

## Fixed Income Concepts

In this part of the book, we will study the pricing and hedging of fixed-income derivatives, that is, derivatives that can be viewed as written on bonds or on interest rates. The complexities of this subject stem from the fact that the underlying bonds or rates will have time-varying and generally random volatilities and have volatilities that must be linked in some way to each other. We are not free to arbitrarily specify the volatilities and correlations, as we can, for example, with basket equity options, because such a specification may imply there is an arbitrage opportunity available from trading the underlying assets, and, of course, one could not trust a derivative value or a hedging strategy derived from such a model. There are many books that provide a more comprehensive and advanced treatment than we will be able to give here, among which Rebonato [Rebonato98, Rebonato02], James and Webber [JW] and Brigo and Mercurio [BM] seem particularly useful.

We will focus on derivatives and underlyings that have very little default risk and ignore the pricing of default (credit) risk. Credit risk and credit derivatives are booming areas in both theory and practice. For this topic, one can consult the recent books of Bielecki and Rutkowski [BR], Duffie and Singleton [DS], Schönbucher [Schonbucher], and Tavakoli [Tavakoli]. Another important topic that will not be covered is mortgages.

In the first section, we introduce a fundamental construct: the yield curve, by which we mean the yields of (possibly theoretical) zero-coupon bonds of various maturities. The last two sections of the chapter (on principal components) are optional—nothing else in the book builds upon them.

### The Yield Curve

Given prices of discount (zero-coupon) bonds of all maturities, any coupon paying bond can be priced as a portfolio of discount bonds. The relationship between time to maturity and the price of the corresponding discount bond with \$1 face value is sometimes called the discount function. An equivalent concept is the yield curve, which is the relationship between time to maturity and the yield of the corresponding discount bond. Yields can be quoted according to different compounding conventions (semi-annual being the most common), but we will continue to use continuous compounding. [The relationship between the annually compounded rate  $y_a$ , the semi-annually compounded rate  $y_s$  and the continuously compounded rate  $y_c$  is  $e^{y_c} = 1 + y_a = (1 + y_s/2)^2$ .] With this convention, the yield of a zero-coupon bond with \$1 face value having  $\tau$  years to maturity and price  $P$  is defined to be the rate  $y$  such that  $P = e^{-y\tau}$ . Denoting this yield  $y$  at maturity  $\tau$  by  $y(\tau)$ , the yield curve is the collection of yields  $\{y(\tau) | \tau > 0\}$ , conventionally plotted with maturity  $\tau$  on the horizontal axis and the yields  $y(\tau)$  on the vertical. Usually (but certainly not always) this is an upward sloping curve, meaning that rates at longer maturities are higher than rates at shorter maturities.

Some amount of estimation is necessary to compute the yield curve. We would like to know yields at arbitrary maturities, but there are not enough actively traded zero-coupon bonds to provide this information. Thus, we have to fill in the missing maturities. We may also use coupon-paying bonds (or swap rates, as we will discuss later) to estimate the yields.

The most popular method of estimating the yield curve from bond prices is to fit a cubic spline, using the prices of a finite set of actively traded bonds. Given the set of bonds, let  $P_i$  denote the price of bond  $i$ ,  $N_i$  the number of dates at which bond  $i$  pays a coupon or its face value, and  $\{\tau_{ij} | j = 1, \dots, N_i\}$  the dates at which bond  $i$  pays a coupon or its face value. Finally, let  $C_{ij}$  denote the cash flow paid by bond  $i$  at date  $\tau_{ij}$  for  $j = 1, \dots, N_i$ . Then for each  $i$  it should be the case that

$$P_i = \sum_{j=1}^{N_i} e^{-y(\tau_{ij})\tau_{ij}} C_{ij}. \quad (1)$$

This simply says that the price should be the present value of the cash flows.

However, in practice, we will typically be unable to find yields  $y(\tau_{ij})$  such that Equation 1 holds exactly for each bond  $i$ . This is due to measurement errors in the form of bid-ask spreads and nonsynchronous pricing. Furthermore, even if we can find such yields, we still face the issue of estimating the yields at other maturities  $\tau \notin \{\tau_{ij}\}$ . The cubic spline is one way to address these issues.

A cubic spline consists of several cubic polynomials spliced together at a set of knot points. Specifically, it consists of maturities  $\tau_1, \dots, \tau_n$  (the knot points) and coefficients  $(a_i, b_i, c_i, d_i)$  for  $i = 0, \dots, n$ , and the yield curve is modeled as

$$y(\tau) = \begin{cases} a_0\tau^3 + b_0\tau^2 + c_0\tau + d_0 & \text{for } 0 < \tau \leq \tau_1, \\ a_1\tau^3 + b_1\tau^2 + c_1\tau + d_1 & \text{for } \tau_1 < \tau \leq \tau_2, \\ \dots & \dots \\ a_n\tau^3 + b_n\tau^2 + c_n\tau + d_n & \text{for } \tau_n < \tau \leq T, \end{cases}$$

where  $T$  is the maximum maturity considered. By changing the constants  $d_1, \dots, d_n$ , we can write this in an equivalent and more convenient form:

$$y(\tau) = \begin{cases} a_0\tau^3 + b_0\tau^2 + c_0\tau + d_0 & \text{for } 0 < \tau \leq \tau_1, \\ a_1(\tau - \tau_1)^3 + b_1(\tau - \tau_1)^2 + c_1(\tau - \tau_1) + d_1 & \text{for } \tau_1 < \tau \leq \tau_2, \\ \dots & \dots \\ a_n(\tau - \tau_n)^3 + b_n(\tau - \tau_n)^2 + c_n(\tau - \tau_n) + d_n & \text{for } \tau_n < \tau \leq T. \end{cases} \quad (2)$$

To splice these polynomials together at the knot points means to choose the coefficients so that the two polynomials that meet at a knot point have the same value and the same first

and second derivatives at the knot point. For example, at the first knot point  $\tau_1$ , we want the adjacent polynomials to satisfy

$$\begin{aligned} \text{Equality of yields:} & \quad a_0\tau_1^3 + b_0\tau_1^2 + c_0\tau_1 + d_0 = d_1, \\ \text{Equality of first derivatives:} & \quad 3a_0\tau_1^2 + 2b_0\tau_1 + c_0 = c_1, \\ \text{Equality of second derivatives:} & \quad 6a_0\tau_1 + 2b_0 = 2b_1. \end{aligned}$$

Thus, given the coefficients  $a_0, b_0, c_0, d_0$ , the only free coefficient for the second polynomial is  $a_1$ . Likewise, at the second knot point  $\tau_2$ , we want the adjacent polynomials to agree with regard to:

$$\begin{aligned} \text{Yields:} & \quad a_1(\tau_2 - \tau_1)^3 + b_1(\tau_2 - \tau_1)^2 + c_1(\tau_2 - \tau_1) + d_1 = d_2, \\ \text{First derivatives:} & \quad 3a_1(\tau_2 - \tau_1)^2 + 2b_1(\tau_2 - \tau_1) + c_1 = c_2, \\ \text{Second derivatives:} & \quad 6a_1(\tau_2 - \tau_1) + 2b_1 = 2b_2. \end{aligned}$$

Thus, the only free coefficient for the third polynomial is  $a_2$ . Continuing in this way, we see that the cubic spline is defined by the knot points and the coefficients  $a_0, b_0, c_0, d_0, a_1, a_2, \dots, a_n$ . One wants to choose the knot points (hopefully, not too many) and these coefficients so that the relations Equation 1 hold as closely as possible in some sense. For more on this subject, a good reference is James and Webber [JW].

## LIBOR

Many fixed-income instruments have cash flows tied to interest rate indices that are quoted as simple (i.e., noncompounded) interest rates. If you deposit \$1 at an annualized simple interest rate of  $\mathcal{R}$  for a period of time  $\Delta t$ , then at the end of the period you will have  $1 + \mathcal{R} \Delta t$  dollars. [The calligraphic symbol  $\mathcal{R}$  will be used only for simple interest rates and hopefully will not be confused with the symbol  $R$  used for the accumulation factor  $R(t) = e^{\int_0^t r(s) ds}$  and the corresponding risk-neutral expectation  $E^R$ .] The most common interest rate index is LIBOR (London Inter-Bank Offered Rate) which is an average rate offered by large London banks on large deposits for a specific term. For example, if  $\mathcal{R}$  denotes six-month LIBOR, a \$1 million deposit at this rate will grow in six months to \$1 million times  $1 + \mathcal{R} \Delta t$ , where  $\Delta t = 1/2$ . We will use LIBOR as a generic term for such indices.

We will frequently find it convenient to express LIBOR rates in terms of equivalent bond prices. As before, let  $P(t, u)$  denote the price at date  $t$  of a zero-coupon bond with a face value of \$1 maturing at date  $u$ , having  $\Delta t = u - t$  years to maturity. Previously, we used this notation only for default-free bonds such as Treasuries. However, there is a small amount of credit risk in LIBOR rates because of the possibility of a bank failure. In discussing derivatives linked to LIBOR rates, such as swaps, caps, and floors, we will use the notation  $P(t, u)$  for the price of a bond having the same default risk as a LIBOR deposit, but our models will ignore the possibility of default. An investment of \$1 at date  $t$  in the bond will purchase  $1/P(t, u)$  units

of the bond, which, in the absence of default, will be worth  $1/P(t, u)$  dollars at maturity. We will assume

$$\frac{1}{P(t, u)} = 1 + \mathcal{R} \Delta t . \quad (3)$$

When necessary for clarity, we call the rate  $\mathcal{R}$  a spot rate (at date  $t$  for the time period  $\Delta t$ ), to distinguish it from forward rates to be defined later. The spot rate is also called a floating rate, because it changes with market conditions.

## Swaps

A plain vanilla interest rate swap involves the swap of a fixed interest rate for a floating interest rate on a given notional principal. Let  $\bar{\mathcal{R}}$  denote the fixed rate on a swap. The floating rate will be LIBOR (or some other interest rate index). In addition to  $\bar{\mathcal{R}}$  and the floating rate index, the swap is defined by payment dates, which we will denote by  $t_1, \dots, t_N$ , with  $t_{i+1} - t_i = \Delta t$ . In the most common form, the reset dates are  $t_0, \dots, t_{N-1}$ , with  $t_1 - t_0 = \Delta t$  also.

At each reset date  $t_i$ , the simple interest rate  $\mathcal{R}_i$  for period  $\Delta t$  is observed. This rate determines a payment at the following date  $t_{i+1}$ . In terms of bond prices,  $\mathcal{R}_i$  is defined in accord with Equation 3, substituting date  $t_i$  for date  $t$  and date  $t_{i+1}$  for date  $u$ ; i.e.,

$$\frac{1}{P(t_i, t_{i+1})} = 1 + \mathcal{R}_i \Delta t . \quad (4)$$

One can enter a swap as the fixed-rate payer, normally called simply the payer or as the fixed-rate receiver, normally called the receiver. The payer pays the fixed rate  $\bar{\mathcal{R}}$  and receives the spot rate  $\mathcal{R}_i$  at each payment date  $t_{i+1}$ . Only the net payment is exchanged. If  $\bar{\mathcal{R}} > \mathcal{R}_i$  then the amount  $(\bar{\mathcal{R}} - \mathcal{R}_i) \Delta t$  is paid at date  $t_{i+1}$  by the payer to the receiver for each \$1 of notional principal. If  $\bar{\mathcal{R}} < \mathcal{R}_i$ , then the payer receives  $(\mathcal{R}_i - \bar{\mathcal{R}}) \Delta t$  from the receiver for each \$1 of notional principal at date  $t_{i+1}$ . To state this more simply, the cash flow to the payer is  $(\mathcal{R}_i - \bar{\mathcal{R}}) \Delta t$  and the cash flow to the receiver is  $(\bar{\mathcal{R}} - \mathcal{R}_i) \Delta t$ , with the usual understanding that a negative cash flow is an outflow. Note that there is no exchange of principal at initiation, and there is no return of principal at the end of the swap. The principal is notional because it is used only to define the interest payments.

The value of a swap to the fixed-rate payer at any date  $t \leq t_0$  is

$$P(t, t_0) - P(t, t_N) - \bar{\mathcal{R}} \Delta t \sum_{i=1}^N P(t, t_i) . \quad (5)$$

To see this, note that  $P(t, t_0)$  is the cost at date  $t$  of receiving \$1 at  $t_0$ . This \$1 can be invested at  $t_0$  at the rate  $\mathcal{R}_0$  and the amount  $\mathcal{R}_0 \Delta t$  withdrawn at  $t_1$  with the \$1 principal rolled over at the new rate  $\mathcal{R}_1$ . Continuing in this way, one obtains the cash flow  $\mathcal{R}_i \Delta t$  at each payment date

$t_{i+1}$  and the recovery of the \$1 principal at date  $t_N$ . The value of the \$1 principal at date  $t_N$  is negated in expression Equation 5 by the term  $-P(t, t_N)$ . Thus,  $P(t, t_0) - P(t, t_N)$  is the value of the floating rate payments on the notional principal. On the other hand,  $\bar{\mathcal{R}} \Delta t \sum_{i=1}^N P(t, t_i)$  is the value of the fixed-rate payments. Therefore, expression Equation 5 is the difference in the values of the floating and fixed-rate legs.

As with a forward price, the swap rate is usually set so that the value of the swap is zero at initiation. A swap initiated at date  $t \leq t_0$  has zero value at initiation if the fixed rate  $\bar{\mathcal{R}}$  equates the expression Equation 5 to zero. This means that  $\bar{\mathcal{R}} = \mathcal{R}(t)$ , where  $\mathcal{R}(t)$  is defined by

$$P(t, t_0) = P(t, t_N) + \mathcal{R}(t) \Delta t \sum_{i=1}^N P(t, t_i) . \quad (6)$$

If  $t = t_0$  the rate  $\mathcal{R}(t)$  is a spot swap rate, and if  $t < t_0$  the rate  $\mathcal{R}(t)$  is a forward swap rate. The concept of forward swap rates will be important in the discussion of swaptions in **Section 5.2 Swaptions**. Of course, there are many spot and forward swap rates at any date, corresponding to swaps with different maturities and different payment (and reset) dates.

The swap yield curve or simply swap curve is the relation between time-to-maturity and the yields of discount bonds, where the discount bond prices and yields are inferred from market swap rates. To explain this in a manner consistent with Section 5, consider date  $t = 0$  and consider swaps with  $t_0 = 0$  (i.e., spot swaps). In the notation of Section 5, and noting that  $P(0, 0) = 1$ , Equation 6 can be written in terms of yields as

$$1 = e^{-y(t_N)t_N} + \sum_{i=1}^N e^{-y(t_i)t_i} \mathcal{R}(0) \Delta t . \quad (7)$$

Consider for example a collection of nineteen swaps at date 0 with semi-annual payments and maturity dates  $t_N = 1.0, t_N = 1.5, \dots, t_N = 10.0$ . Each market swap rate (a different rate  $\mathcal{R}(0)$  for each maturity) can be considered to satisfy Equation 7. In these equations we have the twenty yields  $y(0.5), y(1.0), y(10.0)$ . The yield  $y(0.5)$  will be given by six-month LIBOR according to Equation 3. The other nineteen yields can be obtained by simultaneously solving the system of nineteen equations of the form Equation 7, given the nineteen market swap rates. In practice, there are missing maturities and the swap curve is estimated using a cubic spline or some other technique, as discussed in Section 5.

## Yield to Maturity, Duration, and Convexity

Consider a bond with cash flows  $C_1, \dots, C_N$  at dates  $u_1 < \dots < u_N$  and price  $P$  at date  $t$ , where  $t < u_1$ . Write  $\tau_j = u_j - t$  as the time remaining until the  $j$ -th cash flow is paid. The

(continuously compounded) yield to maturity of the bond at date  $t$  is defined to be the rate  $\mathbf{y}$  such that

$$P = \sum_{j=1}^N e^{-\mathbf{y}\tau_j} C_j . \quad (8)$$

The bold character  $\mathbf{y}$  is meant to distinguish this from the yield  $y$  of a discount bond. Viewing the right-hand side of Equation 8 as a function of  $\mathbf{y}$ , we can express the first derivative in differential form as

$$dP = - \sum_{j=1}^N \tau_j e^{-\mathbf{y}\tau_j} C_j d\mathbf{y} ,$$

or, equivalently,

$$\frac{dP}{P} = - \sum_{j=1}^N \frac{e^{-\mathbf{y}\tau_j} C_j}{P} \tau_j d\mathbf{y} .$$

The factor

$$\sum_{j=1}^N \frac{e^{-\mathbf{y}\tau_j} C_j}{P} \tau_j$$

is called the Macaulay duration of the bond, and we will simplify this to duration. It is a weighted average of the times to maturity  $\tau_j$  of the cash flows, the weight on each time  $\tau_j$  being the fraction of the bond value that the cash flow constitutes (using the same rate  $\mathbf{y}$  to discount all of the cash flows). Thus, we have

$$\frac{dP}{P} = -\text{Duration} \times d\mathbf{y} . \quad (9)$$

Given the initial yield  $\mathbf{y}$  and a change in the yield to  $\mathbf{y}'$ , this equation suggests

the following approximation for the return  $\Delta P/P$ :

$$\frac{\Delta P}{P} \approx -\text{Duration} \times \Delta \mathbf{y} , \quad (10)$$

where  $\Delta \mathbf{y} = \mathbf{y}' - \mathbf{y}$ .

The relationship Equation 10 is the foundation for duration hedging. For example, to duration hedge a liability with a present value of  $x_L$  dollars and a duration of  $D_L$  years, one needs an asset with a value of  $x_A$  dollars and a duration of  $D_A$  satisfying  $x_A D_A = x_L D_L$ . If the change in the yields to maturity of the asset and liability are the same number  $\Delta \mathbf{y}$ , then by Equation 10 the change in the value of the liability will be approximately  $-D_L \Delta \mathbf{y}$  per dollar of initial value, for a total change in value of  $-x_L D_L \Delta \mathbf{y}$  dollars. The change in the value of the asset will approximately offset the change in the value of the liability.

Actually, because of the convexity of the bond price Equation 8 as a function of the yield  $\mathbf{y}$ , the approximation Equation 10 will overstate the loss on an asset when the yield rises and understate the gain when the yield falls. Given a change in yield  $\Delta \mathbf{y}$ , the change in price  $\Delta P = P' - P$  would actually satisfy

$$\Delta P > -\text{Duration} \times P \times \Delta \mathbf{y} .$$

Thus, if a liability is duration hedged, and the asset value is a more convex function of its yield than is the liability value, then an equal change  $\Delta \mathbf{y}$  in their yields will lead to a net gain, the asset value falling less than the liability if  $\Delta \mathbf{y} > 0$  and gaining more than the liability if  $\Delta \mathbf{y} < 0$ . The value of convexity (as a function of the yield to maturity) in a bond portfolio is the same as the value of convexity (as a function of the price of the underlying) in option hedging—cf. [@sec-s\\_deltahedging](#).

In general, the changes in the yields of an asset and a liability (or two different coupon bonds) will not be equal. To understand how the changes will be related, note that Equation 8 of the yield to maturity and Equation 1 relating a bond price to the yields of discount bonds imply

$$\sum_{j=1}^N e^{-\mathbf{y}\tau_j} C_j = \sum_{j=1}^N e^{-y(\tau_j)\tau_j} C_j .$$

Taking differentials of both sides, we have

$$dP = -P \times \text{Duration} \times d\mathbf{y} = -\sum_{j=1}^N \tau_j e^{-y(\tau_j)\tau_j} C_j dy(\tau_j) .$$

If we suppose that the changes  $dy(\tau_j)$  in the yields of the discount bonds are equal, to, say,  $dy$ , then we have

$$dP = -P \times \text{Duration} \times d\mathbf{y} = -P \times \text{Duration}' \times dy ,$$

where we define

$$\text{Duration}' = \sum_{j=1}^N \frac{e^{-y(\tau_j)\tau_j} C_j}{P} \tau_j. \quad (11)$$

This new definition of duration differs from the previous by using the yields of discount bonds to define the fraction of the bond value that each cash flow contributes, rather than the yield to maturity. If the changes in the yields of discount bonds are equal, then we say that there has been a parallel shift in the yield curve—it has moved up or down with the new curve being at each point the same distance from the old. Our previous discussion shows that duration hedging works if we use the new Equation 11 of duration and if the yield curve shifts in a parallel fashion. Of course, parallel shifts in the yield curve are not the only, or even most common, types of shifts. In the next two sections, we discuss hedging against more general types of shifts in the yield curve.

Something very similar to duration hedging works if we can continuously rebalance the hedge and there is only a single factor determining the yield curve (meaning a single Brownian motion driving all yields). To understand this, we must first note that the expressions given in this section for  $dP$  are not the Ito differential, which explains how the bond price evolves over time and would include a second-derivative term. For example, in Equation 9 we are simply asking how different the price would be if the yield to maturity had been different at a given point in time. To define the Ito differential, let now  $\mathbf{y}(t)$  denote the yield to maturity of the bond at date  $t$ . Equation Equation 8 can be restated as

$$P(t) = f(t, \mathbf{y}(t)) = \sum_{j=1}^N e^{\mathbf{y}(t)(u_j - t)} C_j. \quad (12)$$

Note that even if the yield to maturity were constant over time, the bond price would change with  $t$  as a result of the changes in the times to maturity  $u_j - t$  of the cash flows. This creates the dependence on  $t$  in the function  $f(t, \mathbf{y})$ .

From Ito's formula, we have

$$dP = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial \mathbf{y}} d\mathbf{y} + \frac{1}{2} \frac{\partial^2 f}{\partial \mathbf{y}^2} (d\mathbf{y})^2.$$

As explained above, the factor  $\partial f / \partial \mathbf{y}$  equals  $\text{\$} - \text{Duration} \times P\text{\$}$ . The value of convexity appears here in the last term, the derivative  $\partial^2 f / \partial \mathbf{y}^2$  being positive as a result of convexity and analogous to the gamma of an option. Assuming  $d\mathbf{y}(t) = \alpha(t) dt + \sigma(t) dB(t)$  for a Brownian motion  $B$  and some  $\alpha$  and  $\sigma$ , we have

$$\frac{dP}{P} = \frac{1}{P} \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{y}} \alpha + \frac{1}{2} \frac{\partial^2 f}{\partial \mathbf{y}^2} \sigma^2 \right) dt - \text{Duration} \times \sigma dB.$$

If the yields of an asset and liability are driven by the same Brownian motion and the duration hedge is adjusted for the relative volatilities of the asset and liability (holding more



of the asset if its yield volatility is lower), then a duration hedge will hedge the risky part of the change in the liability value. If adjusted continuously, it will provide a perfect hedge, exactly analogous to a delta hedge of an option position. The Vasicek model we will discuss in [?@sec-c\\_\\_vasicek](#) and the Cox-Ingersoll-Ross model we will discuss in [?@sec-c\\_\\_survey](#) are examples of continuous-time models that assume a single Brownian motion driving all yields (i.e., they are single-factor models). The Ho-Lee, Black-Derman-Toy, and Black-Karasinski binomial models that will be discussed in [?@sec-c\\_\\_survey](#) are also single-factor models and have the same implication for hedging. In the following section, we will discuss the fact that, empirically, there appears to be more than one factor determining the yield curve.

## Principal Components

This section will describe a popular statistical method for determining the factors that have the most impact on the yield curve. We consider yields at fixed maturities  $\tau_1, \dots, \tau_N$ , the yield for maturity  $\tau_j$  at date  $t$  being denoted  $y(t, \tau_j)$ . We assume that we have a sample of past yields at dates  $t_0, \dots, t_M$  at equally spaced dates. Thus,  $t_i - t_{i-1} = \Delta t$  for some  $\Delta t$  and each  $i$ . We compute the changes in yields:

$$\Delta_{ij} = y(t_i, \tau_j) - y(t_{i-1}, \tau_j) .$$

Thus we are looking at the changes in the yield curve over time periods of length  $\Delta t$ , focusing on  $N$  points on the yield curve defined by the maturities  $\tau_j$ . Let  $V$  denote the sample covariance matrix of the changes in yields: the element in row  $j$  and column  $k$  of  $V$  is the sample covariance of the changes in yields at maturities  $\tau_j$  and  $\tau_k$ ; thus, the diagonal elements are the sample variances.<sup>1</sup> to estimate the covariance matrix. The following applies equally well to other estimates  $V$  of the covariance matrix.

The method of principal components is to compute the eigenvectors and eigenvalues of the estimated covariance matrix  $V$ . An eigenvector is a vector  $x$  for which there corresponds a number  $\lambda$  such that  $Vx = \lambda x$ . The number  $\lambda$  is called the eigenvalue corresponding to the eigenvector  $x$ . Given the  $N \times N$  symmetric matrix  $V$ , we can construct an  $N \times N$  matrix  $C$  whose columns are eigenvectors of  $V$  and an  $N \times N$  diagonal matrix  $D$  containing the eigenvalues of  $V$  on the diagonal. The eigenvectors can be normalized to have unit length and to be mutually orthogonal, which means that the matrix  $C$  of eigenvectors has the property that  $C^{-1} = C^\top$ , where  $C^{-1}$  denotes the inverse of  $C$  and  $C^\top$  its transpose. The property  $Vx = \lambda x$  for the columns  $x$  of  $C$  implies that  $VC = CD$ . Hence  $C^\top VC = C^\top CD = D$ .

This can be understood as a factor model for the changes in yields, where there are as many

---

<sup>1</sup>The sample covariance matrix  $V$  is an estimate of the unconditional covariances. It is a common finding that variances and covariances change over time. Thus, we could (and probably should) use methods such as those described in [?@sec-c\\_\\_stochasticvolatility](#)

factors as maturities. At date  $t_i$ , the vector of factor realizations  $z_{ij}$  is computed as

$$\begin{pmatrix} z_{i1} \\ \vdots \\ z_{iN} \end{pmatrix} = C^\top \begin{pmatrix} \Delta_{i1} \\ \vdots \\ \Delta_{iN} \end{pmatrix}. \quad (13)$$

Thus,

$$\begin{pmatrix} \Delta_{i1} \\ \vdots \\ \Delta_{iN} \end{pmatrix} = C \begin{pmatrix} z_{i1} \\ \vdots \\ z_{iN} \end{pmatrix}. \quad (14)$$

Given any random vector  $\xi$  with covariance matrix  $\Sigma$  and a linear transformation  $\xi' = L\xi$ , the covariance matrix of  $\xi'$  is  $L\Sigma L^\top$ . Therefore, Equation 13 implies that the covariance matrix of the factors is  $C^\top VC$ , and we observed in the previous paragraph that  $C^\top VC = D$ . Therefore, the factors are uncorrelated, and the factor variances are the eigenvalues of  $V$ .

Let  $\beta_{k\ell}$  denote the  $(k, \ell)$ -th element of  $C$ . Then we can write Equation 14 as

$$\begin{aligned} \Delta_{i1} &= \beta_{11}z_{i1} + \cdots + \beta_{1N}z_{iN}, \\ \vdots &\quad \quad \quad \vdots \\ \Delta_{iN} &= \beta_{N1}z_{i1} + \cdots + \beta_{NN}z_{iN}. \end{aligned} \quad (15)$$

As in any factor model, the factors are common to all of the maturities. Each factor is random, taking a different value at each date  $t_i$ . The  $\beta$ 's represent the loadings of the yield changes on the factors,  $\beta_{k\ell}$  being the loading of the change in the yield at maturity  $\tau_k$  on the  $\ell$ -th factor.

In a normal factor model, there are fewer factors than variables being explained. It serves no point to have a factor model with as many factors as there are variables to be explained (in our case, as many factors as maturities). We can improve the usefulness of the above by omitting some factors. We will omit the factors that are least important in explaining the changes in yields. For example, if we omit all but the first three factors, we will have

$$\begin{aligned} \Delta_{i1} &= \beta_{11}z_{i1} + \beta_{12}z_{i2} + \beta_{13}z_{i3} + \varepsilon_{i1}, \\ \vdots &\quad \quad \quad \vdots \\ \Delta_{iN} &= \beta_{N1}z_{i1} + \beta_{N2}z_{i2} + \beta_{N3}z_{i3} + \varepsilon_{iN}, \end{aligned} \quad (16)$$

where

$$\varepsilon_{ij} = \beta_{j4}z_{i4} + \cdots + \beta_{jN}z_{iN} \quad (17)$$

is interpreted as the residual part. [Frequently, the definition of factor model requires the residuals to be uncorrelated, in which case they are called idiosyncratic risks. of  $\Delta_{ij}$ .

The importance of a factor depends on the factor loadings and the variance of the factor. The loadings on the  $j$ -th factor are the elements in the  $j$ -th column of  $C$ , which is the  $j$ -th

eigenvector of  $V$ . Because the eigenvectors all have unit length and are mutually orthogonal, each vector of loadings has the same importance as any other for explaining the changes in yields. Thus, the importance of a factor in this model depends on the variance of the factor. The variance of the  $j$ -th factor is the  $j$ -th element on the diagonal of  $D$ , which is the eigenvalue corresponding to the  $j$ -th eigenvector. The factors that we should omit are clearly those with small eigenvalues.

As an example, an analysis of monthly changes in (continuously compounded) U.S. Treasury yields from 1992 through 2002 at maturities of 1 month, 3 months, 1 year, 2 years, 3 years, 4 years, and 5 years<sup>2</sup> produces seven eigenvalues (corresponding to the seven maturities) that sum to  $5.526 \times 10^{-5}$ . This sum is the total variance of the seven factors. The largest eigenvalue is 74% of the total, the two largest eigenvalues constitute 94% of the total, and the three largest constitute 98% of the total. Thus, three factors contribute 98% of the total factor variance for this data set, so a factor model with three (or even two) factors explains a very high percentage of the changes in yields for this data set.

The factors can be interpreted by examining the corresponding eigenvectors. The eigenvectors corresponding to the three largest eigenvalues in this data set are the columns below (the first column corresponding to the largest eigenvalue, etc.):

0.1967	-0.8512	0.4782
0.2234	-0.3740	-0.6389
0.3775	-0.1077	-0.4783
0.4415	0.0855	-0.0530
0.4528	0.1532	0.0980
0.4428	0.2013	0.2043
0.4158	0.2294	0.2834

We can interpret these as follows. A positive value in a given month for the factor with the highest variance will lead to an increase in all of the yields, because all of the elements in the first column are positive (it will also lead to a slight increase in the slope due to the loadings at longer maturities being generally slightly larger than the loadings at smaller maturities). A positive value for the next factor will decrease yields at short maturities and increase the yields at longer maturities, thus leading to an increase in the slope of the yield curve. A positive value for the third factor will lead to an increase in yields at short and long maturities and a decrease in yields at intermediate maturities, thus affecting the curvature of the yield curve. Results of this sort are common for data sets containing longer maturities also, leading to the conclusion that the most important factor is the level of the yield curve, the second most important factor is the slope of the yield curve, and the third most important factor is the curvature of the yield curve, with three factors explaining nearly all of the variations in yields.

---

<sup>2</sup>These were computed from discount bond price and yield data from the Center for Research in Security Prices (CRSP) at the University of Chicago.

## Hedging Principal Components

Consider again a bond with cash flows  $C_1, \dots, C_N$  at dates  $\tau_1 < \dots < \tau_N$  and price  $P$  and recall that the price at date  $0 < \tau_1$  should be

$$P = \sum_{j=1}^N e^{-y_j \tau_j} C_j ,$$

where for convenience we are writing  $y_j$  for the yield  $y(0, \tau_j)$  of the discount bond maturing at  $\tau_j$ . Viewing the price as a function of  $y_1, \dots, y_N$ , we can write the differential as

$$dP = - \sum_{j=1}^N \tau_j e^{-y_j \tau_j} C_j dy_j .$$

Equivalently,

$$\frac{dP}{P} = - \sum_{j=1}^N \frac{e^{-y_j \tau_j} C_j \tau_j}{P} dy_j .$$

Given discrete changes  $\Delta y_j$  in the yields, this implies the following approximation for the return:

$$\frac{\Delta P}{P} \approx - \sum_{j=1}^N \frac{e^{-y_j \tau_j} C_j \tau_j}{P} \Delta y_j . \quad (18)$$

As in Section , because of convexity, the approximation understates the new price  $P' = P + \Delta P$ . In Section we considered this equation assuming equal changes in the yields (a parallel shift in the yield curve). Here, we will discuss the more general case.

The approximation Equation 18 suggests how we can hedge against the factors identified in the previous section. For example, let  $\beta_{1\ell}, \dots, \beta_{N\ell}$  denote the loadings of the yields at maturities  $\tau_1, \dots, \tau_N$  on the factor with the  $\ell$ -th greatest variance, for  $\ell = 1, 2, 3$ .<sup>3</sup> Then the factor model suggests

$$\Delta y_j \approx \beta_{j1} z_1 + \beta_{j2} z_2 + \beta_{j3} z_3 ,$$

where the  $z_k$  denote the realizations of the factors. Combining this with Equation 18 results in

$$\begin{aligned} \frac{\Delta P}{P} &\approx - \sum_{j=1}^N \left( \frac{e^{-y_j \tau_j} C_j \tau_j}{P} \sum_{k=1}^3 \beta_{jk} z_k \right) \\ &= - \sum_{k=1}^3 \left( \sum_{j=1}^N \frac{e^{-y_j \tau_j} C_j \tau_j \beta_{jk}}{P} \right) z_k . \end{aligned}$$

---

<sup>3</sup>It is unlikely that these specific maturities would have been used in the principal components algorithm, so the loadings would have to be estimated by interpolating or fitting some type of curve to the loadings of the maturities that were used.

This means that the bond return is approximately a linear combination of the factors, with the coefficient (loading) on the  $k$ -th factor being

$$-\sum_{j=1}^N \frac{e^{-y_j \tau_j} C_j \tau_j \beta_{jk}}{P} . \quad (19)$$

To hedge a liability against the factor, we want this coefficient for the asset multiplied by the dollar value of the asset to equal the corresponding coefficient for the liability multiplied by its dollar value. There are three conditions of this type in a three-factor model, which means that a portfolio of three distinct assets is necessary to hedge a liability. A similar application is in bond portfolio management. If we want to avoid exposing ourselves to factor risk relative to a benchmark portfolio, then we should match the loadings Equation 19 for our portfolio with the corresponding loadings of the benchmark.

## Exercises

Assume the following discount bond prices for each of the following exercises:

$$P(0, 0.5) = 0.995$$

$$P(0, 1.0) = 0.988$$

$$P(0, 1.5) = 0.978$$

$$P(0, 2.0) = 0.966$$

$$P(0, 2.5) = 0.951$$

$$P(0, 3.0) = 0.935$$

$$P(0, 3.5) = 0.916$$

$$P(0, 4.0) = 0.896$$

$$P(0, 4.5) = 0.874$$

$$P(0, 5.0) = 0.850$$

**Exercise 0.1.** Compute the six-month and one-year LIBOR rates.

**Exercise 0.2.** Compute the swap rate for a two-year swap with semi-annual payments.

**Exercise 0.3.** Compute the forward swap rate for a two-year swap with semi-annual payments beginning in

1. one year,
2. two years,
3. three years.

**Exercise 0.4.** Compute and plot the continuously-compounded discount-bond yields at maturities 0.5, 1.0, 1.5, ..., 5.0.

**Exercise 0.5.** Consider a cubic spline as defined in Equation 2 with knot points at two and four years and the following coefficients:

$$a_0 = -0.00163$$

$$b_0 = 0.00812$$

$$c_0 = -0.00676$$

$$d_0 = 0.01184$$

$$a_1 = 0.00052$$

$$b_1 = -0.00169$$

$$c_1 = 0.00609$$

$$d_1 = 0.01770$$

$$a_2 = -0.00175$$

$$b_2 = 0.00141$$

$$c_2 = 0.00552$$

$$d_2 = 0.02725$$

1. Plot the cubic spline and compare it to the plot from the previous exercise.
2. Confirm that the adjacent polynomials have the same values, first derivatives, and second derivatives at each of the two knot points.

**Exercise 0.6.** Create a Python function `DiscountBondPrice` that takes the time  $\tau$  to maturity as an input and returns an estimated price for a discount bond maturing at  $\tau$ , using the cubic spline Equation 2, knot points at 2 and 4 years, and the coefficients in the previous exercise. Confirm that it gives approximately the same discount bond prices as those at the beginning of this set of exercises. *A Warning about Extrapolating to Longer Maturities:* What price does the function give for a discount bond maturing in ten years?

**Exercise 0.7.** Consider a two-year coupon bond with \$1 face value and a semi-annual coupon of \$0.03, with the first coupon being six months away.

1. Compute the price of the bond.
2. Compute the yield to maturity of the bond. Note: There are many ways to do this, including the Excel IRR function.
3. Compute the duration of the bond.

**Exercise 0.8.** Repeat the previous exercise for a one-year coupon bond with \$1 face value and a semi-annual coupon of \$0.04.

**Exercise 0.9.** Repeat the previous exercise for a five-year coupon bond with \$1 face value and a semi-annual coupon of \$0.02.

**Exercise 0.10.** Suppose you have shorted \$100 million of the two-year bond with the semi-annual coupon of \$0.03. How much of the five-year bond with the semi-annual coupon of \$0.02 should you hold to duration hedge the short position?

**Exercise 0.11.** Suppose you have shorted \$100 million of the two-year bond with the semi-annual coupon of \$0.03. Using the data from the principal components example in Section , find a portfolio of the one-year bond with a coupon of \$0.04 and the five-year bond with a coupon of \$0.02 that will hedge the first two principal components of the short position. Assume the loadings of the six-month yield on the two factors are the averages of the loadings of the one-month and one-year yields, the loadings of the 1.5 year yield on the two factors are the averages of the loadings of the 1.0 and 2.0 years, the loadings of the 2.5 year yield are the averages of the loadings of the 2.0 and 3.0 year yields, etc.