2}$$

This equation for the behavior of the interest rate is general, in that the drift and standard deviation are functions of r.[1] Given equation (24.2), what is the bond price? We will see that different bond price models arise from different versions of this interest rate process.

An Impossible Bond Pricing Model

We will first look at a bond pricing model that is intuitive, appealing in its simplicity, and widely used informally as a way to think about bonds. We will assume that the yield curve is flat; that is, at any point in time, zero-coupon bonds at all maturities have the same yield to maturity. If the interest rate changes, yields for all bonds change uniformly so that the yield curve remains flat. Unfortunately, this model of the yield curve gives rise to arbitrage opportunities. It can be instructive, however, to see what doesn't work in order to better appreciate what does.

To analyze the flat-yield curve assumption, we assume that the interest rate follows equation (24.2). The price of zero-coupon bonds is given by

$$P(t, T) = e^{-r(T-t)} \tag{24.3}$$

In this specification, every bond has yield to maturity r.

We now analyze the delta-hedging problem. If we buy one bond maturing at time T_2, hedge by buying N bonds maturing at time T_1, and finance the difference at the short-term interest rate, the bond portfolio has value

$$I = NP(t, T_1) + P(t, T_2) + W = 0 \tag{24.4}$$

[1] The sign of dZ differs in (24.1) and (24.2) because interest rates and bond prices move inversely.

Since W is invested in short-term bonds, we have

$$dW = rWdt \tag{24.5}$$

By Itô's Lemma, and using the formula for the bond price, equation (24.3), the change in the value of the portfolio is

$$\begin{aligned} dI &= NdP(t, T_1) + dP(t, T_2) + dW \\ &= N\left(-(T_1 - t)P(t, T_1)dr + \frac{1}{2}(T_1 - t)^2\sigma^2 P(t, T_1)dt + rP(t, T_1)dt\right) \\ &\quad + \left(-(T_2 - t)P(t, T_2)dr + \frac{1}{2}(T_2 - t)^2\sigma^2 P(t, T_2)dt + rP(t, T_2)dt\right) + rWdt \end{aligned} \tag{24.6}$$

We pick N to eliminate the effect of interest rate changes, dr, on the value of the portfolio. Thus, we set

$$N = -\frac{(T_2 - t)P(t, T_2)}{(T_1 - t)P(t, T_1)} \tag{24.7}$$

The delta-hedged portfolio has no risk and no investment; it should therefore earn zero:

$$dI = 0 \tag{24.8}$$

Combining equations (24.4), (24.6), (24.7), and (24.8), and then simplifying, gives us

$$\tfrac{1}{2}(T_2 - T_1)\sigma^2 = 0 \tag{24.9}$$

This equation cannot hold unless $T_1 = T_2$. Thus, we conclude that *the bond valuation model implied by equations* (24.2) *and* (24.3) *is impossible*, in the sense that arbitrage is possible if the yield curve is stochastic and always flat.

This example demonstrates the difficulties of bond pricing: A casually specified model may give rise to arbitrage opportunities. A crucial feature of bond prices is the nonlinearity of prices as a function of interest rates, a characteristic implicitly ignored in equation (24.3). The same issue arises in pricing stock options: The nonlinearity of the option price with respect to the stock price is critical in pricing options. This is another example of Jensen's inequality.

The example also illustrates that, in general, *hedging a bond portfolio based on duration does not result in a perfect hedge*. Recall that the duration of a zero-coupon bond is the bond's time to maturity. The hedge ratio, equation (24.7), is *exactly* the same as equation (7.13) in Chapter 7. The use of duration to compute hedge ratios assumes that the yield to maturity of all bonds shifts by the same amount, which is what we assumed in equation (24.3). However, this assumption gives rise to arbitrage opportunities. The use of duration to compute hedge ratios can be a useful approximation; however, bonds in equilibrium *must* be priced in such a way that duration-based hedging does not work exactly.

An Equilibrium Equation for Bonds

Let's consider again the bond-hedging problem, only this time we will not assume a particular bond pricing model. Instead we view the bond as a general function of the short-term interest rate, r, which follows equation (24.2).[2]

[2]The discussion in this section follows Vasicek (1977).

First, let's see how the bond behaves. From Itô's Lemma, the bond, which is a function of the interest rate and time, follows the process

$$\begin{aligned} dP(r,t,T) &= \frac{\partial P}{\partial r}dr + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}(dr)^2 + \frac{\partial P}{\partial t}dt \\ &= \left[a(r)\frac{\partial P}{\partial r} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma(r)^2 + \frac{\partial P}{\partial t}\right]dt + \frac{\partial P}{\partial r}\sigma(r)dZ \end{aligned} \tag{24.10}$$

This equation does not look like equation (24.1), but we can define terms so that it does. Let

$$\alpha(r,t,T) = \frac{1}{P(r,t,T)}\left[a(r)\frac{\partial P}{\partial r} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma(r)^2 + \frac{\partial P}{\partial t}\right] \tag{24.11}$$

$$q(r,t,T) = -\frac{1}{P(r,t,T)}\frac{\partial P}{\partial r}\sigma(r) \tag{24.12}$$

We can now rewrite equation (24.10) as

$$\frac{dP(r,t,T)}{P(r,t,T)} = \alpha(r,t,T)dt - q(r,t,T)dZ \tag{24.13}$$

By using equations (24.11) and (24.12) to define α and q, equations (24.1) and (24.13) are the same. Note that α and q depend on both the interest rate and on the time to maturity of the bond.

Now we consider again the delta-hedged bond portfolio, the value of which is given by equation (24.4). From Itô's Lemma, we have

$$\begin{aligned} dI = N\left[\alpha(r,t,T_1)dt - q(r,t,T_1)dZ\right]P(r,t,T_1) \\ + \left[\alpha(r,t,T_2)dt - q(r,t,T_2)dZ\right]P(r,t,T_2) + rWdt \end{aligned} \tag{24.14}$$

In order to eliminate interest rate risk, we set

$$N = -\frac{P(r,t,T_2)}{P(r,t,T_1)}\frac{q(r,t,T_2)}{q(r,t,T_1)} \tag{24.15}$$

Note that by using the definition of q, equation (24.12), this can be rewritten

$$N = -\frac{P_r(r,t,T_2)}{P_r(r,t,T_1)}$$

If you compare this expression to equation (7.13), you will see that $-P_r(r,t,T)$ replaces duration when computing the hedge ratio, N.

Substituting equation (24.15) into equation (24.14), and setting $dI = 0$ (equation (24.8)), we obtain

$$\frac{\alpha(r,t,T_1) - r}{q(r,t,T_1)} = \frac{\alpha(r,t,T_2) - r}{q(r,t,T_2)} \tag{24.16}$$

This equation says that *the Sharpe ratio for the two bonds is equal.* Since both bond prices are driven by the same random term, dZ, they must have the same Sharpe ratio if they are fairly priced. (We demonstrated this proposition in Chapter 20.)

Denote the Sharpe ratio for dZ as $\phi(r, t)$. For any bond we then have

$$\frac{\alpha(r, t, T) - r}{q(r, t, T)} = \phi(r, t) \tag{24.17}$$

Substituting equations (24.11) and (24.12) for α and q then gives us

$$\boxed{\frac{1}{2}\sigma(r)^2\frac{\partial^2 P}{\partial r^2} + [a(r) + \sigma(r)\phi(r, t)]\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0} \tag{24.18}$$

When the short-term interest rate is the only source of uncertainty, *this partial differential equation must be satisfied by any zero-coupon bond.* Different bonds will have different maturity dates and therefore different boundary conditions. All bonds solve the same PDE, however. The Black-Scholes equation, equation (21.11), characterizes claims that are a function of the stock price. Equation (24.18) is the analogous equation for derivative claims that are a function of the interest rate.

A difference between equation (24.18) and equation (21.11) is the explicit appearance of the risk premium, $\sigma(r, t)\phi(r, t)$, in the bond equation. Let's talk about why that happens.

In the context of stock options, the Black-Scholes problem entails hedging an option with a stock, which is an investment asset. The stock is expected to earn its risk premium, which we will call $\phi'\sigma$. Thus, for the stock, the drift term, which is analogous to $a(r)$, equals $r + \phi'\sigma$. The Black-Scholes delta-hedging procedure eliminates the risk premium on the stock. By subtracting the risk premium, we are left with the risk-free rate, r, as a coefficient on the $\partial V/\partial S$ term in equation (21.11).

The interest rate, by contrast, is not the price of an investment asset. The interest rate is a *characteristic* of an asset, not an asset by itself. The risk-neutral process for the interest rate is obtained by subtracting the risk premium from the drift. The risk-neutral process for the interest rate is therefore

$$\boxed{dr = [a(r) + \sigma(r)\phi(r, t)]dt + \sigma(r)d\widetilde{Z}} \tag{24.19}$$

The drift in this equation is what appears in equation (24.18). You can also confirm that equation (24.18) is the same as equation (24.17).

Given a zero-coupon bond (which has a terminal boundary condition that the bond is worth \$1 at maturity), Cox et al. (1985b) show that the solution to equation (24.18) is

$$P[t, T, r(t)] = E_t^*\left[e^{-R(t,T)}\right] \tag{24.20}$$

where E^* represents the expectation taken with respect to risk-neutral probabilities and $R(t, T)$ is the random variable representing the cumulative interest rate over time:

$$R(t, T) = \int_t^T r(s)ds \tag{24.21}$$

Thus, to value a zero-coupon bond, we take the expectation over all the discount factors implied by these paths. We will see the discrete time analogue of this equation when we examine binomial models.

Keep in mind that it is *not* correct to value the bond by discounting the bond payoff by the average interest rate, $\bar{R} = E^*[R(t, T)]$:

$$P(t, T, r) \neq e^{-\bar{R}}$$

Because of Jensen's inequality, this seemingly reasonable procedure gives a different bond price than equation (24.20).

Different bond price models solve equation (24.20), differing only in the details of how r behaves and the modeling of the risk premium.

To summarize, a consistent approach to modeling bonds is to begin with a model of the interest rate and then use equation (24.18) to obtain a partial differential equation that describes the bond price (this equation is really the same as the Black-Scholes equation), but with a time-varying interest rate. Using the PDE together with boundary conditions, we can determine the price of the bond. If this seems familiar, it should: It is *exactly* the procedure we used to price options on stock.

The derivation of equation (24.18) assumes that bond prices are a function of a single state variable, the short-term interest rate r. It is possible to allow bond prices to depend on additional state variables, and there is empirical support for having bond prices depend on more than one state variable. Litterman and Scheinkman (1991) estimate a factor model for Treasury bond returns and find that a three-factor model typically explains more than 95% of the variability in a bond's return. They identify the three factors as level, steepness, and curvature of the yield curve. The single most important factor, the level of interest rates, accounts for almost 90% of the movement in bond returns. The overwhelming importance of the level of interest rates explains why duration-based hedging, despite its conceptual problems, is widely used. We will focus on single-variable models in this chapter.

Delta-Gamma Approximations for Bonds

One interpretation of equation (24.18) is familiar from Chapter 21. Using Itô's Lemma, the expected change in the bond price under the risk-neutral distribtion of the interest rate, equation (24.19), is

$$\frac{1}{dt}E^*(dP) = \frac{1}{2}\sigma(r)^2\frac{\partial^2 P}{\partial r^2} + [a(r) + \sigma(r)\phi(r, t)]\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t}$$

Equation (24.18) therefore says that

$$\frac{1}{dt}E^*(dP) = rP \tag{24.22}$$

This is the same as equation (21.31) for options: Using the risk-neutral distribution, bonds are priced to earn the risk-free rate.

The fact that bonds satisfy equation (24.22) means that, as in Chapter 13, *the delta-gamma-theta approximation for the change in a bond price holds exactly if the interest rate moves one standard deviation*. However, the Greeks for a bond are not exactly the same as duration and convexity.

We discussed bond duration and convexity in Chapter 7. For a zero-coupon bond, duration is time to maturity and convexity is squared time to maturity. Conceptually it seems as if duration should be the delta of a bond and convexity should be gamma. However, this is true only in the "impossible" bond pricing model of equation (24.3). For any correct bond pricing model, duration and convexity will be different than P_r/P and P_{rr}/P. We will see examples of this in the next section.

24.2 EQUILIBRIUM SHORT-RATE BOND PRICE MODELS

In this section we discuss several bond pricing models based on equation (24.18), in which all bond prices are driven by the short-term interest rate, r. The three pricing models we discuss—Rendleman-Bartter, Vasicek, and Cox-Ingersoll-Ross—differ in their specification of $\alpha(r)$, $\sigma(r)$, and $\phi(r)$. These differences can result in very different pricing implications.

The Rendelman-Bartter Model

The simplest models of the short-term interest rate are those in which the interest rate follows arithmetic or geometric Brownian motion. For example, we could write

$$dr = a dt + \sigma dZ \tag{24.23}$$

In this specification, the short-rate is normally distributed with mean $r_0 + at$ and variance $\sigma^2 t$. There are several objections to this model:

- The short-rate can be negative. It is not reasonable to think the *nominal* short-rate can be negative, since if it were, investors would prefer holding cash under a mattress to holding bonds.
- The drift in the short-rate is constant. If $a > 0$, for example, the short-rate will drift up over time forever. In practice if the short-rate rises, we expect it to fall; i.e., it is *mean-reverting*.
- The volatility of the short-rate is the same whether the rate is high or low. In practice, we expect the short-rate to be more volatile if rates are high.

The Rendleman and Bartter (1980) model, by contrast, assumes that the short-rate follows geometric Brownian motion:

$$dr = a r dt + \sigma r dz \tag{24.24}$$

While interest rates can never be negative in this model, one objection to equation (24.24) is that interest rates can be arbitrarily high. In practice we would expect rates to exhibit mean reversion; if rates are high, we expect them on average to decrease. The Rendleman-Bartter model, on the other hand, says that the probability of rates going up or down is the same whether rates are 100% or 1%.

TABLE 24.1 Expected change in the interest rate in the Vasicek model. Assumes $a = 0.2$, $b = 0.1$, and $\sigma = 0.01$.

Short-Rate	Expected Change in Short-Rate
5%	0.01
10%	0
15%	−0.01
20%	−0.02

The Vasicek Model

The Vasicek model incorporates mean reversion:

$$\boxed{dr = a(b - r)dt + \sigma dz} \tag{24.25}$$

This is an Ornstein-Uhlenbeck process (see Chapter 20). The $a(b - r)dt$ term induces mean reversion. Suppose we set $a = 20\%$, $b = 10\%$, and $\sigma = 1\%$. These parameters imply that a one-standard-deviation move for the short-rate is 100 basis points. The parameter b is the level to which short-term interest rates revert. If $r > b$, the short-rate is expected to decrease. If $r < b$, the short-rate is expected to rise. Table 24.1 illustrates mean reversion.

The parameter a reflects the speed with which the interest rate adjusts to b. If $a = 0$, then the short-rate is a random walk. If $a = 1$, the gap between the short-rate and b is expected to be closed in a year. If $a = 20\%$, we expect the rate to decrease in the first year by 20% of the gap.

Note also that the term multiplying dz is simply σ, independent of the level of the interest rate. This formulation implies that it is possible for interest rates to become negative and that the variability of interest rates is independent of the level of rates.

In the Rendleman-Bartter model, interest rates could not be negative because both the mean and variance in that model are proportional to the level of the interest rate. Thus, as the short-rate approaches zero, both the mean and variance also approach zero, and it is never possible for the rate to fall below zero. In the Vasicek model, by contrast, rates can become negative because the variance does not vanish as r approaches zero.

Why would anyone construct a model that permitted negative interest rates? Vasicek used equation (24.25) to illustrate the more general pricing methodology outlined in Section 24.1, not because it was a compelling empirical description of interest rates. The Vasicek model does in fact have some unreasonable pricing implications, in particular negative yields for long-term bonds.

We can solve for the price of a pure discount bond in the Vasicek model. Let the Sharpe ratio for interest rate risk be ϕ. With the Vasicek interest rate dynamics, equation

(24.25), equation (24.18) becomes

$$\frac{1}{2}\sigma^2\frac{\partial^2 P}{\partial r^2} + [a(b-r) + \sigma\phi]\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0$$

The bond price formula that solves this equation subject to the boundary condition $P(T, T, r) = 1$, and assuming $a \neq 0$, is[3]

$$P[t, T, r(t)] = A(t, T)e^{-B(t,T)r(t)} \tag{24.26}$$

where

$$A(t, T) = e^{\bar{r}(B(t,T)+t-T)-B(t,T)^2\sigma^2/4a}$$
$$B(t, T) = (1 - e^{-a(T-t)})/a$$
$$\bar{r} = b + \sigma\phi/a - 0.5\sigma^2/a^2$$

with $\bar{r}$ being the yield to maturity on an infinitely lived bond.

The Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross (CIR) model (Cox et al., 1985b) assumes a short-term interest rate model of the form

$$\boxed{dr = a(b-r)dt + \sigma\sqrt{r}dz} \tag{24.27}$$

The standard deviation of the interest rate is proportional to the square root of the interest rate, instead of being constant as in the Vasicek model. Because of this subtle difference, the CIR model satisfies all the objections to the earlier models:

- It is impossible for interest rates to be negative. If $r = 0$, the drift in the rate is positive and the variance is zero, so the rate will become positive.
- As the short-rate rises, the volatility of the short-rate also rises.
- The short-rate exhibits mean reversion.

The assumption that the variance is proportional to $\sqrt{r}$ also turns out to be convenient analytically—Cox, Ingersoll, and Ross (CIR) derive bond and option pricing formulas using this model. The Sharpe ratio in the CIR model takes the form

$$\phi(r, t) = \bar{\phi}\sqrt{r}/\sigma \tag{24.28}$$

[3]When $a = 0$, the solution is equation (24.26), with

$$A = e^{0.5\sigma\phi(T-t)^2+\sigma^2(T-t)^3/6}$$
$$B = T - t$$

When $a = 0$ the interest rate follows a random walk; therefore, $\bar{r}$ is undefined.

With this specification for the risk premium and equation (24.27), the CIR interest rate dynamics, the partial differential equation for the bond price is

$$\frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} + [a(b-r) + r\bar{\phi}]\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} - rP = 0$$

The CIR bond price looks similar to that for the Vasicek dynamics, equation (24.26), but with $A(t, T)$ and $B(t, T)$ defined differently:

$$P[t, T, r(t)] = A(t, T)e^{-B(t,T)r(t)} \tag{24.29}$$

where

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

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

$$\gamma = \sqrt{(a-\bar{\phi})^2 + 2\sigma^2}$$

With the CIR process, the yield on a long-term bond approaches the value $\bar{r} = 2ab/(a - \bar{\phi} + \gamma)$ as time to maturity goes to infinity.

Comparing Vasicek and CIR

How different are the prices generated by the CIR and Vasicek models? What is the role of the different variance specifications in the two models?

Figure 24.1 illustrates the yield curves generated by the Vasicek and by the CIR models, assuming that the current short-term rate, r, is 5%, $a = 0.2$ and $b = 10\%$. Volatility in the Vasicek model is 2% in the top panel and 10% in the bottom panel. The volatility, σ, has a different interpretation in each model. In the Vasicek model, volatility is absolute, whereas in the CIR model, volatility is scaled by the square root of the current interest rate. To make the CIR volatility comparable at the initial interest rate, it is set so that $\sigma_{\text{CIR}}\sqrt{r} = \sigma_{\text{Vasicek}}$, or 0.0894 in the top panel and 0.447 in the bottom panel. The interest rate risk premium is assumed to be zero.

The two models can exhibit very different behavior. The bottom panel has a relatively high volatility. For short-term bonds—with a maturity extending to about 2.5 years—the yield curves look similar. This is a result of setting the CIR volatility to match the Vasicek volatility. Beyond that point the two diverge, with Vasicek yields below CIR yields. The long-run interest rate in the Vasicek model is −0.025, whereas that in the CIR model is 0.0463. This difference is evident in Figure 24.1 as the Vasicek yields approach zero (in the long run approaching −0.025).

What accounts for the difference in medium to long-term bonds? As discussed earlier, the pricing formulas are based on *averages* of interest rate paths, as in equation (24.20). Some of the interest paths in the Vasicek model will be negative. Although