Modeling the Term Structure of Interest Rates: An Introduction

MARK FISHER

The author is a senior economist on the financial team of the Atlanta Fed's research department. He thanks Jerry Dwyer, Scott Frame, Cesare Robotti, Paula Tkac, and Dan Waggoner for helpful comments.

The term structure of interest rates (also known as the yield curve) plays a central role—both theoretically and practically—in the economy. The Federal Open Market Committee (FOMC) conducts monetary policy by targeting interest rates at the short end of the yield curve. Longer-term yields reflect expectations of future changes in the funds rate by the FOMC. When the expectations about the FOMC's future funds rate moves change, the yield curve reacts as market participants reprice bonds to reflect these changes. In addition, longer-term yields reflect risk premia and convexity premia. (The convexity premia exist because bond yields depend on bond prices in a nonlinear way.) Consequently, the movement of longer-term yields reflects not only changes in expectations but also changes in these other forces.

In an earlier *Economic Review* article, Fisher (2001a), I examined these forces in the context of an extremely simple model in which all uncertainty was resolved by the single flip of a coin. Notwithstanding its simplicity, the model allowed me to present those ideas in an internally consistent way that provided a first view of the issues involved. This article is in some ways a sequel to Fisher (2001a). In the model presented here, some uncertainty is resolved each period, but additional uncertainty about future periods always remains. This article thus represents a big step in terms of the complexity of the analysis.

A model of the term structure is nothing more or less than a model of asset prices specialized to zero-coupon bonds. The central paradigm is this: Assets make payouts in the uncertain future. It is useful to think of the payouts as contingent on future states of the world in which the price of one dollar in each state of the world (that is, the state prices) is given. The value of an asset is the sum of the values of the state-contingent payouts:

$$
\begin{aligned}
\text{(1) asset value} &= \sum_{s\in S} \text{payout}_s \times \text{state price}_s \\
&= \sum_{s\in S} \text{payout}_s \times \overbrace{(\text{state-price deflator}_s \times \text{probability}_s)}^{\text{state price}_s} \\
&= \sum_{s\in S} \underbrace{(\text{payout}_s \times \text{state-price deflator}_s)}_{\text{deflated payout}_s} \times \text{probability}_s,
\end{aligned}
$$

where S is the set of states of the world at all future times. As indicated in equation (1), the state-price deflator equals the state price divided by the probability of the state; in other words, it is the value of a unit payout in a given state conditional on the occurrence of the state. The state-price deflator can be thought of as a stochastic process that evolves through time. The dynamics of the state-price deflator are intimately related to the interest rate and the price of risk, the two components of the price system (for asset pricing). The interest rate characterizes the expected change in the state prices while the price of risk characterizes the volatility of state prices.

Equation (1) reveals the relation between the absence of arbitrage opportunities and the existence of the state-price deflator. An arbitrage is a trading strategy that produces something for nothing. It can be shown that if all state prices are positive, then there are no arbitrage opportunities. Since the probabilities of the states are all positive (by definition), the absence of arbitrage opportunities implies the existence of a (strictly positive) state-price deflator.

The solution to the model of the term structure presented here illustrates a number of important features present to one extent or another in essentially all term structure models.

In a finance model of asset prices, there is no need to go any deeper. Indeed, the article will set up, solve, and calibrate a model of the term structure without any knowledge of, for example, how the nominal interest is related to expected inflation or how bond risk premia are related to investors' attitudes toward risk. Nevertheless, looking beneath the hood to see the connections can be informative. Therefore, Appendix A presents an economics model of asset prices in which the state-price deflator for real dividends (measured in units of consumption) can be identified with the marginal utility of a representative agent. The state-price deflator for nominal dividends can be immediately derived with the introduction of the price level (the price of consumption in terms of dollars). The dynamics of the two state-price deflators reveal the relationship between the real and nominal interest rates, a relationship that explicitly includes expected inflation and the agent's preferences (including risk aversion).

The purpose of this article is to show how to use absence-of-arbitrage conditions to solve for the term structure of interest rates in a discrete-time setting and to do so in a way that is largely independent of the time step. The contribution of this article is the exposition; the article presents no new results from the literature. Elsewhere one may find discrete-time models of asset pricing and the term structure that are essentially the same as the one presented here. The current exposition features two main novelties. First, as alluded to above, this article focuses on modeling the dynamics of the state-price deflator.[1] Second, the model keeps track of the length of the discrete time period. This step complicates the notation a bit, but it has a distinct advantage: By keeping track of the size of the time step, one can see what happens as it becomes arbitrarily small. Consequently, one can see what many of the continuous-time limits look like. In other words, this article provides a bridge from discrete-time models to continuous-time models without requiring the technical overhead necessary to directly perform a continuous-time analysis.[2]

Randomness and uncertainty play a central role in modeling the term structure, and the proper vocabulary is required to treat the subject cogently. The reader is assumed to be familiar with the notions of expectation, mean, variance, and covariance. The article will deal extensively with normal and lognormal random variables. Lognormality plays a very important role in the analysis. The important properties of lognormal random variables are outlined in Appendix B.

Knowledge of calculus is not required to follow the main argument. Calculus is referred to explicitly only in the footnotes; it is used implicitly in approximations.

Finally, the reader should be prepared to become familiar with a certain amount of notation, which is unavoidable when discussing the term structure of interest rates. Equation (6) exemplifies the notational complexity involved. A fairly comprehensive list of notations is presented in Table 1.[3]

Bond Prices and Yields

A discrete time model is adopted in which observations are made at discrete points in time. Time is measured in years. The step size, denoted h, is the length of time between observations (also referred to as the length of the period). For example, if $h = 1/12$ then the step size corresponds to one month.

Bond prices. Default-free zero-coupon bonds are the building blocks for the term structure of interest rates. A zero-coupon bond pays \$1 when it matures at time T. Let t denote the current time. Assume t ″ T. The bond's maturity (measured in years) is $\tau = T - t$. By contrast, the number of steps until the bond matures is given by $n = \tau/h$. For example, if $\tau = T - t = 2$, the bond has a maturity of two years. If, in addition, $h = 1/12$, the bond will mature after $\tau/h = 24$ steps of time. Upon occasion, we will imagine the step size h getting smaller and smaller

TABLE 1

Notation

h	the step size (i.e., length of a time step)
τ	remaining time until maturity of a bond: $\tau = T - t$
$p(t, T)$	price at time t of a (zero-coupon) bond that matures at time T
$\mu(t, T)$	expected return at time t of holding a bond that matures at time T
$y(t, T)$	yield at time t of a bond that matures at time T
$r(t)$	one-period risk-free interest rate at time t (nominal)
$f(t, T)$	forward rate at time t for a loan from time $T - h$ to T
$v(t)$	value of an asset at time t
$d(t)$	rate of dividend flow at time t
$\pi(t)$	state-price deflator at time t
κ_r	speed of mean reversion for the interest rate
θ_r	long-run mean of the interest rate
σ_r	interest rate volatility
$\varepsilon(t)$	interest rate shock
λ	price of risk (nominal)
$\sigma(t, T)$	relative volatility at time t of a bond that matures at time T
ξ	a combination of parameters: $\xi = \kappa_r\theta_r - \lambda\sigma_r$
ϕ_τ	term premium for forward rates at maturity τ
Φ_τ	term premium for zero-coupon yields at maturity τ

while holding both t and T fixed. In such a case, the number of steps until maturity will increase.

Let $p(t, T)$ denote the value at time t of a bond that matures at time T. When the bond matures (at time T) it will be worth its face value: $p(T, T) = 1$.[4] The discount function shows the relation between bond prices and maturity at a fixed point in time; it is obtained by plotting $p(t, t + \tau)$ versus τ for fixed t and $\tau \geq 0$. See Figure 1 for the discount function computed from bond prices on July 29, 1994.

Zero-coupon yields and forward rates. It is natural to express bond prices in terms of their implied yields. The yield to maturity on a zero-coupon bond (that has not yet matured) is defined as

$$y(t, T) := -\log[p(t, T)]/(T - t).$$

The yield to maturity is also known as the zero-coupon yield, or zero-coupon rate, or simply the yield. The yield is the continuously compounded annualized return that would be earned from holding the bond until maturity.[5] The yield curve shows the relation between yields and maturity at a fixed point in time; it is obtained by plotting $y(t, t + \tau)$ versus τ for fixed t and $\tau \geq h$. See Figure 2 for the yield curve computed from bond prices on July 29, 1994.[6] The (short-term, risk-free) interest rate, $r(t)$, is the yield on a one-period bond:

$$(2) \quad r(t) = y(t, t + h) = -\log[p(t, t + h)]/h.$$

Let me emphasize that, for many purposes, bond yields should be thought of simply as a way of

1. See Cochrane (2001) for a good example of a more standard approach to discrete-time asset pricing. Cochrane treats the stochastic discount factor as the central object of analysis. In fact, he treats the state-price deflator (which he calls the state-price density) as a synonym for the stochastic discount factor. Indeed, there is little point in distinguishing between the two in a setting in which the length of the period is always unity. However, for scenarios in which the length of the time step becomes shorter and shorter, the state-price deflator has a useful limit while the stochastic discount factor does not.
2. For an introduction to no-arbitrage conditions and modeling the term structure, consult Fisher (2001a); the companion working paper (Fisher 2001b) contains additional material in Part 2. For introductions to asset pricing in continuous time, see Baxter and Rennie (1996) and Neftci (1996). For a more advanced treatment, see Duffie (1996).
3. Additional notation used in Appendix A is shown in the appendix table.
4. If the bond were subject to default risk, then it might not be worth $1 when it matured.
5. See Fisher (2001a) for a discussion of continuous compounding and its relation to simple compounding.
6. Yield curves reported in the press are typically drawn using the yields on coupon bonds. Consequently, they do not display the characteristic downward curvature evident at the long end of the yield curve in Figure 2.

FIGURE 1

Discount Function Computed from Bond Prices on July 29, 1994

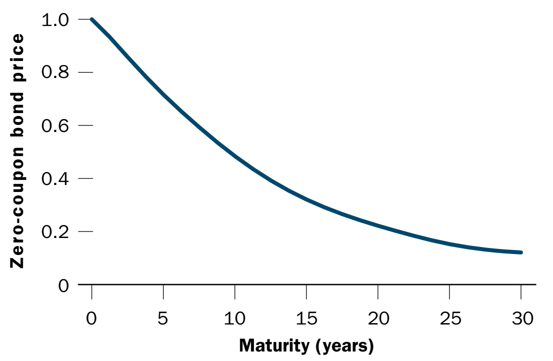


FIGURE 2

Zero-Coupon Yield Curve Computed from Bond Prices on July 29, 1994

expressing bond prices. If a bond's price is known, its yield can be computed; conversely, if a bond's yield is known, its price can be computed. The relation between the two is $p(t, T) = e^{-(T-t)y(t,T)}$.

The forward rate is defined as[7]

$$(3)\quad f(t,T) := -\left(\frac{\log[p(t,T)] - \log[p(t,T-h)]}{h}\right).$$

The forward rate is the yield that one can obtain at time t for a commitment to lend from time $T - h$ to time T.[8] Forward rates are closely related to zero-coupon yields. In particular, the zero-coupon yield $y(t, T)$ can be expressed as the average of the forward rates from $f(t, t + h)$ to $f(t, T)$:[9]

$$(4)\quad y(t,T) = \frac{1}{(T-t)} \sum_{i=1}^{(T-t)/h} f(t, t+ih)h.$$

As a special case, $r(t) = f(t, t + h)$. The relation between forward rates and zero-coupon yields can be inverted; for $T > t + h$, the forward rate $f(t, T)$ equals the yield $y(t, T)$ plus a term that depends on whether the yield curve is rising or falling from maturity $T - h$ to maturity T:[10]

$$(5)\quad f(t,T) = y(t,T) + (T-t-h)\left(\frac{y(t,T) - y(t,T-h)}{h}\right).$$

In economics terminology, the forward rate is the marginal yield obtained by extending the maturity by one period.

Expected return. Thus far, bond prices $p(t, T)$, yields $y(t, T)$, and forward rates $f(t, T)$ have all been considered from the perspective of the current time t when their values are known for sure. As time passes from t to $t + h$, new information arrives and bond prices, yields, and forward rates will change to $p(t + h, T)$, $y(t + h, T)$, and $f(t + h, T)$. When time $t + h$ arrives, these new values will be known. However, from the perspective of time t, they are uncertain—they are random variables.

In dealing with random variables, the conditional expectation operator $E_t[\cdot]$ is a tool that will be used repeatedly. Let x be a random variable. Then $E_t[x]$ is the mean of x, where the mean is computed using all the information available at time t. $E_t[x]$ can be referred to as the expected value of x conditional on the information available at time t.[11]

The expectation operator will be used to define the expected return on a zero-coupon bond over the next step of time. As time passes from t to $t + h$, the price of a given zero-coupon bond changes from $p(t, T)$ to $p(t + h, T)$.[12] As of time $t + h$, the realized return from holding the bond over the period will be $p(t + h, T)/p(t, T)$. At time t, the price next period (assuming $T > t + h$) is uncertain, and consequently the return is uncertain as well.[13] The expected return is $E_t[p(t + h, T)/p(t, T)]$. It is convenient to express the expected return over the period in terms of continuous compounding at an annualized rate. Let $\mu(t, T)$ denote the continuously compounded expected return at time t on a bond that matures at time T (expressed as an annual rate):[14]

$$(6)\quad \mu(t,T) := \frac{\log\left(E_t\left[\frac{p(t+h,T)}{p(t,T)}\right]\right)}{h}.$$

The term $\mu(t, T)$ will be referred to simply as the expected return. The expected return plays a central role in bond pricing and the term structure (as will be seen in a later section).[15]

Asset Pricing and the Absence of Arbitrage

This section first discusses asset pricing in general and then specializes the results to bond prices in particular.

The central idea in asset pricing is that the value of an asset—whether it is a share of stock, a bond, or a portfolio of other assets—is the present value of the expected "dividend" payments, where a dividend should be understood to include any sort of payment, positive or negative. The rate at which dividends are received is distinguished from the amount of dividends received. The total amount of dividends received during the current period is $d(t)h$, which equals the dividend rate $d(t)$ times the length of the period h.

The value of an asset at time t, $v(t)$, equals the value of the current dividend plus the present value of expected future dividends:[16]

$$(7)\quad v(t) = E_t\left[\sum_{i=1}^{\infty}\left(\frac{\pi(t+ih)}{\pi(t)}\right)d(t+ih)h\right],$$

where $\pi(t + ih)$ is the value of the *state-price deflator* at time $t + ih$. Equation (7) is a restatement of equation (1): The dividends correspond to the payouts, the expectations operator $E_t[\cdot]$ embodies the probabilities, and the dependence on the states has been suppressed. In equation (1), the value of the state-price deflator at time 0 was assumed to be one; $\pi(0) = 1$. From the perspective of time 0, the value of the asset at time $t > 0$ is $\pi(t)v(t)$. Dividing this value by $\pi(t)$ produces equation (7).

The state-price deflator plays the central role in asset pricing and consequently in bond pricing.[17] The state-price deflator is always positive, and it shrinks over time on average, making the discount factor less than one. The exposition will show that the positivity is closely related to the absence of arbitrage opportunities and that the shrinkage is closely related to the interest rate.[18]

The ratio $\pi(s)/\pi(t)$ for $s > t$ is called the *stochastic discount factor*. In many discrete-time presentations, it is the stochastic discount factor that is emphasized (for example, see Cochrane 2001). However, the state-price deflator is more useful for the purposes of this exposition because it is better behaved when the size of the time step is allowed to go to zero.

Absence of arbitrage. Equation (7) not only represents the value of an asset but also characterizes the absence of arbitrage opportunities. To demonstrate this, let us introduce self-financing trading strategies. A trading strategy involves buying at time t a portfolio of assets (which could simply be a single asset) and liquidating the portfolio at time $T > t$. During the intervening interval, the trading strategy could involve changing the composition of the portfolio (buying and/or selling assets); in addition, the assets in the portfolio could pay dividends or require additional financing (that is, negative dividends). A trading strategy is self-financing if the portfolio generates no net cash flows during the intervening interval. In order to make a trading strategy self-financing, the investor must (1) reinvest

7. Taking the limit as the step size h goes to zero in equation (3) produces $f(t, T) = -\frac{\partial}{\partial T}\log[p(t, T)]$.
8. Consider the following transaction. At time t buy one bond that matures at time T. The cost of this bond is, of course, $p(t, T)$. To pay for this bond, sell some bonds that mature at time $T - h$. (If you do not already own some of these bonds, you sell them short.) The number of bonds to be sold is $p(t, T)/p(t, T - h)$. Your net cash flow at time t is zero. The future cash flows are as follows. At time $T - h$ pay \$1 for every bond sold, while at time T receive a payment of \$1. Even though these cash flows are in the future, there is no uncertainty about them at time t, and therefore the continuously compounded annualized return on the money invested can be computed at time $T - h$ to time T: $(1/h)\log\{1/[p(t, T)/p(t, T - h)]\} = f(t, T)$.
9. Taking the limit as the step size h goes to zero in equation (4) produces $y(t,T) = (1/T - t)\int_0^{T-t} f(t,t+s)ds = (1/T - t)\int_t^T f(t, s)ds$.
10. Taking the limit as the step size h goes to zero in equation (5) produces $f(t, T) = y(t, T) + (T - t)\frac{\partial}{\partial T}y(t, T)$.
11. Obviously, if x is known for sure at time t, then $E_t[x] = x$. Moreover, if x and y are random variables and a and b are known at time t, then $E_t[ax + by] = aE_t[x] + bE_t[y]$.
12. Concurrently, the maturity of the bond becomes shorter: A bond with maturity $T - t = \tau$ at time t becomes (after one period of time) a bond with maturity $T - (t + h) = \tau - h$ at time $t + h$.
13. For a bond with one only one period remaining to maturity (that is, for which $T = t + h$), the price next period is known: $p(t + h, T) = p(T, T) = 1$.
14. The continuously compounded ex post return is $\log[p(t + h, T)/p(t, T)]/h$. Note that $E_t\{\log[p(t + h, T)/p(t, T)]/h\} \neq \mu(t, T)$ unless there is no uncertainty. This inequality is an example of what is known as Jensen's inequality.
15. The expected return from holding a one-period bond involves no uncertainty. By comparing equations (6) and (2), it can be seen that the expected return from holding a one-period bond is simply the interest rate; that is, $\mu(t, t + h) = r(t)$.
16. The existence of a well-defined price in equation (7) depends on the convergence of an infinite sum. If dividends grow too fast asymptotically relative to the discount factor, the sum will not converge.
17. See Duffie (1996) for an extensive discussion of the state-price deflator.
18. Appendix A shows how to compute the state-price deflator from the marginal utility of consumption. In this setting, the stochastic discount factor is the intertemporal marginal rate of substitution.

any dividends paid during the intervening interval in one or more assets, (2) sell assets (or borrow) to finance any additional required financing, and (3) finance any change in the composition of the portfolio in such a way as to generate no net cash flows at only two points in time: at inception (time t) and at liquidation (time T).

In equation (7), let $v(t)$ represent the cost of a self-financing trading strategy at inception. Upon liquidation, the portfolio generated by the self-financing trading strategy pays a lump-sum dividend equal to the value of the portfolio: $d(T)h = v(T)$. Equation (7) becomes, after both sides are multiplied by $\pi(t)$:

$$\text{(8)} \quad \pi(t)v(t) = E_t[\pi(T)v(T)].$$

Equation (8) embodies the *martingale* (or random-walk) property of deflated asset prices (that is, asset prices multiplied by the state-price deflator). The martingale property says that the expected deflated value of any self-financing portfolio equals the current deflated value. The martingale property implies the absence of arbitrage opportunities.

In general, an arbitrage amounts to getting something for nothing. In this setting, an arbitrage is a self-financing trading strategy that costs nothing ($v(t) = 0$), has no chance of losing money ($v(T) \geq 0$ for all possible outcomes), and has some chance of making money ($v(T) > 0$ for some possible outcome).

For a self-financing trading strategy that costs nothing, equation (8) can be written as

$$\text{(9)} \quad 0 = E_t[\pi(T)v(T)].$$

For equation (9) to hold, if there is any chance of making money (that is, if $v(T) > 0$ for some possible outcome), there must also be an offsetting chance of losing money (that is, $v(T) < 0$ for some possible outcome). Therefore, if equation (9) holds, there are no arbitrage opportunities.[19]

Bond pricing. Let us apply (7) to zero-coupon bonds. Recall that $p(t, T)$ is the current price (at time t) of a bond that pays a single unit dividend on the maturity date, time T; in other words, $v(t) = p(t, T)$ and $d(T)h = 1$.[20] Consequently, in this application, equation (7) becomes

$$\text{(10)} \quad p(t, T) = E_t[\pi(T)/\pi(t)].$$

In modeling zero-coupon bonds, it is convenient to have the value of the bond converge to its face value on its maturity date, so that $p(T, T) = 1$. Thus, the value of the bond will be assumed to include the current dividend on its maturity date.

It is worth emphasizing that equation (10) says that bond prices depend on the dynamics of the state-price deflator—and on nothing else. Therefore, a model of bond prices is nothing more or less than a model of the state-price deflator.[21]

The value of a perpetuity. A *perpetuity* is an asset that pays dividends at a constant rate in perpetuity (that is, forever). Formally, the dividend rate for a perpetuity is $d(t + ih) = 1$ for all $i \geq 0$. Therefore, the value of a perpetuity is given by

$$\text{(11)} \quad v(t) = E_t\left[\sum_{i=1}^{\infty}\left(\frac{\pi(t+ih)}{\pi(t)}\right)h\right] = \sum_{i=1}^{\infty} E_t\left[\frac{\pi(t+ih)}{\pi(t)}\right]h = \sum_{i=1}^{\infty} p(t,t+ih)h.$$

Thus, the value of a perpetuity is the sum of the values of all zero-coupon bonds, treating the bond prices as rates of flow.[22] For example, if the yields on all bonds were equal to a constant r, then $p(t, t + ih) = e^{-rih}$ and the value of a perpetuity would be

$$\sum_{i=1}^{\infty} e^{-rih}h = \frac{h}{e^{rh}-1} \approx \frac{1}{r},$$

where the approximation is accurate for small h.[23]

Dynamics of the State-Price Deflator

As noted above, modeling the term structure of interest rates is nothing more or less than modeling the dynamics of the state-price deflator, a task undertaken in this section. It is convenient to begin by modeling the dynamics of the interest rate.

The interest rate is assumed to evolve through time according to the following *stochastic difference equation*:[24]

$$\text{(12)} \quad r(t + h) - r(t) = \kappa_r[\theta_r - r(t)]h + \sigma_r\varepsilon(t + h).$$

Equation (12) expresses the change in the interest rate as the sum of two components: the expected change, $\kappa_r[\theta_r - r(t)]h$, and the unexpected change, $\sigma_r\varepsilon(t + h)$.[25]

In equation (12), κ_r, θ_r, and σ_r are fixed parameters, while $\varepsilon(t + h)$ is a random variable that is independent of what is known at time t. In particular,

$$\varepsilon(t + h) \sim N(0, h),$$

where $x \sim N(m, v)$ means the random variable x is distributed normally with mean m and variance v.

As a consequence, the conditional mean and variance of the unexpected change are

$$E_t[\sigma_r \varepsilon(t+h)] = \sigma_r E_t[\varepsilon(t+h)] = 0$$
$$E_t[(\sigma_r \varepsilon(t+h))^2] = \sigma_r^2 E_t[\varepsilon(t+h)^2] = \sigma_r^2 h.$$

Thus, the change of the interest rate, $r(t+h) - r(t)$, is normally distributed with mean $\kappa_r[\theta_r - r(t)]h$ and variance $\sigma_r^2 h$. Consequently, the rate of change (per unit of time) of the interest rate is normally distributed:

$$(13)\quad \frac{r(t+h)-r(t)}{h} \sim N\left(\kappa_r[\theta_r - r(t)], \sigma_r^2\right).$$

In Appendix C, it is shown that θ_r is the long-run (unconditional) mean of r, κ_r is the speed of reversion to the long-run mean, and $\sigma_r^2/(2\kappa_r)$ is (approximately) the long-run (unconditional) variance of r.

The bank account. Let $B(t)$ denote the value at time t of the "bank account" (sometimes called the money market account). Funds invested in the bank account earn the interest rate each period. Therefore, assuming no additional funds are deposited or withdrawn, the value of the bank account grows as follows: $B(t+h) = e^{r(t)h}B(t)$, or

$$(14)\quad \log[B(t+h) - \log[B(t)] = r(t)h.$$

There is no uncertainty about the rate of change of the value of the bank account from t to $t+h$. Nevertheless, there is uncertainty about future rates of change because there is uncertainty about the future interest rate. In other words, the return on the bank account is risk free over the next period, but it is not risk free over longer horizons.

Now we turn to modeling the dynamics of the state-price deflator, which first may appear to be completely arbitrary, as if produced out of thin air. It will turn out that the form of the dynamics (in terms of the interest rate and the price of risk) is completely determined by the structure of asset pricing. So the plan is first to "pull a rabbit out of a hat" and then to explain how (and why) the trick works.

For tractability, $\pi(t+h)/\pi(t)$ is assumed to be lognormally distributed.[26] Let

$$(15)\quad \log[\pi(t+h)] - \log[\pi(t)] = -\left(r(t) + \frac{1}{2}\lambda^2\right)h - \lambda\varepsilon(t+h),$$

where λ is the price of risk,[27] which is a fixed parameter in equation (15). Equation (15) implies that the rate of change of the log of the state-price deflator is normally distributed:

$$(16)\quad \frac{\log[\pi(t+h)] - \log[\pi(t)]}{h} \sim N\left(-r(t) - \frac{1}{2}\lambda^2, \lambda^2\right).$$

The dynamics of the state-price deflator are completely specified by the interest rate $r(t)$ and the

19. Note that this argument would not work if the state-price deflator were not always positive. In fact, the absence of arbitrage opportunities can be shown to imply the existence of a (strictly positive) state-price deflator.
20. In terms of the notation of equation (7), this means $d(s) = 1/h$ if $s = T$ and $d(s) = 0$ otherwise. As the length of the period h gets shorter and shorter, the dividend rate per period $d(T)$ must get larger and larger. The reader may be comforted to know that math that allows a sensible limit does exist.
21. Even though bond prices depend on nothing but the state-price deflator, they do not necessarily depend on "all" of the state-price deflator. To see this, suppose the discount factor can be expressed as the product of two factors: xy. In general, $E_t[xy] = E_t[x]E_t[y] + \text{Cov}_t[x, y]$. If $E_t[y] = 1$ and $\text{Cov}_t[x, y] = 0$, then $E_t[xy] = E_t[x]$. In this case, bond prices will depend only on x, and modeling y is irrelevant for bond prices. Nevertheless, other asset prices (such as equity prices) will depend on both x and y via their product.
22. Taking the limit as the step size h goes to zero in equation (11) produces $v(t) = \int_0^\infty p(t, t+s)ds = \int_t^\infty p(t, s)ds$.
23. The phrase "accurate for small h" means that the error due to the approximation can be made as small as desired by making h sufficiently small and that in the limit, as h approaches zero, the expression is exact.
24. Equation (12) can be expressed in the form of a first-order autoregressive process: $r(t+h) = \alpha + \beta r(t) + e(t+h)$ where $\alpha = \theta_r\kappa_r h$, $\beta = 1 - \kappa_r h$, and $e(t+h) = \sigma_r\varepsilon(t+h)$.
25. In the continuous-time limit, the two components are called the *drift* and the *diffusion*. To make the similarities with continuous time more apparent, we could write equation (12) this way: $d_h r(t) = \kappa_r[\theta_r - r(t)]d_h t + \sigma_r d_h W(t)$, where $d_h r(t) = r(t+h) - r(t)$, $d_h t = (t+h) - t = h$, and $d_h W(t) = W(t+h) - W(t) = \varepsilon(t+h)$. When equation (12) is interpreted as an approximation to the continuous-time limit, it is called the *Euler approximation*. It can be shown that the Euler approximation converges (as h goes to zero) in distribution and pathwise to a stochastic differential equation (see Kloeden and Platen 1995).
26. See Appendix B for a discussion of lognormality.
27. At this stage, the "price of risk" is simply the name of the negative of the relative volatility of the state-price deflator. In the discussion that follows, when the absence-of-arbitrage condition is reexpressed in terms of risk and return, it will become apparent why this name is appropriate. The minus sign on λ in the diffusion in equation (15) is merely a convenience (part of the trick).

price of risk λ. The fact that λ is a fixed parameter is a very special assumption that greatly simplifies the model but also greatly restricts its realism. Note that changes in the state-price deflator depend on the same shock that drives changes in the interest rate, $\varepsilon(t+h)$. In other words, there is a single source of uncertainty, which is another simplifying and restrictive assumption.

Having pulled the rabbit out of the hat, let us turn to the explanation of the trick, which amounts to demonstrating that the symbolic expressions $r(t)$ and λ in equation (15) do in fact correspond to the interest rate and the price of risk. We begin with the interest rate. Combining the definition of the interest rate given in equation (2) with the formula for pricing a one-period bond given by $p(t, t+h) = E_t[\pi(t+h)/\pi(t)]$ (see equation [10]), we obtain $E_t[\pi(t+h)/\pi(t)] = e^{-r(t)h}$, or

$$(17)\quad r(t) = \frac{-\log(E_t[\pi(t+h)/\pi(t)])}{h}.$$

Using the algebra of lognormal random variables, one can confirm that equation (17) does indeed hold given equation (15).

To explain how the trick works with regard to the price of risk, we need to derive the risk-return relation, which in turn requires a formal specification of the dynamics of bond prices.

Risk and return. If bond prices are assumed to be conditionally lognormally distributed, then the dynamics of the log of a bond's price can be expressed as follows:

$$(18)\quad \log[p(t+h,T)] - \log[p(t,T)] = \left(\mu(t,T) - \frac{1}{2}\sigma(t,T)^2\right)h + \sigma(t,T)\varepsilon(t+h).$$

Equation (18) implies that the rate of change of the log of bond prices is normally distributed:

$$\frac{\log[p(t+h,T)] - \log[p(t,T)]}{h} \sim N\left(\mu(t,T) - \frac{1}{2}\sigma(t,T)^2, \sigma(t,T)^2\right).$$

Using the algebra of lognormal random variables, one can confirm that

$$\frac{\log(E_t[p(t+h,T)/p(t,T)])}{h} = \mu(t,T),$$

as required by equation (6). Finally, note that

$$\mathrm{Cov}_t\left[\log\left(\frac{\pi(t+h)}{\pi(t)}\right), \log\left(\frac{p(t+h,T)}{p(t,T)}\right)\right] = -\lambda\sigma(t,T)h.$$

Here we see the first connection with λ.

Now that the dynamics of the state-price deflator have been established, the martingale property of deflated bond prices can be used to reexpress the condition for the absence of arbitrage in terms of risk and return.

Equation (10) can be written in terms of deflated bond prices:

$$(19)\quad \pi(t)p(t,T) = E_t[\pi(T)p(T,T)],$$

where, of course, $p(T, T) = 1$. Equation (19) holds for all t ″ T. This martingale property implies that the expected rate of change of $\pi(t)p(t, T)$ is always zero:

$$(20)\quad E_t\left[\frac{\pi(t+h)p(t+h,T) - \pi(t)p(t,T)}{h}\right] = 0.$$

Equation (20) can be expressed as

$$(21)\quad E_t\left[\left(\frac{\pi(t+h)}{\pi(t)}\right)\left(\frac{p(t+h,T)}{p(t,T)}\right)\right] = 1.$$

Equations (15) and (18) and the algebra of lognormally distributed random variables[28] can be used to express (21) as

$$(22)\quad \mu(t,T) = r(t) + \lambda\sigma(t,T).$$

Equation (22) expresses the well-known relation between risk and return: The expected return $\mu(t, T)$ equals the risk-free rate $r(t)$ plus the covariance-based risk premium $\lambda\sigma(t, T)$, which demonstrates why λ is called the price of risk.

Equation (22) is the same relation derived in Fisher (2001a). In that paper, this relation was obtained as a condition for the absence of arbitrage by directly examining potential arbitrage portfolios. In this article, equation (22) is derived as an implication of a very powerful proposition (that is, the martingale property implies no arbitrage opportunities) that does not require actually looking at any portfolios.

Thus far, $\mu(t, T)$ and $\sigma(t, T)$ are purely formal: They are merely place holders. The task of a model of the term structure is to solve equation (22) for $\mu(t, T)$ and $\sigma(t, T)$ in terms of the parameters that determine the dynamics of the interest rate and the state-price deflator (κ_r, θ_r, σ_r, and λ). This task is addressed in the following section.

Solving the Term Structure Model

The solution for bond prices $p(t, T)$ will depend on the term to maturity $\tau = T - t$ and any state variables that appear (either directly or indirectly) in the risk and return condition of equation (22) via the interest rate and the price of risk. In the model in this article, the interest rate is itself a state variable.[29] In fact, it is the only state variable. This statement can be confirmed as follows: First, since the price of risk λ is a constant parameter in this model, no state variables are required to describe its evolution. Second, the evolution of the interest rate depends only on the current value of the interest rate itself, as shown in equation (13).

Since the interest rate is the only state variable in this model, the solution for bond prices will depend only on maturity and the interest rate. Suppose that the solution for bond prices can be expressed as $p(t, T) = P(r(t), T - t)$, where

$$(23) \quad P(r, \tau) = e^{-a(\tau)-b(\tau)r}.$$

(This conjecture will be confirmed below once the model is actually solved). For each value of τ, there is a pair of coefficients $a(\tau)$ and $b(\tau)$. A solution to the term structure model amounts to computing the sequences $\{a(\tau)\}_{\tau=0}^{\infty}$ and $\{b(\tau)\}_{\tau=0}^{\infty}$ in terms of the parameters κ_r, θ_r, σ_r, and λ, where $\tau = 0, h, 2h, 3h\ldots$ The condition $p(T, T) = 1$ implies $P(r, 0) = 1$ for all values of r, which in turn implies $a(0) = 0$ and $b(0) = 0$. These two conditions provide the initial conditions for the sequences $\{a(\tau)\}_{\tau=0}^{\infty}$ and $\{b(\tau)\}_{\tau=0}^{\infty}$.

Equation (23) makes a very strong statement about bond prices. First, it asserts a specific functional form.[30] Second, it says that absolute time does not matter: t and T enter only through their difference, $T - t$. Third, it says the effect of the interest rate on bond prices is independent of what may have occurred in the past.

Before equation (23) is used to express the dynamics of bond prices, it is first used to express the forward rate (see equation [3]):

$$(24) \quad f(t, T) = A_f(T - t) + B_f(T - t)r(t),$$

where

$$A_f(\tau) = \frac{a(\tau)-a(\tau-h)}{h} \quad \text{and}$$

$$B_f(\tau) = \frac{b(\tau)-b(\tau-h)}{h}.$$

As we will see shortly, the absence-of-arbitrage condition amounts to imposing restrictions on the sequences of forward rate coefficients $\{A_f^{\tau}\}_{\tau=1}^{\infty}$ and $\{B_f^{\tau}\}_{\tau=1}^{\infty}$.[31]

Now we turn to the dynamics of bond prices. Using equation (23), we can write[32]

$$(25) \quad \begin{aligned} &\log[p(t+h,T)] - \log[p(t,T)] \\ &= (-a_{\tau-h} - b_{\tau-h} r_{t+h}) - (-a_{\tau} - b_{\tau} r_t) \\ &= \underbrace{(a_{\tau} - a_{\tau-h}) + (b_{\tau} - b_{\tau-h}) r_t}_{f(t,t+\tau)h} - b_{\tau-h}(r_{t+h} - r_t), \end{aligned}$$

where $\tau = T - t$. As indicated in equation (25), the change in the bond price is composed of two parts, one of which depends on the forward rate and the other of which depends on the change in the interest rate (that is, the change in the state variable) as given by equation (13). Collecting terms into drift and diffusion, we have

$$\begin{aligned} &\log[p(t + h, T)] - \log[p(t, T)] \\ &= \{[A_f(T - t) + B_f(T - t)r(t)] \\ &\quad - [b_{\tau-h}\kappa_r(\theta_r - r(t)]\}h - b_{\tau-h}\sigma_r \varepsilon(t + h). \end{aligned}$$

Therefore, by matching coefficients with equation (18), we see that

$$(26) \quad \begin{aligned} \mu(t,t+\tau) = &[A_f(\tau) + B_f(\tau)r(t)] \\ &- b(\tau-h)\kappa_r[\theta_r - r(t)] \\ &+ \frac{1}{2} b(\tau-h)^2 \sigma_r^2; \end{aligned}$$

$$(27) \quad \sigma(t, t + \tau) = -b(\tau - h)\,\sigma_r.$$

Equation (26) indicates that the expected return is composed of three parts: (1) the part due to the passage of time (the forward rate), (2) the part due to the expected change in the interest rate, and (3) the part

28. See Appendix B. Let $x_1 = \log[\pi(t + h)/\pi(t)]$ and $x_2 = \log[p(t + h, T)/p(t, T)]$.

29. In other models, the interest rate may be a function of the state variables. For example, we could have $r(t) = x_1(t) + x_2(t)$, where x_1 and x_2 are the state variables.

30. This form of bond prices is called *exponential affine* because according to equation (23), $\log[p(t, T)] = -a(\tau) - b(\tau)r(t)$ is affine in the state variable $r(t)$ (that is, linear in $r(t)$ plus a constant). There is a broad class of exponential affine models of the term structure. See Duffie and Kan (1996).

31. In the limit as h goes to zero, $A_f(\tau) = a'(\tau)$ and $B_f(\tau) = b'(\tau)$, where the prime sign indicates differentiation.

32. To make the notation more compact for space considerations, subscripts will be used on occasion in place of parentheses to denote arguments for time and maturity. In particular, $a_{\tau} \equiv a(\tau)$, $b_{\tau} \equiv b(\tau)$, $r_t \equiv r(t)$, and so forth.

due to the nonlinearity of the relation between the interest rate and bond prices (Jensen's inequality).[33]

Now these expressions are inserted into equation (22) and rearranged to obtain

$$(28)\quad \left[A_f(\tau)-b(\tau-h)\xi+\frac{1}{2}\sigma_r^2 b(\tau-h)^2\right] +\left[B_f(\tau)+b(\tau-h)\kappa_r-1\right]r(t)=0,$$

where

$$\xi := \kappa_r\theta_r - \lambda\sigma_r.$$

Equation (28) is the absence-of-arbitrage condition for our simple model. In order for equation (28) to hold for every possible value of the state variable $r(t)$, each term in square brackets must be equal to zero. Setting each term equal to zero and rearranging produces the following pair of difference equations:

$$(29)\quad \frac{a(\tau)-a(\tau-h)}{h}=b(\tau-h)\xi-\frac{1}{2}\sigma_r^2 b(\tau-h)^2,$$

$$(30)\quad \frac{b(\tau)-b(\tau-h)}{h}=1-b(\tau-h)\kappa_r,$$

subject to the initial conditions $a(0) = b(0) = 0$.[34] (The left-hand sides of equations (29) and (30) are simply $A_f(\tau)$ and $B_f(\tau)$.)

The solution to equation (29) – equation (30) is

$$(31)\quad a(\tau)=K_1\xi+K_2\sigma_r^2,$$

$$(32)\quad b(\tau)=\frac{1-(1-\kappa_r h)^{\tau/h}}{\kappa_r}\approx\frac{1-e^{-\kappa_r\tau}}{\kappa_r},$$

where

$$K_1:=\frac{\tau}{\kappa_r}-\frac{1-\left(1-h\kappa_r\right)^{\frac{\tau}{h}}}{\kappa_r^{\,2}}\approx\frac{\tau}{\kappa_r}-\frac{1-e^{-\kappa_r\tau}}{\kappa_r^2}$$

$$K_2:=-\frac{\left[\left(1-h\kappa_r\right)^{\frac{\tau}{h}}-1\right]\left[2h\kappa_r+\left(1-h\kappa_r\right)^{\frac{\tau}{h}}-3\right]}{2\kappa_r^{\,3}\left(h\kappa_r-2\right)}-\frac{\tau}{2\kappa_r^{\,2}}$$

$$\approx\frac{3+e^{-2\kappa_r\tau}-4e^{-\kappa_r\tau}}{4\kappa_r^{\,3}}-\frac{\tau}{2\kappa_r^{\,2}}.$$

(The approximations are accurate for small h.[35]) Note that the only parameter K_1 and K_2 depend on is κ_r. At this point, we can substitute the solution for $a(\tau)$ and $b(\tau)$ into $p(t, t+\tau) = e^{-a(\tau)-b(\tau)r(t)}$ and confirm that this expression satisfies the absence-of-arbitrage condition given the way the state-price deflator has been modeled.

Although there are four parameters in the dynamics of the interest rate and the state-price deflator (κ_r, θ_r, σ_r, and λ), the solution for $a(\tau)$ and $b(\tau)$ depends on only three parameters: κ_r, σ_r, and ξ. Suppose we fix κ_r, σ_r, and ξ. Then for any given θ_r we can find a value for λ that is consistent (and vice versa). In order to pin down the value of θ_r (and hence the value of λ), we must turn to the dynamics of the interest rate.[36]

Forces That Shape the Yield Curve

The three forces that shape the yield curve are (1) the expected future interest rate, (2) the risk premium, and (3) the *convexity* effect due to the nonlinear relation between bond prices and rates (Jensen's inequality). This section examines various shapes of the yield curve and how they depend on those forces as expressed in terms of the parameters of the dynamics of the interest rate and the state-price deflator.

We begin by examining the forward rate curve, which displays most clearly the various forces. In Appendix C the dynamics of the interest rate are used to compute an expression for the expected future interest rate in terms of the current interest rate (see equation [C.6]):

$$\begin{aligned} E_t[r_{t+\tau}] &= \theta_r + (1-\kappa_r h)^{\tau/h}(r_t-\theta_r) \\ &= \theta_r + (1-\kappa_r b_\tau)(r_t-\theta_r), \end{aligned}$$

where the solution for b_τ in equation (32) is used in the second line. Given this expression and equation (30), the forward rate can be expressed as

(33)

$$\begin{aligned} f(t,t+\tau) &= A_f^\tau + B_f^\tau r_t \\ &= [\overbrace{(\kappa_r\theta_r-\lambda\sigma_r)}^{\xi} b_{\tau-h}-\frac{1}{2}\sigma_r^2 b_{\tau-h}^2]+(1-\kappa_r b_{\tau-h})r_t \\ &= \underbrace{[\theta_r+(1-\kappa_r b_{\tau-h})(r_t-\theta_r)]}_{E_t[r_{t+\tau-h}]}+\underbrace{[-\lambda\sigma_t^\tau-\frac{1}{2}(\sigma_t^\tau)^2]}_{\phi_\tau}. \end{aligned}$$

A similar decomposition for zero-coupon yields can be obtained recalling equation (4):

$$\begin{aligned} y(t,t+\tau) &= \frac{1}{\tau/h}\sum_{i=1}^{\tau/h} f(t,t+ih) \\ &= \frac{1}{\tau/h}\sum_{i=1}^{\tau/h} E_t[r_{t+(i-1)h}]+\Phi_\tau, \end{aligned}$$

where $\Phi_\tau := \frac{1}{\tau/h} \sum_{i=1}^{\tau/h} \phi_{ih}$.

At this point, we note that our simple term structure model satisfies the weak form of the expectations hypothesis. By way of background, the *strong form* of the expectations hypothesis says that the current forward rate equals the expected future interest rate:

$$(34)\ \ f(t, t+\tau) = E_t[r_{t+\tau-h}].$$

When there is no uncertainty, the strong form of the expectations hypothesis is equivalent to the absence of arbitrage opportunities.[37] The *weak form* of the expectations hypothesis allows for the addition of a constant term premium (denoted by ϕ_τ) that depends on the forecast horizon τ (but not on time t):

$$(35)\ \ f(t, t+\tau) = E_t[r_{t+\tau-h}] + \phi_\tau.$$

The term premium ϕ_τ, as given in equation (33), decomposes into two parts: one that depends on the risk premium and the other that represents a convexity term. These two forces affect the shape of the yield curve in distinctly different ways. To highlight the difference, an approximation that is accurate for small κ_r can be used:

$$\sigma_t^\tau \approx \sigma_r\, \tau \quad \text{and} \quad (\sigma_t^\tau)^2 \approx \sigma_r^2\, \tau^2.$$

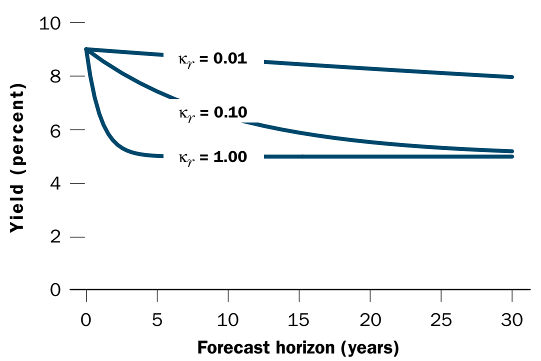
Thus equation (33) can be approximated by

$$(36)\ \ f(t, t+\tau) \approx E_t[r_{t+\tau}] - \lambda\, \sigma_r\, \tau - \frac{1}{2} \sigma_r^2\, \tau^2,$$

where it is also assumed that h is small. Equation (36) displays clearly the three forces that shape the forward-rate curve. The average slope (at the short end) is determined by the covariance between the state-price deflator and the interest rate. Therefore, to capture the average upward slope seen in the data, this covariance must be negative. In particular, if one assumes $\sigma_r > 0$, then one must have $\lambda < 0$.[38]

FIGURE 3

Paths for the Expected Future Interest Rate



Note: Paths assume a long-run mean of 5 percent and a current value of 1 percent. The paths differ by the speed of mean reversion κ_r.

The convexity term tends to offset the upward slope imparted by the risk premium. Although its effect disappears as maturities go to zero, it becomes large as maturities increase and can in fact dominate at the long end and actually cause the term structure to slope downward beyond about twenty years.

Here we examine how changes in the parameters of the dynamics of the interest rate and the state-price deflator affect the yield curve. First, we examine how the speed of mean reversion affects the path of expectations, $E_t[r_{t+\tau}]$. Figure 3 shows the effect of κ_r on the speed at which expectations revert to the mean. The effects of the other two parameters are straightforward. An increase in the risk premium stemming from the price of risk will increase the slope of the yield curve while an increase in the volatility of the interest rate will increase its curvature.

Calibrating the model. Now this simple model will be calibrated to the data to show what the term structure looks like. First the parameters of the interest rate dynamics, κ_r, σ_r, and θ_r, are estimated. Then, given these values, a value for λ (which pins down ξ) can be chosen to try to match actual yield curves.

33. In effect, we have applied *Ito's lemma*, the main workhorse for computing the dynamics of the transformation of a stochastic process in continuous time. Note that the so-called extra term in equation (26) from Jensen's inequality is present in discrete time as well.
34. In the limit as h goes to zero, equations (29) and (30) become a pair of ordinary differential equations: $a'(\tau) = b(\tau)\xi - (1/2)\sigma_r^2 b(\tau)^2$ and $b'(\tau) = 1 - b(\tau)\kappa_r$.
35. The approximations are the exact solutions to the differential equations described in footnote 34.
36. It is a general feature of term structure models that only a subset of the parameters that determine the dynamics of the state-price deflator and the state variables is identified in the cross-sectional bond-pricing solution.
37. For more on the implications of the expectations hypothesis when there is no uncertainty, see Fisher (2001a).
38. This result can be computed directly from equation (33); it does not depend on the approximations used in equation (36).

TABLE 2

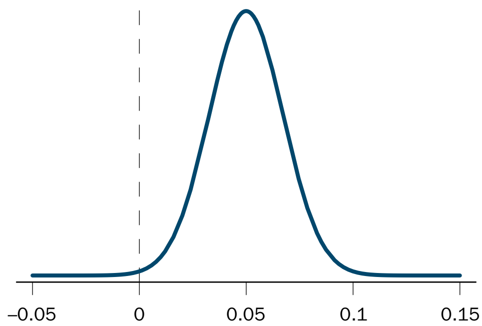
Estimated Parameter Values and Asymptotic Standard Errors

Parameters	θ_r	κ_r	σ_r
Estimated values	0.050	0.124	0.0086
Standard errors	(0.030)	(0.217)	(0.0007)

Note: Standard errors are computed using a Newey-West correction with five lags.

FIGURE 4

Estimated Unconditional Distribution for the Interest Rate



The *method of moments* is one way to estimate the parameters.[39] Let e_{t+h} denote the one-step-ahead forecast error e_t^h (see equation [C.9]). Then

$$\begin{aligned} e_{t+h} &= r_{t+h} - E_t[r_{t+h}] \\ &= r_{t+h} - [r_t + \kappa_r(\theta_r - r_t)h] \\ &= \sigma_r \varepsilon(t+h). \end{aligned}$$

The properties of $\varepsilon(t + h)$ imply the following *moment conditions*:

$$(37)\quad E_t[e_{t+h}] = 0, E_t[r_t e_{t+h}] = 0, \text{ and } E_t[(e_{t+h})^2] = \sigma_r^2 h.$$

These three moment conditions can be used to estimate the three parameters of the interest rate process: κ_r, θ_r, and σ_r.

Suppose there are $N + 1$ observations of the interest rate sampled with time step h: $r_0, r_h, r_{2h},\ldots, r_{Nh}$ and the total time span is $(N + 1)h$. Let $\hat{\kappa}_r$, $\hat{\theta}_r$, and $\hat{\sigma}_r$ denote the estimated values. Then, for $i = 1, 2,\ldots, N$, the estimated forecast error in terms of estimated coefficients is defined to be

$$(38)\quad \hat{e}_{ih} := r_{ih} - [r_{(i-1)h} + \hat{\kappa}_r(\hat{\theta}_r - r_{(i-1)h})h].$$

The sample moments that correspond to equation (37) are

$$(39)\quad \frac{1}{N}\sum_{i=1}^{N}\hat{e}_{ih} = 0, \quad \frac{1}{N}\sum_{i=1}^{N} r_{(i-1)h}\,\hat{e}_{ih} = 0, \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N}(\hat{e}_{ih})^2 = \hat{\sigma}_r^2 h.$$

The method of moments estimators $\hat{\theta}_r$, $\hat{\kappa}_r$, and $\hat{\sigma}_r$ are the solutions to equation (39).

Table 2 presents estimates of the interest rate parameters based on the estimation technique described above using ten years of monthly data for zero-coupon bond prices from December 1987 to November 1997 ($N = 119$ and $h = 1/12$).[40] The estimated unconditional mean of 5 percent is consistent with the average value of the interest rate over this period. The estimated half-life (the length of time it takes the expectation to return halfway to the long-run mean; see equation [C.7]) is $\log(2)/k_r = 5.6$ years, which—although it agrees with what others find—seems quite long. It is an indication of just how persistent the interest rate is on average.

The estimated unconditional distribution is shown in Figure 4. This distribution gives some probability to negative interest rates, but nominal interest rates (which is what we are dealing with here) cannot be negative.[41] The probability that the interest rate is negative is only 0.002, so this problem is not terribly important.

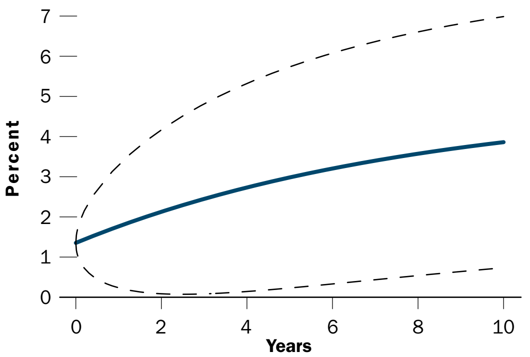
Figure 5 shows the conditional expectation of the interest rate, assuming that the current interest rate is at 1.25 percent, and error bands of two standard errors. After 5.6 years, the expected rate is 3.125 percent, which is halfway from 1.25 percent to the long-run average of 5.0 percent. The error band (at that point) is from 0.4 percent to 6.4 percent, effectively 3 percentage points on each side of the expectation.

Figure 6 shows what the zero-coupon yield curve looks like with the estimated parameters and various values for the price of risk λ. The curve with $\lambda = -0.5$ produces a yield curve that is not entirely unreasonable. Nevertheless, the curve does not display the characteristic downward slope beyond about twenty years. To obtain this feature, one must deviate somewhat from the estimated values. This is understandable: Since this term structure model is so simple, it cannot fit all of the major features of the term structure.

The way we have proceeded thus far is to fit the dynamics of the interest rate as well as possible (given the assumed form of the interest rate dynam-

FIGURE 5

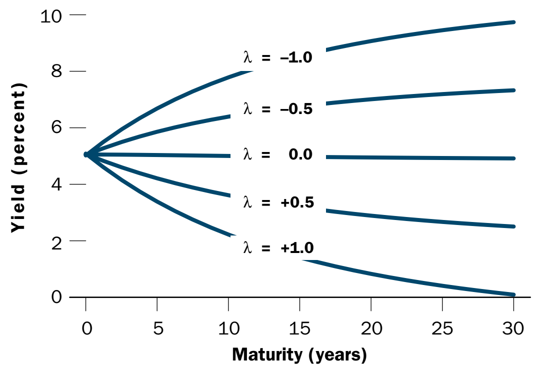
The Expected Path of the Short Interest Rate and Error Bands



Note: Each of the two error bands is two standard deviations from the expectation. The interest rate is set at 1.25 percent.

FIGURE 6

The Zero-Coupon Yield Curve Using Estimated Values for the Price of Risk



Note: The curve assumes r = 0.050. See Table 2 for various values of λ.

ics) and force all the errors on the yield curve. An alternative way to proceed is to fit the parameters to the average yield curve and then let the dynamics of the interest rate take up the slack.[42] Figures 7 and 8 show the zero-coupon yield curve when the parameters are fitted to the average yield curve. In Figure 7, the current interest rate r_T is set at its long-run average of 5 percent while in Figure 8 it is set at 1 percent.

Conclusion and Extensions

In the model of the term structure presented in this article, the interest rate follows a simple stochastic process, and the price of risk is a fixed parameter. Because the price of risk is fixed, the weak form of the expectations hypothesis holds in this model.[43] For this reason (and for others), the model cannot match many important features of the term structure for U.S. data. Nevertheless, the solution to the model presented here illustrates a number of important features present to one extent or another in essentially all term structure models. In addition, the steps followed to solve the model are the same steps one would follow to solve a more general model. To a large extent, therefore, the journey we have taken is itself the goal.

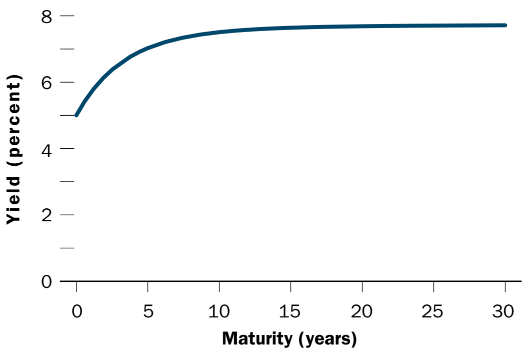
As noted, the model solved in this article satisfies the weak form of the expectations hypothesis. The expectations hypothesis has held a central position in the theory of the term structure of interest rates over the years. However, a substantial amount of evidence shows that actual data do not satisfy the expectations hypothesis.[44] Why does it fail to hold? At its core, the expectations hypothesis asserts that changes in forward rates solely reflect changes in forecasts of future interest rates. Indeed, forward rates do reflect expectations of future rates; but they also reflect other features that are essentially independent of expected future rates. However, in the model solved in this article, the only source of variation in forward rates is changes in expectations of the interest rate. To see this, note that equation (35) can be written as $f(t, T) = E_t[r_{T-h}] + \phi_{T-t}$. Therefore, the dynamics for the forward rate can be expressed as

$$f(t+h, T) - f(t, T)$$
$$= (E_{t+h}[r_{T-h}] - E_t[r_{T-h}]) + (\phi_{T-t-h} - \phi_{T-t}),$$

39. See Greene (2000) for an introduction to the method of moments.
40. The data are taken from yield curves that are computed from the U.S. Treasury bond files of the Center for Research in Security Prices (CRSP).
41. One can always hold currency, which earns a nominal rate of zero. If the nominal interest rate were negative, then bondholders would sell their bonds for currency, putting downward pressure on bond prices and driving the nominal interest rate back up at least to zero.
42. There is yet a third way that involves estimating the cross-sectional yields simultaneously with the interest rate dynamics, treating the yield curve data as a panel.
43. See Vasicek (1977), who provided the first derivation of the continuous-time version of this model using absence-of-arbitrage arguments. Campbell (1986) derived the same model in a representative agent setting.
44. The regression results in Campbell and Shiller (1991) demonstrate that some implications of the expectations hypothesis are very much at odds with the data.

FIGURE 7

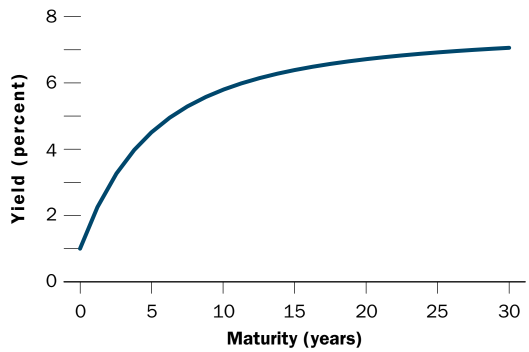
The Zero-Coupon Yield Curve with Parameters Fitted to the Average Yield Curve, Interest Rate of 5 Percent



Note: The parameter values are κ_r = 0.203, σ_r = 0.0041, and (assuming θ_r = 0.050) λ = –0.245.

FIGURE 8

The Zero-Coupon Yield Curve with Parameters Fitted to the Average Yield Curve, Interest Rate of 1 Percent



Note: The parameter values are κ_r = 0.203, σ_r = 0.0041, and (assuming θ_r = 0.050) λ = –0.245.

which shows that all of the random variation in the forward rate is attributable to variation in the expected future interest rate.

To model the failure of the expectations hypothesis one observes in U.S. data, one can expand the model by adding a second state variable that is independent of the interest rate. (This exercise would require a second source of uncertainty in addition to $\varepsilon(t + h)$). In order for a state variable that is independent of the interest rate to affect bond prices (and hence forward rates), it would have to help determine the price of risk. Such a state variable could not help forecast the interest rate (by construction); for this purpose it would simply be noise. Consequently, one would be faced with what is known as an error-in-variables problem were one to use forward rates to forecast future interest rates. It is this feature that would produce the failure of the expectations hypothesis.

A fuller treatment of this subject must be left to a future article.

APPENDIX A

The State-Price Deflator in a Representative Agent Model

This appendix demonstrates how the state-price deflator can be derived from the marginal utility of a representative agent.[1] The reader should be familiar with the discussion on pages 42–48.

We begin by focusing on the utility an agent derives from the flow of consumption. For the moment, we ignore the timing of the consumption flow. Let c denote the rate of consumption (a flow); then ch is the amount of consumption over the period. Let $u(c)$ be the rate of utility (a flow) derived from the rate of consumption; then $u(c)h$ is the amount of utility obtained over the period. We will restrict ourselves to the following utility function:

$$\text{(A.1)} \quad u(c) = \begin{cases} \dfrac{c^{1-\gamma}-1}{1-\gamma} & \gamma \neq 1 \\ \log(c) & \gamma = 1 \end{cases}.$$

This utility function is defined only for positive consumption $c > 0$. The parameter γ is called the coefficient of relative risk aversion; it measures the agent's attitude toward risk. A larger value of γ indicates greater aversion to risk.[2]

The additional utility from a small additional amount of consumption is called marginal utility. Let Δc be a small increment in the rate of consumption. Then the change in consumption is Δch and the change in utility is

$$u(c+\Delta c)h - u(c)h = \left(\frac{u(c+\Delta c)-u(c)}{\Delta c}\right)\Delta ch \approx c^{-\gamma}\Delta ch,$$

where the form of the utility function given in equation (A.1) is used in deriving the approximation. The approximation is accurate for small Δc.[3]

Now we consider the notion of expected lifetime utility, the expectation of discounted future utility from future consumption flows. Let $c(t + ih)$ denote the rate of consumption at time $t + ih$. The expected lifetime utility as of time t for the agent is[4]

$$\text{(A.2)} \quad U(t) = E_t\left\{\sum_{i=0}^{\infty} e^{-\delta ih}\, u[c(t+ih)]h\right\},$$

where $\delta > 0$ (the rate of time preference) measures the investor's impatience. The larger δ is, the less future consumption counts in expected lifetime utility.

Let $\Delta c(t + ih)$ denote the change to the rate of consumption at time $t + ih$. A set of changes in the rates of consumption $\{\Delta c(t+ih)\}_{i=0}^{\infty}$ could be the result of revised investment decisions: Sell government bonds, buy technology stocks, invest in real estate, build a factory, go back to college, etc. From the perspective of time t, the change in lifetime utility attributable to a set of changes in the rates of consumption is the expected value of the discounted sum of all the individual changes in utility:[5]

TABLE

Additional Notation

$c(t)$	rate of consumption flow at time t
δ	rate of time preference
γ	coefficient of relative risk aversion
μ_c	expected growth rate of consumption
σ_c	volatility of growth rate of consumption
$\hat{\pi}(t)$	real state-price deflator
$z(t)$	price level at time t
$x(t)$	expected rate of inflation at time t
σ_z	volatility of the inflation rate
κ_x	speed of mean reversion for the expected inflation rate
θ_x	long-run mean of expected inflation
σ_x	expected inflation volatility
$\hat{r}$	real interest rate
$\hat{\lambda}$	real price of risk

1. The table contains the notation used in this appendix.
2. Here is a thumbnail sketch of the meaning of risk aversion. Suppose there are two equally likely payouts to a gamble, c_1 and c_2, where $c_1 \neq c_2$. The utility of the payouts is $u(c_1)$ and $u(c_2)$. The average utility is $[u(c_1) + u(c_2)]/2$. The average payout is $(c_1 + c_2)/2$. The utility of the average payout is $u[(c_1 + c_2)/2]$. The agent is risk averse if $[u(c_1) + u(c_2)]/2 < u[(c_1 + c_2)/2]$. In other words, the agent is risk averse if the average utility from the payouts of the gamble is less than the utility from the average of the payouts. Given the utility function in equation (A.1), the agent is risk averse if $\gamma > 0$.
3. In the limit as Δc goes to zero, the term in parentheses goes to the derivative $u'(c)$.
4. For simplicity, the agent is assumed to live forever.
5. $\nabla U(t)$ is called the *utility gradient*, and the right-hand side of equation (A.3) is the inner product representation of the utility gradient. In the limit as h goes to zero, $\nabla U(t) = E_t\left[\int_t^{\infty} e^{-\delta s}\, c(s)^{-\gamma}\Delta c(s)ds\right]$. See Duffie and Skiadas (1994) for a general abstract treatment in continuous time.

(A.3) $$\nabla U(t)=E_t\left[\sum_{i=0}^{\infty} e^{-\delta ih} c(t+ih)^{-\gamma}\Delta c(t+ih)h\right].$$

If there were any feasible set of changes in consumption for which $\nabla U(t) \neq 0$, then the agent could use that set of changes to increase utility. Therefore, if the agent has maximized lifetime expected utility, then $\nabla U(t)=0$, or (pulling the first term in the sum out of the expectations operator)

$$0=c(t)^{-\gamma}\Delta c(t)h + E_t\left[\sum_{i=1}^{\infty} e^{-\delta ih} c(t+ih)^{-\gamma}\Delta c(t+ih)h\right].$$

This condition (the so-called first-order condition) can be expressed as

(A.4) $$-\Delta c(t)h = E_t\left[\sum_{i=1}^{\infty}\left(\frac{e^{-\delta(t+ih)}c(t+ih)^{-\gamma}}{e^{-\delta t}c(t)^{-\gamma}}\right)\Delta c(t+ih)h\right].$$

The left-hand side of equation (A.4) is the amount of current consumption the agent is willing to forgo in order to obtain the extra future consumption given on the right-hand side.[6] In other words, $-\Delta c(t)h$ is the value of the *consumption dividends* $\{\Delta c(t+ih)h\}_{i=1}^{\infty}$.

The dividends in equation (A.4) are measured in units of consumption; in other words, they are *real dividends*. Similarly, $-\Delta c(t)h$ is the real value of those real dividends. In order to distinguish real dividends and prices from nominal dividends and prices, we will use a "hat" to denote real values. In particular, let

$$\hat{d}(t+ih) := \Delta c(t+ih) \text{ and } \hat{v}(t) := -\Delta c(t)h.$$

In addition, if we simply define

(A.5) $$\hat{\pi}(t) := e^{-\delta t}c(t)^{-\gamma},$$

then we can write equation (A.4) as

(A.6) $$\hat{v}(t)=E_t\left[\sum_{i=1}^{\infty}\left(\frac{\hat{\pi}(t+ih)}{\hat{\pi}(t)}\right)\hat{d}(t+ih)h\right],$$

which is identical in form to equation (7).

The price level (that is, the price of consumption in terms of dollars) can be used to transform equation (A.6) into an expression for nominal values of nominal dividend streams. Let $z(t)$ denote the price level at time t, and let

(A.7a) $$d(t+ih)=z(t)\hat{d}(t+ih)=z(t)\Delta c(t+ih);$$

(A.7b) $$v(t)=z(t)\hat{v}(t)=-z(t)\Delta c(t)h;$$

(A.7c) $$\pi(t)=\hat{\pi}(t)/z(t)=e^{-\delta t}c(t)^{-\gamma}z(t)^{-1}.$$

Real prices and real dividends have been multiplied by the price level to convert them into nominal values. By contrast, the real state-price deflator has been divided by the price level to convert it into a nominal state-price deflator; this division is required in order to maintain the validity of equation (A.6), which now becomes

$$v(t)=E_t\left[\sum_{i=1}^{\infty}\left(\frac{\pi(t+ih)}{\pi(t)}\right)d(t+ih)h\right].$$

Dynamics of Consumption and the Price Level

The dynamics of the state-price deflator in equation (A.7c) can be computed from the dynamics of consumption and the price level.

First, consider the dynamics of consumption:

(A.8) $$\log[c(t+h)]-\log[c(t)]=\mu_c h+\sigma_c\varepsilon(t+h),$$

where μ_c and σ_c are fixed parameters.[7] In equation (A.8) the expected growth rate of consumption is constant although the actual growth rate is random.

Next we turn to the dynamics of the price level, which we allow to be more complex. In particular, let the expected growth rate of the price level $x(t)$ (the expected inflation rate) be a stochastic process that follows a first-order autoregressive process:[8]

(A.9a) $$\log[z(t+h)]-\log[z(t)]=x(t)h+\sigma_z\varepsilon(t+h)$$

(A.9b) $$x(t+h)-x(t)=\kappa_x[\theta_x-x(t)]h+\sigma_x\varepsilon(t+h).$$

6. For concreteness, assume $\Delta c(t+ih)\geq 0$ for $i \geq 1$. It follows (given $\nabla U(t)=0$) that $\Delta c(t)<0$ and thus $-\Delta c(t)h>0$.

7. Recall that $\varepsilon(t+h)\sim N(0,h)$.

8. The expected rate of inflation is (implicitly) defined as $E_t\{\log[z(t+h)/z(t)]\}/h$. The reader should be aware that some authors define the expected rate of inflation as $\log\{E_t[z(t+h)/z(t)]\}/h$. The difference between the two notions of expected inflation is $\sigma_z^2/2$.

Now we compute the dynamics of the state-price deflator. Given equation (A.7c), we have $\log[\pi(t)] = -\delta t - \gamma \log[c(t)] - z(t)$, from which we compute the dynamics:

$$\text{(A.10)}\quad \log[\pi(t+h)] - \log[\pi(t)] = -\delta h - \gamma\{\log[c(t+h)] - \log[c(t)]\} - \{\log[z(t+h)] - \log[z(t)]\} = -\underbrace{[\delta + \gamma\mu_c + x(t)]}_{r(t)+\frac{1}{2}\lambda^2} h - \underbrace{(\gamma\sigma_c + \sigma_z)}_{\lambda}\varepsilon(t+h).$$

From the dynamics of the state-price deflator we can identify the interest rate and the price of risk. Matching coefficients in the second line of equation (A.10) with equation (15), we obtain

$$\text{(A.11a)}\quad r(t) = \delta + \gamma\mu_c + x(t) - \frac{1}{2}(\gamma\sigma_c + \sigma_z)^2;$$

$$\text{(A.11b)}\quad \lambda = \gamma\sigma_c + \sigma_z.$$

The dynamics of the interest rate follow directly from the dynamics of the expected rate of inflation. Using equations (A.11a) and (A.9b), note that

$$\text{(A.12)}\quad r(t+h) - r(t) = x(t+h) - x(t) = \kappa_x[\theta_x - x(t)]h + \sigma_x\varepsilon(t+h) = \underbrace{\kappa_x}_{\kappa_r}\{\underbrace{[\delta + \gamma\mu_c + \theta_x - \frac{1}{2}(\gamma\sigma_c + \sigma_z)^2]}_{\theta_r} - r(t)\}h + \underbrace{\sigma_x}_{\sigma_r}\varepsilon(t+h),$$

where (in going from the second line to the third line in equation [A.12]) we have used equation (A.11a) a second time to eliminate $x(t)$. We conclude that $\kappa_r = \kappa_x$, $\sigma_r = \sigma_x$, and

$$\theta_r = \delta + \gamma\mu_c + \theta_x - \frac{1}{2}(\gamma\sigma_c + \sigma_z)^2.$$

The Real Interest Rate and the Fisher Equation

In the preceding discussion, the expression for the nominal state-price deflator was used to derive the nominal interest rate and the nominal price of risk. This section follows the same steps with respect to the real state-price deflator to derive expressions for the real interest rate and the real price of risk. These expressions will provide insight into both the real and nominal interest rates.

Given equation (A.5), we have $\log[\hat{\pi}(t)] = -\delta t - \gamma \log[c(t)]$, and thus

$$\log[\hat{\pi}(t+h)] - \log[\hat{\pi}(t)] = -\delta h - \gamma\{\log[c(t+h)] - \log[c(t)]\} = -\underbrace{(\delta + \gamma\mu_c)}_{\hat{r}+\frac{1}{2}\hat{\lambda}^2} h - \underbrace{\gamma\sigma_c}_{\hat{\lambda}}\varepsilon(t+h).$$

These dynamics imply

$$\text{(A.13a)}\quad \hat{r} = \delta + \gamma\mu_c - \frac{1}{2}(\gamma\sigma_c)^2$$

$$\text{(A.13b)}\quad \hat{\lambda} = \gamma\sigma_c$$

where $\hat{r}$ is the real interest rate and $\hat{\lambda}$ is the real price of risk.

Equation (A.13a) shows that the real interest rate is composed of three terms. The first term is the rate of time preference, δ: Greater impatience is associated with higher real interest rates. The second and third terms each involve γ, but γ plays a different role in each term. In the third term, which reflects Jensen's inequality, γ measures risk aversion, as one would expect from the discussion of the utility function in equation (A.1).

However, γ does not represent risk aversion in the second term in equation (A.13a). In the lifetime expected utility function in equation (A.2), not only does γ measure the agent's risk aversion, but $1/\gamma$ measures the agent's *elasticity of intertemporal substitution*.[9] If there were no uncertainty about the growth rate of consumption (if $\sigma_c = 0$), then equation (A.13a) could be reexpressed as

$$\text{(A.14)}\quad -\mu_c = \gamma^{-1}(\delta - \hat{r}).$$

Equation (A.14) can be thought of as expressing the demand for $-\mu_c$ (the log of current consumption relative to future consumption) in terms of the real interest rate (the log of the price of current consumption in terms of future consumption), so that γ^{-1} corresponds to the elasticity of demand.[10] The more responsive current consumption is to the real

9. More general intertemporal utility functions such as *recursive utility* allow for separate parameters for the two effects. See Epstein and Zin (1991) and Duffie and Epstein (1992).

10. Note that the rate of time preference δ equals the real interest rate that makes current and future consumption equal ($\mu_c = 0$). This feature is essentially the definition of the rate of time preference.

interest rate (that is, the bigger γ^{-1} is), the less responsive the real interest rate is in equation (A.13a) to the growth rate of consumption.

Thus far we have examined the role of preferences (in the form of the parameters δ and γ) in determining the real interest rate. Technology (in the form of the marginal product of capital) also plays an important role in determining the real interest rate, but in equation (A.13a) its effects occur indirectly through μ_c and σ_c. A more complete model would make these effects explicit. However, for our purposes, making these effects explicit is not necessary: No matter what the effects of technology are, we are able to correctly infer the price system from marginal utility because we have assumed the agent has optimized, and consequently the price system is proportional to the utility gradient.

We now return to consideration of the nominal interest rate. The expressions for the real interest rate and the real price of risk in equation (A.13) can be used to gain additional insight into the expression for the nominal interest rate in equation (A.11a). The nominal interest rate can be expressed as a so-called Fisher equation (Irving Fisher 1930) that incorporates uncertainty:

$$\text{(A.15)} \quad r(t) = \hat{r} + x(t) - \hat{\lambda}\sigma_z - \frac{1}{2}\sigma_z^2.$$

Equation (A.15) says that the nominal rate equals the real rate plus the expected rate of inflation plus two additional terms: (1) a risk premium that depends on the covariance between the consumption growth and inflation and (2) a Jensen's inequality term that depends on the variance of inflation. The decomposition of the nominal rate into the four components on the right-hand side of equation (A.15) is quite general. However, in a more general model, all four components would be stochastic and contribute to the variation of the nominal rate.[11]

Solving the Real Term Structure

Because the real interest rate $\hat{r}$ is deterministic in this example, the no-arbitrage condition is equivalent to the expectations hypothesis. Moreover, since the real rate is constant, the term structure of real zero-coupon yields is flat, which we can confirm by following the steps used on pages 49–50 to solve the nominal term structure.

The no-arbitrage condition for the real price of a real bond is given by

$$\hat{p}(t,T) = E_t[\hat{\pi}(T)/\hat{\pi}(t)].$$

This condition can be expressed in terms of risk and return:

$$\text{(A.16)} \quad \hat{\mu}(t,T) = \hat{r} + \hat{\lambda}\hat{\sigma}(t,T),$$

where $\hat{\mu}(t,T)$ is the real expected return on a real bond and $\hat{\sigma}(t,T)$ is the volatility of the real return on a real bond. Since both $\hat{r}$ and $\hat{\lambda}$ are constant parameters, there are no state variables; consequently, the real price of a real bond will depend only on the maturity of the bond. In particular, assume $\hat{p}(t,T) = \hat{P}(T-t)$, where $\hat{P}(\tau) = {}^{-\hat{a}(\tau)}$. Then

$$\log[p(t+h,T)] - \log[p(t,T)] = \underbrace{\left(\frac{\hat{a}(\tau) - \hat{a}(\tau-h)}{h}\right)}_{\hat{\mu}(t,T)} h,$$

where $\tau = T - t$ and $\hat{\sigma}(t,T) = 0$. Therefore, the no-arbitrage condition in equation (A.16) becomes

$$\frac{\hat{a}(\tau) - \hat{a}(\tau-h)}{h} = \hat{r},$$

subject to $\hat{a}(0) = 0$. The solution is $\hat{a}(\tau) = \hat{r}\tau$, which implies $\hat{p}(t,T) = e^{-\hat{r}(T-t)}$, which in turn implies the solution for real zero-coupon yields: $\hat{y}(t,T) = \hat{r}$.

11. Also, the nominal price of risk can be expressed as the real price of risk plus the volatility of the inflation rate: $\lambda = \hat{\lambda} + \sigma_z$.

APPENDIX B

Facts about Lognormal Random Variables

Lognormal random variables are built from normal random variables. Let x be a random variable that is normally distributed with (population) mean m and (population) variance $v > 0$. This variable is denoted as $x \sim N(m, v)$. The probability density function (PDF) for x is

$$f(x) = \frac{e^{-(x-m)^2/(2v)}}{\sqrt{2\pi}\sqrt{v}}.$$

The graph of $f(x)$ is the well-known bell curve. The mean and variance can be expressed in terms of the expectation operator:[1]

$$E[x] = \int_{-\infty}^{\infty} x f(x) dx = m,$$

and $$\text{Var}[x] = E[(x-m)^2] = \int_{-\infty}^{\infty} (x-m)^2 f(x) dx = v.$$

A lognormally distributed random variable can be constructed from a normally distributed random variable in the following way: If x is normally distributed, then $z = e^x$ is lognormally distributed. In other words, the log of z is normally distributed: $\log(z) = x \sim N(m, v)$.

The most important fact about a lognormally distributed random variable is that the (population) mean of z depends on both m and v:

(B.1) $$E[z] = \int_{-\infty}^{\infty} e^x f(x) dx = e^{m+\frac{1}{2}v}.$$

Note that $\log(E[z]) = m + \frac{1}{2}v$. It is worth emphasizing that $\log(E[z]) \neq E(\log[z])$ unless $v = 0$. This inequality is an example of *Jensen's inequality*, which states that given a suitable function g, $g(E[x]) \neq E(g[x])$ unless $g[x]$ is linear in x (or unless x is deterministic).[2]

Two Random Variables and One Shock

Suppose there are two normally distributed random variables x_1 and x_2, where $x_i \sim N(m_i, v_i)$, for $i = 1, 2$. In particular, suppose

$$x_i = \left(\mu_i - \frac{1}{2}\sigma_i^2\right)h + \sigma_i \varepsilon,$$

where $\varepsilon \sim N(0, h)$. Then,

$$m_i = E[x_i] = \left(\mu_i - \frac{1}{2}\sigma_i^2\right)h$$

$$v_i = \text{Var}[x_i] = E[(x_i - m_i)^2] = E[(\sigma_i \varepsilon)^2] = \sigma_i^2 E[\varepsilon^2] = \sigma_i^2 h.$$

In addition,

$$\begin{aligned}\text{Cov}[x_1, x_2] &= E[(x_1 - m_1)(x_2 - m_2)] \\ &= E\left[(\sigma_1 \varepsilon)(\sigma_2 \varepsilon)\right] \\ &= \sigma_1 \sigma_2 E_t[\varepsilon^2] \\ &= \sigma_1 \sigma_2 h.\end{aligned}$$

Note that $x_1 + x_2 \sim N(M, V)$, where

$$M = E[x_1] + E[x_2] = [m_1 + m_2 - \frac{1}{2}(\sigma_1^2 + \sigma_2^2)]h$$

$$V = \text{Var}[x_1] + \text{Var}[x_2] + 2\text{Cov}[x_1, x_2] = (\sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2)h.$$

If $z_i = e^{x_i}$, then $E[z_i] = e^{\mu_i h}$ and

$$E[z_1 z_2] = E[e^{x_1 + x_2}] = e^{M + \frac{1}{2}V} = e^{(\mu_1 + \mu_2 + \sigma_1\sigma_2)h}.$$

Therefore, $E[z_1 z_2] = 1$ implies

(B.2) $\mu_1 + \mu_2 + \sigma_1\sigma_2 = 0.$

Additional information. It may be of interest to note that $\text{Var}[z_i] \approx \text{Var}[x_i]$ and $\text{Cov}[z_1, z_2] \approx \text{Cov}[x_1, x_2]$, where the approximations are accurate for small h. (These approximations are not used elsewhere in the article.) To see where these approximations come from, we use the following facts. First, given any two random variables y_1 and y_2, $\text{Cov}[y_1, y_2] = E[y_1 y_2] - E[y_1]E[y_2]$. A special case of this fact is $\text{Var}[y_i] = \text{E}[y_i^2] - E[y_i]^2$. Next, note that $2x_i \sim N[2(\mu_i - \sigma_i^2/2), 4\sigma_i^2]$, so that $E[z_i^2] = E[e^{2x_i}] = e^{(2\mu_i + \sigma_i^2)h}$. Then

$$\begin{aligned}\text{Var}[z_i] &= E[z_i^2] - E[z_i]^2 = E[e^{2x_i}] - E[e^{x_i}]^2 \\ &= e^{(2\mu_i + \sigma_i^2)h} - e^{2\mu_i h} = e^{2\mu_i h}(e^{\sigma_i^2 h} - 1) \\ &\approx \sigma_i^2 h = \text{Var}[x_i]\end{aligned}$$

and $$\begin{aligned}\text{Cov}[z_1, z_2] &= E[z_1 z_2] - E[z_1]E[z_2] \\ &= E[e^{x_1 + x_2}] - E[e^{x_1}]E[e^{x_2}] \\ &= e^{(\mu_1 + \mu_2 + \sigma_1\sigma_2)h} - e^{(\mu_1 + \mu_2)h} \\ &= e^{(\mu_1 + \mu_2)h}(e^{\sigma_1\sigma_2 h} - 1) \\ &\approx \sigma_1 \sigma_2 h = \text{Cov}[x_1, x_2].\end{aligned}$$

1. Unconditional expectations are used in the appendix. The same relations hold for conditional expectations.
2. Here is another example of Jensen's inequality. Let $z = e^x$, where $x \sim N(m, v)$. Note that $z^{-1} = e^{-x}$ and $-x \sim N(-m, v)$. Therefore, $E[z^{-1}] = e^{-m+\frac{1}{2}v}$ and $E[z]\, \text{E}[z^{-1}] = e^{m+\frac{1}{2}v} e^{-m+\frac{1}{2}v} = e^v \neq 1 = \text{E}[z\, z^{-1}]$, unless $v = 0$.

APPENDIX C

Forecasting the Interest Rate

Equation (13) will be used to produce forecasts of $r(T)$ given the information available at time t, $E_t[r(T)]$.

Note that the future interest rate $r(T)$ is the current interest rate $r(t)$ plus all the changes in the interest rate between t and T, which can be written[1]

$$r_T = r_t + (r_{t+h} - r_t) + (r_{t+2h} - r_{t+h}) + \cdots + (r_T - r_{T-h})$$
$$= r_t + \sum_{i=1}^{\tau/h} (r_{t+ih} - r_{t+(i-1)h}).$$

Therefore the expected future interest rate equals the current interest rate plus the sum of the expected changes in the interest rate:

$$\text{(C.1)} \quad E_t[r_T] = r_t + \sum_{i=1}^{\tau/h} E_t[r_{t+ih} - r_{t+(i-1)h}]$$
$$= r_t + \sum_{i=1}^{\tau/h} \kappa_r \{\theta_r - E_t[r_{t+(i-1)h}]\} h.$$

Equation (13) is used in the second line in equation (C.1). Fixing t and given r_t, consider the forecast of r_T as a function of the forecast horizon $\tau = T - t$. To this end, let $g(\tau) := E_t[r_{t+\tau}]$. Then equation (C.1) can be written as

$$\text{(C.2)} \quad g(\tau) = g(0) + \sum_{i=1}^{\tau/h} \kappa_r \{\theta_r - g[(i-1)h]\} h,$$

where $g(0) = r_t$. Equation (C.2) can be expressed as a first-order linear difference equation in $g(\tau)$:

$$\text{(C.3)} \quad \frac{g(\tau+h) - g(\tau)}{h}$$
$$= \kappa_r \theta_r - \kappa_r g(\tau), \quad \text{subject to } g(0) = r_t.$$

The solution to (C.3) is

$$\text{(C.4)} \quad E_t[r_{t+\tau}] = g(\tau) = A_r^\tau + B_r^\tau r_t,$$

where

$$\text{(C.5)} \quad A_r^\tau = (1 - B_r^\tau)\theta_r \quad \text{and} \quad B_r^\tau = \left(1 - \kappa_r h\right)^{\tau/h}.$$

Note that A_r^τ and B_r^τ depend only on the forecast horizon τ and the parameters θ_r and κ_r. As τ gets larger, B_r^τ gets closer to zero. B_r^τ displays *exponential decay*. When h is small, $B_r^\tau \approx e^{-\kappa_r \tau}$. The approximation is quite good for monthly data ($h = 1/12$) given the estimated values for κ_r (as we will see).

Given equation (C.5), equation (C.4) can be written as

$$\text{(C.6)} \quad E_t[r_{t+\tau}] = \theta_r + B_r^\tau (r_t - \theta_r)$$
$$= \theta_r + (1 - \kappa_r h)^{\tau/h} (r_t - \theta_r).$$

An arbitrarily large forecast horizon implies

$$\lim_{\tau \to \infty} E_t[r_{t+\tau}] = \theta_r,$$

which confirms that θ_r is the long-run mean. The expected half-life of a given deviation from the mean $(r_t - \theta_r)$ is obtained by solving $B_r^\tau = 1/2$ for τ:

$$\text{(C.7)} \quad \tau = \frac{\log(2)}{\kappa_r}\left(\frac{-\kappa_r h}{\log(1 - \kappa_r h)}\right) \approx \frac{\log(2)}{\kappa_r}.$$

Again, the approximation is accurate when h is small. The half-life is inversely proportional to κ_r, confirming the interpretation of κ_r as the speed of mean reversion. To get a feel for the magnitudes, note that $\kappa_r = 7$ corresponds to a half-life of about 5.2 weeks, $\kappa_r = 0.7$ corresponds to a half-life of about one year, and $\kappa_r = 0.07$ corresponds to a half-life of about ten years.

Forecast Revisions and Forecast Errors

In this section, forecast revisions and forecast errors are used to compute the conditional and unconditional variance of the interest rate.

As time passes and new information arrives, the forecast of r_T will be revised from $E_t[r_T]$ to $E_{t+h}[r_T]$. The forecast revision is given by

$$\text{(C.8)} \quad E_{t+h}[r_T] - E_t[r_T] = E_{t+h}[r_{t+\tau}] - E_t[r_{t+\tau}]$$
$$= \left(A_r^{\tau-h} + B_r^{\tau-h} r_{t+h}\right)$$
$$- \left(A_r^\tau + B_r^\tau r_t\right)$$
$$= B_r^{\tau-h} \sigma_r \varepsilon(t+h),$$

where $\tau = T - t$. The second equality in (C.8) follows from the solution for $E_t[r_{t+\tau}]$ given in equation (C.4). The third equality follows from the dynamics of r given in equation (13), replacing r_{t+h} with $r_t + \kappa_r(\theta_r - r_t)h + \sigma_r\varepsilon(t + h)$ and collecting terms. (This latter simplification involves a substantial amount of rearrangement—so much that one probably cannot confirm it in one's head.) Using equation (C.8), the following three

1. To make the notation more compact for space considerations, subscripts will be used in place of parentheses to denote arguments for time and maturity. In particular, $r_t \equiv r(t)$ and so forth.

properties of the forecast revision can all be inferred from the properties of $\varepsilon(t+h)$. First, the forecast revision is zero on average; second, it is independent of previous forecast revisions; and, third, its (conditional) variance is

$$E_t\left[\left(B_r^{\tau-h}\sigma_r\,\varepsilon(t+h)\right)^2\right]=\left(B_r^{\tau-h}\right)^2\sigma_r^2\,h,$$

which is $\left(B_r^{\tau-h}\right)^2$ times the conditional variance of $r_{t+h}-r_t$.

The forecast error is defined as the difference between the actual value $r_{t+\tau}$ and its conditional expectation as of time t:

(C.9) $e_t^\tau := r_{t+\tau}-E_t[r_{t+\tau}].$

Here t does not indicate when e_t^τ is known but rather when the forecast was made; e_t^τ is not known for certain until time $t+\tau$. The forecast error can be expressed in terms of the sum of revisions in the forecasts:

$$\begin{aligned}e_t^\tau &= r_{t+\tau}-E_t[r_{t+\tau}]=(E_{t+h}[r_{t+\tau}]-E_t[r_{t+\tau}])\\ &\quad+(E_{t+2h}[r_{t+\tau}]-E_{t+h}[r_{t+\tau}])+\cdots\\ &\quad+(E_{t+\tau-h}[r_{t+\tau}]-E_{t+\tau-2h}[r_{t+\tau}])\\ &\quad+(\underbrace{E_{t+\tau}[r_{t+\tau}]}_{r_{t+\tau}}-E_{t+\tau-h}[r_{t+\tau}]).\end{aligned}$$

Since each of the forecast revisions is zero on average, so is the forecast error:

$$E_t[e_t^\tau]=0.$$

(If the forecast error were not zero on average, then one could make better forecasts; in other words, if the forecast error were not zero on average, the conditional expectation must not have been computed properly.) Since each successive forecast revision is independent of those that precede it, the variance of the sum of the revisions equals the sum of the variances of the revisions:

(C.10)

$$\begin{aligned}E_t\left[(e_t^\tau)^2\right]&=\sum_{i=1}^{\tau/h}E_t\left[(E_{t+ih}[r_{t+\tau}]-E_{t+(i-1)h}[r_{t+\tau}])^2\right]\\ &=\sigma_r^2\sum_{i=1}^{\tau/h}\left(B_r^{\tau-ih}\right)^2 h\\ &=\frac{\sigma_r^2\left[1-(1-\kappa_r h)^{2\tau/h}\right]}{2\kappa_r(1-\frac{1}{2}\kappa_r h)}\approx\frac{\sigma_r^2\left(1-e^{-2\kappa_r\tau}\right)}{2\kappa_r}.\end{aligned}$$

The expectation of the squared forecast error is also known as the conditional variance:

$$\mathrm{Var}_t[r_{t+\tau}]=E_t\left[(e_t^\tau)^2\right].$$

The unconditional variance can be computed as follows:

$$\mathrm{Var}[r]=\lim_{\tau\to\infty}\mathrm{Var}_t[r_{t+\tau}]=\frac{\sigma_r^2}{2\kappa_r(1-\frac{1}{2}\kappa_r h)}\approx\frac{\sigma_r^2}{2\kappa_r}$$

for small h.

REFERENCES

Baxter, Martin, and Andrew Rennie. 1996. *Financial calculus: An introduction to derivative pricing.* Cambridge, New York, and Melbourne: Cambridge University Press.

Campbell, John Y. 1986. A defense of traditional hypotheses about the term structure of interest rates. *Journal of Finance* 41, no. 1:183–93.

Campbell, John Y., and Robert J. Shiller. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58, no. 195:495–514.

Cochrane, John H. 2001. *Asset pricing.* Princeton, N.J.: Princeton University Press.

Duffie, Darrell. 1996. *Dynamic asset pricing theory.* 2nd ed. Princeton, N.J.: Princeton University Press.

Duffie, Darrell, and Larry G. Epstein. 1992. Asset pricing with stochastic differential utility. *Review of Financial Studies* 5, no. 3:411–36.

Duffie, Darrell, and Rui Kan. 1996. A yield-factor model of interest rates. *Mathematical Finance* 6, no. 4:379–406.

Duffie, Darrell, and Costis Skiadas. 1994. Continuous-time security pricing: A utility gradient approach. *Journal of Mathematical Economics* 23, no. 2:107–31.

Epstein, Larry G., and Stanley E. Zin. 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy* 99, no. 2:263–86.

Fisher, Irving. 1930. *The theory of interest*. New York: Macmillan.

Fisher, Mark. 2001a. Forces that shape the yield curve. Federal Reserve Bank of Atlanta *Economic Review* 86, no. 1:1–15.

———. 2001b. Forces that shape the yield curve: Parts 1 and 2. Federal Reserve Bank of Atlanta Working Paper 2001-3, March.

Greene, William H. 2000. *Econometric analysis*. 4th ed. Upper Saddle River, N.J.: Prentice-Hall.

Kloeden, Peter E., and Eckhard Platen. 1995. 2nd corrected printing. *Numerical solution of stochastic differential equations.* New York: Springer-Verlag.

Neftci, Salih N. 1996. *An introduction to the mathematics of financial derivatives.* San Diego, Calif.: Academic Press.

Vasicek, Oldrich. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, no. 2:177–88.