Advanced Financial Risk Management:

Tools and Techniques for Integrated Credit Risk, Market Risk, Liquidity Risk and Interest Rate Risk Management

April 2012

Donald R. van Deventer
Kenji Imai
Mark Mesler

Part 2: Risk Management Techniques for Interest Rate Analytics

Chapter 13:
Special Cases of Heath Jarrow and Morton Interesting Rate Modeling

In this chapter, we continue our review of the history of interest rate analytics, but this time we focus on term structure models. We start with the example of parallel yield curves shifts and duration as a naïve term structure model to illustrate the approach academics have typically taken to interest rate modeling:

Step 1: Make an assumption about how interest rates vary
Step 2: Impose no-arbitrage restrictions
Step 3: Derive what the shape of the yield curve must be

The approach taken by Heath Jarrow and Morton was the exact opposite:

Step 1: Observe what the shape of the yield curve actually is
Step 2: Add assumptions about the volatility of rates
Step 3: Impose no-arbitrage conditions on yield curve movement from its current shape
Step 4: Derive the nature of yield curve movements from the existing yield curve shape
The power of the HJM approach was illustrated in Chapters 6 through 9, with four different sets of assumptions about interest rate volatility. In this chapter, we demonstrate that the early academic

work on term structure model, in every case, is a special case of HJM. By special case, we mean a specific assumption in Step 2 of the general HJM process above.

In Chapter 12, we were introduced to Frederick Macaulay's duration concept. At the heart of the duration concept is the implicit assumption that "rates" at all maturities move at the same time in the same direction by the same absolute amount. This assumption about how rates move can be labeled a "term structure model." At the same time, it's known on trader floors all over the world as "PVBP," the present value of a basis point, a common measure of risk.

The purpose of this chapter is to introduce the traditional academic approach to term structure models and the procedures for evaluating the reasonableness of a given term structure model. We show in each case that the resulting model is a special case of the Heath Jarrow and Morton approach. We review historical developments in term structure modeling to show readers and adherents to these traditional models that a more realistic approach to interest rate modeling using 6 to 10 risk factors in the Heath Jarrow and Morton framework is a logical step forward from what has been done in the past. We go on to fully implement the HJM term structure model concept in later chapters so that we can calculate the credit-adjusted value of the Jarrow-Merton put option, the best measure of risk, as outlined in Chapter 1.

What is an academic term structure model and why were they developed?

When analyzing fixed income portfolios, the value of a balance sheet, or a fixed income option it is not enough to know where interest rates currently are. We demonstrated this clearly in Chapters 6 through 9. It's not possible to know where rates will be in the future with any certainty. Instead, we can analyze a portfolio or a specific security by making an assumption about the random process by which rates will move in the future. Once we have made an assumption about the random (or "stochastic") process by which rates move, we can then derive the probability distribution of interest rates at any point in the future. Without this kind of assumption, we can't analyze interest rate risk effectively. As we noted above, the early academic work on term structure modeling ignored the current shape of the yield curve in this derivation.

The vocabulary of term structure models

The mathematics for analyzing something whose value changes continuously but in a random process is heavily used in physics, and these tools from physics have been responsible for most of the derivatives valuation formulas used in finance in the last five decades. At the risk of sending some readers searching for the remote control device, we want to spend a little time on these mathematical techniques because of the power they bring to a clear understanding of risk and what to do about it. The tools we describe in this chapter were employed by Heath, Jarrow and Morton to derive the no arbitrage n factor term structure model framework that we employed in Chapters 6 through 9.

In the "stochastic process" mathematics used in physics, the change in a random variable x over the next instant is written dx. The most common assumption (but not the only assumption) about the random jumps in the size of a variable x is that they are normally distributed. The "noise" or shock to the random variable is assumed to come from a random number generator called a Wiener process. It is a random variable, say z, whose value changes from instant to instant but which has a normal distribution with mean zero and standard deviation of 1. For example, if we want to assume that the continuously compounded zero coupon bond yield y(t) for a given maturity t changes continuously with a mean change of zero and a standard deviation of 1, we would write

$$dy = dz$$

In stochastic process mathematics, all interest rates are written as decimals like .03 instead of 3, so the assumption that rates jump each instant by 1 (which means 100%) is too extreme. We can scale the jumps of interest rates by multiplying the noise generating Wiener process by a scale factor. In the case of interest rates, this scale factor is the "volatility" of the stochastic process driving interest rate movements. This "volatility" is related to but different from the "volatility" of stock prices assumed by Black and Scholes to be lognormally distributed. In the Black-Scholes model, the return on stock prices is assumed to be normally distributed, not stock prices themselves. In our simple interest rate model, the change in interest rates is assumed to be driven by this normally distributed "shock term" z with changes dz. We label this interest rate volatility σ2. In that case, the stochastic process for the movement in yields is

$$dy = \sigma dz$$

The volatility term allows us to scale the size of the shocks to interest rates up and down. We can set this interest rate volatility scale factor to any level we want, although normally it would be set to levels which most realistically reflect the movements of interest rates. We can calculate this implied interest rate volatility from the prices of observable securities like interest rate caps, interest rate floors, or callable bonds. We will show the relationship between interest rate volatility and the volatility of the Black-Scholes model in Chapter 18 on the random interest rates version of the Merton credit model.

The yield that we have modeled above is a random walk, a stochastic process where the random variable drifts randomly with no trend up or down. As time passes, ultimately the yield may rise to infinity or fall to negative infinity. This isn't a very realistic assumption about interest rates, which usually move in cycles, depending on the country and the central bank's approach to monetary policy. How can we introduce interest rate cycles to our model? We need to introduce some form of drift in interest rates over time. One form of drift is to assume that interest rates change by some formula as time passes:

$$dy = \alpha(t)dt + \sigma dz$$

The formula above assumes that on average, the change in interest rates over a given instant will be the change given by the function α 5(t), with random shocks in the amount of σ6dz. This formula is closely linked to the term structure models proposed by Ho and Lee [1986], with is the simplest of the single factor models within the Heath, Jarrow and Morton [1992] framework. The Ho and Lee model, generalized by Heath Jarrow and Morton, was the first term structure model to fit actual observable yield curve data exactly. The function α 7(t) can be chosen to fit a yield curve smoothed using the techniques of Chapter 5 exactly. We go into more detail on the Ho and Lee special case of HJM later in this chapter.

The assumption about yield movements above isn't satisfactory in the sense that it is still a random walk of interest rates, although there is the drift term α 8(t) built into the random walk. The best way to build the interest rate cycle into the random movement of interest rates is to assume that the interest rate drifts back to some long run level, except for the random shocks from the dz term. One process which does this is

$$dy = \alpha(\gamma - y)dt + \sigma dz$$

This is the Ornstein-Uhlenbeck process which is at the heart of the Vasicek [1977] model which we introduce below. As we note below, the Vasicek model is another special case of the HJM framework. The term α 10 above is called the "speed of mean reversion." It is assumed to be a positive number. The larger it is, the faster y drifts back toward its long run mean γ 11 and the shorter and more violent interest rate cycles will be. Since α 12 is positive, when y is above γ 13, it will be pulled down except for the impact of the shocks which emanate from the dz term. When y is less than γ 14, it will be pulled up.

In almost every early term structure model, the speed of mean reversion and the volatility of interest rates play a key role in determining the value of securities which either are interest options or which have interest rate options embedded in them. Many of the early term structure models, in spite of their simplicity, were powerful enough and analytically attractive enough that they were used in both the random interest rate version of the Merton credit model (Chapter 18) and the more modern reduced form credit models of Duffie and Singleton [1999] and Jarrow[1999, 2001] which we discuss in Chapter 16.

For that reason, it's important to understand a little more about how these interest rate models are derived.

Ito's lemma

Once a term structure model has been chosen, we need to be able to draw conclusions about how securities whose values depend on interest rates should move around. For instance, if a given interest rate y is assumed to be the random factor than determines the price of a zero coupon bond P, it is logical to ask how the zero coupon bond price moves as time passes and y varies. The formula used to do this is called Ito's lemma.[1] Ito's lemma puts the movement in the bond price P in stochastic process terms like this:

[1] For more on Ito's lemma, see Shimko [1992] for implementations in applied finance.

$$dP = P_y dy + \frac{1}{2} P_{yy} (dy)^2 + P_t$$

where the subscripts denote partial derivatives. This formula looks very much like the Taylor series expansion that we used in our discussion on convexity in Chapter 12. The change in the price of a zero coupon bond equals its first derivative times the change in y plus one-half of its second derivative times the volatility of y plus the "drift" in prices over time. The terms dy and $(dy)^2$ depend on the stochastic process chosen for y. If the stochastic process is the Ornstein-Uhlenbeck process given above, we can substitute in the values for dy and $(dy)^2$. If we do this, then the movements in P can be rewritten

$$\begin{aligned} dP &= P_y\left[\alpha(\gamma - y)dt + \sigma dz\right] + \frac{1}{2} P_{yy}\sigma^2 dt + P_t dt \\ &= \left[P_y \alpha(\gamma - y) + \frac{1}{2} P_{yy}\sigma^2 + P_t\right]dt + P_y \sigma dz \\ &= g(y,t)dt + h(y,t)dz \end{aligned}$$

For any stochastic process, the dy term is the stochastic process itself. The term $(dy)^2$ is the instantaneous variance of y. In the case of the Ornstein-Uhlenbeck process, it is the square of the coefficient of dz, $\sigma 17^2$. The term g(y,t), which depends on the level of rates y and time t, is the "drift" in the bond price. The term h(y,t) is the bond price volatility in the term structure model sense, not the Black-Scholes sense. We have very neatly divided the expression for movements in the bond's price into two pieces. The first term depends on the level of interest rates and time, and the second term depends on the dz term which introduces random shocks to the interest rate market. If dz or h were always zero, then interest rates would not be random. They would drift over time, but there would be no surprises. Not surprisingly, we will devote a lot of time to discussing the hedging that eliminates these interest rate shocks.

Ito's Lemma for More Than One Random Variable

What if the zero coupon bond price depended on two random variables, x and y? Then we would be dealing with a "two factor" term structure model, like the HJM two factor model in Chapter 8. The formula for the movement in the bond's price for such a two factor term structure model is given by Ito's lemma as follows:

$$dP = P_x dx + P_y dy + \frac{1}{2} P_{xx}(dx)^2 + \frac{1}{2} P_{yy}(dy)^2 + P_{xy}(dxdy) + P_t$$

The instantaneous correlation between the two Wiener processes Z_x and Z_y which provide the random shocks to x and y is reflected in the formula for the random movement of P by the instantaneous correlation coefficient rho; (dxdy) is defined as $\rho\sigma_x\sigma_y$. The sigmas are the instantaneous variances of the two variables driving this term structure model, parallel to the variance of our original "single factor" term structure model above. David Shimko [1992] has an excellent discussion of how to apply this kind of arithmetic to many problems in finance. Our need is restricted to term structure models, so we now turn to the construction of a single factor term structure model based on the duration concept.

Using Ito's Lemma to build a term structure model

In their pioneering work, Heath, Jarrow and Morton[1992] clearly outline the steps that must be followed in developing a term structure model. There are potentially hundreds of alternative term structure models but their derivation has a common core. The steps involved in building a term structure model almost always follow this process:

1. Make an assumption about the random (stochastic) process that interest rates follow that is consistent with market behavior in general.
2. Use Ito's lemma to specify how zero coupon bond prices move
3. Impose the constraint that there not be riskless arbitrage in the bond market via the following steps:
    a. Use a bond of one maturity (in the case of a "one factor" model) to hedge the risk of a bond which has another maturity[2]
    b. After eliminating the risk of this "portfolio," impose the constraint that this hedged portfolio earn the instantaneous (short term) risk free rate of interest.
4. Solve the resulting partial differential equation for zero coupon bond prices.
5. Examine whether this implies reasonable or unreasonable conditions on the market
6. If the economic implications are reasonable, proceed to value other securities with the model. If they are unreasonable, reject the assumption about how rates move as a reasonable basis for a term structure model.

We will illustrate this process by analyzing the economic implications of the assumption of parallel yield curve shifts underlying the traditional duration analysis.

Duration as a term structure model

[2] If there are two random factors driving movements in interest rates, two bonds would be necessary to eliminate the risk of these random factors. N bonds would be necessary for an N factor model. This parallels the HJM hedging example in Chapter 12.

In this section, we use the continuous compounding formulas in Chapters 4 and 12 to analyze the yield to maturity y_i on a zero coupon bond with maturity t_i and its value P.

$$P(t_i) = e^{-y_i t_i}$$

We want to examine what happens when all yields at time zero y_0 shift by a parallel amount x, with x being the same for all maturities. After the shift of x, the price of a zero coupon bond with maturity τ 19=T-t (t is the current time in years and T is the maturity date of the bond in years, for a net maturity of T-t) will be

$$P(\tau) = e^{-(y_0+x)\tau}$$

where $y=y_0+x$. We assume that x is initially zero and that we are given today's yield curve and know the values of y for all maturities.

We now make our assumption about the term structure model for interest rates. We assume that the movements in x follow a random walk, that is we assume changes in x have no drift term and that the change in x is normally distributed as follows:

$$dx = \sigma dz$$

The scale factor σ22 controls the volatility of the parallel shifts x. We now proceed to step 2 of the previous section by using Ito's lemma to specify how the zero coupon bond price for a bond with time to maturity T-t moves. We use the formulas for the first and second derivatives of the bond price P and our knowledge of what dx and $(dx)^2$ are. Ito's lemma says

$$\begin{aligned} dP &= P_x dx + \frac{1}{2} P_{xx} (dx)^2 + P_t \\ &= -\tau P \sigma dz + \frac{1}{2} \tau^2 P \sigma^2 dt + yPdt \\ &= \left[\frac{1}{2} \tau^2 P \sigma^2 + yP \right] dt - \tau P \sigma dz \\ &= g(y,t)Pdt + h(y,t)Pdz \end{aligned}$$

because

$$y = y_0 + x$$
$$\tau = T - t$$
$$P = \exp^{-y\tau}$$
$$P_x = -\tau P$$
$$P_{xx} = \tau^2 P$$
$$P_t = yP$$
$$g(y,t) = \frac{1}{2}\tau^2\sigma^2 + y$$
$$h(y,t) = -\tau\sigma$$

We now go to step 3 as outlined in the previous section. We want to form a portfolio of 1 unit of bond 1 with maturity τ 25$_1$ and w units of the bond 2 with maturity τ 26$_2$. This is exactly the same kind of hedge construction that we did in Chapter 12. The value of this portfolio is

$$W = P_1 + w\,P_2$$

We then apply Ito's lemma to changes in the value of the portfolio as the parallel shift amount x moves:

$$\begin{aligned} dW &= dP_1 + w\,dP_2 \\ &= g_1 P_1 dt + h_1 P_1 dz + w\left[g_2 P_2 dt + h_2 P_2 dz\right] \\ &= \left[g_1 P_1 + w\,g_2 P_2\right]dt + \left[h_1 P_1 + w\,h_2 P_2\right]dz \end{aligned}$$

As in step 3b above, we want to eliminate the interest rate risk in this portfolio. Since all interest rate risk comes from the random shock term dz, "eliminating the interest rate risk" means choosing w such that the coefficient of dz is zero. This means that the proper hedge ratio w for zero interest rate risk is

$$w = \frac{-h_1 P_1}{h_2 P_2} = \frac{\tau_1 P_1}{\tau_2 P_2}$$

Substituting this into the equation above means

$$dW = \left[g_1 P_1 + \frac{\tau_1 P_1}{\tau_2 P_2} g_2 P_2\right]dt$$

Now we impose the no arbitrage condition. Since the interest rate risk has been eliminated from this portfolio, the instantaneous return on the portfolio dW should equal the riskless[3] short term rate (the rate y with maturity 0, y(0)) times the value of the portfolio W:

$$dW = r(P_1 + w P_2)dt$$
$$= \left[g_1 P_1 + w g_2 P_2 \right] dt$$

Rewritten, this means that

$$\frac{r P_1 - \left(\frac{1}{2} \tau_1^2 \sigma^2 + y_1 \right) P_1}{\tau_1 P_1} = \frac{r P_2 - \left(\frac{1}{2} \tau_2^2 \sigma^2 + y_2 \right) P_2}{\tau_2 P_2} = k$$

This ratio has to be equal for any two maturities or there will be riskless arbitrage opportunities in this bond market. We define this ratio k as the "market price of risk". Choosing bond 1 and dropping the subscript 1, this means that the yield y must satisfy the following relationship:

$$rP - \frac{1}{2} \tau^2 \sigma^2 P - yP = k \tau P$$

Rearranging means that the yield y for any maturity must adhere to the following relationship:

$$y(\tau) = r - k\tau - \frac{1}{2} \tau^2 \sigma^2$$

[3] We have implicitly assumed in this chapter that we are modeling the risk free term structure. We will relax this assumption in Chapter 17.

This equation comes from step 4 in the section above. For a no arbitrage equilibrium in the bond market under a parallel shift in the yield curve, the yield y must be a quadratic function of time to maturity τ 35. If the function is not quadratic, there would be riskless arbitrage in the market. This is a surprising and little understood implication of the assumption of parallel yield curve shifts in the market. Clearly, the implication that yields are quadratic is good news and bad news. The good news is that the yield curve math is simple and can fit a lot of observable yield curves over some portion of their range. The bad news is the basic shape of the quadratic function!

Conclusions About The Use of Duration's Parallel Shift Assumptions

Unfortunately, the assumption that yields move in parallel fashion results in a number of conclusions that don't make sense if we impose the no-riskless-arbitrage conditions: First, interest rates can have only one "hump" in the yield curve. This hump will occur at the maturity

$$\tau = -\frac{k}{\sigma^2}$$

Much more important, beyond a certain maturity, yields will turn ***negative*** since sigma is positive. Yields will zero for two values of T defined by setting y equal to zero and solving for the maturities consistent with zero yield:

$$T = -\frac{k}{\sigma^2} \pm \sqrt{\frac{2r}{\sigma^2} + \frac{k^2}{\sigma^4}}$$

For normal values of k, these values of T define a range over which y will be positive. Outside of that range, yields to maturity will be negative. Forward rates are also quadratic and have the form

$$f(\tau) = r - 2k\tau - \frac{3}{2}\tau^2\sigma^2$$

because forward rates are related to yields by the formula

$$f(\tau) = y(\tau) + \tau y'(\tau)$$

as shown in Chapter 5. Zero coupon bond prices are given by

$$P(\tau) = e^{-y(\tau)\tau} = e^{-r\tau + k\tau^2 + \frac{1}{2}\sigma^2\tau^3}$$

These features of a parallel shift-based term structure model pose very serious problems. If the parallel yield curve shift assumption is used, as it is when one uses the traditional duration approach to hedging, the resulting bond market is either

(a) inconsistent with the real world because of negative yields, if we impose the no arbitrage conditions, or

(b) allowing of riskless arbitrage if we let yields have a shape that is not quadratic in the term to maturity.

In either case, the model is unacceptable for use by market participants because its flaws are so serious. The most important point of this chapter, however, is to show how we can move in a very systematic way from an assumption about how interest rates move to explicit formulas for forward rates, zero coupon bond prices, zero coupon bond yields, and many other complex securities. We have laid the groundwork for “extending” the models to be consistent with observable yield curves, which we demonstrated with HJM 1, 2 and 3 factor models in Chapters 6 through 9. Most importantly, we have laid the groundwork for valuing the Jarrow-Merton put option that quantifies interest rate risk, market risk, liquidity risk and credit risk in a fully integrated way.

We now use the same sort of analysis to identify some more attractive alternative models of the term structure that were developed prior to the publication of the Heath, Jarrow and Morton approach (1990, 1990, 1992).

The Vasicek and Extended Vasicek Models

In this section, we introduce four term structure models which are based on increasingly realistic assumptions about the random movement of interest rates. All of the models are special cases of the Heath Jarrow and Morton framework. As we saw in Chapters 8 and 9, the legacy assumption that the yield curve shifts in a parallel fashion is inconsistent with actual movements of interest rates over 50 years of daily data in the United States. Parallel yield curve shifts can be consistent with a no arbitrage economy, but only if the yield curve is quadratic and only if yields become negative beyond a given point on the yield curve. For this reason, academics sought a richer set of assumptions about yield curve movements that would allow us to derive a theoretical yield curve whose shape is as close as possible, if not identical, to observable market yields and whose other properties are realistic.

This chapter will be too mathematical for many practitioners and too simple for many hard-core "rocket scientists." We ask both groups to hang in there as our objective in this volume is to show clearly the links between good theory and best practice in risk management. Along the way, we will point the readers to complementary sources that are either more practical or more theoretical, depending on the point of view of the reader.

We start with the simplest possible model. We will use the same assumptions for each of the four models we review in this chapter. We will assume bond prices in all four models depend solely on a single random factor, the short rate of interest. This is a simpler framework than we employed using the two one factor HJM models in Chapters 6 and 7. We keep it simple for expository purposes. By the end of the chapter, the reader will be grateful that the insights of HJM are so powerful that the sorts of derivations we now turn to are no longer necessary.

As we go through each of these four single factor term structure models, we go through the same steps as in the prior section to derive the form that zero coupon bond prices and yields must have for the risk-free bond market to be a no arbitrage market. The HJM framework, by contrast, derives the no arbitrage framework for n factors, with 1 factor as the simplest special case. We go through the normal academic term structure model process:

- We specify the stochastic process for the movements in interest rates
- We use Ito's lemma to derive the stochastic movements in zero coupon bond prices
- We impose no arbitrage restrictions to generate the partial differential equation that determines zero coupon bond prices
- We solve this equation for bond prices, subject to the boundary condition that the price of a zero coupon bond with zero years to maturity must be 1.

The four models we will discuss all assume that changes in the short term rate of interest are normally distributed. This assumption is a little more questionable in low rate environments like Japan since the mid-1990s and the U.S. in the wake of the 2007-2009 credit crisis. The HJM framework does not require this specification, but we use it in the remainder of this chapter for expository reasons. The four models we will review here can be summarized like this:

1. Parallel rate shock model, where changes in the short term risk free rate r are a random walk with no drift over time. This is the same as the continuous time duration model reviewed in the prior section. It was originally introduced by Robert Merton. Because the model has zero drift in rates over time, this model generally will not exactly match the observable yield curve.

2. The extended parallel rate shock model which matches the observable yield curve exactly. Changes in the short term risk free rate of interest r are a random walk with non-zero drift, allowing a perfect fit to an observable yield curve. We label this model the "extended" Merton model or Ho and Lee model.

3. A non-parallel rate shift model, in which changes in the risk free short term rate of interest r allow for interest rate cycles. Analytically, we would say that changes in r follow the “Ornstein-Uhlenbeck” process with a drift term, and r is reverting to a constant long term value plus normally distributed shocks (the Vasicek [1977] model). It is this “mean reversion” which causes cycles in interest rates.

4. A non-parallel rate shift model which fits the observable yield curve exactly. In this model changes in r follow the Ornstein-Uhlenbeck process again. The drift in r is such that we match the observable yield curve. This model is called the "extended" Vasicek or Hull and White model.

Each of these models has their strengths and weaknesses. All are special cases of the HJM approach and less realistic than any of the HJM implementations in Chapters 6 through 9. We discuss each in term. We use these models to fit the risk-free yield curves in the countries we care about. We build the credit models of Chapters 16 and 17 on top of them to get the yield curves for risky issuers. The term structure models in this chapter should NOT be used for yield curves with credit risk.

The Merton Term Structure Model: Parallel Yield Curve Shifts

One simple assumption about the short term risk free interest rate r is that it follows a simple random walk[4] with a zero drift over time. In the earlier part of this chapter, we applied this assumption to zero coupon bond yields of all maturities and required that this random shift in rates be the same for all maturities. Robert Merton [1970] applied the random walk assumption to the short term rate of interest r and derived what this means for the rest of the yield curve. We can write the change in r as

$$dr = \sigma dZ$$

[4] To be precise, a Gaussian random walk.

The change in the short rate of interest r equals a constant sigma times a random shock term, where Z represents a standard Wiener process with mean zero and standard deviation of 1. The constant sigma is the instantaneous volatility of interest rates which we can scale up and down to maximize the goodness of fit versus the observable yield curve. We will be careful to distinguish between interest rate volatility and the volatility of zero coupon bond prices throughout the rest of this book. From above, we know that Ito's lemma can be used to write the stochastic process for the random changes in the price of a zero coupon bond price as of time t with maturity T:

$$dP = P_r\,dr + \frac{1}{2}P_{rr}(dr)^2 + P_t$$

The change in the bond's price equals its first derivative with respect to r times the change in r, plus a term multiplied by the second derivative of the bond's price and a drift term, the derivative of the bond's price with respect to the current time t. In this random walk model, the variance of the interest r is $(dr)^2$ = sigma^2, so the expression above can be expanded to read

$$dP = P_r\,\sigma dZ + \frac{1}{2}P_{rr}\,\sigma^2 + P_t$$

Our goal is to find the formula for P as a function of the short rate r. Once we do this, we can vary the time to maturity T and derive the shape of the entire yield curve and how the yield curve moves in response to changes in r.

The next step in doing this is to impose the condition that no arbitrage is possible in the bond market. Since we have only one random factor in the economy, we know that we can eliminate this risk factor by hedging with only one instrument, as we pointed out in Chapter 12. If we had n independent risk factors, we would need n instruments to eliminate the n risks. Let's assume that an investor holds one unit of a zero coupon bond, bond 1, with maturity T_1. The investor forms a portfolio of amount W which consists of the one unit of bond 1 and w units of bond 2, which has a maturity T_2. The bond with maturity T_2 is the hedging instrument. The value of the portfolio is

$$W = P_1 + w\,P_2$$

What is the proper hedge ratio w, and what are the implications of a perfect hedge for bond pricing? We know from Ito's lemma that the change in the value of this hedged portfolio is

$$dW = dP_1 + w\,dP_2$$
$$= P_{1_r}\sigma dZ + \left(\frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_t}\right)dt + w\left[P_{2_r}\sigma dZ + \left(\frac{1}{2}P_{2_{rr}}\sigma^2 + P_{2_t}\right)dt\right]$$

Gathering the coefficients of the random shock term dZ together gives

$$dW = \left(\frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_t}\right)dt + w\left(\frac{1}{2}P_{2_{rr}}\sigma^2 + P_{2_t}\right)dt + \left[P_{1_r} + wP_{2_r}\right]\sigma dZ$$

We repeat the process we followed above to arrive at the zero risk hedge. If we choose w, the hedge amount of bond 2, such that the coefficient of the random shock term dZ is zero, then we have a perfect hedge and a riskless portfolio. The value of w for which this is true is

$$w = \frac{-P_{1_r}}{P_{2_r}}$$

Note that this is identical to the hedge ratio we found for the duration-based parallel shift model. If we use this hedge ratio, then the instantaneous return on the portfolio should exactly equal the short term rate of interest times the value of the portfolio:

$$dW = rWdt = (r\,P_1 + rw\,P_2)dt$$

If this is not true, then riskless arbitrage will be possible. Imposing this condition and rearranging the equation above gives us the following relationship:

$$dW = \left(r\,P_1 + rw\,P_2\right)dt = \left(\frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_t}\right)dt + w\left(\frac{1}{2}P_{2_{rr}}\sigma^2 + P_{2_t}\right)dt$$

We then substitute the expression for w above into this equation, eliminate the dt coefficient from both sides, rearrange, and divide by interest rate volatility sigma to get the following relationship.

$$-\lambda = \frac{\frac{1}{2} P_{1_{rr}} \sigma^2 + P_{1_t} - r\, P_1}{\sigma\, P_{1_r}} = \frac{\frac{1}{2} P_{2_{rr}} \sigma^2 + P_{2_t} - r\, P_2}{\sigma\, P_{2_r}}$$

For any two maturities T_1 and T_2, the no arbitrage condition requires that this ratio be equal. We call the negative of this ratio λ,[5] the market price of risk. For normal risk aversion, the market price of risk should be positive because in a risk averse market riskier (longer maturity) bonds should have a higher expected return than the short rate r and because the rate sensitivity of all bonds P_r is negative. Since the choice of T_1 and T_2 is arbitrary, the market price of risk ratio must be constant for all maturities. ***It is the fixed income equivalent of the Sharpe ratio***, which measures "excess return" per unit of risk with the following statistic:

$$Sharpe\ Ratio = \frac{Expected\ Return\ -\ Risk\ Free\ Return}{Standard\ Deviation\ of\ Return}$$

We will return to the Sharpe ratio and the strong parallels between equilibrium in the fixed income and the equity markets frequently in this book. In our case, the numerator is made up of the drift in the bond's price

$$drift\ =\ expected\ return = \frac{1}{2} P_{rr} \sigma^2 + P_t$$

which also equals the expected return since the shock term or stochastic component of the bond's price change on average has zero expected value.

We now solve the equation above for any given bond with maturity T and remaining time to maturity T-t=τ as of time t, under the assumption that lambda (the market price of risk) is constant over time. Rearranging the equation above gives the following partial differential equation

$$\lambda\sigma P_r + \frac{1}{2} P_{rr} \sigma^2 + P_t - rP = 0$$

which must be solved subject to the boundary condition that the price of the zero coupon bond upon maturity T must equal its principal amount 1:

$$P(r,T,T) = 1$$

---

5 We insert the minus sign in front of λ**Error! Main Document Only.** to be consistent with other authors' notation and subsequent chapters.

We use a common method of solving partial differential equations, the educated guess. We guess that P has the solution

$$P(r,t,T) = P(r,\tau) = e^{rF(\tau)+G(\tau)}$$

where T-t= τ and we need to solve for the forms of the functions F and G. We know that

$$P_r = F(\tau)P$$

$$P_{rr} = F(\tau)^2$$

$$P_t = -P_\tau = (-rF'-G')P$$

Substituting these values into the partial differential equation and dividing by the bond price P gives the following:

$$\lambda\sigma F + \frac{1}{2}\sigma^2 F^2 - rF' - G' - r = 0$$

We know from Merton [1970] and the similar equation above that the solution to this equation is

$$F(\tau) = -\tau$$

$$G(\tau) = -\frac{\lambda\sigma\tau^2}{2} + \frac{1}{6}\sigma^2\tau^3$$

and therefore the formula for the price of a zero coupon bond is given by

$$P(r,t,T) = P(r,\tau) = e^{-r\tau - \frac{\lambda\sigma\tau^2}{2} + \frac{1}{6}\sigma^2\tau^3}$$

This is nearly the same bond pricing equation that we obtained under the assumption of parallel shifts in bond prices. It has the same virtues and the same liabilities of the duration approach:

- It is a simple analytical formula
- Zero coupon bond yields are a quadratic function of time to maturity
- Yields turn negative (and zero coupon bond prices rise above one) beyond a certain point
- If interest rate volatility σ is zero, zero coupon bond yields are constant for all maturities and equal to r

Using continuous compounding,

$$P(r,\tau) = e^{-y(\tau)\tau}$$

and the zero coupon bond yield formula can be calculated as

$$y = r + \frac{1}{2}\lambda\sigma\tau - \frac{1}{6}\sigma^2\tau^2$$

The last term in this formula ultimately becomes so large as time to maturity increases that yields become negative. Setting y' = 0 and solving for the maximum level of y (the "hump" in the yield curve) shows that this occurs at a time to maturity of

$$\tau^* = \frac{3\lambda}{2\sigma}$$

and that the yield to maturity at that point is

$$y(\tau^*) = r + \frac{3}{8}\lambda^2$$

Note that the peak in the yield curve is independent of interest rate volatility, although the location of this hump is affected by rate volatility. As volatility increases, the hump moves to shorter and shorter maturity points on the yield curve. We can also determine the point at which the zero coupon bond yield equals zero using the quadratic formula. This occurs when the yield curve reaches a time to maturity of

$$\tau^* = \frac{3\lambda}{2\sigma} + \frac{3}{\sigma}\sqrt{\frac{\lambda^2}{4} + \frac{2r}{3}}$$

The market price of risk is normally expected to be positive so this formula shows rates will turn negative at a positive term to maturity τ^*.

The price volatility of zero coupon bonds is given by

$$\sigma P_r = -\tau\sigma P$$

Another useful tool that we can derive from this term structure model is the percentage change in the price of zero coupon bonds for small changes in the short rate r. Under the Merton term structure model this sensitivity to r is

$$r\text{-}duration = -\frac{P_r}{P} = -\frac{-\tau P}{P} = \tau$$

so the rate sensitivity of a zero coupon bond in this model is its time to maturity in years. We label this percentage change "r-duration" to contrast it with Macaulay's measure of price sensitivity discussed in earlier chapters, where price is differentiated with respect to its continuous yield to maturity. We can calculate the "r-duration" of coupon bearing bonds using the same logic. The hedge ratio w given above for hedging of a bond with maturity T_1 with a bond maturing at T_2 is

$$w = \frac{-P_{1_r}}{P_{2_r}} = -\frac{\tau_1 P_1}{\tau_2 P_2}$$

The Merton term structure model gives us very important insights into the process of deriving a term structure model. Its simple formulas make it a useful expository tool, but the negative yields that result from the formula are a major concern. It leads to a logical question: can we "extend" the Merton model to fit the actual yield curve perfectly? If so, this superficially at least would allow us to avoid the negative yield problems associated with the model. We turn to this task next.

The Extended Merton Model

Ho and Lee [1986] extended the Merton model to fit a given initial yield curve perfectly in a discrete time framework. The Ho and Lee/Extended Merton model is a special case of Heath Jarrow and Morton where interest volatility at any current time t for future time T in state s_t $\sigma(t,T;s_t)$ is a deterministic constant independent of state s_t: $\sigma(t,T;s_t)=\sigma$. As Jarrow (2002, Chapter 15, page 288) notes, this implies that forward rates and spot rates can be negative. In this section, we derive the Ho and Lee/Extended Merton model using the same approach above. This basic "extension" technique is common to all term structure models developed prior to publication of the HJM model. In the extended Merton case, we assume that the short rate of interest is again the single stochastic factor driving movements in the yield curve. Instead of assuming that the short rate r is a random walk, however, we assume that it has a time dependent drift term a(t):

$$dr = a(t)dt + \sigma dZ$$

As before, Z represents a standard Wiener process that provides random shocks to r; Z has mean zero and standard deviation of 1. The instantaneous standard deviation of interest rates is again the constant sigma, which we can scale up and down to fit observable market yields. Appendix A shows how to derive zero coupon bond prices in this model using exactly the same procedures as the previous section. The only difference is that we allow the change in the short rate of interest to follow the drift term a(t). When we apply the no arbitrage conditions and derive zero coupon bond prices, we get the following formula:

$$P(r,t,T) = e^{-r\tau - \frac{\lambda\sigma\tau^2}{2} + \frac{1}{6}\sigma^2\tau^3 - \int_t^T a(s)(T-s)ds}$$

Note that this is identical to the formula for the zero coupon bond price in the Merton term structure model with the exception of the last term. The value of the zero coupon bond is 1 when t=T as the boundary condition demands.

Like the Merton term structure model, the price volatility of zero coupon bonds is given by[6]

$$\sigma P_r = -\tau\sigma P$$

and the percentage change in price of zero coupon bonds for small changes in the short rate r is

$$r\text{-}duration = -\frac{P_r}{P} = -\frac{-\tau P}{P} = \tau$$

The behavior of the yield to maturity on a zero coupon bond in the extended Merton-Ho and Lee model is always consistent with the observable yield curve since the expression for yield to maturity

$$y(\tau) = r + \frac{\lambda\sigma\tau}{2} - \frac{1}{6}\sigma^2\tau^2 + \frac{1}{\tau}\int_t^T a(s)(T-s)ds$$

contains the "extension term"

$$extension(\tau) = \frac{1}{\tau}\int_t^T a(s)(T-s)ds$$

The function a(s) is chosen such that the theoretical zero coupon yield to maturity y and the actual zero coupon yield are exactly the same. The function a(s) is the "plug" which makes the model fit and

[6] We could change the sign of this expression to make it positive since the sign of the Wiener process dZ is arbitrary.

corrects for "model error" which would otherwise cause the model to give implausible results. Since any functional form for yields can be adapted to fit a yield curve precisely,[7] it is critical in examining any model for plausibility to minimize the impact of this "extension" term. Why? Because the extension term itself contains no economic content.

In the case of the Ho and Lee model, the underlying model would otherwise cause interest rates to sink to negative infinity, just as in the Merton model. The extension term's magnitude, therefore, must offset the negative interest zero coupon bond yields that would otherwise be predicted by the model. As maturities get infinitely long, the magnitude of the extension term will become infinite in size. This is a significant cause for concern, even in the extended form of the model. One of the most important attributes of term structure model selection is to find one which best captures the underlying economics of the yield curve so that this “extension” is minimized for the entire maturity spectrum.

[7] For example, consider this term structure model, where zero coupon bond yields are a linear function of years to maturity: $y(\tau)=.05+.01\tau$. It is clearly a ridiculous model, but it can be made to fit the yield curve exactly in a manner similar to the Ho and Lee model.

The Vasicek Model

Both the Merton model and its extended counterpart the Ho and Lee model are based on an assumption about random interest rate movements that imply that, for any positive interest rate volatility, zero coupon bond yields will be negative at every single instant in time for long maturities beyond a critical maturity τ^*. The extended version of the Merton model, the Ho and Lee model, offsets the negative yields with an "extension" factor that must grow larger and larger as maturities lengthen. Vasicek [1977] proposed a model that avoids the certainty of negative yields and eliminates the need for a potentially infinitely large "extension" factor. Even more important, the Vasicek model produces realistic interest rate cycles. Vasicek accomplishes this by assuming that the short rate r has a constant interest rate volatility sigma like the models above, with an important twist: the short rate exhibits "mean reversion" around the long run average level of interest rates. The Vasicek model and its extended form are special cases of HJM with a single risk factor and a time (but not state) dependent volatility function with takes the form $\sigma(t,T;s_t)=ke^{-a(T-t)}$ where k and a are constants. Vasicek calls the long run average level of interest rates μ:

$$dr = \alpha(\mu - r)dt + \sigma dZ$$

The change in the short term rate of interest r is a function of a “drift” term and a shock term where

r is the instantaneous short rate of interest

α is the speed of mean reversion

μ is the long run expected value for r, and

σ is the instantaneous standard deviation of r

Z is the standard Wiener process with mean zero and standard deviation of 1, exactly like we have used in the first two examples. The difference from previous models comes in the drift term, which gives this stochastic process its name: the Ornstein-Uhlenbeck process. The drift term in the stochastic process proposed by Vasicek pulls the short rate r back toward μ, so μ can be thought of as the long-run level of the short rate. When the short rate r is above μ, the first term tends to pull r downward since a is assumed to be positive. When the short rate r is below μ, r tends to drift upward. The second term of the stochastic process, of course, applies random shocks to the short rate which may temporarily offset the tendencies toward mean reversion of the underlying stochastic process. The impact of mean reversion is to create realistic interest rate cycles, with the level of alpha determining the length and violence of rises and falls in interest rates. What are the implications of this model for the pricing of bonds? We derive the zero coupon bond pricing formula in Appendix B. The zero coupon bond pricing formula is an exponential function with two terms, one of which is a function of the short term rate of interest r:

$$P(r,t,T) = e^{-F(t,T)r - G(t,T)}$$

$$= \exp\left[-rF(\tau) - \left(\mu + \frac{\lambda\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2} \right)[\tau - F(\tau)] - \frac{\sigma^2 F^2(\tau)}{4\alpha} \right]$$

where the function F is defined in Appendix B as

$$F(t,T) = F(\tau) = \frac{1}{\alpha}\left(1 - e^{-\alpha\tau}\right)$$

The continuously compounded zero coupon bond yield in the Vasicek model uses the definition of function G in Appendix B:

$$y(\tau) = -\frac{1}{\tau}\ln[P(\tau)] = -\frac{1}{\tau}[-rF(\tau) - G(\tau)]$$
$$= \frac{F(\tau)}{\tau}r + \frac{G(\tau)}{\tau}$$

As time to maturity gets infinitely long, one can calculate the infinitely long zero coupon bond yield maturity as

$$y(\infty) = \mu + \frac{\lambda\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2}$$

This yield to maturity is positive for almost all realistic sets of parameter values, correcting one of the major objections to the Merton term structure model, without the necessity to "extend" the yield curve with a time dependent drift in the short rate of interest r.

What does duration mean in the context of the Vasicek model? Macaulay defined duration, as we saw in Chapter 12, as the percentage change in bond prices with respect to [the parallel shift in] yield to maturity in a continuous time context. The parallel shift in the Macaulay model was the single stochastic factor driving the yield curve. In the Vasicek model, the short rate r is the stochastic factor. We define "r-duration" as above: the percentage change in the price of a bond with respect to changes in the short rate r:

$$r\text{-}duration = \frac{-P_r}{P} = \frac{F(\tau)P}{P} = F(\tau) = \frac{1}{\alpha}\left[1 - e^{-\alpha\tau}\right]$$

This function F is a powerful tool one can use for hedging in the Vasicek model and its extended version, which we cover in the next section. The hedge ratio necessary to hedge one unit of a zero coupon bond with a remaining maturity of τ_1 using a bond with remaining maturity of τ_2 is

$$w = -\frac{P_{1_r}}{P_{2_r}} = -\frac{F(\tau_1)P[\tau_1]}{F(\tau_2)P[\tau_2]}$$

a hedge ratio substantially different from that using the Merton or Ho and Lee models:

$$w = -\frac{\tau_1 P[\tau_1]}{\tau_2 P[\tau_2]}$$

In practical use, it is this difference in hedge ratios that allows us to distinguish between different models. The ability to extend a model, as in the previous section, renders all extendible models equally good in the sense of fitting observable data. For example, the HJM framework allows the analyst to fit an observable term structure of zero coupon bond yields with (almost) any volatility structure. In reality, however, the explanatory power of various term structure models and volatility assumptions can be substantially different. Ultimately, the relative performance of each model's hedge ratios and its ability to explain price movements of traded securities with the fewest parameters are what differentiates the best models from the others.

The stochastic process proposed by Vasicek allows us to calculate the expected value and variance of the short rate at any time in the future s from the perspective of current time t. Denoting the short rate at time t by r(t), the expected value of the short rate at future time s is

$$E_t[r(s)] = \mu + [r(t) - \mu]e^{-\alpha(s-t)}$$

The standard deviation of the potential values of r around this mean value is

$$Standard\ \ Deviation_t\,[r(s)] = \sqrt{\frac{\sigma^2}{2\alpha}\left[1 - \exp^{-2\alpha(s-t)}\right]}$$

Because r(s) is normally distributed, there is a positive probability that r(s) can be negative. As pointed out by Black [1995], this is inconsistent with a no arbitrage economy in the special sense that consumers hold an option, which Black assumed is costless, to hold cash instead of investing at negative interest rates. The magnitude of this theoretical problem with the Vasicek model[8] depends on the level of interest rates and the parameters chosen. The biggest problem with the Vasicek model is not the positive probability of negative rates. Like all of the term structure models described in this chapter, the biggest problem is the assumption that only one random factor drives the yield curve, which is inconsistent with the evidence of Dickler, Jarrow and van Deventer (2011) that 6 to 10 random factors drive the movements of the risk-free term structure.

The Extended Vasicek-Hull and White Model

Hull and White [1990] bridged the gap between the observable yield curve and the theoretical yield curve implied by the one factor Vasicek model by "extending" or stretching the theoretical yield curve to fit the actual market data. The Hull and White extension was published contemporaneously with the Heath Jarrow and Morton (1990, 1990, 1992) extension for n-factor models. Hull and White apply the identical logic as the previous section, but they allow the market price of risk term λ to drift over time, instead of assuming it is constant as in the Vasicek model. If we assume this term drifts over time, what are the implications for the pricing formula for zero coupon bonds?

Hull and White use the time dependent drift term in interest rates theta, where theta in turn depends on a time dependent market price of risk:

$$\theta(t) = \alpha\mu + \lambda(t)\sigma$$

Hull and White derive the zero coupon bond price in the Extended Vasicek model using this assumption as shown in Appendix C:

$$P(r,t,T) = e^{-F(t,T)r - G(t,T)}$$

$$= \exp\left[-rF(\tau) - \int_t^T F(s,T)\theta(s)ds + \left(\frac{\sigma^2}{2\alpha^2}\right)[\tau - F(\tau)] - \frac{\sigma^2 F^2(\tau)}{4\alpha}\right]$$

---

[8] The same objection applies to the Merton and Ho and Lee models and a wide range of other model which assume a constant volatility of interest rates, regardless of the level of short term interest rates. Since the storage of cash is not costless, as we discussed earlier in this book, negative rates are becoming increasingly common in countries like Japan and the U.S., which are recovering from serious financial crises.

The zero coupon bond price is again an exponential function with two terms, one which is multiplied by the random short term rate of interest r and the other which contains the functions F and G. Function F is defined the same way as in the previous section. Function G contains the extension term and is defined in Appendix C. As in the Vasicek model, the price sensitivity of a zero coupon bond is given by the formula

$$r\text{-}duration = \frac{-P_r}{P} = \frac{F(\tau)P}{P} = F(\tau) = \frac{1}{\alpha}\left[1-e^{-\alpha\tau}\right]$$

This is the same formula as in the regular Vasicek model. We now turn to a brief historical review of other term structure models proposed in the academic literature leading up to the publication of the HJM framework.

Alternative Term Structure Models

A complete review of the vast array of term structure models that have been proposed in the academic literature is beyond the scope of this chapter. We refer the serious student of term structure model history to these excellent references by Dai, Duffie, Kan, Singleton and Jarrow:

Dai, Qiang and Kenneth J. Singleton, "Specification Analysis of Affine Term Structure Models," *The Journal of Finance*, Volume LV, Number 5, October 2000.

Dai, Qiang and Kenneth J. Singleton, "Term Structure Dynamics in Theory and Reality," *The Review of Financial Studies*, Volume 16, Number 3, Fall 2003.

Duffie, Darrell and Rui Kan, "Multi-factor Term Structure Models," *Philosophical Transactions: Physical Sciences and Engineering*, Volume 347, Number 1684, Mathematical Models in Finance, June 1994.

Duffie, Darrell and Kenneth J. Singleton, "An Econometric Model of the Term Structure of Interest-Rate Swap Yields," *The Journal of Finance*, Volume LII, Number 4, September 1997.

Duffie, Darrell and Kenneth J. Singleton, "Modeling Term Structures of Defaultable Bonds," The Review of Financial Studies, Volume 12, Number 4, 1999.

Jarrow, Robert A., "The Term Structure of Interest Rates," *Annual Review of Financial Economics*, 1, 2009.

In the remainder of this chapter, we will briefly a small subset of the term structure models of the last five decades, with occasional asides to remind the reader that each of the models mentioned is a special case of Heath Jarrow and Morton. None of the models reviewed by the authors above has more than the 3 random factors that we illustrated using the HJM framework in Chapter 9. Consistent with Dickler, Jarrow and van Deventer (2011), we remind the reader that 6 to 10 factors are needed for both realism and regulatory purposes, per the Bank for International Settlements regulations cited in prior chapters.

In the remainder of this chapter, we consider alternative term structure models that offer a richer array of potential movements in the term structure of interest rates, while preserving the implicit assumption that default will not occur. In Chapters 16 through 17, we broaden our analysis to include multi-factor models of the yield curve that explicitly incorporate the probability of default. For example, the Shimko, Tejima and van Deventer extension of the Merton model, which we discuss in Chapter 16, is a two-factor term structure model in the guise of a discussion of the valuation of risky debt. The two risky factors are the short rate of interest r, which drives the risk free yield curve's movements, and the value of the underlying assets being financed.

Alternative One-Factor Interest Rate Models

The Cox, Ingersoll and Ross (CIR) [1985] model, which the authors derive in a general equilibrium framework, has been as popular in the academic community to the same degree as the Vasicek model has been among financial market participants. It is therefore appropriate that we begin our discussion of one factor model alternatives to the Vasicek model with a review of the CIR approach.

The CIR Model

The Cox, Ingersoll and Ross model has been particularly influential because the original version of the paper was in circulation in the academic community at least since 1977, even though the paper wasn't formally published until 1985. CIR assume that the short term interest rate is the single stochastic factor driving interest rate movements, and that the variance of the short rate of interest is proportional to the level of interest rates. This is a special case of HJM, which Jarrow (2002, Chapter 15, page 288) labels “nearly proportional volatility such that the interest rate volatility function σ(t,T;st)=k(t,T) log [F(t,T)], subject to an upper bound of a constant M. Below this upper bound, volatility is proportional to the continuously compounded forward rate so that forward rates are always non-negative. This implies that interest rate volatility is higher in periods of high interest rates than it is in periods of low interest rates. The realism of this assumption depends on which financial market is being studied, but casual empiricism would lead one to believe that it's a more desirable property when modeling the United States (1978-1985), Brazil, or Mexico than it would perhaps be for Japanese financial markets today. This observation is consistent with the rising levels of interest rate volatility that we found in the U.S. Treasury market from 1962-2011 in Chapters 6 through 9, but the CIR assumption is much simpler than the pattern of volatility levels reflected in the actual data.

The authors assume that stochastic movements in the short rate take the form

$$dr = k(\mu - r)dt + \sigma\sqrt{r}dz$$

The change in interest rates will have a mean reverting drift term, which causes interest rate cycles, just like the extended Vasicek and Vasicek models. The interest rate shock term has the same volatility as the Vasicek model but it is multiplied by the square root of the short term rate of interest. CIR show that the value of a zero coupon bond with maturity $\tau = T - t$ takes the form

$$P(r,\tau) = A(\tau)e^{-B(\tau)r}$$

We refer the reader to the original CIR paper for the definitions of A and B. The authors also derive an analytical solution for the price of a European call option on a zero coupon bond.

Dothan [1978] provides a model of short rate movements where the short term riskless rate of interest r follows a geometric Wiener process. The short rate has a lognormal distribution and will therefore always be positive:

$$dr = \sigma r dz$$

The resulting analytical solution for zero coupon bond prices is quite complex. The Dothan model, while sharing one of the attractive properties of the CIR model, lacks the "mean reversion" term which causes interest rate cycles, one of the key features necessary in a realistic term structure model.

Longstaff [1989] proposes a model in which the variance of the short rate is proportional to the level of the short rate, like the CIR model, and the mean version of the short rate is a function of its square root:

$$dr = k(\mu - \sqrt{r})dt + \sigma\sqrt{r}dz$$

The resulting pricing equation for zero coupon bonds is very unique in comparison to other term structure models in that the implied yield to maturity on zero coupon bonds is a non-linear function of the short rate of interest:

$$P(r,\tau) = A(\tau)e^{B(\tau)r + C(\tau)\sqrt{r}}$$

A, B, and C are complex functions of the term structure model parameters and are described in Longstaff.

Black, Derman and Toy [1990] suggest another model which avoids the problem of negative interest rates and allows for time dependent parameters. The stochastic process specifies the percentage change in the short term rate of interest is a mean reverting function of the natural logarithm of interest rates:

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

The model has many virtues from the perspective of financial market participants. It combines the ability to fit the observable yield curve (like the Ho and Lee and extended Vasicek models) with the non-negative restriction on interest rates and the ability to model the volatility curve observable in the market. The model's liability is the lack of tractable analytical solutions, which are very useful in (a) confirming the accuracy of numerical techniques and (b) valuing large portfolios where speed is essential.

Black and Karasinski [1991] further refine the Black, Derman and Toy approach with the explicit incorporation of time-dependent mean reversion:

$$d[(\ln(r)] = \phi(t)\big[\ln[\mu(t)] - \ln(r)\big]dt + \sigma(t)dz$$

This modification allows the model to fit observable cap prices, one of the richest sources of observable market data incorporating interest rate volatility information if one ignores the possibility of manipulation of the short term London interbank offered rate, which we discuss in Chapter 17. The authors describe in detail how to model bond prices and interest derivatives using a lattice approach. The model, like its predecessor the Black, Derman and Toy model, has been quite popular among financial market participants.

Two Factor Interest Rate Models

As we discuss below, one of the challenges in specifying a two-factor model is selecting which two factors are the most appropriate. In Chapters 16 and 17, we discuss credit models where the two factors of the risky debt term structure model are the riskless short rate of interest and either (a) the value of the asset being financed (in the Shimko, Tejima, and van Deventer extension of the Merton credit model) or (b) a macro factor driving default, as in the Jarrow credit model. In this section, we review a number of two-factor models that are based on various assumptions about the two risky factors on the assumption that we are modeling a risk-free yield curve.

The Brennan and Schwartz Model

Brennan and Schwartz [1979] introduced a two factor model where both a long term rate and a short term rate follow a joint Gauss-Markov process. The long term rate is defined as the yield on a consol (perpetual) bond. Brennan and Schwartz assume that the log of the short rate has the following stochastic process:

$$d[\ln(r)] = \alpha[\ln(l) - \ln(p) - \ln(r)]dt + \sigma_1 dz_1$$

The short rate r moves in response to the level of the consol rate l and a parameter p relating the "target value" of ln [r]) relative to the level of ln (l). Brennan and Schwartz show that the stochastic process for the consol rate can be written

$$dl = l[l - r + \sigma_2^2 + \lambda_2 \sigma_2]dt + l\,\sigma_2 dz_2$$

Lambda in this expression is the market price of long term interest rate risk. Longstaff and Schwartz proceed to test the model on Canadian government bond data, with good results.

Chen and Scott [1992] derive a two-factor model in which the nominal rate of interest I is the sum of two independent variables y_1 and y_2, both of which follow the stochastic process specified by CIR:

$$dy_i = k_i(\theta_i - y_i)dt + \sigma_i \sqrt{y_i}\, dz_i$$

Chen and Scott show that the price of a discount bond in this model is

$$P(y_1, y_2, t, T) = A_1 A_2 e^{-B_1 y_1 - B_2 y_2}$$

where A and B have the same definition as in the CIR model, with the addition of the appropriate subscripts. The authors go on to value a wide range of interest rate derivatives using this model. The end result is a powerful model with highly desirable properties and a wealth of analytical solutions.

Hull and White [1993] show that there is a similar extension for the Vasicek model when the nominal interest rate I is the sum of two factors r_1 and r_2. The value of a zero coupon bond with maturity tau is simply the product of two factors P_1 and P_2 which have exactly the same functional form as the single factor Vasicek model, except that one is driven by r_1 and the other by r_2:

$$V(r_1,r_2,\tau) = P_1(r_1,\tau)P_2(r_2,\tau)$$

Both stochastic factors are assumed to follow stochastic processes identical to the normal Vasicek model:

$$dr_i = \alpha_i(\mu_i - r_i)dt + \sigma_i dz_i$$

Longstaff and Schwartz [1992] propose a model in which two stochastic factors, which are assumed to be uncorrelated, drive interest rate movements. The factors x and y are assumed to follow the stochastic processes

$$dx = (\gamma - \delta x)dt + \sqrt{x}\, dz_1$$

$$dy = (\eta - vy)dt + \sqrt{y}\, dz_2$$

Longstaff and Schwartz derive the value of a discount bond in this economy to be

$$V(x,y,\tau) = E_1(\tau)e^{E_2(\tau)x+E_3(\tau)y}$$

Chen's Three Factor Term Structure Model

As factors are added to a term structure model, it becomes more realistic and more complex. Chen [1994] introduces a three factor term structure model where the short term rate of interest is random and mean reverting around a level which is also random. Moreover, Chen also assumes that the volatility of the short rate is random:

$$dr = k(\theta - r)dt + \sqrt{\sigma}\sqrt{r}\,dz_1$$
$$d\theta = v(\bar{\theta} - \theta)dt + \zeta\sqrt{\theta}dz_2$$
$$d\sigma = \mu(\bar{\sigma} - \sigma)dt + \eta\sqrt{\sigma}dz_3$$

The volatility of the short rate is written by Chen as ?. The long-term level of interest rates theta is itself random and mean-reverting around a long-term level. The volatility of the short rate is also mean-reverting with a form much like the CIR specification. Chen goes on to derive, with considerable effort, closed form solutions for this richly-specified model.

Chen's model demonstrates that there is a rich array of choices for financial market participants who require the additional explanatory power of a two or three factor model and who are willing to incur the costs, which we discuss below, of such models.

Reprising the Heath, Jarrow and Morton Approach

Realism in term structure modeling requires 6 to 10 random factors driving the risk free term structure of interest rates. The review of 1, 2 and 3 factor models above has probably left the reader with the sense that models can quickly become intractable as the number of factors rises without some unifying approach. That unifying approach is the Heath Jarrow and Morton framework. The reason for presenting the HJM approach first, in Chapters 6 through 9, is to show that it is much easier to use in practice than many of the "simpler" term structure models discussed above that have fewer risk factors and, therefore, less explanatory power. When we say "HJM" in this context, we include its popular variant the "Libor Market Model" which has been widely employed in valuing securities that are a function of the Libor-swap curve. We discuss the Libor market model in Chapter 17, where we focus on the modeling of term structures where the underlying yield curve embeds the possibility of default. For risk free term structures, the HJM framework is so rich that the expansion of risk factors from the 3 factors of Chapter 9 to 4, 5, 5, 10 or 20 factors is routine. Tedious, but routine. That is what a sophisticated risk management software solution is for--to remove the "tedious" from risk management so a talented risk manager can spend his or her time on the art and judgment element of the discipline.

In the next chapter, we turn to term structure implementation questions. What volatility function is best if we want to employ a multi-factor HJM interest rate modeling framework?

Appendix A: Deriving Zero Coupon Bond Prices in the Extended Merton Model

Ito's lemma gives us the same formula for changes in the value of a zero coupon bond P, except for substitution of the slightly more complex expression for dr:

$$dP = P_r\,dr + \frac{1}{2}P_{rr}(dr)^2 + P_t$$

$$= P_r\,[a(t) + \sigma dZ] + \frac{1}{2}P_{rr}\sigma^2 + P_t$$

We form a no-arbitrage portfolio with value W as in the first section in this chapter so that the coefficient of the dZ term is zero. We get the same hedge ratio as given in the original Merton model:

$$w = \frac{-P_{1_r}}{P_{2_r}}$$

By applying the no-arbitrage condition that dW = rW, we are led to a no arbitrage condition closely related to that of the original Merton model:

$$-\lambda = \frac{P_{1_r}a(t) + \frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_t} - r\,P_1}{\sigma P_{1_r}} = \frac{P_{2_r}a(t) + \frac{1}{2}P_{2_{rr}}\sigma^2 + P_{2_t} - r\,P_2}{\sigma P_{2_r}}$$

In the Ho and Lee model, the market price of risk must be equal for any two zero coupon bonds with arbitrary maturities T_1 and T_2. In the Ho and Lee case, the market price of risk again is the fixed income counterpart of the Sharpe ratio, with expected return on each bond equal to

$$drift = expected\ return = P_r\,a(t) + \frac{1}{2}P_{rr}\sigma^2 + P_t$$

Now the value of the zero coupon bond price P for any given maturity is fixed by the shape of the yield curve which Ho and Lee seek to match perfectly. Our mission in solving the partial differential equation in the Ho and Lee model is to find the relationship between the drift term a(t) and the bond price P. The partial differential equation which must be solved comes from rearranging the no arbitrage condition above, and our continued assumption that lambda is constant:

$$P_r\,[a(t) + \lambda\sigma] + \frac{1}{2}P_{rr}\sigma^2 + P_t - rP = 0$$

which must hold subject to the fact that the value of a zero coupon bond must equal 1 at maturity.

$$P(r,T,T) = 1$$

We use an educated guess to postulate a solution and then see what must be true for our guess to be correct (if it is possible to make it correct). We guess that the solution P is closely related to the Merton model:

$$P(r,t,T) = e^{-r\tau + G(t,T)}$$

We seek to find the function G which satisfies the partial differential equation. We take the partial derivatives of P and substitute them into the partial differential equation:

$$P_r = -\tau P$$
$$P_{rr} = \tau^2 P$$
$$P_t = (r + G')P$$

When we substitute these partial derivatives into the partial differential equation and simplify, we get the following relationship:

$$-[a(t) + \lambda\sigma]\tau + \frac{1}{2}\sigma^2\tau^2 + G' = 0$$

We solve this differential equation in G by translating the derivative of G with respect to the current time t (not τ) into this expression:

$$G(t,T) = -\frac{\lambda\sigma\tau^2}{2} + \frac{1}{6}\sigma^2\tau^3 - \int_t^T a(s)(T-s)ds$$

Therefore, the value of a zero coupon bond in the extended Merton-Ho and Lee model is given by the equation

$$P(r,t,T) = e^{-r\tau - \frac{\lambda\sigma\tau^2}{2} + \frac{1}{6}\sigma^2\tau^3 - \int_t^T a(s)(T-s)ds}$$

Appendix B: Deriving Zero Coupon Bond Prices in the Vasicek Model

We use the notation of Chen [1992] to answer this question using the same process as in the Merton and Ho and Lee models. By Ito's lemma, movements in the price of a zero coupon bond is

$$dP = P_r dr + \frac{1}{2}P_{rr}(dr)^2 + P_t$$
$$= P_r[\alpha(\mu - r)dt + \sigma dZ] + \frac{1}{2}P_{rr}\sigma^2 + P_t$$

Using exactly the same no arbitrage argument that we used above, we can eliminate the dZ term by choosing the hedge ratio necessary to eliminate interest rate risk (which is what dZ represents). The partial differential equation consistent with a no arbitrage bond market in the Vasicek model is

$$P_r\,[\alpha(\mu - r) + \lambda\sigma] + \frac{1}{2}P_{rr}\sigma^2 + P_t - rP = 0$$

and, as above, it must be solved subject to the boundary condition that a zero coupon bond's price at maturity equals its principal amount, 1:

$$P(r,T,T) = 1$$

The market price of risk λ is assumed to be constant. As a working assumption, we guess that the zero coupon bond price has the solution

$$P(r,t,T) = P(r,\tau) = e^{-rF(\tau)-G(\tau)}$$

where F and G are unknown functions of τ = T - t. If our working assumption is correct, we will be able to obtain solutions for F and G. We know that

$$P_r = -FP$$
$$P_{rr} = F^2 P$$
$$P_t = (-rF_t - G_t)P$$

By replacing the derivatives of the zero coupon bond price P in the partial differential equation above and rearranging, we know that the following relationship must hold:

$$r[\alpha F - F_t - 1] + \left[\frac{1}{2}F^2\sigma^2 - F(\alpha\mu + \lambda\sigma) - G_t\right] = 0$$

This relationship must hold for all values of r, so the coefficient of r must be zero.

$$\alpha F - F_t - 1 = 0$$

We can solve this partial differential equation by rearranging it until we have

$$\frac{\alpha\, F_t}{1 - \alpha F} = -\alpha$$

We then take the integral of both sides such that

$$\int_t^T \frac{\alpha\, F_t(s,T)}{1 - \alpha F(s,T)}ds = -\int_t^T \alpha ds$$

Evaluating the integrals on both sides of the equation leaves the relationship

$$\ln[1 - \alpha F(t,T)] = -\alpha(T - t)$$

Calculating the exponential of both sides defines F:

$$F(t,T) = F(\tau) = \frac{1}{\alpha}\left(1 - e^{-\alpha\tau}\right)$$

To determine the value of the function G, we must solve the partial differential equation

$$\frac{1}{2}F^2\sigma^2 - F(\alpha\mu + \lambda\sigma) - G_t = 0$$

or

$$G_t = \frac{1}{2}F^2\sigma^2 - F(\alpha\mu + \lambda\sigma)$$

We can calculate that

$$\int_t^T F(s,T)ds = \frac{1}{\alpha}\left[\tau - F(\tau)\right]$$

and that

$$\int_t^T F^2(s,T)ds = \frac{1}{\alpha^2}\left[\tau - F(\tau)\right] - \frac{F^2(\tau)}{2\alpha}$$

We can take the integral of both sides of the equation above to solve for G:

$$\int_t^T G_s(s,T)ds = \int_t^T\left[\frac{1}{2}F(s,T)^2\sigma^2 - F(s,T)(\alpha\mu + \lambda\sigma)\right]ds$$

$$= \frac{1}{2}\sigma^2\left[\frac{1}{\alpha^2}\left(\tau - F(\tau)\right) - \frac{F^2(\tau)}{2\alpha}\right] - \frac{\alpha\mu + \lambda\sigma}{\alpha}[\tau - F(\tau)]$$

$$= \left[\frac{\sigma^2}{2\alpha^2} - \mu - \frac{\lambda\sigma}{\alpha}\right][\tau - F(\tau)] - \frac{\sigma^2}{4\alpha}F^2(\tau)$$

Since

$$\int_t^T G_s(s,T)ds = G(T,T) - G(t,T)$$

and since G(T,T) must be zero for the boundary condition that P(T,T)=1,

$$G(t,T) = G(\tau) = \left[\mu + \frac{\lambda\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2}\right][\tau - F(\tau)] + \frac{\sigma^2}{4\alpha}F^2(\tau)$$

This means that the value of a zero coupon bond in the Vasicek model is

$$P(r,t,T) = e^{-F(t,T)r - G(t,T)}$$

$$= \exp\left[-rF(\tau) - \left(\mu + \frac{\lambda\sigma}{\alpha} - \frac{\sigma^2}{2\alpha^2} \right)[\tau - F(\tau)] - \frac{\sigma^2 F^2(\tau)}{4\alpha} \right]$$

Appendix C: Valuing Zero Coupon Bonds in the Extended Vasicek Model

The partial differential equation changes only very slightly from the one we used in the third section of this chapter:

$$P_r\,[\,\alpha(\,\mu - r) + \lambda(t)\sigma\,] + \frac{1}{2} P_{rr}\,\sigma^2 + P_t - rP = 0$$

subject to the usual requirement that the bond's price equals 1 at maturity:

$$P(r,T,T) = 1$$

We can rewrite the partial differential equation as

$$P_r\,[(\,\alpha\mu + \lambda(t)\sigma\,) - \alpha r] + \frac{1}{2} P_{rr}\,\sigma^2 + P_t - rP = 0$$

and using the definition

$$\theta(t) = \alpha\mu + \lambda(t)\sigma$$

we can simplify the partial differential equation to the point that it looks almost identical to that of the Ho and Lee model earlier in the chapter:

$$P_r\,[\,\theta(t) - \alpha r] + \frac{1}{2} P_{rr}\,\sigma^2 + P_t - rP = 0$$

As in the third section in this chapter, we assume that the solution to the pricing of zero coupon bonds takes the form

$$P(r,t,T) = P(r,\tau) = e^{-rF(\tau) - G(\tau)}$$

where F and G are unknown functions of τ = T - t. Using the partial derivatives of P under the assumption that we have guessed the functional form correctly, we know that the following relationship must hold:

$$r[\alpha F - F_t - 1] + \left[\frac{1}{2}F^2\sigma^2 - F\theta(t) - G_t\right] = 0$$

From Appendix B, we can prove

$$F(t,T) = F(\tau) = \frac{1}{\alpha}\left(1 - e^{-\alpha\tau}\right)$$

We now must solve for G given the following equation:

$$G_t = \frac{1}{2}F^2\sigma^2 - F\theta(t)$$

We do this by taking the integral of both sides and making use of the integral of F^2 which was given in Appendix B to arrive at the solution for G:

$$G(t,T) = G(\tau) = \int_t^T F(s,T)\theta(s)ds - \frac{\sigma^2}{2\alpha^2}[\tau - F(\tau)] + \frac{\sigma^2}{4\alpha}F^2(\tau)$$

Therefore, under the extended Vasicek model, the value of a zero coupon bond is given by the formula

$$P(r,t,T) = e^{-F(t,T)r - G(t,T)}$$

$$= \exp\left[-rF(\tau) - \int_t^T F(s,T)\theta(s)ds + \left(\frac{\sigma^2}{2\alpha^2}\right)[\tau - F(\tau)] - \frac{\sigma^2 F^2(\tau)}{4\alpha}\right]$$