

## ***Steven Russell***

*Steven Russell is an economist at the Federal Reserve Bank of St. Louis. Lynn Dietrich provided research assistance.*

# Understanding the Term Structure of Interest Rates: The Expectations Theory

**T**HE INTEREST RATES on loans and securities provide basic summary measures of their attractiveness to lenders. The role played by interest rates in allocating funds across financial markets is very similar to the role played by prices in allocating resources in markets for goods and services. Just as a relatively high price of a particular good tends to draw physical resources into its production, a relatively high interest rate on a particular type of security tends to draw funds into the activities that type of security is issued to finance. And just as identifying the factors that help determine prices is a key area of inquiry among economists who study goods markets, identifying the factors that help determine interest rates is a key area of inquiry for those who study financial markets.

Economic theory suggests that one important factor explaining the differences in the interest rates on different securities may be differences in their *terms*—that is, in the lengths of time before they mature. The relationship between the terms of securities and their market rates of interest is known as the *term structure* of interest rates. To display the term structure of interest rates on securities of a particular type at a particular point in time, economists use a diagram

called a *yield curve*. As a result, term structure theory is often described as the theory of the yield curve.

Economists are interested in term structure theory for a number of reasons. One reason is that since the actual term structure of interest rates is easy to observe, the accuracy of the predictions of different term structure theories is relatively easy to evaluate. These theories are usually based on assumptions and principles that have applications in other branches of economic theory. If such principles prove useful in explaining the term structure, they might also prove useful in contexts in which their relevance is less easy to evaluate. One theory of the term structure that will be described here, for example, suggests that a behavioral trait called *risk aversion* may play a major role in determining the shape of the yield curve. If subsequent research lends credence to this theory, economists may give more emphasis to risk aversion in constructing theories of other aspects of financial market operation.<sup>1</sup>

A second reason why economists are interested in term structure theories is that they help explain the ways in which changes in short-term interest

<sup>1</sup>Examples include the role of financial intermediaries and the pricing of claims to physical assets (such as stocks).

rates—rates on securities with relatively short terms—affect the levels of long-term interest rates. Economic theory suggests that monetary policy may have a direct effect on short-term interest rates, but little, if any, direct effect on long-term rates. It also suggests that long-term rates play a critical role in a number of important economic decisions, such as firms' decisions about investment, and households' decisions about purchases of homes and other durable goods. Theories of the term structure may help explain the mechanism by which monetary policy affects these decisions.<sup>2</sup>

A third reason economists are interested in the term structure is that it may provide information about the *expectations* of participants in financial markets. These expectations are of considerable interest to forecasters and policy-makers. Market participants' beliefs about what may happen in the future influence their current decisions; these decisions, in turn, help determine what actually happens in the future. Thus, knowledge of participants' expectations is critical to forecasting future events or determining the effects of different policies.

Many economists believe that the people best able to forecast events in a market are in fact the participants in that market. If this is true, interest rate forecasting and inferring the nature of financial market participants' expectations amount to the same thing. The term structure theory that will be described in this article, which is called the *expectations theory*, suggests that the observed term structure can indeed be used to infer market participants' expectations about future interest rates—and through them, what actual future rates might be, and how events that tend to influence these rates may unfold. These events could include changes in the rate of economic growth or changes in monetary policy, for example.

The goal of this article is to provide a simple but thorough description of the expectations theory. The first section of the article lays the groundwork by explaining the basic concept and principles of interest rates and securities pricing. The presentation emphasizes issues that are particularly relevant to understanding how

the financial market goes about assigning different interest rates to securities with different terms. The second part of the article presents the expectations theory itself. The presentation is oriented around two widely noted observations about the term structure: (1) that yield curves are usually upward-sloping, and (2) that the steepness and/or direction of their slopes tends to change systematically as interest rates rise and fall.

## BUILDING BLOCKS OF THE TERM STRUCTURE

### *Prices, Interest Rates and Time*

Since the expectations theory tries to explain certain aspects of the way interest rates are determined, it is impossible to understand the theory without a thorough understanding of the nature and role of interest rates. A good starting point is the analogy we drew earlier between the prices of goods and services and the interest rates on securities. In our economy, purchasers of goods or services almost always pay with money, so the "price" of a given quantity of goods is simply the number of dollars paid for it. In markets where the goods are readily divisible and more or less uniform in quality, such as markets for agricultural commodities, the price is usually thought of as a number of dollars *per unit* of goods. This way of thinking about prices reflects what economists call the Law of One Price: when information is readily available and the number of buyers and sellers is large, each transaction involving a particular good tends to take place at the same unit price, regardless of the quantity of the good exchanged.

*Discount and Return Ratios*—In the securities market, one can think of lenders as buyers, and of future payments as the items they purchase. People lend to the federal government, for instance, by buying U.S. Treasury securities, which are government promises to repay the loans by making one or more future payments. The direct securities market counterpart of a price in a goods market would be the number

<sup>2</sup>Term structure theories are traditionally stated in terms of nominal or money interest rates. Economic theory predicts, however, that it is primarily real interest rates—interest rates net of expected inflation—that influence the decisions of households and firms. It is possible to formulate versions of most term-structure theories, including the theory described in this article, that apply specifically to real interest rates. Since we cannot observe inflation expectations, however,

we cannot measure real interest rates directly. This makes it difficult to describe real-interest-rate versions of the theories in terms non-economists are likely to understand.

of dollars lent (paid) today per dollar repaid in the future (future dollar purchased).<sup>3</sup> A security that cost \$10,000 and returned \$12,500 at a later date, for example, would have a unit price of 0.80. This price might be called a *discount ratio*.<sup>4</sup>

Economists usually conform to financial market practice by thinking about securities in terms of return rather than discount ratios—that is, ratios of amounts repaid to the amounts lent, rather than the reverse. We can define the *return ratio* on a single-payment security as the ratio of its maturity payment to its price (that is, the amount lent). The return ratio on the security just described would be 1.25—the reciprocal of its discount ratio.

*Accounting for the Time Dimension*—The return ratio, it turns out, is not a very good analogue to the market price: it suffers from a serious problem that is directly connected to the topic of this article. In a competitive market, we think of the unit price as capturing all the price information a prospective buyer needs to allow him to decide whether to buy a particular good. Stated differently, a buyer should be indifferent between two purchases that take place at the same price.<sup>5</sup> This raises the question of whether a lender will actually be indifferent between making two loans (purchasing two securities) that have the same return ratio. Suppose, for instance, that a lender has a choice between making a \$10,000 loan that repays \$12,500 at the end of two years, and a \$10,000 loan that repays \$12,500 at the end of five years. Each of these loans has the same return ratio. Which is he likely to choose?

It seems fairly obvious that our hypothetical lender will prefer the former of these loans to the latter: the former loan repays the same amount at an earlier date. The fact that the two loans have identical return ratios is not enough to make this lender indifferent between them.

The return ratio is flawed because it neglects an important aspect of securities transactions that is absent from most goods transactions. This aspect is the *time dimension*. A securities transaction is an exchange that takes place over an interval of time, and the length of the interval is important to the parties in the transaction. Lenders are likely to be less interested in the total amount to be repaid than in the amount to be repaid per unit of time.

How can we adjust the return ratio to take the time dimension into account? If all loans had the same term, no adjustment would be needed. Fortunately, any loan with a term of more than one period can be expressed as a sequence of one-period loans with identical one-period return ratios. A five-year loan, for example, can be expressed as a sequence of five one-year loans with a common annual return ratio. We can use these annual-equivalent return ratios to compare the returns on loans with different terms.

In order to be more concrete about this statement, we need to define some notation. Let's call the current date "date 0" and the maturity date of a given security "date N," so that the term of the security is N periods. From now on we will think of the periods as years; this is convenient, but not essential. Let  $V_0$  represent the amount lent and  $V_N$  the amount repaid. The return ratio on the loan is thus  $V_N/V_0$ , and the *per-period* (usually annual) *return ratio* is:<sup>6</sup>

$$R \equiv \sqrt[N]{\frac{V_N}{V_0}}$$

We can compute this ratio for any single-payment loan, as long as we know the amount lent, the amount repaid and the term. It provides us with exactly what we are looking for: a numerical yardstick that can be used to

<sup>3</sup>For the moment, we will make the (inaccurate) assumption that all loans/securities return a single payment at a fixed maturity date.

<sup>4</sup>Since prospective lenders always have the option of storing their money, the discount ratio should always be less than one. (No lender with this option will make a loan that returns less money than he lent.)

<sup>5</sup>We must assume that the goods do not differ in quality, and that price information is freely available. We must also assume that the goods are readily divisible, so that any quantity can be purchased at the given unit price. These are standard assumptions in the theory of competitive markets.

<sup>6</sup>The symbol "≡" should be read "is equal, by definition, to."

compare the returns on any two loans, regardless of their terms.<sup>7</sup>

To conform to financial market practice, we must modify the annual return ratio a little further. Market participants like to divide the repayment on loans into two components: one equal to the amount lent, which is called the *principal*, and another representing the remainder, which is called the *interest*.<sup>8</sup> They measure the return on loans as ratios of the interest to the principal. In our notation, market participants think of these returns in terms of *net return ratios*

$$r \equiv \frac{V_N - V_0}{V_0} = \frac{V_N}{V_0} - 1.$$

Unfortunately, the net return ratio suffers from the same problems of term comparison as the return ratio. However, we can define a *net per-period* (again, usually annual) *interest rate* by

$$r \equiv \sqrt[N]{\frac{V_N}{V_0}} - 1 = R - 1,$$

which is a per-period version of  $r$ . The annual interest rate serves as the financial market's basic measure of the attractiveness of the returns on securities. Very often it is converted into a percentage by multiplying it by 100.

If the annual interest rate truly serves as the analogue of the market price for securities, we can expect that in a competitive market it will be determined by the interaction of supply and demand. Financial market participants will face a *market interest rate*  $r^*$ , which they will view as beyond their power to influence, and will make their borrowing and lending decisions accordingly.<sup>9</sup>

### Pricing Securities

The annual interest rate formula can be used to determine the price of a security: the amount

a person who comes to the market offering to make a fixed repayment, at a fixed date in the future, will be able to borrow. If we let  $V_N$  represent the repayment a borrower promises to make exactly  $N$  years in the future, then he will be able to borrow (sell his security for) an amount  $V_0$ , where

$$V_0 = \frac{V_N}{(1+r^*)^N}.$$

This is the basic formula for "pricing" (or discounting) securities.

So far, we have assumed that all loans/securities return a single payment at a fixed maturity date. We know that in practice, however, most securities return multiple payments at multiple future dates. As long as the amounts and dates of these payments are known, we can simply price them separately and sum them to obtain the security's total price, or *present value*

$$V_0 = \frac{V_1}{1+r^*} + \frac{V_2}{(1+r^*)^2} + \dots + \frac{V_N}{(1+r^*)^N} = \sum_{t=1}^N \frac{V_t}{(1+r^*)^t}.$$

The present value of a sequence of future payments is the current market value of those payments, where the market value is determined by discounting the future payments back to the present at the market interest rate. Here,  $V_t$  represents the payment at the end of any date  $t$  (if there is no payment at a particular date  $\hat{t}$ , we say that  $V_{\hat{t}} = 0$ ) and  $1/(1+r^*)^t$  represents the discount factor applied to that payment.

*Secondary Market Pricing*—We are now ready to confront a pair of questions that are crucial in understanding the term structure. First, suppose the owner of a security wants to sell it before it comes due—that is, in the *secondary* market. How much can he expect to receive for it?

<sup>7</sup>Suppose we construct a sequence of one-period loans  $\{(V_0, V_1), (V_1, V_2), \dots, (V_{N-1}, V_N)\}$ , where  $V_j$  represents the amount lent at date  $j$ , and  $V_{j+1}$  is the amount repaid one period later. This sequence has the properties that (1) the amount lent at date 0 is  $V_0$ , (2) the amount repaid at date  $N$  is  $V_N$  and (3) the amount repaid on the  $t^{\text{th}}$  loan in the sequence, at any intermediate date  $t+1$ , is identical to the amount lent on the  $t+1^{\text{st}}$  loan, which is extended at the same date. (Thus, the loans are "rolled over" from date to date.) Properties (1) through (3) guarantee that, from the lender's point of view, this sequence of one-period loans is identical to the multiperiod loan. It turns out that only one sequence of loans satisfies these three properties and is consistent with our requirement that the return ratios on each loan be

identical. This is the sequence produced when each successive one-period loan is extended at a return ratio of  $R$ , as defined above.

<sup>8</sup>Part of the reason for this is that, as was noted above, anyone contemplating making a loan has the option of "lending to himself" by simply storing the money. As a result, people are unlikely to make loans unless the dollar repayment exceeds the dollar principal—that is, unless they receive interest.

<sup>9</sup>Hereafter, the "\*" superscript signifies that this particular value of the annual interest rate  $r$  is the one selected by the market.

The key to answering this question is to recognize that from a lender's point of view, a security purchased in the secondary market is essentially identical to (is a perfect substitute for) a security he might purchase in the *primary* or new issue market. The primary-market substitute would have a term equal to the *remaining term* on the secondary security—the number of years the security has left to run. It would return payments in the same amounts, and at the same dates, as the remaining payments on the secondary security—those that have yet to be made and would consequently be collected by the security's purchaser.

We can use this substitution principle, along with what we have just learned about primary-market pricing, to price a security sold in the secondary market. We will call the date at which the security is sold date  $T$ , and the price of the security at that date  $V_T$ . The remaining term of the security is then  $N-T$ , and its remaining payments are due at dates  $T+1, T+2, \dots, N-1, N$ .<sup>10</sup> The payments are consequently due  $1, 2, \dots, N-T-1, N-T$  periods in the future, relative to date  $T$ . (We'll assume that the payment due at date  $T$  has already been made.) Continuing our notational convention that subscripts represent dates, we'll let  $r_T^*$  denote the market interest rate at date  $T$ . We can then write

$$\begin{aligned} V_T &= \frac{V_{T+1}}{1+r_T^*} + \frac{V_{T+2}}{(1+r_T^*)^2} + \dots + \frac{V_N}{(1+r_T^*)^{N-T}} \\ &= \sum_{i=1}^{N-T} \frac{V_{T+i}}{(1+r_T^*)^i} \end{aligned}$$

It is important to note that  $r_T^*$ , the market rate on the date when the security is sold, may be different from the market rate when the security was issued (which we will call  $r_0^*$ ). If  $r_T^*$  is relatively low then the secondary market price  $V_T$  will be relatively high, and vice versa. This dependence of current secondary market prices on current interest rates (and of future secondary market prices on future interest rates) will

play a key role in our ultimate explanation for the slope of the yield curve.

*Interest Rates and Yields*—The securities pricing formula just presented can be used to help us tackle a second important question. Suppose we have a multiple payment security that is selling in the market at a known price. This could be either a newly issued security or a security sold in the secondary market. What is the annual interest rate on the security?

Since this security returns multiple payments, we cannot apply the annual interest rate formula that was presented on page 39. We can, however, exploit the fact that the annual interest rate on this security must be the rate that gives it its current market price—that is, the rate that makes the present value of its stream of future payments equal to its market price. Consequently, the market interest rate  $r_T^*$  must solve the equation

$$V_T = \frac{V_{T+1}}{1+r_T} + \frac{V_{T+2}}{(1+r_T)^2} + \dots + \frac{V_N}{(1+r_T)^{N-T}}$$

Here,  $V_T$  is the price of the security—which we are now assuming that we know—and  $V_{T+1}, V_{T+2}, \dots, V_N$  are the remaining payments on the security.

Since this equation has only the single unknown  $r_T$ , we might expect to be able to solve it to obtain  $r_T^*$ .<sup>11</sup> This is usually accomplished using numerical methods. These methods proceed by starting with a guess for  $r_T^*$ , computing the associated present value, and adjusting the guess according to the sign and size of the difference between this value and the actual market price. An annual interest rate computed in this manner—that is, as the solution to a present value equation—is called a *yield*.<sup>12</sup>

We have now—finally!—learned enough to begin investigating the term structure of interest rates. One way to start is by constructing a *yield curve*: a diagram which, as noted above, displays the relationship between the remaining terms of, and the yields on, different securities.

<sup>10</sup>Some of these payments may be zero. In the case of a single-payment security, for example, there is only one remaining payment; it is received at date  $N$ .

<sup>11</sup>The fact that this equation is not linear rules out standard algebraic solution methods. If the security in question has only two payments remaining (if  $N-T=2$ ), the equation can be transformed into a second-order polynomial equation and solved using the quadratic formula.

<sup>12</sup>For most of the rest of this paper the terms "interest rate" and "yield" will be used interchangeably. Unfortunately, participants in financial markets compute what they call interest rates on securities in a variety of ways, and some of them are significantly different from yields. These differences can be particularly important for securities with terms of less than a year. For details, see Mishkin (1989), pp. 82-92.

A problem we must confront in doing this is that many factors other than different remaining terms can cause differences in the yields on securities. These include differences in credit risk (that is, in the likelihood of default by the borrower) and in tax treatment. To isolate yield differences that are due solely to term differences, we need to compare the yields on securities that do not differ in these other characteristics. One simple way to do this is to compare the yields on securities issued by the U.S. Treasury. Treasury securities are issued with a wide variety of terms and are traded in a large and active secondary market—a fact that makes it possible to obtain a secondary market yield quotation for virtually any term. Treasury securities can also be thought of as essentially riskless, since the federal government is the only organization in the United States that can legally print money to cover its debts. Finally, the interest on all these securities is taxed on the same basis.

### THE EXPECTATIONS THEORY

What does economic theory have to say about the term structure? As with most questions in economics, there are a number of differing views. The theory described below, however, is accepted, at least in part, by most economists interested in monetary and financial issues. It is called the expectations theory.<sup>13</sup>

A basic challenge for term structure theory is to explain two empirical regularities, or “stylized facts,” of the interest rate term structure. These regularities can be described as facts about the slope or steepness of the yield curve at different points in time. One of them involves the direction the yield curve usually slopes: most of the time, the yield curve is gently upward-sloping. Another involves circumstances that seem to produce curves with unusual slopes: when short-term interest rates are relatively high, the yield curve is often downward-sloping; when short-term rates are relatively low, the curve is often steeply upward-sloping.

### *Linking Short-Term and Long-Term Interest Rates*

A point of departure for the expectations theory is the role of secondary markets in transforming the effective terms of securities. Suppose, for example, that a lender owns a five-year Treasury bond which he purchased in the primary market. The bond is maturing, but the lender now wishes he had lent for 10 years. If he takes the maturity payment on his five-year bond and uses it to purchase a second five-year bond, he will, in effect, have lent for 10 years. The only difference between this and the single 10-year loan is that the rate of return the lender receives over the coming five years will be determined by current market conditions, rather than conditions five years ago.

Suppose, conversely, that this lender owns a 10-year Treasury bond which he purchased five years ago. He has now decided that he needs his money and would have preferred to have lent for five years. If there were no secondary market, he would be stuck: he would not be repaid by the Treasury until the bond matured five years in the future. The secondary market allows him to receive early repayment indirectly, by selling his bond to another lender. If he chooses to sell the bond, he will, in effect, have lent for five rather than 10 years. The only difference between this and a true five-year loan is that the amount of the repayment (the sale price of the bond) will depend on current market conditions, rather than conditions five years ago.

Now suppose (rather unrealistically) that there is no uncertainty about future interest rates, so that lenders today know exactly what market yields on securities with different terms will be five years in the future. Suppose further that they know that the future five-year Treasury yield will be identical to the current five-year yield—say, 7½ percent. How will this affect the current yield on 10-year Treasury securities?

<sup>13</sup>Early statements of the expectations theory include various works of Irving Fisher [see the citations listed by Wood (1964), p. 457, footnote 1]. The theory was elaborated by Keynes (1930), Lutz (1940) and Hicks (1946); these authors proposed a variant of the expectations theory that has become known as the liquidity premium theory. This variant will be described at some length below. The most prominent alternative to the expectations theory is the market segmentation theory of Culbertson (1957). Another variant of the expectations theory, which combines elements of both the liquidity premium and market segmentation

theories, is the preferred habitat theory of Modigliani and Sutch (1966).

We can answer this question by process of elimination, ruling out possibilities that are clearly wrong until we are left with a single one that must be right. Suppose first that the current yield on 10-year Treasury bonds is higher than  $7\frac{1}{2}$  percent. We have seen that if a lender sells such a bond after five years, the yield to maturity its buyer will receive must be exactly the same as the yield on a newly issued five-year bond he might purchase instead. This future five-year yield, we have assumed, will be exactly  $7\frac{1}{2}$  percent. Consequently, the (five-year) yield the original lender will receive when he sells the 10-year bond, after holding it for five years, must be higher than  $7\frac{1}{2}$  percent: otherwise, the bond's 10-year yield, which is the average of its yields for the first and second five years, could not exceed that figure. But if it is possible to obtain a five-year yield of more than  $7\frac{1}{2}$  percent by purchasing a 10-year bond and selling it after five years, why would any current lender buy a newly issued five-year bond, or a secondary bond with five years left to run—each of which, according to our assumptions, will yield exactly  $7\frac{1}{2}$  percent? Clearly, if five-year bonds are to survive in the current market, the current yield on 10-year bonds must not in fact be higher than  $7\frac{1}{2}$  percent.

Now suppose that the current yield on 10-year bonds is lower than  $7\frac{1}{2}$  percent. Then if a lender buys a five-year bond today, he will receive a yield over five years that is higher than the 10-year yield. If he wants to lend for 10 years, he can use the maturity payment on the first five-year bond to purchase a second five-year bond. Since we have assumed that the yield on this second bond will be exactly  $7\frac{1}{2}$  percent, the average yield he receives over the 10-year period will also be exactly  $7\frac{1}{2}$  percent. This average yield is higher than the 10-year bond yield, however; consequently, no current lender will buy a 10-year bond. If 10-year bonds are to survive in the current market, their yields must not in fact be lower than  $7\frac{1}{2}$  percent.

We have just seen that if five- and 10-year bonds are to coexist in the market, the 10-year bond yield can be neither higher nor lower than the five-year bond yield. This means, of course, that it must be equal to the five-year yield. An argument of the same sort could be applied, with equal ease, to any long term, and

any pair of short terms that sum to it. Thus, under these assumptions, *if lenders know that short-term rates will remain constant in the future, current long-term rates must be equal to current short-term rates, so that the yield curve will be perfectly flat.*

Now suppose that instead of knowing the five-year rate will remain constant for the next 10 years, we know it will remain constant (at  $7\frac{1}{2}$  percent) for five years, and then rise to 10 percent. What must the current rate on a current 10-year security be? Notice that if a lender purchases a five-year bond yielding  $7\frac{1}{2}$  percent today, and rolls it over for a second five-year bond yielding 10 percent, he will receive an average annual rate of  $8\frac{3}{4}$  percent over the 10-year period. Under the circumstances, he would be foolish to lend for ten years at any rate lower than  $8\frac{3}{4}$  percent. Conversely, suppose that the U.S. Treasury wishes to borrow for a period of 10 years. If it borrows by issuing a five-year bond and then rolls the loan over for a second five years, it pays an average annual rate of  $8\frac{3}{4}$  percent. Clearly, it would be foolish to offer more than  $8\frac{3}{4}$  percent on its 10-year bonds.

Extending this argument to different long terms and different combinations of short terms that sum to them leads to the following prediction: *if there is no uncertainty about future interest rates, current long-term rates must be an appropriately weighted average of current and future short-term rates.*

Notice that, for the purposes of this prediction, a "long" term does not have to be long by conventional standards. A two-year rate, for instance, must be a weighted average of current and future one-year rates, while a six-month rate must be a weighted average of current and future three-month rates, etc. Clearly, it would be helpful to have a baseline "very short-term" rate to organize these sorts of predictions around. A natural candidate would be the rate on a riskless security with a term of zero.

What kind of security has a zero term? One example would be a security on which you can get your money back at any time. We have securities like this in the form of *demand deposits* or checking accounts. While these deposits are not issued by the U.S. Treasury, the fact that they are insured by the federal government makes them virtually as safe as Treasury securities.<sup>14</sup> We can consequently

<sup>14</sup>Strictly speaking, this is true only for personal deposits, and only up to a maximum of \$100,000 per deposit.

define the baseline interest rate,  $r^0$ , as the rate on a perfectly safe zero-term security and identify it in practice with the market rate on federally insured bank deposits.<sup>15</sup>

We can now state a mathematical rule for determining the rate of interest on a security with a term of  $N$  as a function of the base rate  $r^0$ . (We must continue to assume that financial market participants know the levels of future rates.) Let  $r_0^0$  represent the current rate of interest on a federally insured demand deposit. Let  $r_K^0$  represent the value of this same rate beginning at date  $K$ , when there will be a one-time, permanent change in the rate. Let  $r_0^N$  represent the current rate on a security with a term of  $N$ . (We will refer to  $r_0^N$  as a term-adjusted rate; the rationale for this usage will become clear later in the paper. Notice that we are letting subscripts represent dates, and superscripts terms to maturity.) Then

$$(*) \quad \begin{aligned} r_0^N &= r_0^0, \quad 0 < N \leq K, \text{ and} \\ r_0^N &= \frac{r_0^0 K + r_K^0 (N - K)}{N}, \quad N > K. \end{aligned}$$

The coefficients  $K$  of the current base rate  $r_0^0$ , and  $N - K$  of the future base rate  $r_K^0$ , are the appropriate weights referred to in the italicized prediction on page 42. Here,  $K$  is the number of years at which the base rate will stay at its original level  $r_0^0$ , and  $N - K$  is the number of years at which it will stay at its new level  $r_K^0$ . While this formula has been stated for the case in which market rates will change only once, it is easy to generalize to cover the case of multiple base rate changes.<sup>16</sup>

A yield curve drawn under the assumption that lenders know that the base rate will fall in the near future (that  $K$  is not very large, and that  $r_K^0 < r_0^0$ ) is displayed in figure 1.<sup>17</sup>

The assumption that lenders have complete and perfect knowledge about future interest rates is not very realistic. A more reasonable assumption might be that there is some uncertainty about future rates, but that lenders know their *expected values*—that is, their best forecasts, given the information available. If lenders base their decisions entirely on these best forecasts, then formula (\*) is still a valid description of the expectations theory provided that the rate  $r_K^0$  is interpreted as the expected value of the term-zero rate at date  $K$ . People who behave like this—those who base their decisions entirely on the forecast provided by the expected value—are said to be *risk neutral*.

*Systematic slope changes*—We can now explain one of the two empirical regularities identified in the introduction: the fact that yield curves tend to be steeply upward-sloping when short-term interest rates are low and often slope downward when short-term rates are high. Before we can do this, however, we need to consider what we mean when we say that interest rates are “low” or “high.” Is a 20 percent short-term rate high, for example? In the United States, the answer to this question is almost certainly “yes.” In Israel, or Argentina, however, the answer to the same question would almost certainly be “no.” This is because in recent U.S. history interest rates have rarely risen as high as 20 percent and, when they have done so, have quickly returned to lower levels. In recent Israeli or Argentinian history,

<sup>15</sup>A complication arises because demand deposit accounts do not pay interest, while functionally equivalent checkable accounts [negotiated order of withdrawal (NOW) accounts and money market deposit accounts (MMDAs), for example] are interest-bearing. Most economists believe that demand deposits pay interest indirectly, since banks that issue them typically do not charge fees that cover the costs of maintaining the accounts and providing funds transfer (checking, etc.) services. These issues are discussed and the implicit interest rates on demand deposits estimated by Klein (1974) and Dotsey (1983), among others. We will interpret  $r^0$  as this implicit demand deposit rate, or, equivalently, as the explicit interest rate on NOW accounts or MMDAs issued by institutions that do charge cost-covering fees. Under this interpretation,  $r^0$  will be a positive number.

<sup>16</sup>Suppose we know that the base rate will change at future dates  $K_1, K_2, \dots, K_J$ , and that the new base rates at these dates will be  $r_{K_1}^0, r_{K_2}^0, \dots, r_{K_J}^0$ . For notational convenience, call the current date (heretofore date 0) date  $K_0$ . Then the

current term-adjusted rate on a security with a term of  $N$  ( $N > K_J$ ) will be given by

$$(**) \quad r_0^N = \frac{\left[ \sum_{i=0}^{J-1} r_{K_i}^0 (K_{i+1} - K_i) \right] + r_{K_J}^0 (N - K_J)}{N}.$$

Both formulas (\*) and (\*\*) are approximations of the exact formulas. For details, see the shaded insert on the following page.

<sup>17</sup>Along the horizontal axis in figure 1,  $N$  represents a particular term longer than  $K$ , and  $r_0^N$  the term-adjusted rate on a security with that term. Since  $N$  is fairly close to  $K$ , the weighted average that determines  $r_0^N$  is strongly influenced by the  $K$  years at which the base rate will remain at its original, high level  $r_0^0$ . As the term lengthens, the influence of this period wanes and the term-adjusted rate gets closer and closer to the new, lower base rate  $r_K^0$ .

## The Exact Formula Linking Short- and Long-Term Rates

Both formula (\*) and the generalized version presented in footnote 16 are linearized approximations of the exact formula. The exact version of formula (\*) states that, if  $r_0^0$  is the current base rate, and  $r_K^0$  is the base rate at date K, then the current N-period term-adjusted rate  $r_0^N$  satisfies the relationship

$$(1+r_0^N)^N = (1+r_0^0)^K (1+r_K^0)^{N-K},$$

which implies

$$r_0^N = \sqrt[N]{(1+r_0^0)^K (1+r_K^0)^{N-K}} - 1.$$

If we know the base rate will change at future dates  $K_1, K_2, \dots, K_J$ , and that the new base rates at these dates will be  $r_{K_1}^0, r_{K_2}^0, \dots, r_{K_J}^0$  [again, for notational convenience, calling the current date (date 0) date  $K_0$ , and the terminal date (date N) date  $K_{J+1}$ ], then the current term-adjusted rate on a security with a term of N ( $N > K_J$ ) satisfies the relationship

$$(1+r_0^N)^N = \prod_{i=1}^{J+1} (1+r_{K_{i-1}}^0)^{K_i - K_{i-1}},$$

[here  $\prod$  is the multiplicative analogue of  $\sum$ ].

This implies

by contrast, rates have rarely fallen as low as 20 percent and, when they have done so, have quickly returned to higher levels.

When we say that interest rates are high or low, what we usually mean is that they are high or low relative to recent historical experience, and that we feel this experience gives us a good deal of guidance about the level of interest rates in the future. Thus, when we say interest rates are high we usually expect them to fall in the future, and vice versa. As we have just seen, the expectations theory predicts that when we expect rates to fall the yield curve

$$r_0^N = \sqrt[N]{\prod_{i=1}^{J+1} (1+r_{K_{i-1}}^0)^{K_i - K_{i-1}}} - 1.$$

Fortunately, the approximations given by the linearized formulas are adequate for most purposes. In the case described on pp. 42 of the text, for instance, the yield given by the exact formula is 8.743 percent, compared to the linearized figure of 8.750 percent.

The expectations theory can also be shown to imply that, if  $r_0^N$  is the current N-period term-adjusted rate, and  $r_K^0$  is the current K-period rate, then  $r_K^{N-K}$ , the term-adjusted rate on (N-K)-period securities that is expected to prevail at date K, satisfies the relationship

$$(1+r_K^{N-K})^{N-K} = (1+r_0^N)^N / (1+r_0^K)^K,$$

which implies that

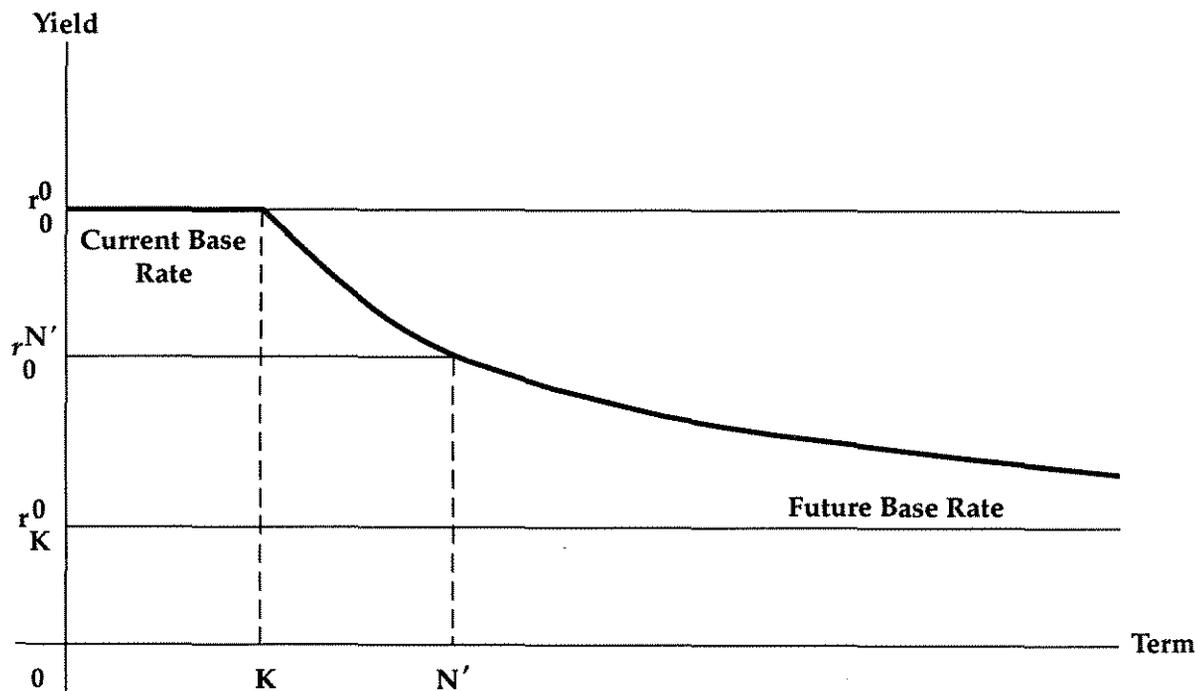
$$r_K^{N-K} = \sqrt[N-K]{(1+r_0^N)^N / (1+r_0^K)^K} - 1.$$

The rate  $r_K^{N-K}$  is often referred to as the "K-period forward rate" on a security with a term of N-K. The expectations theory is often described as a theory that identifies the forward rate with the expected future spot rate.

will slope downward, and that when we expect them to rise the curve will slope upward.

The simple expectations theory has the virtue of great flexibility: if you are willing to make sufficiently artful assumptions about lenders' expectations about the pattern of future interest rates, you can use this theory to explain the shape of virtually any yield curve. The theory provides an explanation for one basic empirical regularity about yield curves that is rather difficult to believe, however. The regularity in question is that most of the time, during the last century at least, yield curves have been distinctly

**Figure 1**  
**Term-adjusted Yield When the Base Rate Will Fall in the Future**



upward-sloping.<sup>18</sup> The simple expectations theory could explain this only by assuming that lenders usually expect rates to rise persistently over time. This assumption does not seem plausible, unless you believe that lenders were extremely poor forecasters. While interest rates have varied considerably during the past century, there is little evidence that they have increased on average, or that market participants had any reason to expect them to do so. Indeed, the evidence suggests that people usually expect future short-term interest rates to remain near current levels.<sup>19</sup> What we need, then, is a modified version of the theory that will predict

an upward-sloping yield curve under this assumption.

### *Interest Risk, Term Premia and the Slope of the Yield Curve*

Any alternative explanation for the fact that yield curves are normally upward-sloping must be based on something about long-term securities that makes them systematically less attractive to lenders than short-term securities. As we have just seen, the expectations theory predicts that, if lenders know for certain that short-term interest rates will remain constant, they should

<sup>18</sup>See Malkiel (1970), pp. 5-6, 12; Kessel (1965), pp. 17-19; and Shiller (1990), p. 629. It is sometimes asserted that yield curves were usually downward sloping during the late nineteenth and early twentieth centuries: see Meiselman (1962), appendix C, and Homer and Sylla (1991), pp. 317-22, 403-09 for descriptions and explanations of this phenomenon.

<sup>19</sup>The simple statistical models of interest rate behavior that explain the data best are based on the assumption that

rates have a long-run average or mean level and tend to return toward that level, rather slowly, after departing from it. These models imply that, if short-term interest rates are currently near their mean level (where they should be most of the time), they should be expected to stay near the current level in both the short and the long run, and that, even if they are far from the mean level, they should be expected to stay near the current level in the short run.

be indifferent between lending by purchasing short-term securities and lending by purchasing long-term ones. Long- and short-term interest rates should consequently be equal, and the yield curve should be flat. This prediction implies that any alternative explanation for the upward slope of the curve must be based on the effects of uncertainty about future interest rates.

*Interest Risk and Capital Losses*—One reason why uncertainty about interest rates may influence the behavior of lenders is that unanticipated changes in interest rates affect the value of their securities in the secondary market. Suppose, to return to an earlier example, that a lender buys a 10-year security that returns a yield of 7½ percent and sells it in the secondary market after five years. If interest rates have remained unchanged in the interim, the secondary market price of his security will give him a five-year yield of 7½ percent. If they have risen, the price will be lower, and he will receive a lower yield.

As we have already noted, the reason for these price and yield changes is that a security sold in the secondary market must compete with primary market securities with the same term as its remaining term. If the market interest rate on primary securities has risen, the yield on secondary securities must rise to the same level; since the remaining payments on these securities are fixed, this rise can be arranged only through a decline in the securities' market price. A formal way to see this is by inspecting the secondary market pricing formula for a single-payment security:

$$V_T = \frac{V_N}{(1+r_T^*)^{N-T}}$$

If  $r_0^* = r_T^*$ , so that interest rates have not changed since this security was issued, its price will be

$$\hat{V}_T = \frac{V_N}{(1+r_0^*)^{N-T}}$$

It is easy to check, by applying the annual interest rate formula, that both the T-year *ex post* yield on this security (the yield from date 0, when it was issued, to date T, when it is sold) and the N-T year *ex ante* yield (the yield from date T, when it is sold, to date N, when it will mature) are equal to the initial rate  $r_0^*$ .

We will call  $\hat{V}_T$  the anticipated price of this security. If the actual price  $V_T$  exceeds the anticipated price  $\hat{V}_T$ , we say the original lender has experienced a *capital gain*. The amount of the gain is simply  $V_T - \hat{V}_T$ . If the anticipated price falls short of the actual price, the lender has experienced a capital loss in the amount  $\hat{V}_T - V_T$ . It is clear from our pricing formula that capital gains occur if  $r_T^*$  falls short of  $r_0^*$  (if market interest rates have fallen), and vice versa. This means that lenders' expectations about future capital gains and losses must be tied to their expectations about future interest rates.

What should we assume about expectations regarding future interest rates? As we noted toward the end of the previous section, it seems reasonable to assume that market participants recognize that interest rates may change, but expect them to remain constant on average.<sup>20</sup>

<sup>20</sup>The expectations theory offers no explanation for the reasons market participants might expect short-term rates to change. It is a theory that attempts to explain the levels of long-term interest rates relative to the current levels of short-term rates, not one that attempts to explain their absolute levels. Stated differently, the expectations theory is not a true theory of the *determination* of interest rates. Market participants may expect short-term interest rates to change because they expect changes in any of the innumerable factors economic theory predicts might influence them.

Economic theory suggests that interest rates of the sort discussed in this article (money or *nominal* interest rates) are sums of real interest rates (rates expressed in terms of the purchasing power of the dollar amounts lent and repaid) and expected rates of inflation. This is the so-called Fisher equation. As a result, the question of interest rate determination is sometimes thought of as two questions: what determines real interest rates, and what determines inflation expectations. Most economists believe that nominal factors (such as changes in the levels or growth

rates of monetary aggregates) play the principal role in driving inflation expectations, while real factors (such as technological changes, changes in the perceived attractiveness of investment opportunities, changes in demographic structure or changes in the nature of financial regulation) play the principal role in real interest rate determination. There is, however, considerable disagreement about the degree of interaction between nominal and real factors, and especially about whether changes in nominal factors can have persistent effects on real interest rates.

Under this assumption, the expected capital gains on future secondary market sales of securities are approximately zero.<sup>21</sup>

It seems conceivable that this situation might not bother lenders. Economists usually assume, however, that the satisfaction a person derives from an extra dollar's worth of expenditures declines as the total value of his expenditures increases. If this is so, he will find the gain in satisfaction provided by the extra goods he can purchase if his returns exceed his expectations to be smaller than the loss in satisfaction from the goods he will have to refrain from purchasing if his returns fall short of his expectations. This should cause actuarially fair (zero expected loss) uncertainty about the future returns on his securities to upset him. A person who behaves like this is said to be *risk averse*.

Since buying term securities exposes lenders to actuarially fair return uncertainty, while buying securities with zero terms (such as demand deposits) does not, risk averse lenders will be reluctant to buy term securities. They will insist on higher expected yields on term securities than on demand deposits to compensate themselves for the uncertainty. The notion that financial decisionmakers are risk averse is widely accepted by economists, and we will adopt it without further discussion.

*Interest Risk and the Term Structure*—We have just explained why term securities tend to have higher yields than demand deposits when both are default-free: term securities carry interest risk, but demand deposits do not. We have not yet explained why securities with longer terms tend to have higher yields than those with shorter ones. Our discussion certainly suggests a possible explanation, however: longer-term securities may carry more interest risk than shorter-term ones. But why should this be the case?

We will begin our investigation of this question by posing another question that is closely related. Suppose we have two single-payment securities with different terms, but the same original (date 0) prices and yields. If market interest rates remain unchanged, their current (date T) prices will also be identical, even

though their maturity payments will not be. But suppose that the market interest rate—specifically, the market “base rate”  $r^0$ —rises by a fixed amount from date 0 to date T (so that  $r_T^0 = r_0^0 + \Delta r$ , with  $\Delta r > 0$ ). Which security will fall furthest in price?

Notice that the remaining term of the short-term security will be smaller than that of the long-term security; if we call the short term  $N_s$ , and the long term  $N_l$ , then the remaining terms of these securities are  $N_s - T$  and  $N_l - T$ , respectively. Since market yields have risen, the short-term secondary security must generate extra interest to compete with newly-issued short-term securities. The amount of extra interest will be approximately  $\Delta r V_T (N_s - T)$ ; this is the rate increase  $\Delta r$ , applied to the (common) secondary market price  $V_T$ , for each year of the remaining term ( $N_s - T$ ). The long-term security must also generate extra interest; in this case, the amount is  $\Delta r V_T (N_l - T)$ . This is the same rate increase, applied to the same base price, but continued for  $N_l - N_s$  additional years.

Of course, neither security can really produce “extra interest” in the conventional sense. The interest is paid indirectly, as part of the maturity payment, and the time and date of that payment are fixed. Instead, the price of each security must decline far enough so that it can increase at the new (and higher) annual rate  $r_T^0$ , while still reaching the fixed maturity payment  $V_N$  at the maturity date N. Since the price of the long-term security will have to increase at this rate for a much longer time, it will have to fall much further than the price of the short-term security. The relative sizes of the two price declines will be approximately equal to the relative sizes of the securities' remaining terms. A security with four years left to run will suffer a price decline approximately double that of a security with the same secondary market price but only two years left to run, and so on.

*The Term Premium*—If the risk of capital loss on securities tends to increase in proportion to their remaining terms, lenders who demand interest compensation for bearing this risk will demand more compensation on long-term

<sup>21</sup>Since the secondary market price is computed by dividing the maturity payment by the gross interest rate  $1+r$ , an increase in the rate by a given percentage causes a fall in the price that is slightly smaller than the rise in the price caused by an equal percentage decrease in the rate.

As a result, the expected price change is slightly positive. Although this effect is never very strong, it becomes more pronounced as the remaining term of the secondary security increases.

securities than on short-term securities.<sup>22</sup> This will tend to make the yields on longer-term securities higher than those on securities with shorter terms—that is, it will tend to make the yield curve upward-sloping.<sup>23</sup>

We can define the *term premium* on Treasury securities of a given term as the difference between the yield on those securities and the yield on federally insured demand deposits. That is,

$$\tau^N = r^N - r^0, \text{ or equivalently } r^N = r^0 + \tau^N,$$

where  $r^N$  represents the yield on N-term Treasury securities, and  $\tau^N$  represents their term premium. We now have a theory that predicts that the term premium should increase systematically with the remaining term, and, more specifically, that it should increase *in proportion* to the remaining term. We can formalize this by writing

$$\tau^N = \tau(N) \equiv mN,$$

where  $m$  is a positive constant of proportionality. A plot of the sort of yield curve consistent with this prediction is displayed in figure 2.

We might refer to the number  $m$  as the *coefficient of risk aversion*. Different values of  $m$  can be thought of as indicating different degrees of lenders' risk aversion. If  $m$  is relatively high, a small increase in the term and, thus, in the risk of capital loss, will cause lenders to demand a good deal of compensation in the form of a large increase in the term premium. This is the kind of behavior we would expect from very risk-averse lenders. If  $m$  is low, on the other hand, it will take a large increase in the term, and, thus, the risk to cause lenders to demand much additional compensation.

This is the kind of behavior we would expect from lenders who are not very risk-averse.<sup>24</sup>

It was pointed out earlier that lenders may not always expect the level of short-term interest rates—in particular, the level of the term-zero rate—to stay constant on average. When they do not, the base rates to which the term premia must be added will also depend on the term. These term-dependent base rates have been referred to as *term-adjusted rates*, and their current values have been denoted  $r_0^N$ . The actual yield should be the sum of the term-adjusted rate and the term premium:

$$r_0^N = r_0^N + \tau^N.$$

*Abnormal Yield Curves*—This latest addition to the expectations theory allows us to consider the role of the term premium in determining the shape of abnormal yield curves—the sort that appear when lenders expect interest rates to change in the future. In this case, the actual yield should be given by the sum of the term-adjusted rate (that is, the weighted-average base rate) and the appropriate term premium. This can produce curves that slope in one direction along one part of their range, but in the opposite direction along another part. If lenders expect interest rates to remain constant for a short period, and then fall sharply, for example, the yield curve will appear humped, sloping upward at very short terms, peaking near the term corresponding to the date at which rates are expected to decline, and sloping downward for a range of terms thereafter (see figure 3).<sup>25</sup> Curves with this shape are frequently observed shortly before economic recessions begin, presumably because interest rates tend to fall sharply during recessions.

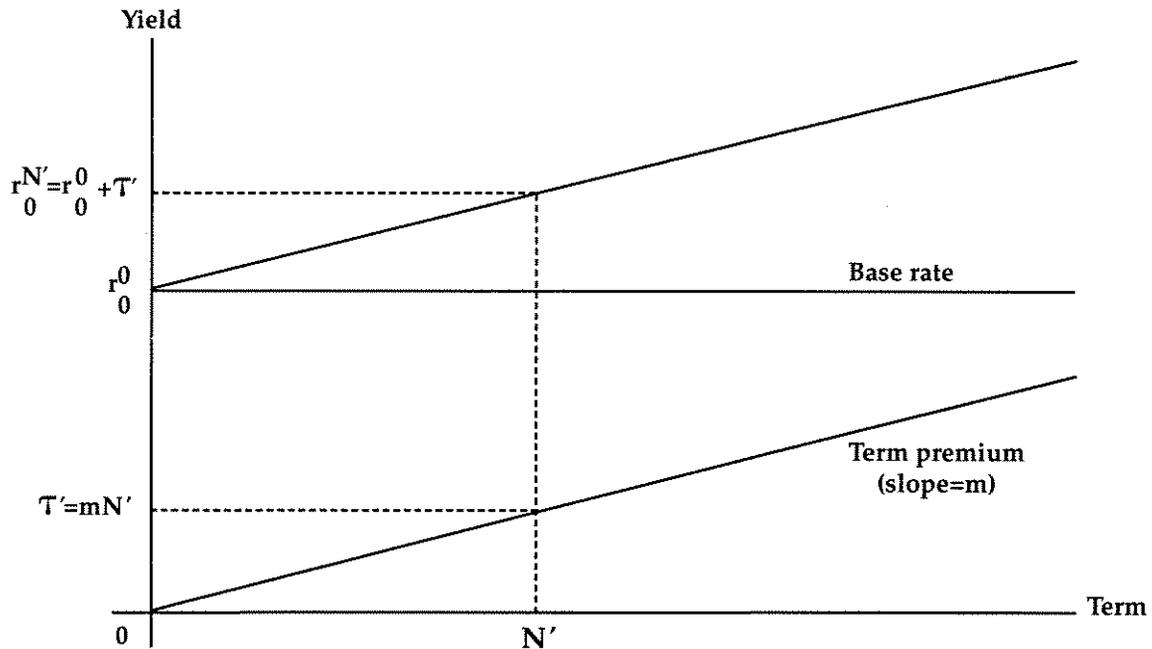
<sup>22</sup>In reality, the increase in risk is slightly less than proportional to the term, but the deviation from exact proportionality is very small. We are implicitly assuming that the change in the base rate, if any, will occur at a known future date, and that the rate, having changed, will remain at its new level permanently. We are also assuming that  $T$ , the date of sale, is fixed and known.

<sup>23</sup>Early statements of the liquidity premium theory include Keynes (1930), Hicks (1946) and Meiselman (1962). The term "liquidity premium" is based on the notion that liquidity—the ability to sell an asset rapidly and without loss—is valuable to lenders, and lenders will charge interest premia on assets that are relatively illiquid. Since the risk of capital loss is the risk that an asset may ultimately be saleable only at a loss, the premium for capital loss risk is in a sense a liquidity premium.

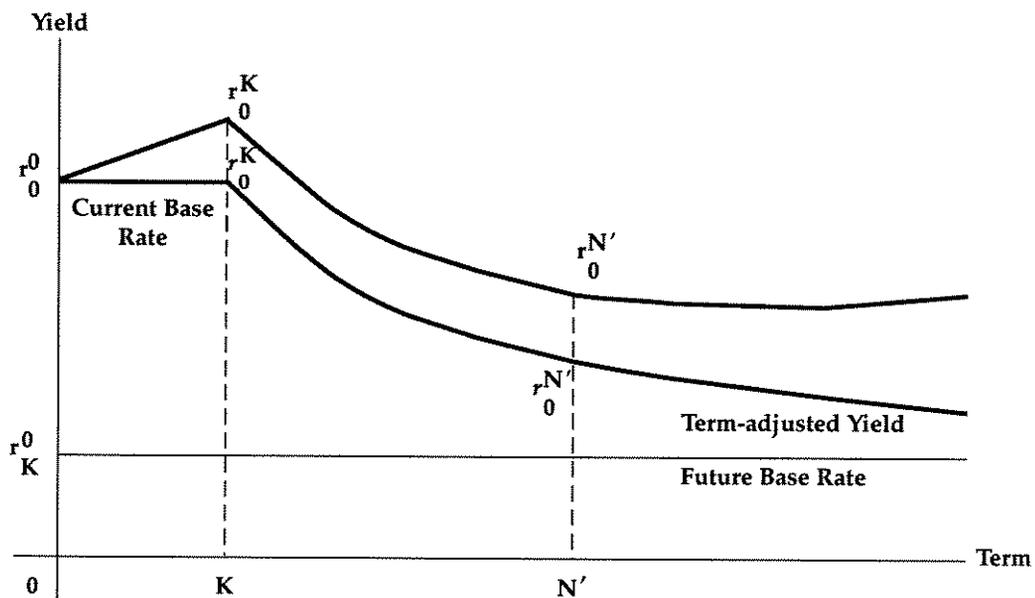
<sup>24</sup>If  $m = 0$ , lenders do not require any compensation for the risk of capital loss. As noted earlier, lenders who behave in this manner are said to be risk-neutral.

<sup>25</sup>Note that if a normal yield curve (a hypothetical curve observed when interest rates are expected to remain constant, on average) is upward-sloping, the expectations theory does not always interpret an upward-sloping yield curve as an indication that the market expects interest rates to rise. To obtain the right directional signal, the slope of the observed yield curve must be compared to the slope of a normal curve. The theory now interprets an observed yield curve that is upward-sloping, but flatter than normal, as a signal that the market expects interest rates to fall slightly.

**Figure 2**  
**Yield Curve When the Base Rate Is Constant**



**Figure 3**  
**Yield Curve When the Base Rate Will Fall in the Future**



## CONCLUDING REMARKS

This article presents a basic description of the concepts and issues involved in the study of the term structure of interest rates. It has also presented a simple version of the expectations theory of the term structure. This theory predicts that the shape of the yield curve is determined by the expectations of financial market participants about the level of future interest rates and by their uncertainty about the accuracy of their expectations.

The analysis presented here suggests that the expectations theory can help explain two important "stylized facts" about yield curves: the fact that the steepness and direction of their slopes tend to vary systematically with the level of short-term interest rates, and the fact that they are usually upward-sloping. The explanation for the former fact is that forward-looking lenders will refuse to purchase term securities unless long-term interest rates are averages of the short-term interest rates that the lenders expect at various points in the future. The explanation for the latter fact is that the interest risk on securities tends to increase with their terms; this causes risk-averse lenders to demand amounts of interest compensation that also increase with the terms.

## REFERENCES

- Culbertson, John M. "The Term Structure of Interest Rates," *Quarterly Journal of Economics* (November 1957), pp. 485-517.
- Dotsey, Michael. "An Examination of Implicit Interest Rates on Demand Deposits," Federal Reserve Bank of Richmond *Economic Review* (September/October 1983), pp. 3-11.
- Hicks, John R. *Value and Capital*, 2d ed. (Oxford University Press, 1946).
- Homer, Sidney, and Richard Sylla. *A History of Interest Rates*, 3d ed. (Rutgers University Press, 1991).
- Kessel, Reuben A. "The Cyclical Behavior of the Term Structure of Interest Rates," National Bureau of Economic Research Occasional Paper #91 (1965).
- Keynes, John M. *A Treatise on Money*, Vol. 1&2 (Harcourt, Brace and Company, 1930).
- Klein, Benjamin. "Competitive Interest Payments on Bank Deposits and the Long-Run Demand for Money," *American Economic Review* (December 1974), pp. 931-49.
- Lutz, Friedrich A. "The Structure of Interest Rates," *Quarterly Journal of Economics* (November 1940), pp. 36-63.
- Malkiel, Burton G. "The Term Structure of Interest Rates: Theory, Empirical Evidence, and Applications," Monograph (General Learning Corporation, Morristown, New Jersey, 1970).
- Meiselman, David. *The Term Structure of Interest Rates* (Prentice Hall, Inc., 1962).
- Mishkin, Frederic S. *The Economics of Money, Banking, and Financial Markets*, 2d ed. (Scott, Foresman and Company, 1989).
- Modigliani, Franco, and Richard Sutch. "Innovations in Interest Rate Policy," *American Economic Review* (May 1966), pp. 178-97.
- Russell, Steven H. "Reference Notes on Interest Rates, Securities Pricing, and Related Topics," Unpublished monograph, 1991.
- Shiller, Robert J. "The Term Structure of Interest Rates," (with an appendix by J. Huston McCulloch) in Benjamin M. Friedman and Frank H. Hahn, eds., *Handbook of Monetary Economics*, Vol. 1 (North Holland: Elsevier Science Publishing Company, Inc., 1990).
- Wood, John H. "The Expectations Hypothesis, the Yield Curve, and Monetary Policy," *Quarterly Journal of Economics* (August 1964), pp. 457-70.