that there are 253 trading days in a year and that market prices have no volatility when markets are closed, i.e., weekends, holidays.

9.4.2 Provide a 95% confidence interval for $\hat{H}(0)$ and an estimate of the number of simulations needed to yield a price estimate that is within \$.05 of the true price.

9.4.3 How does this price compare with the price given by the Goldman-Sosin-Gatto formula? Can you explain the discrepancy? Which price would you use to decide whether to accept or reject CLM's proposal?

10 Fixed-Income Securities

IN THIS CHAPTER and the next we turn our attention to the bond markets. We study bonds that have no call provisions or default risk, so that their payments are fully specified in advance. Such bonds deserve the name *$xed-income securities* that is often used more loosely to describe bonds whose future payments are in fact uncertain. In the US markets, almost all true fixed-income securities are issued by the US Treasury. Conventional Treasury securities make fixed payments in nominal terms, but in early 1996 the Treasury announced plans to issue *indexed* bonds whose nominal payments are indexed to inflation so that their payments are fixed in real terms.[1]

Many of the ideas discussed in earlier chapters can be applied to fixed-income securities as well as to any other asset. But there are several reasons to devote special attention to fixed-income securities. First, the fixed-income markets have developed separately from the equity markets. They have their own institutional structure and their own terminology. Likewise the academic study of fixed-income securities has its own traditions. Second, the markets for Treasury securities are extremely large regardless of whether size is measured by quantities outstanding or quantities traded. Third, fixed-income securities have a special place in finance theory because they have no cash-flow uncertainty, so their prices vary only as discount rates vary. By studying fixed-income securities we can explore the effects of changing discount rates without having to face the complications introduced by changing expectations of future cash flows. The prices of conventional Treasury securities carry information about nominal discount rates, while the prices of indexed securities carry information about real discount rates. Finally, many other assets can be seen as combinations of fixed-income securities and derivative securities; a callable bond, for example, is a fixed-income security less a put option.

[1] Such bonds have already been issued by the UK, Canadian, and several other govemments. See Campbell and Shiller (1996) for a review.

The literature on fixed-income securities is vast.[2] We break it into two main parts. First, in this chapter we introduce basic concepts and discuss empirical work on linear time-series models of bond yields. This work is only loosely motivated by theory and has the practical aim of exploring the forecasting power of the term structure of interest rates. In Chapter **11** we turn to more ambitious, fully specified term-structure models that can be used to price interest-rate derivative securities.

10.1 Basic Concepts

In principle a fixed-income security can promise a stream of future payments of any form, but there are two classic cases.

Zero-coupon bonds, also called *discount bonds,* make a single payment at a date in the future known as the *maturity date.* The size of this payment is the *face value* of the bond. The length of time to the maturity date is the *maturity* of the bond. US Treasury bills (Treasury obligations with maturity at issue of up to **12** months) take this form.

Coupon bonds make *coupon payments* of a given fraction of face value at equally spaced dates up to and including the maturity date, when the face value is also paid. US Treasury notes and bonds (Treasury obligations with maturity at issue above **12** months) take this form. Coupon payments on Treasury notes and bonds are made every six months, but the coupon rates for these instruments are normally quoted at an annual rate; thus a 7% Treasury bond actually pays **3.5%** of face value every six months up to and including maturity.[3]

Coupon bonds can be thought of as packages of discount bonds, one corresponding to each coupon payment and one corresponding to the final coupon payment together with the repayment of principal. This is not merely an academic concept, as the principal and interest components of US Treasury bonds have been traded separately under the Treasury's STRIPS (Separate Trading of Registered Interest and Principal Securities) program since **1985,** and the prices of such Treasury *strips* at all maturities have been reported daily in the *Wall Street Journal* since **1989.**

[2]Fortunately it has increased in quality since Ed Kane's judgement: "It is generally agreed that, ceteris paribus, the fertility of a field is roughly proportional to the quantity of manure that has been dumped upon it in the recent past. By this standard, the term structure of interest rates has become ... an extraordinarily fertile field indeed" (Kane [1970]). See Melino (1988) or Shiller (1990) for excellent recent surveys, and Sundaresan (1996) for a book-length treatment.

[3]See a textbook such as Fabozzi and Fabozzi (1995) or Fabozzi (1996) for further details on the markets for US Treasury securities.

10.1.1 Discount Bonds

We first define and illustrate basic bond market concepts for discount bonds. The *yield to maturity* on a bond is that discount rate which equates the present value of the bond's payments to its price. Thus if P_{nt} is the time t price of a discount bond that makes a single payment of **$1** at time $t+n$, and Y_{nt} is the bond's yield to maturity, we have

$$P_{nt} = \frac{1}{(1+Y_{nt})^n}, \tag{10.1.1}$$

so the yield can be found from the price as

$$(1+Y_{nt}) = P_{nt}^{-(\frac{1}{n})}. \tag{10.1.2}$$

It is common in the empirical finance literature to work with log or continuously compounded variables. This has the usual advantage that it transforms the nonlinear equation **(10.1.2)** into a linear one. Using lowercase letters for logs the relationship between log yield and log price is

$$y_{nt} = -\left(\frac{1}{n}\right) p_{nt}. \tag{10.1.3}$$

The *term structure of interest rates* is the set of yields to maturity, at a given time, on bonds of different maturities. The *yield spread* $S_{nt} \equiv Y_{nt} - Y_{1t}$, or in log terms $s_{nt} \equiv y_{nt} - y_{1t}$, is the difference between the yield on an n-period bond and the yield on a one-period bond, a measure of the shape of the term structure. The *yield curve* is a plot of the term structure, that is, a plot of Y_{nt} or y_{nt} against n on some particular date t. The solid line in Figure **10.1.1** shows the log zero-coupon yield curve for US Treasury securities at the end of January 1987.[4] This particular yield curve rises at first, then falls at longer maturities so that it has a hump shape. This is not unusual, although the yield curve is most commonly upward-sloping over the whole range of maturities. Sometimes the yield curve is *inverted,* sloping down over the whole range of maturities.

Holding-Period Returns

The *holding-period return* on a bond is the return over some holding period less than the bond's maturity. In order to economize on notation, we specialize at once to the case where the holding period is a single period.[5] We

[4]This curve is not based on quoted strip prices, which are readily available only for recent years, but is estimated from the prices of coupon-bearing Treasury bonds. Figure 10.1.1 is due to McCulloch and Kwon (1993) and uses McCulloch's (1971, 1975) estimation method as discussed in section 10.1.3 below.

[5]Shiller (1990) gives a much more comprehensive treatment, which requires more complicated notation.

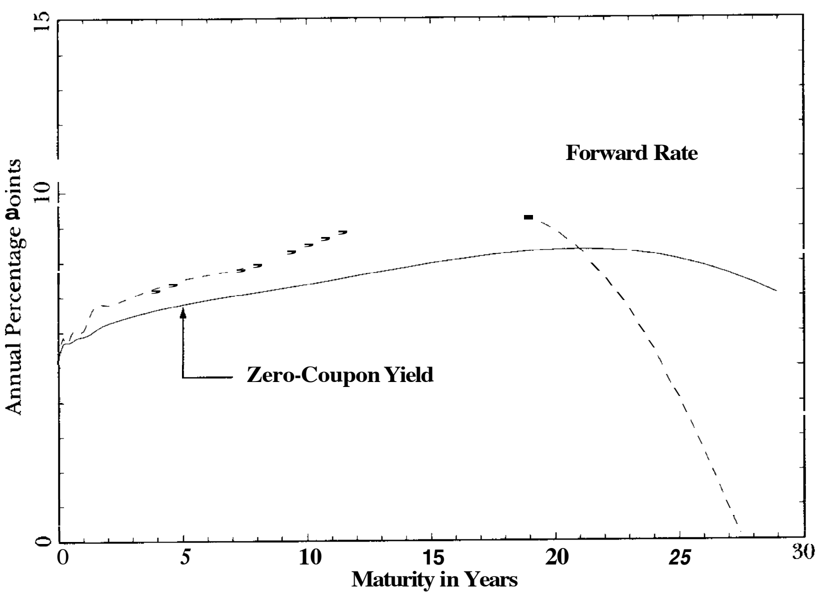

Figure 10.1. *Zero-Coupon* Yield and Forward-Rate Curves in January 1987

define $R_{n,t+1}$ as the one-period holding-period return on an n-period bond purchased at time t and sold at time $t+1$. Since the bond will be an $(n-1)$-period bond when it is sold, the sale price is $P_{n-1,t+1}$ and the holding-period return is

$$(1+R_{n,t+1}) = \frac{P_{n-1,t+1}}{P_{nt}} = \frac{(1+Y_{nt})^n}{(1+Y_{n-1,t+1})^{n-1}}. \tag{10.1.4}$$

The holding-period return in (10.1.4) is high if the bond has a high yield when it is purchased at time t, and if it has a low yield when it is sold at time $t+1$ (since a low yield corresponds to a high price).

Moving to logs for simplicity, the log holding-period return, $r_{n,t+1} \equiv \log(1+R_{n,t+1})$, is

$$\begin{aligned} r_{n,t+1} = p_{n-1,t+1} - p_{nt} &= n\,y_{nt} - (n-1)y_{n-1,t+1} \\ &= y_{nt} - (n-1)(y_{n-1,t+1} - y_{nt}). \end{aligned} \tag{10.1.5}$$

The last equality in (10.1.5) shows how the holding-period return is determined by the beginning-of-period yield (positively) and the change in the yield over the holding period (negatively).

Equation (10.1.5) can be rearranged so that it relates the log bond price today to the log price tomorrow and the return over the next period: $p_{nt} = r_{n,t+1} + p_{n-1,t+1}$. One can solve this difference equation forward, substituting out future log bond prices until the maturity date is reached (and noting that the log price at maturity equals zero) to obtain $p_{nt} = -\sum_{i=0}^{n-1} r_{n-i,t+1+i}$, or in terms of the yield

$$y_{nt} = \left(\frac{1}{n}\right)\sum_{i=0}^{n-1} r_{n-i,t+1+i}. \tag{10.1.6}$$

This equation shows that the log yield to maturity on a zero-coupon bond equals the average log return per period if the bond is held to maturity.

Forward Rates

Bonds of different maturities can be combined to guarantee an interest rate on a fixed-income investment to be made in the future; the interest rate on this investment is called a forward *rate*.[6]

To guarantee at time t an interest rate on a one-period investment to be made at time $t+n$, an investor can proceed as follows. The desired future investment will pay \$1 at time $t+n+1$ so she first buys one $(n+1)$-period bond; this costs $P_{n+1,t}$ at time t and pays \$1 at time $t+n+1$. The investor wants to transfer the cost of this investment from time t to time $t+n$; to do this she sells $P_{n+1,t}/P_{nt}$ n-period bonds. This produces a positive cash flow of $P_{nt}(P_{n+1,t}/P_{nt}) = P_{n+1,t}$ at time t, exactly enough to offset the negative time t cash flow from the first transaction. The sale of n-period bonds implies a negative cash flow of $P_{n+1,t}/P_{nt}$ at time $t+n$. This can be thought of as the cost of the one-period investment to be made at time $t+n$. The cash flows resulting from these transactions are illustrated in Figure 10.2.

The forward rate is defined to be the return on the time $t+n$ investment of $P_{n+1,t}/P_{nt}$:

$$(1+F_{nt}) = \frac{1}{(P_{n+1,t}/P_{nt})} = \frac{(1+Y_{n+1,t})^{n+1}}{(1+Y_{nt})^n} \tag{10.1.7}$$

In the notation F_{nt} the first subscript refers to the number of periods ahead that the one-period investment is to be made, and the second subscript refers to the date at which the forward rate is set. At the cost of additional complexity in notation we could also define forward rates for multiperiod investments, but we do not pursue this further here.

[6] An example of forward trading is the *when-issued* market in US Treasury securities. After an auction of new securities is announced but before the securities are issued, the securities are traded in the when-issued market, with settlement to occur when the securities are issued.

Transactions	time t	$t+n$	$t+n+1$
Buy 1 $(n+1)$-period bond	$-P_{n+1,t}$		1
Sell $P_{n+1,t}/P_{nt}$ n-period bonds	$\left(\frac{P_{n+1,t}}{P_{nt}}\right) P_{nt}$	$-\frac{P_{n+1,t}}{P_{nt}}$	
Net	0	$-\frac{P_{n+1,t}}{P_{nt}}$	1

Figure 10.2. Cash Flows in a *Forward* Transaction

Moving to logs for simplicity, the n-period-ahead log forward rate is

$$\begin{aligned} f_{nt} &= p_{nt} - p_{n+1,t} \\ &= (n+1)\, y_{n+1,t} - n\, y_{nt} \\ &= y_{n+1,t} + n(y_{n+1,t} - y_{nt}) \\ &= y_{nt} + (n+1)(y_{n+1,t} - y_{nt}). \end{aligned} \tag{10.1.8}$$

Equation (10.1.8) shows that the forward rate is positive whenever discount bond prices fall with maturity. Also, the forward rate is above both the n-period and the $(n+1)$-period discount bond yields when the $(n+1)$-period yield is above the n-period yield, that is, when the yield curve is upward-sloping.[7] This relation between a yield to maturity and the instantaneous forward rate at that maturity is analogous to the relation between marginal and average cost. The yield to maturity is the average cost of borrowing for n periods, while the forward rate is the marginal cost of extending the time period of the loan.

Figure 10.1 illustrates the relation between the forward-rate curve (shown as a dashed line) and the yield curve (a solid line). The forward-rate curve lies above the yield curve when the yield curve is upward sloping, and below it when the yield curve is downward-sloping. The two curves cross when the yield curve is flat. These are the standard properties of marginal and average cost curves. When the cost of a marginal unit exceeds the cost of an average unit then the average cost increases with the addition of the

[7] As the time unit shrinks relative to the bond maturity n, the formula (10.1.8) approaches $f_{nt} = y_{nt} + n\,\partial y_{nt}/\partial n$, the n-period yield plus n times the slope of the yield curve at maturity n.

marginal unit, so the average cost rises when the marginal cost is above the average cost. Conversely, the average cost falls when the marginal cost is below the average cost.

10.1.2 Coupon Bonds

As we have already emphasized, a coupon bond can be viewed as a package of discount bonds, one with face value equal to the coupon for each date at which a coupon is paid, and one with the same face value and maturity as the coupon bond itself. Figure 10.3 gives a time line to illustrate the time pattern of payments on a coupon bond.

The price of a coupon bond depends not only on its maturity n and the date t, but also on its coupon rate. To keep notation as simple as possible, we define a period as the time interval between coupon payments and C as the coupon rate *per period.* In the case of US Treasury bonds a period is six months, and C is one half the conventionally quoted annual coupon rate. We write the price of a coupon bond as P_{cnt} to show its dependence on the coupon rate.

The per-period yield to maturity on a coupon bond, Y_{cnt}, is defined as that discount rate which equates the present value of the bond's payments to its price, so we have

$$P_{cnt} = \frac{C}{(1+Y_{cnt})} + \frac{C}{(1+Y_{cnt})^2} + \cdots + \frac{1+C}{(1+Y_{cnt})^n}. \tag{10.1.9}$$

In the case of US Treasury bonds, where a period is six months, Y_{cnt} is the six-month yield and the annual yield is conventionally quoted as twice Y_{cnt}.

Equation (10.1.9) cannot be inverted to get an analytical solution for Y_{cnt}. Instead it must be solved numerically, but the procedure is straightforward since all future payments are positive so there is a unique positive real solution for Y_{cnt}.[8] Unlike the yield to maturity on a discount bond, the yield to maturity on a coupon bond does not necessarily equal the per-period return if the bond is held to maturity. That return is not even defined until one specifies the reinvestment strategy for coupons received prior to maturity. The yield to maturity equals the per-period return on the coupon bond held to maturity only if coupons are reinvested at a rate equal to the yield to maturity.

The implicit yield formula (10.1.9) simplifies in two important special cases. First, when $P_{cnt} = 1$, the bond is said to be selling at par. In this case the yield just equals the coupon rate: $Y_{cnt} = C$. Second, when maturity n

[8] With negative future payments, there can be multiple positive real solutions to (10.1.9). In the analysis of investment projects, the discount rate that equates the present value of a project to its cost is known as the *internal rate of return.* When projects have some negative cash flows in the future, there can be multiple solutions for the internal rate of return.

Time t	$t+1$	$t+2$	$\cdots$	$t+n-1$	$t+n$
(1) Maturity	1	2		$n-1$	n
(2) Face value	C	C		C	$1+C$
(3) Present value discounted at Y_{cnt}	$\frac{C}{(1+Y_{cnt})}$	$\frac{C}{(1+Y_{cnt})^2}$		$\frac{C}{(1+Y_{cnt})^{n-1}}$	$\frac{(1+C)}{(1+Y_{cnt})^n}$
(1) × (3)	$\frac{C}{(1+Y_{cnt})}$	$\frac{2C}{(1+Y_{cnt})^2}$		$\frac{(n-1)C}{(1+Y_{cnt})^{n-1}}$	$\frac{n(1+C)}{(1+Y_{cnt})^n}$

$$D_{cnt} = \frac{\Sigma[(1)\times(3)]}{\Sigma(3)}$$

Figure 10.3. *Calculation of Duration for a Coupon Bond*

is infinite, the bond is called a ***consol*** or ***perpetuity.*** In this case the yield just equals the ratio of the bond price to the coupon rate: $Y_{c\infty t} = C/P_{c\infty t}$.

Duration and Immunization

For discount bonds, maturity measures the length of time that a bondholder has invested money. But for coupon bonds, maturity is an imperfect measure of this length of time because much of a coupon bond's value comes from payments that are made before maturity. ***Macaulay's duration,*** due to Macaulay (1938), is intended to be a better measure; like maturity, its units are time periods. To understand Macaulay's duration, think of a coupon bond as a package of discount bonds. Macaulay's duration is a weighted average of the maturities of the underlying discount bonds, where the weight on each maturity is the present value of the corresponding discount bond calculated using the coupon bond's yield as the discount rate:[9]

$$D_{cnt} = \frac{\frac{C}{(1+Y_{cnt})} + 2\frac{C}{(1+Y_{cnt})^2} + \cdots + n\frac{(1+C)}{(1+Y_{cnt})^n}}{P_{cnt}} = \frac{C\sum_{i=1}^{n}\frac{i}{(1+Y_{cnt})^i} + \frac{n}{(1+Y_{cnt})^n}}{P_{cnt}}. \qquad (10.1.10)$$

The maturity of the first component discount bond is one period and this receives a weight of $C/(1+Y_{cnt})$, the present value of this bond when Y_{cnt} is the discount rate; the maturity of the second discount bond is two and this receives a weight of $C/(1+Y_{cnt})^2$; and so on until the last discount bond of maturity n gets a weight of $(1+C)/(1+Y_{cnt})^n$. To convert this into an average, we divide by the sum of the weights $C/(1+Y_{cnt}) + C/(1+Y_{cnt})^2 + \cdots + (1+C)/(1+Y_{cnt})^n$, which from *(10.1.9)* is just the bond price P_{cnt}. These calculations are illustrated graphically in Figure *10.3.*

When $C = 0$, the bond is a discount bond and Macaulay's duration equals maturity. When $C > 0$, Macaulay's duration is less than maturity and it declines with the coupon rate. For a given coupon rate, duration declines with the bond yield because a higher yield reduces the weight on more distant payments in the average *(10.1.10)*. The duration formula simplifies when a coupon bond is selling at par or has an infinite maturity. A par bond has price $P_{cnt} = 1$ and yield $Y_{cnt} = C$, so duration becomes $D_{cnt} = (1-(1+Y_{cnt})^{-n})/(1-(1+Y_{cnt})^{-1})$. A consol bond with infinite maturity has yield $Y_{c\infty t} = C/P_{c\infty t}$ so duration becomes $D_{c\infty t} = (1+Y_{c\infty t})/Y_{c\infty t}$.

Numerical examples that illustrate these properties are given in Table *10.1.* The table shows Macaulay's duration (and modified duration, defined in *(10.1.12)* below, in parentheses) for bonds with yields and coupon

[9]Macaulay also suggests that one could use yields on discount bonds rather than the yield on the coupon bond to calculate the present value of each coupon payment. However this approach requires that one measure a complete zero-coupon term structure.

Table 10.1. *Macaulay's and modified duration for selected bonds.*

		Maturity (years) 1	2	5	10	30	∞
Coupon rate	0%						
Yield	0%	1.000	2.000	5.000	10.000	30.000	—
		(1.000)	(2.000)	(5.000)	(10.000)	(30.000)	
	5%	1.000	2.000	5.000	10.000	30.000	—
		(0.976)	(1.951)	(4.878)	(9.756)	(29.268)	
	10%	1.000	2.000	5.000	10.000	30.000	—
		(0.952)	(1.905)	(4.762)	(9.524)	(28.571)	
Coupon rate	5%						
Yield	0%	0.988	1.932	4.550	8.417	21.150	—
		(0.988)	(1.932)	(4.550)	(8.417)	(21.150)	
	5%	0.988	1.928	4.485	7.989	15.841	20.500
		(0.964)	(1.881)	(4.376)	(7.795)	(15.454)	(20.000)
	10%	0.988	1.924	4.414	7.489	10.957	10.500
		(0.940)	(1.832)	(4.204)	(7.132)	(10.436)	(10.000)
Coupon rate	10%						
Yield	0%	0.977	1.875	4.250	7.625	18.938	—
		(0.977)	(1.875)	(4.250)	(7.625)	(18.938)	
	5%	0.977	1.868	4.156	7.107	14.025	20.500
		(0.953)	(1.823)	(4.054)	(6.933)	(13.683)	(20.000)
	10%	0.976	1.862	4.054	6.543	9.938	10.500
		(0.930)	(1.773)	(3.861)	(6.231)	(9.465)	(10.000)

The table reports Macaulay's duration and, in parentheses, modified duration for bonds with selected yields and maturities. Duration, yield, and maturity are stated in annual units but the underlying calculations assume that bond payments are made at six-month intervals.

rates of 0%, 5%, and 10%, and maturities ranging from one year to infinity. Duration is given in years but is calculated using six-month periods as would be appropriate for US Treasury bonds.

If we take the derivative of (10.1.9) with respect to Y_{cnt}, or equivalently with respect to $(1+Y_{cnt})$, we find that Macaulay's duration has another very

important property. It is the negative of the elasticity of a coupon bond's price with respect to its gross yield $(1+Y_{cnt})$:[10]

$$D_{cnt} = -\frac{dP_{cnt}}{d(1+Y_{cnt})}\frac{(1+Y_{cnt})}{P_{cnt}}. \tag{10.1.11}$$

In industry applications, Macaulay's duration is often divided by the gross yield $(1+Y_{cnt})$ to get what is called *modified duration:*

$$\frac{D_{cnt}}{(1+Y_{cnt})} = -\frac{dP_{cnt}}{dY_{cnt}}\frac{1}{P_{cnt}}. \tag{10.1.12}$$

Modified duration measures the proportional sensitivity of a bond's price to a small absolute change in its yield. Thus if modified duration is 10, an increase in the yield of 1 basis point (say from 3.00% to 3.01%) will cause a 10 basis point or 0.10% drop in the bond price.[11]

Macaulay's duration and modified duration are sometimes used to answer the following question: What single coupon bond best approximates the return on a zero-coupon bond with a given maturity? This question is of practical interest because many financial intermediaries have long-term zero-coupon liabilities, such as pension obligations, and they may wish to match or *immunize* these liabilities with coupon-bearing Treasury bonds.[12] Although today stripped zero-coupon Treasury bonds are available, they may be unattractive because of tax clientele and liquidity effects, so the immunization problem remains relevant. If there is a *parallel* shift in the yield curve so that bond yields of all maturities move by the same amount, then a change in the zero-coupon yield is accompanied by an equal change in the coupon bond yield. In this case equation (10.1.11) shows that a coupon bond whose Macaulay duration equals the maturity of the zero-coupon liability (equivalently, a coupon bond whose modified duration equals the modified duration of the zero-coupon liability) has, to a first-order approximation, the same return as the zero-coupon liability. This bond—or any portfolio of bonds with the same duration—solves the immunization problem for small, parallel shifts in the term structure.

Although this approach is attractively simple, there are several reasons why it must be used with caution. First, it assumes that yields of all maturities move by the same amount, in a parallel shift of the term structure. We

[10]The elasticity of a variable B with respect to a variable A is defined to be the derivative of B with respect to A, times A/B: $(dB/dA)(A/B)$. Equivalently, it is the derivative of $\log(B)$ with respect to $\log(A)$.

[11]Note that if duration is measured in six-month time units, then yields should be measured on a six-month basis. One can convert to an annual basis by halving duration and doubling yields. The numbers in Table 10.1 have been annualized in this way.

[12]Immunization was originally defined by Reddington (1952) as "the investment of the assets in such a way that the existing business is immune to a general change in the rate of interest". Fabozzi and Fabozzi (1995), Chapter 42, gives a comprehensive discussion.

show in Section 10.2.1 that historically, movements in short-term interest rates have tended to be larger than movements in longer-term bond yields. Some modified approaches have been developed to handle the more realistic case where short yields move more than long yields, so that there are nonparallel shifts in the term structure (see Bierwag, Kaufman, and Toevs [1983], Granito [1984], Ingersoll, Skelton, and Weil [1978]).

Second, (10.1.11) and (10.1.12) give first-order derivatives so they apply only to infinitesimally small changes in yields. Figure 10.4 illustrates the fact that the relationship between the log price and the yield on a bond is convex rather than linear. The slope of this relationship, modified duration, increases as yields fall (a fact shown also in Table 10.1). This may be taken into account by using a second—rder derivative. The *convexity* of a bond is defined as

$$\text{Convexity} \equiv \frac{\partial^2 P_{cnt}}{\partial Y_{cnt}^2}\frac{1}{P_{cnt}} = \frac{C\sum_{i=1}^{n}\frac{i(i+1)}{(1+Y_{cnt})^{i+2}} + \frac{n(n+1)}{(1+Y_{cnt})^{n+2}}}{P_{cnt}}, \tag{10.1.13}$$

and convexity can be used in a second-order Taylor series approximation of the price impact of a change in yield:

$$\begin{aligned}\frac{dP_{cn}(Y_{cn})}{P_{cn}} &\approx \frac{dP_{cn}}{dY_{cn}}\frac{1}{P_{cn}}\,dY_{cn} + \frac{1}{2}\frac{d^2P_{cn}}{dY_{cn}^2}\frac{1}{P_{cn}}(dY_{cn})^2 \\ &= (-\text{ modified duration})\ dY_{cn} \\ &\quad + \frac{1}{2}\text{ convexity }(dY_{cn})^2. \end{aligned} \tag{10.1.14}$$

Finally, both Macaulay's duration and modified duration assume that cash flows are fixed and do not change when interest rates change. This assumption is appropriate for Treasury securities but not for callable securities such as corporate bonds or mortgage-backed securities, or for securities with default risk if the probability of default varies with the level of interest rates. By modelling the way in which cash flows vary with interest rates, it is possible to calculate the sensitivity of prices to interest rates for these more complicated securities; this sensitivity is known as *effective duration.*[13]

A Loglinear Model for Coupon Bonds

The idea of duration has also been used in the academic literature to find approximate linear relationships between log coupon bond yields, holding-period returns, and forward rates that are analogous to the exact relationships for zero-coupon bonds. To understand this approach, start from the

[13]See Fabozzi and Fabozzi (1995), Chapters 28–30, and Fabozzi (1996) for a discussion of various methods used by fixed-income analysts to calculate effective duration.

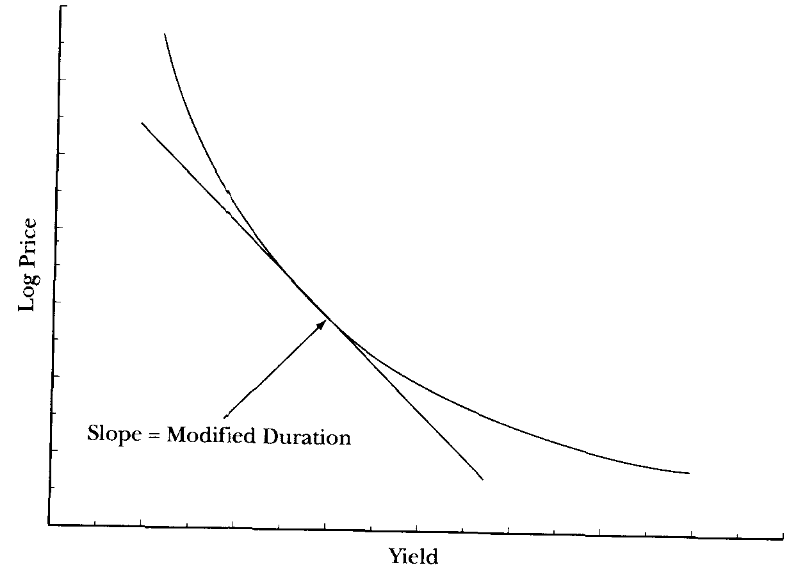

Figure 10.4. *The Price-Yield Relationship*

loglinear approximate return formula (7.1.19) derived in Chapter 7, and apply it to the one-period return $r_{c,n,t+1}$ on an n-period coupon bond:

$$r_{c,n,t+1} \approx k + \rho p_{c,n-1,t+1} + (1-\rho)c - p_{cnt}. \tag{10.1.15}$$

Here the log nominal coupon c plays the role of the dividend on stock, but of course it is fixed rather than random. The parameters p and k are given by $\rho \equiv 1/(1+\exp(\overline{c-p}))$ and $k \equiv -\log(\rho) - (1-\text{p})\log(1/\rho - 1)$. When the bond is selling at par, then its price is \$1 so its log price is zero and $\rho = 1/(1+C) = (1+Y_{cnt})^{-1}$. It is standard to use this value for p, which gives a good approximation for returns on bonds selling close to par.

One can treat (10.1.15), like the analogous zero-coupon expression (10.1.5), as a difference equation in the log bond price. Solving forward to the maturity date one obtains

$$p_{cnt} = \sum_{i=0}^{n-1}\rho^i[k + (1-\rho)c - r_{c,n-i,t+1+i}]. \tag{10.1.16}$$

This equation relates the price of a coupon bond to its stream of coupon payments and the future returns on the bond. A similar approximation of the log yield to maturity y_{cnt} shows that it satisfies an equation of the same form:

$$p_{cnt} \approx \sum_{i=0}^{n-1} \rho^i [k + (1-\rho)c - y_{cnt}]$$
$$= \frac{(1-\rho^n)}{(1-\rho)} [k + (1-\rho)c - y_{cnt}]. \qquad (10.1.17)$$

Equations (10.1.16) and (10.1.17) together imply that the n-period coupon bond yield satisfies $y_{cnt} \approx ((1-\rho)/(1-\rho^n)) \sum_{i=0}^{n-1} \rho^i r_{c,n-i,t+1+i}$. Thus although there is no exact relationship there is an approximate equality between the log yield to maturity on a coupon bond and a weighted average of the returns on the bond when it is held to maturity.

Equation (10.1.11) tells us that Macaulay's duration for a coupon bond is the derivative of its log price with respect to its log yield. Equation (10.1.17) gives this derivative as

$$D_{cn} \approx \frac{(1-\rho^n)}{(1-\rho)} = \frac{1-(1+Y_{cnt})^{-n}}{1-(1+Y_{cnt})^{-1}}, \qquad (10.1.18)$$

where the second equality uses $\rho = (1+Y_{cnt})^{-1}$. As noted above, this relation between duration and yield holds exactly for a bond selling at par.

Substituting (10.1.17) and (10.1.18) into (10.1.15), we obtain a loglinear relation between holding-period returns and yields for coupon bonds:

$$r_{c,n,t+1} \approx D_{cn}\, y_{cnt} - (D_{cn}-1) y_{c,n-1,t+1}. \qquad (10.1.19)$$

This equation was first derived by Shiller, Campbell, and Schoenholtz (1983).[14] It is analogous to (10.1.5) for zero-coupon bonds; maturity in that equation is replaced by duration here, and of course the two equations are consistent with one another for a zero-coupon bond whose duration equals its maturity.

A similar analysis for forward rates shows that the n-period-ahead 1-period forward rate implicit in the coupon-bearing term structure is

$$f_{nt} \approx \frac{D_{c,n+1}\, y_{c,n+1,t} - D_{cn}\, y_{cnt}}{D_{c,n+1} - D_{cn}} \qquad (10.1.20)$$

This formula, which is also due to Shiller, Campbell, and Schoenholtz (1983), reduces to the discount bond formula (10.1.8) when duration equals maturity.

[14] **Shiller, Campbell, and Schoenholtz use Y_{cnt} instead of y_{cnt}, but these are equivalent to the same first-order approximation used to derive (10.1.19). They also derive formulas relating multiperiod holding returns to yields.**

10.1.3 Estimating the Zero-Coupon Term Structure

The classic immunization problem is that of finding a coupon bond or portfolio of coupon bonds whose return has the same sensitivity to small interest-rate movements as the return on a given zerocoupon bond. Alternatively, one can try to find a portfolio of coupon bonds whose cash flows exactly match those of a given zero-coupon bond. In general, this portfolio will involve shortselling some bonds. This procedure has academic interest as well; one can extract an implied zero-coupon term structure from the coupon term structure.

If the complete zerocoupon term structure—that is, the prices of discount bonds $P_1 \dots P_n$ maturing at each coupon date—is known, then it is easy to find the price of a coupon bond as

$$P_{cn} = P_1 C + P_2 C + \cdots + P_n(1+C). \qquad (10.1.21)$$

Time subscripts are omitted here and throughout this section to economize on notation.

Similarly, if a complete coupon term structure—the prices of coupon bonds $P_{c1} \dots P_{cn}$ maturing at each coupon date—is available, then (10.1.21) can be used to back out the implied zero-coupon term structure. Starting with a one-period coupon bond, $P_{c1} = P_1(1+C)$ so $P_1 = P_{c1}/(1+C)$. We can then proceed iteratively. Given discount bond prices $P_1, \dots, P_{n-1}$, we can find P_n as

$$P_n = \frac{P_{cn} - P_{n-1}C - \cdots - P_1 C}{1+C}. \qquad (10.1.22)$$

Sometimes the coupon term structure may be more-than-complete in the sense that at least one coupon bond matures on each coupon date and several coupon bonds mature on some coupon dates. In this case (10.1.21) restricts the prices of some coupon bonds to be exact functions of the prices of other coupon bonds. Such restrictions are unlikely to hold in practice because of tax effects and other market frictions. To handle this Carleton and Cooper (1976) suggest adding a bond-specific error term to (10.1.21) and estimating it as a cross-sectional regression with all the bonds outstanding at a particular date. If these bonds are indexed $i = 1 \dots I$, then the regression is

$$P_{c_i n_i} = P_1 C_i + P_2 C_i + \cdots + P_{n_i}(1+C_i) + u_i, \qquad i = 1 \dots I, \qquad (10.1.23)$$

where C_i is the coupon on the ith bond and n_i is the maturity of the ith bond. The regressors are coupon payments at different dates, and the coefficients are the discount bond prices P_j, $j = 1 \dots N$, where $N = \max_i n_i$ is the longest coupon bond maturity. The system can be estimated by OLS provided that the coupon term structure is complete and that $I \geq N$.

Spline Estimation

In practice the term structure of coupon bonds is usually incomplete, and this means that the coefficients in (10.1.23) are not identified without imposing further restrictions. It seems natural to impose that the prices of discount bonds should vary smoothly with maturity. McCulloch (1971, 1975) suggests that a convenient way to do this is to write P_n, regarded as a function of maturity $P(n)$, as a linear combination of certain prespecified functions:

$$P_n = P(n) = 1 + \sum_{j=1}^{J} a_j f_j(n). \tag{10.1.24}$$

McCulloch calls $P(n)$ the discount function. The $f_j(n)$ in (10.1.24) are known functions of maturity n, and the a_j are coefficients to be estimated. Since $P(0) = 1$, we must have $f_j(0) = 0$ for all **j**.

Substituting (10.1.24) into (10.1.23) and rearranging, we obtain a regression equation

$$\Pi_i = \sum_{j=1}^{J} a_j X_{ij} + u_i, \qquad i = 1 \ldots I, \tag{10.1.25}$$

where $\Pi_i \equiv P_{c_i n_i} - 1 - C_i n_i$, the difference between the coupon bond price and the undiscounted value of its future payments, and $X_{ij} \equiv f_j(n_i) + C_i \sum_{l=1}^{n_i} f_j(l)$. Like equation (10.1.23), this equation can be estimated by OLS, but there are now only **J** coefficients rather than N.[15]

A key question is how to specify the functions $f_j(n)$ in (10.1.24). One simple possibility is to make $P(n)$, the discount function, a polynomial. To do this one sets $f_j(n) = n^j$. Although a sufficiently high-order polynomial can approximate any function, in practice one may want to use more parameters to fit the discount function at some maturities rather than others. For example one may want a more flexible approximation in maturity ranges where many bonds are traded.

To meet this need McCulloch suggests that $P(n)$ should be a spline function.[16] An rth-order spline, defined over some finite interval, is a piecewise rth-order polynomial with r − 1 continuous derivatives; its rth derivative is a step function. The points where the rth derivative changes discontinuously (including the points at the beginning and end of the interval over which the spline is defined) are known as knot points. If there are K knot points, there are K − 1 subintervals in each of which the spline is a polynomial. The spline has K − 2 + r free parameters, r for the first subinterval and 1 (that determines the unrestricted rth derivative) for each of the $K-2$ following subintervals. McCulloch suggests that the knot points should be chosen so that each subinterval contains an equal number of bond maturity dates.

If forward rates are to be continuous, the discount function must have at least one continuous derivative. Hence a quadratic spline, estimated by McCulloch (1971), is the lowest-order spline that can fit the discount function. If we require that the forward-rate curve should also be continuously differentiable, then we need to use a cubic spline, estimated by McCulloch (1975) and others. McCulloch's papers give the rather complicated formulas for the functions $f_j(n)$ that make $P(n)$ a quadratic or cubic spline.[17]

Tax Effects

OLS estimation of (10.1.25) chooses the parameters a_j so that the bond pricing errors u_i are uncorrelated with the variables X_{ij} that define the discount function. If a sufficiently flexible spline is used, then the pricing errors will be uncorrelated with maturity or any nonlinear function of maturity. Pricing errors may, however, be correlated with the coupon rate which is the other defining characteristic of a bond. Indeed McCulloch (1971) found that his model tended to underprice bonds that were selling below par because of their low coupon rates.

McCulloch (1975) attributes this to a tax effect. US Treasury bond coupons are taxed as ordinary income while price appreciation on a coupon-bearing bond purchased at a discount is taxed as capital gains. If the capital gains tax rate τ_g is less than the ordinary income tax rate τ (as has often been the case historically), then this can explain a price premium on bonds selling below par. For an investor who holds a bond to maturity the pricing formula (10.1.21) should be modified to

$$P_{cn} = [1 - \tau_g(1 - P_{cn})]P(n) + (1 - \tau)C \sum_{i=1}^{n} P(i). \tag{10.1.26}$$

The spline approach can be modified to handle tax effects like that in (10.1.26), at the cost of some additional complexity in estimation. Once tax effects are included, coupon bond prices must be used to construct the variables X_{ij} on the right-hand side of (10.1.25). This means that the bond pricing errors are correlated with the regressors so the equation must be

[15] The bond pricing errors are unlikely to be homoskedastic. McCulloch argues that the standard deviation of u_i is proportional to the bid-ask spread for bond i, and thus weights each observation by the reciprocal of its spread. This is not required for consistency, but may improve the efficiency of the estimates.

[16] Suits, Mason, and Chan (1978) give an accessible introduction to spline methodology.

[17] Adams and van Deventer (1994) argue for the use of a fourth-order spline, with the cubic term omitted, in order to maximize the 'smoothness" of the forward-rate curve, where smoothness is defined to be minus the average squared second derivative of the forward-rate curve with respect to maturity.

estimated by instrumental variables rather than simple OLS. Litzenberger and Rolfo (1984) apply a tax-adjusted spline model of this sort to bond market data from several different countries.

The tax-adjusted spline model assumes that the same tax rates are relevant for all bonds. The model cannot handle "clientele" effects, in which differently taxed investors specialize in different bonds. Schaefer (1981, 1982) suggests that clientele effects can be handled by first finding a set of taxefficient bonds for an investorin a particular tax bracket, then estimating an implied zero-coupon yield curve from those bonds alone.

Nonlinear Models

Despite the flexibility of the spline approach, spline functions have some unappealing properties. First, since splines are polynomials they imply a discount function which diverges as maturity increases rather than going to zero as required by theory. Implied forward rates also diverge rather than converging to any fixed limit. Second, there is no simple way to ensure that the discount function always declines with maturity (i.e., that all forward rates are positive). The forward curve illustrated in Figure 10.1 goes negative at a maturity of 27 years, and this behavior is not uncommon (see Shea [1984]). These problems are related to the fact that a flat zero-coupon yield curve implies an exponentially declining discount function, which is not easily approximated by a polynomial function. Since any plausible yield curve flattens out at the long end, splines are likely to have difficulties with longer-maturity bonds.

These difficulties have led some authors to suggest nonlinear alternatives to the linear specification (10.1.24). One alternative, suggested by Vasicek and Fong (1982), is to use an ***exponential spline,*** a spline applied to a negative exponential transformation of maturity. The exponential spline has the desirable property that forward rates and zero-coupon yields converge to a fixed limit as maturity increases. More generally, a flat yield curve is easy to fit with an exponential spline.

Although the exponential spline is appealing in theory, it is not clear that it performs better than the standard spline in practice (see Shea [1985]). The exponential spline does not make it easier to restrict forward rates to be positive. As for its long-maturity behavior, it is important to remember that forward rates cannot be directly estimated beyond the maturity of the longest coupon bond; they can only be identified by restricting the relation between long-horizon forward rates and shorter-horizon forward rates. The exponential spline, like the standard spline, fits the observed maturity range flexibly, leaving the limiting forward rate and speed of convergence to this rate to be determined more by the restrictions of the spline than by any characteristics of the long-horizon data. Since the exponential spline involves nonlinear estimation of a parameter used to transform maturity, it is more difficult to use than the standard spline and this cost may outweigh the exponential spline's desirable long-horizon properties. In any case, forward rate and yield curves should be treated with caution if they are extrapolated beyond the maturity of the longest traded bond.

Some other authors have solved the problem of negative forward rates by restricting the shape of the zero-coupon yield curve. Nelson and Siegel (1987), for example, model the instantaneous forward rate at maturity n as the solution to a second-order differential equation with equal roots: $f(n) = \beta_0 + \beta_1 \exp(-\alpha n) + \alpha n \beta_2 \exp(-\alpha n)$. This implies that the discount function is double-exponential:

$$P(n) = \exp[-\beta_0 n + (\beta_1 + \beta_2)(1 - \exp(-\alpha n))/\alpha - n\beta_2 \exp(-\alpha n)]$$

This specification generates forward-rate and yield curves with a desirable range of shapes, iqcluding upward-sloping, inverted, and hump-shaped. Svensson (1994) has developed this specificationfurther. Other recent work has generated bond-price formulas from fully specified generalequilibrium models of the term structure, which we discuss in Chapter 11.

10.2 Interpreting the Term Structure of Interest Rates

There is a large empirical literature which tests statements about expected-return relationships among bonds without deriving these statements from a fully specified equilibrium model. For simplicity we discuss this literature assuming that zero-coupon bond prices are observed or can be estimated from coupon bond prices.

10.2.1 The Expectations Hypothesis

The most popular simple model of the term structure is known as the ***expectations hypothesis.*** We distinguish the ***pure expectations hypothesis (PEH)*** (PEH), which says that expected excess returns on long-term over short-term bonds are zero, from the ***expectations hypothesis*** (EH), which says that expected excess returns are constant over time. This terminology is due to Lutz (1940).

Different Forms of the Pure Expectations Hypothesis

We also distinguish different forms of the PEH, according to the time horizon over which expected excess returns are zero. A first form of the PEH equates the one-period expected returns on one-period and n-period bonds. The oneperiod return on a one-period bond is known in advance to be $(1 + Y_{1t})$, so this form of the PEH implies

$$(1 + Y_{1t}) = \mathrm{E}_t[1 + R_{n,t+1}] = (1 + Y_{nt})^n E_t\left[(1 + Y_{n-1,t+1})^{-(n-1)}\right], \quad (10.2.1)$$

where the second equality follows from the definition of holding-period return and the fact that $(1+Y_{nt})$ is known at time t.

A second form of the PEH equates the n-period expected returns on one-period and n-period bonds:

$$(1+Y_{nt})^n = \mathrm{E}_t\left[(1+Y_{1t})(1+Y_{1,t+1})\ldots(1+Y_{1,t+n-1})\right]. \qquad (10.2.2)$$

Here $(1+Y_{nt})^n$ is the n-period return on an n-period bond, which equals the expected return from rolling over one-period bonds for n periods. It is straightforward to show that if (10.2.2) holds for all n, it implies

$$1+F_{n-1,t} \equiv \frac{(1+Y_{nt})^n}{(1+Y_{n-1,t})^{n-1}} = \mathrm{E}_t[1+Y_{1,t+n-1}]. \qquad (10.2.3)$$

Under this form of the PEH, the (n – 1)-period-ahead one-period forward rate equals the expected (n – 1)-period-ahead spot rate.

It is also straightforward to show that if (10.2.2) holds for all n, it implies

$$(1+Y_{nt})^n = (1+Y_{1t})\,\mathrm{E}_t\left[(1+Y_{n-1,t+1})^{n-1}\right]. \qquad (10.2.4)$$

But (10.2.4) is inconsistent with (10.2.1) whenever interest rates are random. The problem is that by Jensen's Inequality, the expectation of the reciprocal of a random variable is not the reciprocal of the expectation of that random variable. Thus the pure expectations hypothesis cannot hold in both its one-period form and its n-period form.[18]

One can understand this problem more clearly by assuming that interest rates are lognormal and homoskedastic and taking logs of the one-period PEH equation (10.2.1) and the n-period PEH equation (10.2.4). Noting that from equation (10.1.5) the excess one-period log return on an n-period bond is

$$r_{n,t+1} - y_{1t} = (y_{nt} - y_{1t}) - (n-1)(y_{n-1,t+1} - y_{nt}), \qquad (10.2.5)$$

equation (10.2.1) implies that

$$\mathrm{E}[r_{n,t+1} - y_{1t}] = -(1/2)\,\mathrm{Var}[r_{n,t+1} - y_{1t}], \qquad (10.2.6)$$

while (10.2.4) implies that

$$\mathrm{E}[r_{n,t+1} - y_{1t}] = (1/2)\,\mathrm{Var}[r_{n,t+1} - y_{1t}]. \qquad (10.2.7)$$

The difference between the right-hand sides of (10.2.6) and (10.2.7) is

[18]Cox, Ingersoll, and Ross (1981a) make this point very clearly. They also argue that in continuous time, only expected equality of instantaneous returns (a model corresponding to (10.2.1)) is consistent with the absence of arbitrage. But McCulloch (1993) has shown that this result depends on restrictive assumptions and does not hold in general.

Table 10.2. *Means and standard deviations of term-structure variables.*

Variable	Long bond maturity (n)						
	2	3	6	12	24	48	120
Excess return	0.385	0.564	0.848	0.917	0.709	0.614	−0.048
$r_{n,t+1} - y_{1t}$	(0.644)	(1.222)	(2.954)	(6.218)	(11.33)	(19.40)	(37.08)
Change in yield	0.010	0.010	0.010	0.010	0.011	0.011	0.012
$y_{n,t+1} - y_{nt}$	(0.592)	(0.576)	(0.570)	(0.547)	(0.488)	(0.410)	(0.310)
Change in yield	−0.188	−0.119	−0.056	−0.014	0.011	0.011	0.012
$y_{n-1,t+1} - y_{nt}$	(0.608)	(0.586)	(0.573)	(0.555)	(0.488)	(0.410)	(0.310)
Yield spread	0.197	0.326	0.570	0.765	0.958	1.153	1.367
$y_{nt} - y_{1t}$	(0.212)	(0.303)	(0.438)	(0.594)	(0.797)	(1.012)	(1.237)

Long bond maturities are measured in months. For each variable the table reports the sample mean and sample standard deviation (in parentheses) using monthly data over the period 1952:1-1991:2. The units are annualized percentage points. The underlying data are zero-coupon bond yields from McCulloch and Kwon (1993).

$\mathrm{Var}[r_{n,t+1} - y_{1t}]$, which measures the quantitative importance of the Jensen's Inequality effect in a lognormal homoskedastic model.

Table 10.2 reports unconditional sample means and standard deviations for several term-structure variables over the period 1952:1 to 1991:2.[19] All data are monthly, but are measured in annualized percentage points; that is, the raw variables are multiplied by 1200. The first row shows the mean and standard deviation of excess returns on n-month zero-coupon bonds over one-month bills. The mean excess return is positive and rising with maturity at first, but it starts to fall at a maturity of one year and is even slightly negative for ten-year zero-coupon bonds.

This pattern can be understood by breaking excess returns into the two components on the right-hand side of equation (10.2.5): the yield spread $(y_{nt} - y_{1t})$ between n-period and one-period bonds, and $-(n-1)$ times the change in yield $(y_{n-1,t+1} - y_{nt})$ on the n-period bond. Interest rates of all fixed maturities rise during the sample period, as shown in the second row of Table 10.2 and illustrated for one-month and ten-year rates in Figure 10.5. At the short end of the term structure this effect is offset by the decline in maturity from n to n – 1 as the bond is held for one month; thus the change

[19]Table 10.2 is an expanded version of a table shown in Campbell (1995). The numbers given here are slightly different from the numbers in that paper because the sample period used in that paper was 1951:1 to 1990:2, although it was erroneously reported to be 1952:1 to 1991:2.

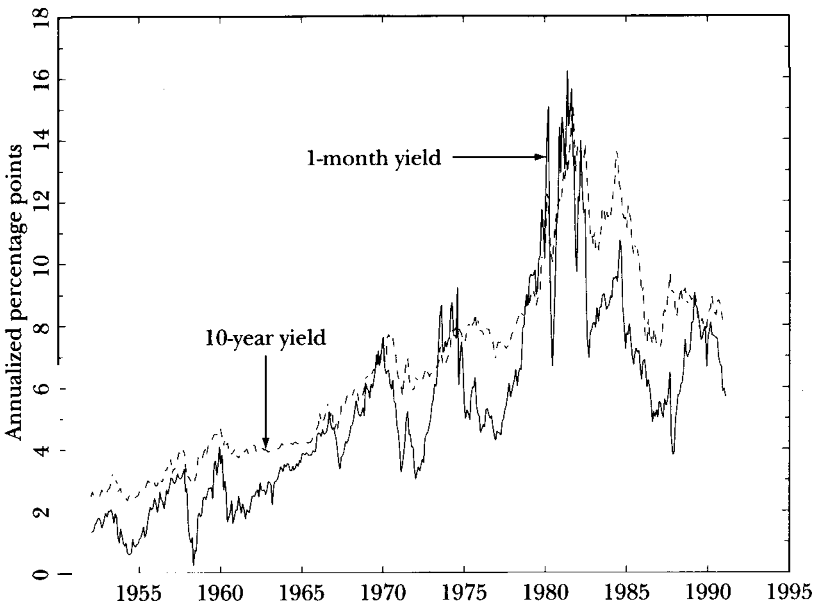

Figure 10.5. Short- and *Long-Term* Interest Rates *1952* to *1991*

in yield $(y_{n-1,t+1} - y_{nt})$, shown in the third row of Table 10.2, is negative for short bonds, contributing positively to their return.[20] At the long end of the term structure, however, the decline in maturity from n to n−1 is negligible, and so the change in yield $(y_{n-1,t+1} - y_{nt})$ is positive, causing capital losses on long zero-coupon bonds which outweigh the higher yields offered by these bonds, shown in the fourth row of Table 10.2.

The standard deviation of excess returns rises rapidly with maturity. If excess bond returns are white noise, then the standard error of the sample mean is the standard deviation divided by the square root of the sample size (469 months). The standard error for n = 2 is only 0.03%, whereas the standard error for n = 120 is 1.71%. Thus the pattern of mean returns is imprecisely estimated at long maturities.

The standard deviation of excess returns also determines the size of the wedge between the one-period and n-period forms of the pure expectations hypothesis. The difference between mean annualized excess returns under (10.2.6) and (10.2.7) is only 0.0003% for n = 2. It is still only 0.11% for

[20]**Investors who seek to profit from this tendency of bond yields to fall as maturity shrinks are said to be "riding the yield curve."**

n = 24. But it rises to 1.15% for n = 120. This calculation shows that the differences between different forms of the PEH are small except for very long-maturity zero-coupon bonds. Since these bonds have the most imprecisely estimated mean returns, the data reject all forms of the PEH at the short end of the term structure, but reject no forms of the PEH at the long end of the term structure. In this sense the distinction between different forms of the PEH is not critical for evaluating this hypothesis.

Most empirical research uses neither the one-period form of the PEH (10.2.6), nor the n-period form (10.2.7), but a log form of the PEH that equates the expected log returns on bonds of all maturities:

$$\mathrm{E}[r_{n,t+1} - y_{1t}] = 0. \tag{10.2.8}$$

This model is halfway between equations (10.2.6) and (10.2.7) and can be justified as an approximation to either of them when variance terms are small. Alternatively, it can be derived directly as in McCulloch (1993).

Implications of the Log Pure Expectations Hypothesis

Once the PEH is formulated in logs, it is comparatively easy to state its implications for longer-term bonds. The log PEH implies, first, that the one-period log yield (which is the same as the one-period return on a one-period bond) should equal the expected log holding return on a longer n-period bond held for one period:

$$y_{1t} = \mathrm{E}_t[r_{n,t+1}]. \tag{10.2.9}$$

Second, a long-term n-period log yield should equal the expected sum of n successive log yields on one-period bonds which are rolled over for n periods:

$$y_{nt} = (1/n)\sum_{i=0}^{n-1} \mathrm{E}_t[y_{1,t+i}]. \tag{10.2.10}$$

Finally, the $(n-1)$-period-ahead one-period log forward rate should equal the expected one-period spot rate (n − 1) periods ahead:

$$f_{n-1,t} = \mathrm{E}_t[y_{1,t+n-1}]. \tag{10.2.11}$$

This implies that the log forward rate for a one-period investment to be made at a particular date in the future should follow a martingale:

$$f_{nt} = \mathrm{E}_t[y_{1,t+n}] = \mathrm{E}_t\left[\mathrm{E}_{t+1}[y_{1,t+n}]\right] = \mathrm{E}_t[f_{n-1,t+1}]. \tag{10.2.12}$$

any of equations (10.2.9), (10.2.10), and (10.2.11) hold for all n and t, en the other equations also hold for all n and t. Also, if any of these equations hold for $n = 2$ at some date t, then the other equations also hold for $n = 2$ and the same date t. Note however that (10.2.9)–(10.2.11) are not generally equivalent for particular n and t.

Alternatives to the Pure Expectations Hypothesis

The expectations hypothesis (EH) is more general than the PEH in that it allows the expected returns on bonds of different maturities to differ by constants, which can depend on maturity but not on time. The differences between expected returns on bonds of different maturities are sometimes called *term premia*. The PEH says that term premia are zero, while the EH says that they are constant through time.[21] Like the PEH, the EH can be formulated for one-period simple returns, for n-period simple returns, or for log returns. If bond returns are lognormal and homoskedastic, as in Singleton (1990), then these formulations are consistent with one another because the Jensen's Inequality effects are constant over time. Recent empirical research typically concentrates on the log form of the EH.

Early discussions of the term structure tended to ignore the possibility that term premia might vary over time, concentrating instead on their sign. Hicks (1946) and Lutz (1940) argued that lenders prefer short maturities while borrowers prefer long maturities, so that long bonds should have higher average returns than short bonds. Modigliani and Sutch (1966) argued that different lenders and borrowers might have different *preferred habitats*, so that term premia might be negative as well as positive. All these authors disputed the PEH but did not explicitly question the EH. More recent work has used intertemporal asset pricing theory to derive both the average sign and the time-variation of term premia; we discuss this work in Chapter 11.

10.2.2 Yield Spreads and Interest Rate Forecasts

We now consider empirical evidence on the expectations hypothesis (EH). Since the EH allows constant differences in the expected returns on short- and long-term bonds, it does not restrict constant terms so for convenience we drop constants from all equations in this section.

So far we have stated the implications of the expectations hypothesis for the levels of nominal interest rates. In post-World War II US data, nominal interest rates seem to follow a highly persistent process with a root very close to unity, so much empirical work uses yield spreads instead of yield levels.[22]

[21]This usage is the most common one in the literature. Fama (1984), Fama (1990), and Fama and Bliss (1987), however, use "term premia" to refer to realized, rather than expected, excess returns on long-term bonds.

[22]See Chapters 2 and 7 for a discussion of unit roots. The persistence of the short-rate process is discussed further in Chapter 11.

Recall that the yield spread between the n-period yield and the one-period yield is $s_{nt} = y_{nt} - y_{1t}$. Equation (10.1.6) implies that

$$s_{nt} = \left(\frac{1}{n}\right) \mathrm{E}_t\left[\sum_{i=1}^{n}\left[(y_{1,t+i} - y_{1t}) + (r_{n+1-i,t+i} - y_{1,t+i})\right]\right]$$

$$= \left(\frac{1}{n}\right) \mathrm{E}_t\left[\sum_{i=1}^{n}\left[(n-i)\Delta y_{1,t+i} + (r_{n+1-i,t+i} - y_{1,t+i})\right]\right]. \quad (10.2.13)$$

The second equality in equation (10.2.13) replaces multiperiod interest rate changes by sums of single-period interest rate changes. The equation says that the yield spread equals a weighted average of expected future interest rate changes, plus an unweighted average of expected future excess returns on long bonds. If changes in interest rates are stationary (that is, if interest rates themselves have one unit root but not two), and if excess returns are stationary (as would be implied by any model in which risk aversion and bonds' risk characteristics are stationary), then the yield spread is also stationary. This means that yields of different maturities are *cointegrated*.[23]

The expectations hypothesis says that the second term on the right-hand side of (10.2.13) is constant. This has important implications for the relation between the yield spread and future interest rates. It means that the yield spread is (up to a constant) the optimal forecaster of the change in the long-bond yield over the life of the short bond, and the optimal forecaster of changes in short rates over the life of the long bond. Recalling that we have dropped all constant terms, the relations are

$$\left(\frac{1}{n-1}\right) s_{nt} = \mathrm{E}_t[y_{n-1,t+1} - y_{nt}], \quad (10.2.14)$$

and

$$s_{nt} = \mathrm{E}_t\left[\sum_{i=1}^{n-1}(1 - i/n)\Delta y_{1,t+i}\right]. \quad (10.2.15)$$

Equation (10.2.14) can be obtained by substituting the definition of $r_{n,t+1}$, (10.1.5), into (10.2.9) and rearranging. It shows that when the yield spread is high, the long rate is expected to rise. This is because a high yield spread gives the long bond a yield advantage which must be offset by an anticipated capital loss. Such a capital loss can only come about through an increase in the long-bond yield. Equation (10.2.15) follows directly from (10.2.13) with constant expected excess returns. It shows that when the yield spread is

[23]See Campbell and Shiller (1987) for a discussion of cointegration in the term structure of interest rates.

Table 10.3. *Regression coefficients $\hat{\beta}_n$ and $\hat{\gamma}_n$.*

Dependent variable	Long bond maturity (n)						
	2	3	6	12	24	48	120
Long-yield changes (10.2.16)	0.003	−0.145	−0.835	−1.435	−1.448	−2.262	−4.226
	(0.191)	(0.282)	(0.442)	(0.599)	(1.004)	(1.458)	(2.076)
Short-rate changes (10.2.18)	0.502	0.467	0.320	0.272	0.363	0.442	1.402
	(0.096)	(0.148)	(0.146)	(0.208)	(0.223)	(0.384)	(0.147)

Long bond maturities are measured in months. The first row reports the estimated regression coefficient $\hat{\beta}_n$ from (10.2.16), with an asymptotic standard error (in parentheses) calculated to allow for heteroskedasticity in the manner described in the Appendix. The second row reports the the estimated regression coefficient $\hat{\gamma}_n$ from (10.2.18), with an asymptotic standard error calculated in the same manner, allowing also for residual autocorrelation. The expectations hypothesis of the term structure implies that both $\hat{\beta}_n$ and $\hat{\gamma}_n$ should equal one. The underlying data are monthly zero-coupon bond yields over the period 1952:1 to 1991:2, from McCulloch and Kwon (1993).

high, short rates are expected to rise so that the average short rate over the life of the long bond equals the initial long-bond yield. Near-term increases in short rates are given greater weight than further-off increases, because they affect the level of short rates during a greater part of the life of the long bond.

Yield Spreads and Future Long Rates

Equation (10.2.14), which says that high yield spreads should forecast increases in long rates, fares poorly in the data. Macaulay (1938) first noted the fact that high yield spreads actually tend to precede decreases in long rates. He wrote: "The yields of bonds of the highest grade should *fall* during a period in which short-term rates are higher than the yields of the bonds and *rise* during a period in which short-term rates are lower. Now experience is more nearly the opposite" (Macaulay [1938, p. 33]).

Table **10.3** reports estimates of the coefficient β_n and its standard error in the regression

$$y_{n-1,t+1} - y_{n,t} \;=\; \alpha_n + \beta_n\left(\frac{s_{nt}}{n-1}\right) + \epsilon_{n,t}. \qquad (10.2.16)$$

The maturity n varies from **3** months to **120** months (**10** years).[24] According

[24]For maturities above one year the table uses the approximation $y_{n-1,t+1} \approx y_{n,t+1}$. Note that this is not the same as approximating $p_{n-1,t+1}$ by $p_{n,t+1}$. The numbers given differ slightly from those in Campbell (1995) because that paper uses the sample period 1951:1 to 1990:2, erroneously reported as 1952:1 to 1991:2.

to the expectations hypothesis, we should find $\beta_n = 1$. In fact all the estimates in Table **10.3** are negative; all are significantly less than one, and some are significantly less than zero. When the long-short yield spread is high the long yield tends to fall, amplifying the yield differential between long and short bonds, rather than rising to offset the yield differential as required by the expectations hypothesis.

The regression equation **(10.2.16)** contains the same information as a regression of the excess one-period return on an n-period bond onto the yield spread s_{nt}. Equation **(10.2.5)** relating excess returns to yields implies that the excess-return regression would have a coefficient of $(1-\beta_n)$. Thus the negative estimates of β_n in Table **10.3** correspond to a strong positive relationship between yield spreads and excess returns on long bonds. This is similar to the positive relationship between dividend yields and stock returns discussed in Chapter 7.[25]

One difficulty with the regression **(10.2.16)** is that it is particularly sensitive to measurement error in the long-term interest rate (see Stambaugh [1988]). Since the long rate appears both in the regressor with a positive sign and in the dependent variable with a negative sign, measurement error would tend to produce the negative signs found in Table **10.3.** Campbell and Shiller **(1991)** point out that this can be handled by using instrumental variables regression where the instruments are correlated with the yield spread but not with the bond yield measurement error. They try a variety of instruments and find that the negative regression coefficients are quite robust.

Yield Spreads and Future Short Rates

There is much more truth in proposition (10.2.15), that high yield spreads should forecast long-term increases in short rates. This can be tested either directly or indirectly. The direct approach is to form the *ex post* value of the short-rate changes that appear on the right-hand side of **(10.2.15)** and to regress this on the yield spread. We define

$$s^*_{nt} \;\equiv\; \sum_{i=1}^{n-1}(1 - i/n)\Delta y_{1,t+i}, \qquad (10.2.17)$$

[25]Campbell and Ammer (1993), Fama and French (1989), and Keim and Stambaugh (1986) show that yield spreads help to forecast excess returns on bonds as well as on other long-term assets. Campbell and Shiller (1991) and Shiller, Campbell, and Schoenholtz (1983) show that yield spreads tend to forecast declines in long-bond yields.

and run the regression

$$s^*_{nt} = \mu_n + \gamma_n s_{nt} + \epsilon_{nt}. \qquad (10.2.18)$$

The expectations hypothesis implies that $\gamma_n = 1$ for all n.[26]

Table *10.3* reports estimated $\hat{\gamma}_n$ coefficients with standard errors, correcting for heteroskedasticity and overlap in the equation errors in the manner discussed in the Appendix. The estimated coefficients have a U shape: For small n they are smaller than one but significantly positive; up to a year or so they decline with n, becoming insignificantly different from zero; beyond one year the coefficients increase and at ten years the coefficient is even significantly greater than one. Thus Table *10.3* shows that yield spreads have forecasting power for short-rate movements over a horizon of two or three months, and again over horizons of several years. Around one year, however, yield-spread variation seems almost unrelated to subsequent movements in short rates.

The regression equation *(10.2.18)* contains the same information as a regression of $(1/n)$ times the excess n-period return on an n-period bond onto the yield spread s_{nt}. The relation between excess returns and yields implies that the excess-return regression would have a coefficient of $(1-\gamma_n)$. Table *10.3* implies that yield spreads forecast excess returns out to horizons of several years, but the forecasting power diminishes towards ten years.

There are several econometric difficulties with the direct approach just described. First, one loses n periods of data at the end of the sample period. This can be quite serious: For example, the ten-year regression in Table *10.3* ends in *1981,* whereas the three-month regression ends in *1991.* This makes a substantial difference to the results, as discussed by Campbell and Shiller *(1991).* Second, the error term ϵ_{nt} is a moving average of order $(n-1)$, so standard errors must be corrected in the manner described in the Appendix. This can lead to finite-sample problems when $(n-1)$ is not small relative to the sample size. Third, the regressor is serially correlated and correlated with lags of the dependent variable, and this too can cause finite-sample problems (see Mankiw and Shapiro [1986], Richardson and Stock [1990], and Stambaugh [1986]).

Although these econometric problems are important, they do not seem to account for the U-shaped pattern of coefficients. Campbell and Shiller *(1991)* find similar results using a vector autoregressive (VAR) methodology like that described in Section *7.2.3* of Chapter *7*. They find that the long-term yield spread is highly correlated with an unrestricted VAR forecast of future short-rate movements, while the intermediate-term yield spread is much more weakly correlated with the VAR forecast.

[26] Fama (1984) and Shiller, Campbell, and Schoenholtz (1983) use this approach at the short end of the term structure, while Fama and Bliss (1987) extend it to the long end. Campbell and Shiller (1991) provide a comprehensive review.

To interpret Table *10.3*, it is helpful to return to equation *(10.2.13)* and rewrite it as

$$s_{nt} = sy_{nt} + sr_{nt}, \qquad (10.2.19)$$

where

$$sy_{nt} \equiv \mathrm{E}_t[s^*_{nt}] = \left(\frac{1}{n}\right)\mathrm{E}_t\left[\sum_{i=1}^{n}(n-i)\Delta y_{1,t+i}\right],$$

and

$$sr_{nt} \equiv \left(\frac{1}{n}\right)\mathrm{E}_t\left[\sum_{i=1}^{n}(r_{n+1-i,t+i} - y_{1,t+i})\right].$$

In general the yield spread is the sum of two components, one that forecasts interest rate changes (sy_{nt}) and one that forecasts excess returns on long bonds (sr_{nt}). This means that the regression coefficient γ_n in equation *(10.2.18)* is

$$\begin{aligned}\gamma_n &= \frac{\mathrm{Cov}[s^*_{nt}, s_{nt}]}{\mathrm{Var}[s_{nt}]}\\ &= \frac{\mathrm{Var}[sy_{nt}] + \mathrm{Cov}[sy_{nt}, sr_{nt}]}{\mathrm{Var}[sy_{nt}] + \mathrm{Var}[sr_{nt}] + 2\,\mathrm{Cov}[sy_{nt}, sr_{nt}]}.\end{aligned} \qquad (10.2.20)$$

For any given variance of excess-return forecasts sr_{nt}, as the variance of interest rate forecasts sy_{nt} goes to zero the coefficient γ_n goes to zero, but as the variance of sy_{nt} increases the coefficient γ_n goes to one. The U-shaped pattern of regression coefficients in Table *10.3* may be explained by reduced forecastability of interest rate movements at horizons around one year. There may be some short-run forecastability arising from Federal Reserve operating procedures, and some long-run forecastability arising from business-cycle effects on interest rates, but at a one-year horizon the Federal Reserve may smooth interest rates so that the variability of sy_{nt} is small. Balduzzi, Bertola, and Foresi (1993), Rudebusch (1995), and Roberts, Runkle, and Whiteman *(1996)* argue for this interpretation of the evidence. Consistent with this explanation, Mankiw and Miron *(1986)* show that the predictions of the expectations hypothesis fit the data better in periods when interest rate movements have been highly forecastable, such as the period mediately before the founding of the Federal Reserve System.

10.3 Conclusion

esults in Table *10.3* imply that naive investors, who judge bonds by their s to maturity and buy long bonds when their yields are relatively high, e tended to earn superior returns in the postwar period in the United

11
Term-Structure Models

THIS CHAPTER EXPLORES the large modern literature on fully specified general-equilibrium models of the term structure of interest rates. Much of this literature is set in continuous time, which simplifies some of the theoretical analysis but complicates empirical implementation. Since we focus on the econometric testing of the models and their empirical implications, we adopt a discrete-time approach; however we take care to relate all our results to their continuous-time equivalents. We follow the literature by first developing models for real bonds, but we discuss in some detail how these models can be used to price nominal bonds.

All the models in this chapter start from the general asset pricing condition introduced as (8.1.3) in Chapter 8: $1 = \mathrm{E}_t[(1+R_{i,t+1})M_{t+1}]$, where $R_{i,t+1}$ is the real return on some asset i and M_{t+1} is the stochastic discount factor. As we explained in Section 8.1 of Chapter 8, this condition implies that the expected return on any asset is negatively related to its covariance with the stochastic discount factor. In models with utility-maximizing investors, the stochastic discount factor measures the marginal utility of investors. Assets whose returns covary positively with the stochastic discount factor tend to pay off when marginal utility is high—they deliver wealth at times when wealth is most valuable to investors. Investors are willing to pay high prices and accept low returns on such assets.

Fixed-income securities are particularly easy to price using this framework. When cash flows are random, the stochastic properties of the cash flows help to determine the covariance of an asset's return with the stochastic discount factor. But a fixed-income security has deterministic cash flows, so it covaries with the stochastic discount factor only because there is time-variation in discount rates. This variation in discount rates is driven by the time-series behavior of the stochastic discount factor, so term-structure models are equivalent to time-series models for the stochastic discount factor.

From (10.1.4) in Chapter 10, we know that returns on n-period real zero-coupon bonds are related to real bond prices in a particularly simple way: $(1 + R_{n,t+1}) = P_{n-1,t+1}/P_{nt}$. Substituting this into (8.1.3), we find that the real price of an n-period real bond, P_{nt}, satisfies

$$P_{nt} = \mathrm{E}_t[P_{n-1,t+1} M_{t+1}]. \tag{11.0.1}$$

This equation lends itself to a recursive approach. We model P_{nt} as a function of those state variables that are relevant for forecasting the M_{t+1} process. Given that process and the function relating $P_{n-1,t}$ to state variables, we can calculate the function relating P_{nt} to state variables. We start the calculation by noting that $P_{0t} = 1$.

Equation (11.0.1) can also be solved forward to express the n-period bond price as the expected product of n stochastic discount factors:

$$P_{nt} = \mathrm{E}_t[M_{t+1} \dots M_{t+n}]. \tag{11.0.2}$$

Although we emphasize the recursive approach, in some models it is more convenient to work directly with (11.0.2).

Section 11.1 explores a class of simple models in which all relevant variables are conditionally lognormal and log bond yields are linear in state variables. These *affine-yield models* include all the most commonly used term-structure models. Section 11.2 shows how these models can be fit to nominal interest rate data, and reviews their strengths and weaknesses. One of the main uses of term-structure models is in pricing interest-rate derivative securities; we discuss this application in Section 11.3. We show how standard term-structure models can be modified so that they fit the current term structure exactly. We then use the models to price forwards, futures, and options on fixed-income securities.

11.1 Affine-Yield Models

To keep matters simple, we assume throughout this section that the distribution of the stochastic discount factor M_{t+1} is conditionally lognormal. We specify models in which bond prices are jointly lognormal with M_{t+1}. We can then take logs of (11.0.1) to obtain

$$p_{nt} = \mathrm{E}_t[m_{t+1} + p_{n-1,t+1}] + (1/2)\,\mathrm{Var}_t[m_{t+1} + p_{n-1,t+1}], \tag{11.1.1}$$

where as usual lowercase letters denote the logs of the corresponding uppercase letters so for example $m_{t+1} = \log(M_{t+1})$. This is the basic equation we shall use.

We begin with two models in which a single state variable forecasts the stochastic discount factor. Section 11.1.1 discusses the first model, in which m_{t+1} is homoskedastic, while Section 11.1.2 discusses the second model, in which the conditional variance of m_{t+1} changes over time. These are

discrete-time versions of the well-known models of Vasicek (1977) and Cox, Ingersoll, and Ross (1985a), respectively. Section 11.1.3 then considers a more general model with two state variables, a discrete-time version of the model of Longstaff and Schwartz (1992). All of these models have the property that log bond prices, and hence log bond yields, are linear or *affine* in the state variables. This ensures the desired joint lognormality of bond prices with the stochastic discount factor. Section 11.1.4 describes the general properties of these *affine-yield* models, and discusses some alternative modelling approaches.[1]

11.1.1 A Homoskedastic *Single-Factor* Model

It is convenient to work with the negative of the log stochastic discount factor, $-m_{t+1}$. Without loss of generality, this can be expressed as the sum of its one-period-ahead conditional expectation x_t and an innovation ϵ_{t+1}:

$$-m_{t+1} = x_t + \epsilon_{t+1}. \tag{11.1.2}$$

We assume that ϵ_{t+1} is normally distributed with constant variance.

Next we assume that x_{t+1} follows the simplest interesting time-series process, a univariate AR(1) process with mean μ and persistence ϕ. The shock to x_{t+1} is written ξ_{t+1}:

$$x_{t+1} = (1-\phi)\mu + \phi x_t + \xi_{t+1}. \tag{11.1.3}$$

The innovations to m_{t+1} and x_{t+1} may be correlated. To capture this, we write ϵ_{t+1} as

$$\epsilon_{t+1} = \beta\xi_{t+1} + \eta_{t+1}, \tag{11.1.4}$$

where ξ_{t+1} and η_{t+1} are normally distributed with constant variances and are uncorrelated with each other.

The presence of the uncorrelated shock η_{t+1} only affects the average level of the term structure and not its average slope or its time-series behavior. To simplify notation, we accordingly drop it and assume that $\epsilon_{t+1} = \beta\xi_{t+1}$. Equation (11.1.2) can then be rewritten as

$$-m_{t+1} = x_t + \beta\xi_{t+1}. \tag{11.1.5}$$

The innovation ξ_{t+1} is now the only shock in the system; accordingly we can write its variance simply as σ^2 without causing confusion.

Equations (11.1.5) and (11.1.3) imply that $-m_{t+1}$ can be written as an ARMA(1,1) process since it is the sum of an AR(1) process and white noise.

[1]Our discrete-time presentation follows Singleton (1990), Sun (1992), and especially Backus (1993). Sun (1992) explores the relation between discrete-time and continuous-time models in more detail.

In fact, $-m_{t+1}$ has the same structure as asset returns did in the example of Chapter 7, Section 7.1.4. As in that example, it is important to realize that $-m_{t+1}$ is not a univariate process even though its conditional expectation x_t is univariate. Thus the univariate autocorrelations of $-m_{t+1}$ do not tell us all we need to know for asset pricing; different sets of parameter values, with different implications for asset pricing, could be consistent with the same set of univariate autocorrelations for $-m_{t+1}$. For example, these autocorrelations could all be zero because $\alpha^2 = 0$, which would make interest rates constant, but they could also be zero for $\alpha^2 \neq 0$ if β takes on a particular value, and in this case interest rates would vary over time.

We can determine the price of a one-period bond by noting that when $n = 1$, $p_{n-1,t+1} = p_{0,t+1} = 0$, so the terms involving $p_{n-1,t+1}$ in equation (11.1.1) drop out. Substituting (11.1.5) and (11.1.3) into (11.1.1), we have

$$p_{1t} = \mathrm{E}_t[m_{t+1}] + (1/2)\mathrm{Var}_t[m_{t+1}] = -x_t + \beta^2\sigma^2/2. \tag{11.1.6}$$

The one-period bond yield $y_{1t} = -p_{1t}$, so

$$y_{1t} = x_t - \beta^2\sigma^2/2. \tag{11.1.7}$$

The short rate equals the state variable less a constant term, so it inherits the AR(1) dynamics of the state variable. Indeed, we can think of the short rate as measuring the state of the economy in this model. Note that there is nothing in equation (11.1.7) that rules out a negative short rate.

We now guess that the form of the price function for an n-period bond is

$$-p_{nt} = A_n + B_n x_t. \tag{11.1.8}$$

Since the n-period bond yield $y_{nt} = -p_{nt}/n$, we are guessing that the yield on a bond of any maturity is linear or *affine* in the state variable x_t (Brown and Schaefer [1991]). We already know that bond prices for $n = 0$ and $n = 1$ satisfy equation (11.1.8), with $A_0 = B_0 = 0$, $A_1 = -\beta^2\sigma^2/2$, and $B_1 = 1$. We proceed to verify our guess by showing that it is consistent with the pricing relation (11.1.1). At the same time we can derive recursive formulas for the coefficients A_n and B_n.

Our guess for the price function (11.1.8) implies that the two terms on the right-hand side of (11.1.1) are

$$\begin{aligned}\mathrm{E}_t[m_{t+1} + p_{n-1,t+1}] &= -x_t - A_{n-1} - B_{n-1}(1-\phi)\mu - B_{n-1}\phi\, x_t,\\ \mathrm{Var}_t[m_{t+1} + p_{n-1,t+1}] &= (\beta + B_{n-1})^2\sigma^2.\end{aligned} \tag{11.1.9}$$

Substituting (11.1.8) and (11.1.9) into (11.1.1), we get

$$\begin{aligned}A_n &+ B_n x_t - x_t - A_{n-1} - B_{n-1}(1-\phi)\mu - B_{n-1}\phi\, x_t\\ &+ (\beta + B_{n-1})^2\sigma^2/2 = 0.\end{aligned} \tag{11.1.10}$$

This must hold for any x_t, so the coefficients on x_t must sum to zero and the remaining coefficients must also sum to zero. This implies

$$\begin{aligned}B_n &= 1 + \phi B_{n-1} = (1-\phi^n)/(1-\phi),\\ A_n - A_{n-1} &= (1-\phi)\mu B_{n-1} - (\beta + B_{n-1})^2\sigma^2/2.\end{aligned} \tag{11.1.11}$$

We have now verified the guess (11.1.8), since with the coefficients in (11.1.11) the price function (11.1.8) satisfies the asset pricing equation (11.1.1) and its assumption that bond returns are conditionally lognormal.

Implications of the Homoskedastic Model

The homoskedastic bond pricing model has several interesting implications. First, the coefficient B_n measures the fall in the log price of an n-period bond when there is an increase in the state variable x_t or equivalently in the one-period interest rate y_{1t}. It therefore measures the sensitivity of the n-period bond return to the one-period interest rate. Equation (11.1.11) shows that the coefficient B_n follows a simple univariate linear difference equation in n, with solution $(1-\phi^n)/(1-\phi)$. As n increases, B_n approaches a limit $B = 1/(1-\phi)$. Thus bond prices fall when short rates rise, and the sensitivity of bond returns to short rates increases with maturity.

Note that B_n is different from duration, defined in Section 10.1.2 of Chapter 10. Duration measures the sensitivity of the n-period bond return to the n-period bond yield, and for zero-coupon bonds duration equals maturity. B_n measures the sensitivity of the n-period bond return to the one-period interest rate; it is always less than maturity because the n-period bond yield moves less than one-for-one with the one-period interest rate.

A second implication of the model is that the expected log excess return on an n-period bond over a one-period bond, $\mathrm{E}_t[r_{n,t+1}] - y_{1t} = \mathrm{E}_t[p_{n-1,t+1}] - p_{nt} + p_{1t}$, is given by

$$\begin{aligned}\mathrm{E}_t[r_{n,t+1}] - y_{1t} &= -\mathrm{Cov}_t[r_{n,t+1}, m_{t+1}] - \mathrm{Var}_t[r_{n,t+1}]/2\\ &= B_{n-1}\mathrm{Cov}_t[x_{t+1}, m_{t+1}] - B_{n-1}^2\mathrm{Var}_t[x_{t+1}]/2\\ &= -B_{n-1}\beta\sigma^2 - B_{n-1}^2\sigma^2/2.\end{aligned} \tag{11.1.12}$$

The first equality in (11.1.12) is a general result, discussed in Chapter 8, that holds for the excess log return on any asset over the riskfree interest rate. It can be obtained by taking logs of the fundamental relation $1 = \mathrm{E}_t[(1 + R_{i,t+1})M_{t+1}]$ for the n-period bond and the short interest rate, and then taking the difference between the two equations. It says that the expected excess log return is the sum of a risk premium term and a Jensen's Inequality term in the own variance which appears because we are working in logs.

The second equality in (11.1.12) uses the fact that the unexpected component of the log return on an n-period bond is just $-B_{n-1}$ times the innovation in the state variable. The third equality in (11.1.12) uses the fact that the conditional variance of x_{t+1} and its conditional covariance with m_{t+1} are constants to show that the expected log excess return on any bond is constant over time, so that the log expectations hypothesis—but not the log pure expectations hypothesis—holds.

$-B_{n-1}$ is the coefficient from a regression of n-period log bond returns on state variable innovations, so we can interpret $-B_{n-1}$ as the bond's loading on the single source of risk and $\beta\sigma^2$ as the reward for bearing a unit of risk. Alternatively, following Vasicek (1977) and others, we might calculate the price of risk as the ratio of the expected excess log return on a bond, plus one half its own variance to adjust for Jensen's Inequality, to the standard deviation of the excess log return on the bond. Defined this way, the price of risk is just $-\beta\sigma$ in this model.

The homoskedastic bond pricing model also has implications for the pattern of forward rates, and hence for the shape of the yield curve. To derive these implications, we note that in any termstructure model the n-period-ahead forward rate f_{nt} satisfies

$$\begin{aligned} f_{nt} &= p_{nt} - p_{n+1,t} \\ &= -p_{1t} + (\mathrm{E}_t[p_{n,t+1}] - p_{n+1,t} + p_{1t}) - (\mathrm{E}_t[p_{n,t+1}] - p_{nt}) \\ &= y_{1t} + (\mathrm{E}_t[r_{n+1,t+1}] - y_{1t}) - (\mathrm{E}_t[p_{n,t+1}] - p_{nt}). \end{aligned} \tag{11.1.13}$$

In this model $\mathrm{E}_t[p_{n,t+1}] - p_{nt} = B_n \mathrm{E}_t[\Delta x_{t+1}]$, and $\mathrm{E}_t[r_{n+1,t+1}] - y_{1t}$ is given by (11.1.12). Substituting into (11.1.13) and using $B_n = (1-\phi^n)/(1-\phi)$, we get

$$\begin{aligned} f_{nt} &= \mu - \left[\beta + \left(\frac{1-\phi^n}{1-\phi}\right)\right]^2 \frac{\sigma^2}{2} + \phi^n(x_t - \mu) \\ &= \left[\mu - \left(\beta + \frac{1}{1-\phi}\right)^2 \frac{\sigma^2}{2}\right] + \left[(x_t - \mu) + \left(\frac{1+\beta(1-\phi)}{(1-\phi)^2}\right)\sigma^2\right]\phi^n \\ &\quad - \left[\left(\frac{1}{1-\phi}\right)^2 \frac{\sigma^2}{2}\right]\phi^{2n}. \end{aligned} \tag{11.1.14}$$

The first equality in (11.1.14) shows that the change in the n-period forward rate is ϕ^n times the change in x_t. Thus movements in the forward rate die out geometrically at rate ϕ. This can be understood by noting that the log expectations hypothesis holds in this model, so forward-rate movements reflect movements in the expected future short rate which are given by ϕ^n times movements in the current short rate.

As maturity n increases, the forward rate approaches

$$\mu - \left(\beta + 1/(1-\phi)\right)^2\sigma^2/2,$$

a constant that does not depend on the current value of the state variable x_t. Equation (11.1.7) implies that the average short rate is $\mu - \beta^2\sigma^2/2$. Thus the difference between the limiting forward rate and the average short rate is

$$-\left(1/(1-\phi)\right)^2\sigma^2/2 - \left(\beta/(1-\phi)\right)\sigma^2.$$

This is the same as the limiting expected log excess return on a long-term bond. Because of the Jensen's Inequality effect, the log forward-rate curve tends to slope downwards towards its limit unless β is sufficiently negative, $\beta < -1/2(1-\phi)$.

As x_t varies, the forward-rate curve may take on different shapes. The second equality in (11.1.14) shows that the forward-rate curve can be written as the sum of a component that does not vary with n, a component that dies out with n at rate ϕ, and a component that dies out with n at rate ϕ^2. The third component has a constant coefficient with a negative sign; thus there is always a steeply rising component of the forward-rate curve. The second component has a coefficient that varies with x_t, so this component may be slowly rising, slowly falling, or flat. Hence the forward-rate curve may be rising throughout, falling throughout *(inverted)*, or may be rising at first and then falling (*hump-shaped*) if the third component initially dominates and then is dominated by the second component further out along the curve. These are the most common shapes for nominal forward-rate curves. Thus, if one is willing to apply the model to nominal interest rates, disregarding the fact that it allows interest rates to go negative, one can fit most observed nominal term structures. However the model cannot generate a forward-rate curve which is falling at first and then rising *(inverted hump-shaped)*, as occasionally seen in the data.

It is worth noting that when $\phi = 1$, the one-period interest rate follows a random walk. In this case the coefficients A_n and B_n never converge as n increases. We have $B_n = n$ and $A_n - A_{n-1} = -(\beta + n - 1)^2\sigma^2/2$. The forward rate becomes $f_{nt} = x_t - (\beta + n)^2\sigma^2/2$, which may increase with maturity at first if β is negative but eventually decreases with maturity forever. Thus the homoskedastic bond pricing model does not allow the limiting forward rate to be both finite and time-varying; either $\phi < 1$, in which case the limiting forward rate is constant over time, or $\phi = 1$, in which case there is no finite limiting forward rate. This restriction may seem rather counterintuitive; in fact it follows from the very general result—derived by Dybvig, Ingersoll, and Koss (1996)—that the limiting forward rate, if it exists, can never fall. In the homoskedastic model with $\phi < 1$ the limiting forward rate never falls because it is constant; in the homoskedastic model with $\phi = 1$ the limiting forward rate does not exist.

The discrete-time model developed in this section is closely related to the continuous-time model of Vasicek (1977). Vasicek specifies a continuous-time AR(1) or Ornstein-Uhlenbeck process for the short interest rate r, given by the following stochastic differential equation:

$$dr = \kappa(\theta - r)dt + \sigma dB, \tag{11.1.15}$$

where κ, 6, and a are constants.* Also, Vasicek assumes that the price *of* interest rate risk—the ratio of the expected excess return on a bond to the standard deviation of the excess return on the bond—is a constant that does not depend on the level of the short interest rate. The model of this section derives an AR(1) process for the short rate and a constant price of risk from primitive assumptions on the stochastic discount factor.

Equilibrium Interpretation of the Model

Our analysis has shown that the sign of the coefficient β determines the sign of all bond risk premia. To understand this, consider the effects of a positive shock ξ_{t+1} which increases the state variable x_{t+1} and lowers all bond prices. When β is positive the shock also drives down m_{t+1}, so bond returns are positively correlated with the stochastic discount factor. This correlation has hedge value, so risk premia on bonds are negative. When β is negative, on the other hand, bond returns are negatively correlated with the stochastic discount factor, and risk premia are positive.

We can get more intuition by considering the case where the stochastic discount factor reflects the power utility function of a representative agent, as in Chapter 8. In this case $M_{t+1} = \delta(C_{t+1}/C_t)^{-\gamma}$, where 6 is the discount factor and γ is the risk-aversion coefficient of the representative agent. Taking logs, we have

$$m_{t+1} = \log(\delta) - \gamma\Delta c_{t+1}. \tag{11.1.16}$$

It follows that $x_t \equiv \mathrm{E}_t[-m_{t+1}] = -\log(\delta) + \gamma\mathrm{E}_t[\Delta c_{t+1}]$, and $\epsilon_{t+1} \equiv -m_{t+1} - \mathrm{E}_t[-m_{t+1}] = \gamma(\Delta c_{t+1} - \mathrm{E}_t[\Delta c_{t+1}])$. x_t is a linear function of expected consumption growth, and ϵ_{t+1} is proportional to the innovation in consumption growth. The term-structure model of this section then implies that expected consumption growth is an AR(1) process, so that realized consumption growth is an ARMA(1,1) process. The coefficient β governs the covariance between consumption innovations and revisions in expected future consumption growth. If β is positive, then a positive consumption shock today drives up expected future consumption growth and increases interest rates; the resulting fall in bond prices makes bonds covary negatively with consumption and gives them negative risk premia. If

[2]As in Chapter 9, dB in (11.1.15) denotes the increment to a Brownian motion; it should not be confused with the bond price coefficients B_n of this section.

β is negative, a positive shock to consumption lowers interest rates so bonds have positive risk premia.

Campbell (1986) explores the relation between bond risk premia and the timeseries properties of consumption in a related model. Campbell's model is similar to the one here in that consumption and asset returns are conditionally lognormal and homoskedastic. It is more restrictive than the model here because it makes consumption growth (rather than expected consumption growth) a univariate stochastic process, but it is more general in that it does not require expected consumption growth to follow an AR(1) process. Campbell shows that the sign of the risk premium for an n-period bond depends on whether a consumption innovation raises or lowers consumption growth expected over (n − 1) periods. Backus and Zin (1994) explore this model in greater detail. Backus, Gregory, and Zin (1989) also relate bond risk premia to the time-series properties of consumption growth and interest rates.

Cox, Ingersoll, and Ross (1985a) show how to derive a continuous-time term-structure model like the one in this section from an underlying production model. Sun (1992) and Backus (1993) restate their results in discrete time. Assume that there is a representative agent with discount factor 6 and time-separable log utility. Suppose that the agent faces a budget constraint of the form

$$K_{t+1} = (K_t - C_t)X_t V_{t+1}, \tag{11.1.17}$$

where K_t is capital at the start of the period, $(K_t - C_t)$ is invested capital, and $X_t V_{t+1}$ is the return on capital. This budget constraint has constant returns to scale because the return on capital does not depend on the level of capital. X_t is the anticipated component of the return and V_{t+1} is an unanticipated technology shock. With log utility it is well-known that the agent chooses $C_t/K_t = (1 - 6)$. Substituting this into (11.1.17) and taking logs we find that

$$\Delta c_{t+1} = \log(\delta) + x_t + v_{t+1}, \tag{11.1.18}$$

where $v_{t+1} \equiv \log(V_{t+1})$, and $-m_{t+1} = -\log(\delta) + \Delta c_{t+1} = x_t + \epsilon_{t+1}$. This derivation allows x_t to follow any process, including the AR(1) assumed by the term-structure model.

11.1.2 A Square-Root *Single-Factor* Model

The homoskedastic model of the previous section is appealing because of its simplicity, but it has several unattractive features. First, it assumes that interest rate changes have constant variance. Second, the model allows interest rates to go negative. This makes it applicable to real interest rates, but less appropriate for nominal interest rates. Third, it implies that risk premia are constant over time, contrary to the evidence presented in Section **10.2.1** of Chapter **10.** One can alter the model to handle these problems, while retaining much of the simplicity of the basic structure, by allowing the state variable x_t to follow a conditionally lognormal but heteroskedastic ***square-root*** process. This change is entirely consistent with the equilibrium foundations for the model given in the previous section.

The square-root model, which is a discrete-time version of the famous Cox, Ingersoll, and Ross (1985a) continuous-time model, replaces **(11.1.5)** and **(11.1.3)** with

$$-m_{t+1} = x_t + x_t^{1/2}\epsilon_{t+1} = x_t + x_t^{1/2}\beta\xi_{t+1}, \qquad (11.1.19)$$

$$x_{t+1} = (1-\phi)\mu + \phi x_t + x_t^{1/2}\xi_{t+1}. \qquad (11.1.20)$$

The new element here is that the shock ξ_{t+1} is multiplied by $x_t^{1/2}$. To understand the importance of this, recall that in the homoskedastic model x_{t+i} and m_{t+i} are normal conditional on x_t for all $i \geq$ **1.** This means that one can analyze the homoskedastic model either by taking logs of **(11.0.1)** to get the recursive equation (11.1.1), or by taking logs of **(11.0.2)** to get an n-period loglinear equation:

$$p_{nt} = \mathrm{E}_t[m_{t+1} + \cdots + m_{t+n}] + (1/2)\,\mathrm{Var}_t[m_{t+1} + \cdots + m_{t+n}]. \qquad (11.1.21)$$

Calculations based on **(11.1.21)** are more cumbersome than the analysis presented in the previous section, but they give the same result. In the square-root model, by contrast, x_{t+1} and m_{t+1} are normal conditional on x_t but x_{t+i} and m_{t+i} are nonnormal conditional on x_t for all $i >$ **1.** This means that one can only analyze the square-root model using the recursive equation **(11.1.1)**; the n-period loglinear relation **(11.1.21)** does not hold in the square-root model.

Proceeding with the recursive analysis as before, we can determine the price of a one-period bond by substituting **(11.1.19)** into **(11.1.1)** to get

$$p_{1t} = \mathrm{E}_t[m_{t+1}] + (1/2)\,\mathrm{Var}_t[m_{t+1}] = -x_t(1-\beta^2\sigma^2/2). \qquad (11.1.22)$$

The one-period bond yield $y_{1t} = -p_{1t}$ is now proportional to the state variable x_t. Once again the short rate measures the state of the economy in the model.

Since the short rate is proportional to the state variable, it inherits the property that its conditional variance is proportional to its level. Many authors have noted that interest rate volatility tends to be higher when interest rates are high; in Section 11.2.2 we discuss the empirical evidence on this point. This property also makes it hard for the interest rate to go negative, since the upward drift in the state variable tends to dominate the random shocks as x_t declines towards zero. Cox, Ingersoll, and Ross (1985a)

show that negative interest rates are ruled out in the continuous-time version of this model, where the instantaneous interest rate follows the process $dr = \kappa(\theta - r)dt + \sigma r^{1/2} dB$.[3] Time-variation in volatility also produces time-variation in term premia, so that the log expectations hypothesis no longer holds in this model.

We now guess that the price function for an n-period bond has the same linear form as before, $-p_{nt} = A_n + B_n x_t$, equation (11.1.8). In this model $A_0 = B_0 = 0$, $A_1 = 0$, and $B_1 = 1 - \beta^2\sigma^2/2$. It is straightforward to verify the guess and to show that A_n and B_n obey

$$B_n = 1 + \phi B_{n-1} - (\beta + B_{n-1})^2\sigma^2/2,$$

$$A_n - A_{n-1} = (1-\phi)\mu B_{n-1}. \qquad (11.1.23)$$

Comparing **(11.1.23)** with **(11.1.11)**, we see that the term in σ^2 has been moved from the equation describing A_n to the equation describing B_n. This is because the variance is now proportional to the state variable, so it affects the slope coefficient rather than the intercept coefficient for the bond price. The limiting value of B_n, which we write as B, is now the solution to a quadratic equation, but for realistic parameter values this solution is close to the limit $1/(1-\phi)$ from the previous model. Thus B_n is positive and increasing in n.

The expected excess log bond return in the square-root model is given by

$$\begin{aligned} \mathrm{E}_t[r_{n,t+1}] - y_{1t} &= -\mathrm{Cov}_t[r_{n,t+1}, m_{t+1}] - \mathrm{Var}_t[r_{n,t+1}]/2 \\ &= B_{n-1}\,\mathrm{Cov}_t[x_{t+1}, m_{t+1}] - B_{n-1}^2\,\mathrm{Var}_t[x_{t+1}]/2 \\ &= (-B_{n-1}\beta\sigma^2 - B_{n-1}^2\sigma^2/2)x_t. \end{aligned} \qquad (11.1.24)$$

The first two equalities here are the same as in the previous model. The third equality is the formula from the previous model, (11.1.12), multiplied by the state variable x_t. Thus the expected log excess return is proportional to the state variable x_t or, equivalently, to the short interest rate y_{1t}. This is the expected result since the conditional variance of interest rates is proportional to x_t. Once again the sign of β determines the sign of the risk premium term in (11.1.24). Since the standard deviation of excess bond returns is proportional to the square root of x_t, the price of interest rate risk—the ratio of the expected excess log return on a bond, plus one half its own variance to adjust for Jensen's Inequality, to the standard deviation of

[3] Depending on the parameter values, it may be possible for the interest rate to be zero in the continuous-time model. Longstaff (1992) discusses alternative ways to model this possibility.

the excess log return on the bond—is also proportional to the square root of x_t.

The forward rate in the square-root model is given by

$$\begin{aligned} f_{nt} &= y_{1t} + B_n(\mathrm{E}_t[\Delta x_{t+1}] - \mathrm{Cov}_t[x_{t+1}, m_{t+1}]) - B_n^2 \mathrm{Var}_t[x_{t+1}]/2 \\ &= (1-\beta^2\sigma^2/2)x_t - B_n[(1-\phi)(x_t-\mu) + x_t\beta\sigma^2] \\ &\quad - B_n^2 x_t\sigma^2/2. \end{aligned} \tag{11.1.25}$$

The first equality in (11.1.25) is the same as in the homoskedastic model, while the second equality multiplies variance terms by x_t where appropriate. It can be shown that the square-root model permits the same range of shapes for the yield curve—upward-sloping, inverted, and humped—as the homoskedastic model.

Pearson and Sun (1994) have shown that the square-root model can be generalized to allow the variance of the state variable to be linear in the level of the state variable, rather than proportional to it. One simply replaces the $x_t^{1/2}$ terms, multiplying the shocks in (11.1.19) and (11.1.20) with terms of the form $(\alpha_0 + \alpha_1 x_t)^{1/2}$. The resulting model is tractable because it remains in the affine-yield class, and it nests both the homoskedastic model (the case $\alpha_0 = 1, \alpha_1 = 0$) and the basic square-root model (the case $\alpha_0 = 0$, $\alpha_1 = 1$).

11.1.3 A Two-Factor Model

So far we have only considered single-factor models. Such models imply that all bond returns are perfectly correlated. While bond returns do tend to be highly correlated, their correlations are certainly not one and so it is natural to ask how this implication can be avoided.

We now present a simple model in which there are two factors rather than one, so that bond returns are no longer perfectly correlated.[4] The model is a discrete-time version of the model of Longstaff and Schwartz (1992). It replaces (11.1.19) with

$$-m_{t+1} = x_{1t} + x_{2t} + x_{1t}^{1/2}\epsilon_{t+1}, \tag{11.1.26}$$

and replaces (11.1.20) with a pair of equations for the state variables:

$$x_{1,t+1} = (1-\phi_1)\mu_1 + \phi_1 x_{1t} + x_{1t}^{1/2}\xi_{1,t+1}, \tag{11.1.27}$$

$$x_{2,t+1} = (1-\phi_2)\mu_2 + \phi_2 x_{2t} + x_{2t}^{1/2}\xi_{2,t+1}. \tag{11.1.28}$$

[4]Although bond returns are not perfectly correlated in this model, the covariance matrix of bond returns has rank two and hence is singular whenever we observe more than two bonds. We discuss this point further in Section 11.1.4.

Finally, the relation between the shocks is

$$\epsilon_{t+1} = \beta\xi_{1,t+1}, \tag{11.1.29}$$

and the shocks $\xi_{1,t+1}$ and $\xi_{2,t+1}$ are uncorrelated with each other. We will write σ_1^2 for the variance of $\xi_{1,t+1}$ and σ_2^2 for the variance of $\xi_{2,t+1}$.

In this model, minus the log stochastic discount factor is forecast by two state variables, x_{1t} and x_{2t}. The variance of the innovation to the log stochastic discount factor is proportional to the level of x_{1t}, as in the square-root model; and each of the two state variables follows a square-root autoregressive process. Finally, the log stochastic discount factor is conditionally correlated with x_1 but not with x_2. This last assumption is required to keep the model in the tractable affine-yield class. Note that the two-factor model nests the single-factor square-root model, which can be obtained by setting $x_{2t} = 0$, but does not nest the single-factor homoskedastic model.

Proceeding in the usual way, we find that the price of a one-period bond is

$$p_{1t} = \mathrm{E}_t[m_{t+1}] + (1/2)\mathrm{Var}_t[m_{t+1}] = -x_{1t} - x_{2t} + x_{1t}\beta^2\sigma_1^2/2. \tag{11.1.30}$$

The one-period bond yield $y_{1t} = -p_{1t}$ is no longer proportional to the state variable x_{1t}, because it depends also on x_{2t}. The short interest rate is no longer sufficient to measure the state of the economy in this model. Longstaff and Schwartz (1992) point out, however, that the conditional variance of the short rate is a different linear function of the two state variables:

$$\mathrm{Var}_t[y_{1,t+1}] = (1-\beta^2\sigma_1^2/2)^2\sigma_1^2 x_{1t} + \sigma_2^2 x_{2t}. \tag{11.1.31}$$

Thus the short rate and its conditional volatility summarize the state of the economy, and one can always state the model in terms of these two variables.

We guess that the price function for an n-period bond is linear in the two state variables: $-p_{nt} = A_n + B_{1n}x_{1t} + B_{2n}x_{2t}$. We already know that $A_0 = B_{10} = B_{20} = 0$, $A_1 = 0$, $B_{11} = 1 - \sigma_\epsilon^2/2$, and $B_{21} = 1$. It is straightforward to show that A_n, B_{1n}, and B_{2n} obey

$$\begin{aligned} B_{1n} &= 1 + \phi_1 B_{1,n-1} - (\beta + B_{1,n-1})^2\sigma_1^2/2, \\ B_{2n} &= 1 + \phi_2 B_{2,n-1} - B_{2,n-1}^2\sigma_2^2/2, \\ A_n - A_{n-1} &= (1-\phi_1)\mu_1 B_{1,n-1} + (1-\phi_2)\mu_2 B_{2,n-1}. \end{aligned} \tag{11.1.32}$$

The difference equation for B_{1n} is the same as in the single-factor square-root model, (11.1.23), but the difference equation for B_{2n} includes only a term in the own variance of x_2 because x_2 is uncorrelated with m and does not affect the variance of m. The difference equation for A, is just the sum of two terms, each of which has the familiar form from the single-factor square-root model.

The expected excess log bond return in the two-factor model is given by

$$\begin{aligned} \mathrm{E}_t[r_{n,t+1}] - y_{1t} &= -\mathrm{Cov}_t[r_{n,t+1}, m_{t+1}] - \mathrm{Var}_t[r_{n,t+1}]/2 \\ &= B_{1,n-1}\,\mathrm{Cov}_t[x_{1,t+1}, m_{t+1}] - B_{1,n-1}^2\,\mathrm{Var}_t[x_{1,t+1}]/2 \\ &\quad - B_{2,n-1}^2\,\mathrm{Var}_t[x_{2,t+1}]/2 \\ &= [-B_{1,n-1}\beta\sigma_1^2 - B_{1,n-1}^2\sigma_1^2/2]x_{1t} \\ &\quad -[B_{2,n-1}^2\sigma_2^2/2]x_{2t}. \end{aligned} \tag{11.1.33}$$

This is the same as in the square-root model, with the addition of an extra term, arising from Jensen's Inequality, in the variance of $x_{2,t+1}$.

The forward rate in the two-factor model is given by

$$\begin{aligned} f_{nt} &= y_{1t} + B_{1n}(\mathrm{E}_t[\Delta x_{1,t+1}] - \mathrm{Cov}_t[x_{1,t+1}, m_{t+1}]) + B_{2n}\mathrm{E}_t[\Delta x_{2,t+1}] \\ &\quad - B_{1n}^2\,\mathrm{Var}_t[x_{1,t+1}]/2 - B_{2n}^2\,\mathrm{Var}_t[x_{2,t+1}]/2 \\ &= (1-\beta^2\sigma_1^2/2)x_{1t} + x_{2t} - B_{1n}(1-\phi_1)(x_{1t}-\mu_1) \\ &\quad -B_{2n}(1-\phi_2)(x_{2t}-\mu_2) - B_{1n}\,x_{1t}\,\beta\sigma_1^2 \\ &\quad -B_{1n}^2\,x_{1t}\,\sigma_1^2/2 - B_{2n}^2\,x_{2t}\,\sigma_2^2/2. \end{aligned} \tag{11.1.34}$$

This is the obvious generalization of the square-root model. Importantly, it can generate more complicated shapes for the yield curve, including inverted hump shapes, as the independent movements of both x_{1t} and x_{2t} affect the term structure.

The analysis of this model illustrates an important principle. As Cox, Ingersoll, and Ross (1985a) and Dybvig (1989) have emphasized, under certain circumstances one can construct multifactor term-structure models simply by "adding up" single-factor models. Whenever the stochastic discount factor m_{t+1} can be written as the sum of two independent processes, then the resulting term structure is the sum of the term structures that would exist under each of these processes. In the Longstaff and Schwartz (1992) model the stochastic discount factor is the sum of $-x_{1t} - x_{1t}^{1/2}\beta\xi_{1,t+1}$ and $-x_{2t}$, and these components are independent of each other. Inspection of (11.1.34) shows that the resulting term structure is just the sum of a general square-root term structure driven by the x_{1t} process and a special square-root term structure with parameter restriction $\beta = 0$ driven by the x_{2t} process.

11.1.4 Beyond Affine-Yield Models

We have considered a sequence of models, each of which turns out to have the property that log bond yields are linear or *affine* in the underlying state variables. Brown and Schaefer (1991) and Duffie and Kan (1993) have clarified the primitive assumptions necessary to get an affine-yield model. In the discrete-time framework used here, these conditions are most easily stated by defining a vector $\mathbf{x}_t$ which contains the log stochastic discount factor m_t and the time t values of the state variables relevant for forecasting future m_{t+i}, $i = 1\ldots n$. If the conditional forecast of x one period ahead, $\mathrm{E}_t[\mathbf{x}_{t+1}]$, is affine in the state variables, and if the conditional distribution of x one period ahead is normal with a variance-covariance matrix $\mathrm{Var}_t[\mathbf{x}_{t+1}]$ which is affine in the state variables, then the resulting term-structure model is an affine-yield model.

To see this, consider the steps we used to derive the implications of each successive term-structure model. We first calculated the log short-term interest rate; this is affine in the underlying state variables if m_{t+1} is conditionally normal and $\mathrm{E}_t[m_{t+1}]$ and $\mathrm{Var}_t[m_{t+1}]$ are affine in the state variables. We next guessed that log bond yields were affine and proceeded to verify the guess. If yields are affine, and if x is conditionally normal with affine variance-covariance matrix, then the risk premium on any bond is affine. Finally we derived log forward rates; these are affine if the short rate, risk premium, and the expected change in the state variable are all affine. Affine forward rates imply affine yields, verifying that the model is in the affine-yield class.

Brown and Schaefer (1991) and Duffie and Kan (1993) state conditions on the short rate which deliver an affine-yield model in a continuous-time setting. They show that the risk-adjusted drift in the short rate—the expected change in the short rate less the covariance of the short rate with the stochastic discount factor—and the variance of the short rate must both be affine to get an affine-yield model. The models of Vasicek (1977), Cox, Ingersoll, and Ross (1985a), and Pearson and Sun (1994) satisfy these requirements, but some other continuous-time models such as that of Brennan and Schwartz (1979) do not.

Affine-yield models have a number of desirable properties which help to explain their appeal. First, log bond yields inherit the conditional normality assumed for the underlying state variables. Second, because log bond yields are linear functions of the state variables we can renormalize the model so that the yields themselves are the state variables. This is obvious in a one-factor model where the short rate is the state variable, but it is equally possible in a model with any number of factors. Longstaff and Schwartz (1992) present their two-factor model as one in which the volatility of the short rate and the level of the short rate are the factors; the model could be written equally well in terms of any two bond yields of fixed maturities. Third, affine-yield models with K state variables imply that the term structure of interest rates can be summarized by the levels of K bond yields at each point in time and the constant coefficients relating other bond yields to the K basis yields. In this sense affine-yield models are linear; their nonlinearity is confined to the process governing the intertemporal evolution of the K basis yields and the relation between the cross-sectional coefficients and the underlying parameters of the model.

Affine-yield models also have some disadvantages. The linear relations among bond yields mean that the covariance matrix of bond returns has rank K--equivalently, we can perfectly fit the return on any bond using a regression on K other contemporaneous bond returns. This implication will always be rejected by a data set containing more than K bonds, unless we add extra error terms to the model. Affine-yield models also limit the way in which interest rate volatility can change with the level of interest rates; for example a model in which volatility is proportional to the square of the interest rate is not affine. Finally, as Constantinides (1992) emphasizes, single-factor affine-yield models imply that risk premia on long-term bonds always have the same sign.

If we move outside the affine-yield class of models, we can no longer work with equation (11.1.1) but must return to the underlying nonlinear difference equation (11.0.1) or its n-period representation (11.0.2). In general these equations must be solved numerically. One common method is to set up a binomial tree for the short-term interest rate. Black, Derman, and Toy (1990) and Black and Karasinski (1991), for example, assume that the simple one-period yield Y_{1t} is conditionally lognormal (as opposed to the assumption of affine-yield models that $(1+Y_{1t})$ is conditionally lognormal). They use a binomial tree to solve their models for the implied term structure of interest rates. Constantinides (1992), however, presents a model that can be solved in closed form. His model makes the log stochastic discount factor a sum of noncentral chi-squared random variables rather than a normal random variable, and Constantinides is then able to calculate the expectations in (11.0.2) analytically.

11.2 Fitting Term-Structure Models to the Data

11.2.1 Real Bonds, Nominal Bonds, and *Inflation*

The term-structure models described so far apply to bonds whose payoffs are riskless in real terms. Almost all actual bonds instead have payoffs that are

riskless in nominal terms.[5] We now discuss how the models can be adapted to deal with this fact.

To study nominal bonds we need to introduce some new notation. We write the nominal price index at time t as Q_t, and the gross rate of inflation from t to t + 1 as $\Pi_{t+1} \equiv Q_{t+1}/Q_t$. We have already defined P_{nt} to be the real price of an n-period real bond which pays one goods unit at time t + n; we now define $P^{\$}_{nt}$ to be the nominal price of an n-period nominal bond which pays \$1 at time t + n. From these definitions it follows that the nominal price of an n-period real bond is $P_{nt}\,Q_t$, and the real price of an n-period nominal bond is $P^{\$}_{nt}/Q_t$. We do not adopt any special notation for these last two concepts.

If we now apply the general asset pricing condition,

$$1 = \mathrm{E}_t[(1+R_{i,t+1})M_{t+1}],$$

to the real return on an n-period nominal bond, we find that

$$\frac{P^{\$}_{nt}}{Q_t} = \mathrm{E}_t\left[\frac{P^{\$}_{n-1,t+1}}{Q_{t+1}}\,M_{t+1}\right]. \tag{11.2.1}$$

Multiplying through by Q_t, we have

$$\begin{aligned} P^{\$}_{nt} &= \mathrm{E}_t\left[P^{\$}_{n-1,t+1}\,M_{t+1}\,\frac{Q_t}{Q_{t+1}}\right] \\ &= \mathrm{E}_t\left[P^{\$}_{n-1,t+1}\,\frac{M_{t+1}}{\Pi_{t+1}}\right] \\ &= \mathrm{E}_t[\,P^{\$}_{n-1,t+1}\,M^{\$}_{t+1}\,], \end{aligned} \tag{11.2.2}$$

where $M^{\$}_{t+1} \equiv M_{t+1}/\Pi_{t+1}$ can be thought of as a nominal stochastic discount factor that prices nominal returns.

The empirical literature on nominal bonds uses this result in one of two ways. The first approach is to take the primitive assumptions that we made about M_{t+1} in Section 11.1 and to apply them instead to $M^{\$}_{t+1}$. The real term-structure models of the last section are then reinterpreted as nominal term-structure models. Brown and Dybvig (1986), for example, do this when

[5]Some governments, notably those of Canada, Israel, and the UK, have issued bonds whose nominal payoffs are linked to a nominal price index. In 1996 the US Treasury is considering issuing similar securities. These index-linked bonds approximate real bonds but are rarely exactly equivalent to real bonds. Brown and Schaefer (1994) give a lucid discussion of the imperfections in the UK indexing system, and apply the Cox, Ingersoll, and Ross (1985a) model to UK index-linked bonds. See also Barr and Campbell (1995) and Campbell and Shiller (1996).

they apply the Cox, Ingersoll, and Ross (1985a) square-root model directly to data on US nominal bond prices. The square-root model restricts interest rates to be positive, and in this respect it is more appropriate for nominal interest rates than for real interest rates.

The second approach is to assume that the two components of the nominal stochastic discount factor, M_{t+1} and $1/\Pi_{t+1}$, are independent of each other. To see how this assumption facilitates empirical work, take logs of the nominal stochastic discount factor to get

$$m^{\$}_{t+1} = m_{t+1} - \pi_{t+1}. \tag{11.2.3}$$

When the components m_{t+1} and π_{t+1} are independent, we can price nominal bonds by using the insights of Cox, Ingersoll, and Ross (1985a) and Dybvig (1989). Recall from Section 11.1.3 their result that the log bond price in a model with two independent components of the stochastic discount factor is the sum of the log bond prices implied by each component. We can, for example, apply the Longstaff and Schwartz (1992) model to nominal bonds by assuming that m_{t+1} is described by a square-root single-factor model, $-m_{t+1} = x_{1t} + x_{1t}^{1/2}\beta\xi_{1,t+1}$, and that π_{t+1} is known at t and equal to a state variable x_{2t}. We then get $-m^{\$}_{t+1} = -m_{t+1} + \pi_{t+1} = x_{1t} + x_{1t}^{1/2}\beta\xi_{1,t+1} + x_{2t}$, and the Longstaff-Schwartz model describes nominal bonds.

More generally, the assumption that M_{t+1} and $1/\Pi_{t+1}$ are independent implies that prices of nominal bonds are just prices of real bonds multiplied by the expectation of the future real value of money, and that expected real returns on nominal bonds are the same as expected real returns on real bonds. To see this, consider equation (11.2.2) with maturity $n = 1$, and note that the independence of M_{t+1} and $1/\Pi_{t+1}$ allows us to replace the expectation of their product by the product of their expectations:

$$P^{\$}_{1t} = \mathrm{E}_t[M^{\$}_{t+1}] = \mathrm{E}_t[M_{t+1}]\,\mathrm{E}_t\left[\frac{1}{\Pi_{t+1}}\right] = P_{1t}\,Q_t\,\mathrm{E}_t\left[\frac{1}{Q_{t+1}}\right], \tag{11.2.4}$$

since $P_{1t} = \mathrm{E}_t[M_{t+1}]$ and $1/\Pi_{t+1} = Q_t/Q_{t+1}$. Thus the nominal price of a bond which pays \$1 tomorrow is the nominal price of a bond which pays one unit of goods tomorrow, times the expectation of the real value of \$1 tomorrow.

We now guess that a similar relationship holds for all maturities n, and we prove this by induction. If the (n−1)-period relationship holds, $P^{\$}_{n-1,t} = P_{n-1,t}\,Q_t\,\mathrm{E}_t[1/Q_{t+n-1}]$, then

$$\begin{aligned} P^{\$}_{nt} &= \mathrm{E}_t\left[P^{\$}_{n-1,t+1}\,M_{t+1}\,\frac{Q_t}{Q_{t+1}}\right] \\ &= \mathrm{E}_t\left[P_{n-1,t+1}\,Q_{t+1}\,\mathrm{E}_{t+1}\left[\frac{1}{Q_{t+n}}\right]M_{t+1}\,\frac{Q_t}{Q_{t+1}}\right] \\ &= Q_t\,\mathrm{E}_t\left[P_{n-1,t+1}\,M_{t+1}\,\mathrm{E}_{t+1}\left[\frac{1}{Q_{t+n}}\right]\right] \\ &= P_{nt}\,Q_t\,\mathrm{E}_t\left[\frac{1}{Q_{t+n}}\right], \end{aligned} \tag{11.2.5}$$

where the last equality uses both the independence of real variables from the price level (which enables us to replace the expectation of a product by the product of expectations), and the fact that $P_{nt} = \mathrm{E}_t[P_{n-1,t+1}\,M_{t+1}]$. Equation (11.2.5) is the desired result that the nominal price of a bond which pays \$1 at time $t+n$ is the nominal price of a bond which pays one unit of goods at time $t+n$, times the expected real value of \$1 at time $t+n$. Dividing (11.2.5) by Q_t, we can see that the same relationship holds between the real prices of nominal bonds and the real prices of real bonds. Further, (11.2.5) implies that the expected real return on a nominal bond equals the expected real return on a real bond:

$$\begin{aligned} \mathrm{E}_t\left[\frac{P^{\$}_{n-1,t+1}}{P^{\$}_{nt}}\,\frac{Q_t}{Q_{t+1}}\right] &= \mathrm{E}_t\left[\frac{\mathrm{E}_{t+1}[1/Q_{t+n}]P_{n-1,t+1}\,Q_{t+1}}{\mathrm{E}_t[1/Q_{t+1}]P_{nt}\,Q_t}\,\frac{Q_t}{Q_{t+1}}\right] \\ &= \mathrm{E}_t\left[\frac{P_{n-1,t+1}}{P_{nt}}\right]. \end{aligned} \tag{11.2.6}$$

Gibbons and Ramaswamy (1993) use these results to test the implications of real term-structure models for econometric forecasts of real returns on nominal bonds.

Although it is extremely convenient to assume that inflation is independent of the real stochastic discount factor, this assumption may be unrealistic. Barr and Campbell (1995), Campbell and Ammer (1993), and Pennacchi (1991), using respectively UK data on indexed and nominal bonds, rationalexpectations methodology applied to US data, and survey data, all find that innovations to expected inflation are negatively correlated in the short run with innovations to expected future real interest rates. More directly, Campbell and Shiller (1996) find that inflation innovations are correlated with stock returns and real consumption growth, proxies for the stochastic discount factor suggested by the traditional CAPM of Chapter 5 and the consumption CAPM of Chapter 8.

11.2.2 Empirical Evidence on *Affine-Yield* Models

All the models we have discussed so far need additional error terms if they are to fit the data. To see why, consider a model in which the real stochastic discount factor is driven by a single state variable. In such a model, returns on all real bonds are perfectly correlated because the model has only a single shock. Similarly, returns on all nominal bonds are perfectly correlated in any