A Two Factor Latent Variable Model of the Term Structure of Interest Rates

Rodney L. Jacobs Robert A. Jones *

February 1986

I INTRODUCTION

Continuous time models of the term structure of interest rates (see bibliography) provide unified theories of the pricing of all interest rate related claims: bills and bonds of different maturities, and options and futures contracts on these instruments. More specifically, they postulate that the prices of all such securities depend on a usefully small common list of underlying 'factors', and that economic agents have rational expectations about how these factors fluctuate over time. This paper explores the empirical validity and usefulness of this view.

Section II of the paper sets out a a stylized continuous time model of general equilibrium, with stochastic production and endogenous capital stock. The model gives rise to a two factor pricing model of interest rate related claims. One state variable is the instantaneous interest rate; the other is the unobservable value to which it is regressing. Security prices are solutions to the appropriate partial differential equations.

Section III addresses three problems of empirically implementing the model of section II. First, the model parameters are those of a continuous time process which must be estimated from discrete time observations. Second, theoretical security prices predicted by the model are only implicitly defined as the solution to a partial differential equation. Finally, only one of the conceptual state variables is observable. The first two problems are handled by solving the state variable equations to obtain the equivalent discrete time process, and by using a numerical solution of the partial differential equation for security prices. The third problem is handled by jointly estimating the time path of the latent state variable along with the other model parameters. This approach permits all parameters of the model, and the latent state variable, to be estimated from a time series cross-section of prices from any interest rate market. Since theoretical models in finance frequently have conceptual variables that

*Wells Fargo Bank and Simon Fraser University respectively. We are grateful to Michael Brennan and Andrew Lo for helpful discussion and comments on an earlier version of this paper presented at the 1985 Winter meetings of the Econometric Society.

cannot be directly observed (e.g., expectations, aggregate wealth, etc.), the method used has potentially broad application.

Section IV presents empirical results. The model is fit to weekly observations of US Treasury Bill and Note prices between 1978 and 1984, and separately to Treasury Bill futures prices over the same period. Hypotheses investigated include whether representative participants in the cash and futures markets held similar beliefs about the process governing interest rate movements, whether their perception of the current state was the same, and whether their beliefs were *ex post* rational. The results suggest that expectations significantly deviated from rationality in a statistical sense, but that the deviation was quantitatively small. In addition, we explore the quantitative importance of the distinction between futures and forward prices arising from the daily settlement of gains and losses on futures contracts. Finally we look at the adequacy of using the approximate discrete time analogue of the continuous time model.

II EQUILIBRIUM INTEREST RATES AND SECURITY PRICES

1 An Equilibrium Pricing Model

The theoretical framework guiding our empirical specification is a specialization of Cox, Ingersoll and Ross (1978). Consider an economy with one good — consumable capital. The stock of this good at time t is w. There is one constant returns to scale stochastic production activity, in which all nonconsumed capital is invested. The capital stock evolves according to a stochastic process

$$dw = wf(x)\,dt + wg(x)\,dz - c\,dt \tag{1}$$

where dz is the increment to a standard n-dimensional brownian motion, x is a vector of technology state variables, and c is the rate of consumption at time t.[1] Technological change is exogenous to the model, and follows a known observable process

$$dx = a(x)\,dt + s(x)\,dz \tag{2}$$

There is one individual. His objective is to maximize $E_0 \int_0^\infty u(c,t)\,dt$. The above elements constitute the real side of the economy. constitute the real side of the economy.

Now introduce a competitive market for a single financial asset. Each unit promises a dividend flow $q(x,w,t)\,dt$ and payment at maturity T of $P(x,w,T)$. Let $P(\underline{x},\underline{w},t)$ be its price at time t, with $\underline{x},\underline{w}$ denoting the entire history of x and w up to time t, or current information state. P is a nonanticipative functional. The individual may allocate wealth as he chooses between real production and this security. Trading is costless; the asset is

[1] $f(x)$ is a scalar valued function; $g(x)$ and $a(x)$ are a n-dimensional row and column vectors respectively; $s(x)$ is a $n \times n$ matrix. The vector x may include any other exogeneous state variables relevant for real security prices (e.g., the price level in a money economy), even though they do not affect the return on real investment.

divisible; its aggregate supply is zero. The problem is to find a $P(\underline{x}, \underline{w}, t)$ that is consistent with equilibrium.

The economy is completely specified by its technology f, g, a, s, tastes u, security characteristics q, $P(., T)$, and initial state $\underline{x}_0, \underline{w}_0$. Let α denote the fraction of wealth invested in the security at time t. An equilibrium consists of a

stochastic process for	w
consumption rule	$c^*(\underline{x}, \underline{w}, t)$
investment rule	$\alpha^*(\underline{x}, \underline{w}, t)$
pricing functional	$P(\underline{x}, \underline{w}, t)$

such that

(a) The individual optimizes: c^*, α^* solve $\max_{c,\alpha} E_0 \int_0^\infty u(c,t)\,dt$ subject to the wealth dynamics $dW = (1-\alpha)[f\,dt + g\,dz]W + \alpha[q\,dt + dP]W/P - c\,dt$

(b) Markets clear: $\alpha^* = 0$

(c) Consistent expectations: $w = W^*$ for all $\underline{x}, \underline{w}, t$.[2]

Though there is only one individual, planned wealth W is distinguished from aggregate wealth w to emphasize his price-taking behaviour in the security market. That is, he takes the process for w, and hence P, as given when selecting c^*, α^*. Together, (a)-(c) describe a Radner (1968) equilibrium of plans, prices and price expectations. The security price is the individual's marginal rate of substitution between current consumption and a security with the given characteristics.

Solving for the equilibrium entails solving the individual's consumption-investment problem. However the optimality condition invoked below requires that the variables x and P follow Ito processes — given their current values, future values cannot depend on history. This would be the case if $P(\underline{x}, \underline{w}, t)$ depended only on the current x, w, t, and if, from inspection of (1), aggregate consumption also depended only on the current state. Assuming the above, one obtains an optimal control for individual consumption that indeed does lead to

[2]As it stands, there is no apparent constraint on consumption. The decision problem as posed in (a) is intended as a shorthand for the limiting utility and consumption/investment policies, if they exist, of a sequence of finite horizon problems

$$\max_{c,\alpha} E_0\{\int_0^{\bar{T}} u(c,t)dt + B(W,\bar{T}\}$$

as $\bar{T} \to \infty$, where $B(W,\bar{T})$ is an appropriate utility of terminal wealth function. The expectation operator is conditional on the information available at time 0 (i.e., $\underline{x}_0, \underline{w}_0$).

these properties, validating the assumption and providing an equilibrium solution. Hence let us write the dynamics of P as

$$\frac{dP}{P} = \beta(x, w, t)\, dt + \sigma(x, w, t)\, dz \tag{3}$$

Define the value function for the individual's consumption-investment problem

$$J(x, W, t) \equiv \max_{c,\alpha} E_t \int_t^\infty u(c, \tau)\, d\tau \tag{4}$$

Under appropriate regularity conditions, a sufficient condition for the feedback rules $c^*(x, W, t), \alpha^*(x, W, t)$ to be optimal within the broader class of nonanticipative controls is that they solve[3]

$$0 = J_t + \max_{c,\alpha}\{u(c, t) + \Psi^{c,\alpha} J(x, W, t)\} \tag{5}$$

$\Psi^{c,\alpha}$ denotes the differential generator of J — the expected rate of change in J for particular values of c, α.[4] Performing the static maximization in (5) with respect to c, α gives first order conditions for an optimum. Substituting in the market clearing condition $\alpha^* = 0$, they are

$$\begin{aligned} 0 &= u_c - J_W \\ 0 &= (f - \beta - q/P)J_W + (gg' - g\sigma')W J_{WW} + (g - \sigma)s' J_{xW} \end{aligned} \tag{6}$$

Subscripts indicate partial differentiation and g' indicates the transpose of g. Various equilibrium yield relationships can be obtained by rearranging the second equation in (6).[5] However to solve for prices, β and σ must be eliminated. Applying Ito's lemma to $P(x, w, t)$,

$$\begin{aligned} \beta P &= P_t + a' P_x + (wf - c^*)P_w + \frac{1}{2}\sum (ss')_{ij}(P_{xx})_{ij} + gs' P_{wx} + \frac{1}{2} gg' P_{ww} \\ \sigma P &= P_w g + P_x' s \end{aligned} \tag{7}$$

The consistent expectations assumption has been invoked here by using the process parameters for w implied by (1). Given a utility function u, the equilibrium can be determined by solving

[3]See Kushner (1967) Th.IV.8, or Fleming and Rishel (1975) Cor. VI.4.2. Their results do not appear to rule out the possibility of non-feedback rules (i.e., rules that depend on past values of the state variable as well) that do just as well as the feedback rule. Thus there may well exist other equilibria for which security prices do depend on how the economy arrived at its current state.

[4]The differential generator is in this case

$$\begin{aligned} \Psi^{c,\alpha} J &= a' J_x + \sum (ss')_{ij}(P_{xx})_{ij}/2 + [(1-\alpha)f + \alpha(\beta + q/P) - c/W]W J_W \\ &\quad + [(1-\alpha)^2 gg' + 2\alpha(1-\alpha)g\sigma' + \alpha^2 \sigma\sigma']W^2 J_{WW}/2 + [(1-\alpha)g + \alpha\sigma]s' W J_{xW} \end{aligned}$$

[5]The instantaneous riskless rate r is the dividend yield required on a security that is riskless ($\sigma = 0$) and offers no capital appreciation ($\beta = 0$). Substitution into (6) gives $r \equiv q/P = f + gg' w J_{ww}/J_w + gs' J_{wx}/J_w$ as the equilibrium rate, permitting the equilibrium yields on other types of securities to be expressed relative to r.

(6a) for c in terms of J_W, substituting this into (5), and solving the resulting partial differential equation for J. Substituting this J plus the expressions for β, σ from (7) into (6b) then gives a partial differential equation for $P(x, w, t)$ that can be solved for prices.

If we assume logarithmic utility with constant rate of time preference λ, the equilibrium can be solved up to the differential equation for P without restrictions on the technology. The solution, which may be verified by substitution, is

$$\begin{aligned} u(c,t) &= e^{-\lambda t}\ln c & \qquad c^*(\underline{x},\underline{w},t) &= \lambda w \\ dw &= w[f-\lambda]\,dt + wg\,dz & \qquad \alpha^*(\underline{x},\underline{w},t) &= 0 \end{aligned} \tag{8}$$

$$J(x,w,t) = \tfrac{e^{-\lambda t}}{\lambda}\ln w + h(x,t)$$

For securities with q and $P(.,T)$ not dependent on w, equilibrium prices are also independent of w, and given by the solution $P(x,t)$ to

$$0 = \frac{1}{2}\sum_{i,j=1}^{n}(ss')_{ij}(P_{xx})_{ij} + (a' - gs')P_x + q + P_t - (f - gg')P \tag{9}$$

subject to the given boundary condition $P(x, T)$.[6]

2 The Riskless Interest Rate

The individual rationally anticipates always holding zero quantity of the financial asset. His marginal rate of substitution between it and current consumption would be unaffected by the availability of additional securities that he anticipates also never holding. Thus (9) applies to all zero aggregate supply financial assets that might be introduced.[7]

[6] If either q or P(.,T) does depend on w, then the additional terms

$$\ldots + gg'w^2P_{ww}/2 + [f - gg' - \lambda]wP_w + gs'wP_{wx}$$

must be added to the pde for P. The function $h(x,t)$ is the solution to

$$0 = h_t + a'h_x + \sum(ss')_{ij}(h_{xx})_{ij}/2 + e^{-\lambda t}(f - \lambda + \lambda\ln\lambda - gg'/2)/\lambda$$

[7] More formally, let α_i denote the fraction of wealth invested in the i^{th} zero aggregate supply security and expand the individual's optimal consumption- investment problem of (a). The first order conditions for an optimum, upon substitution of the market clearing conditions $\alpha_i^* = 0$, are as in (6) except that there is one additional equation like (6b) for each additional security. The β, q, P, σ in the i^{th} such equation are the values for the corresponding security only. Its pricing function can thus be found without reference to the other securities. This piecemeal approach is only valid, however, in a world of homogeneous individuals. With differences in tastes or endowments equilibrium prices would generally be influenced by the menu of securities available for trade, i.e., the degree of completeness of markets.

Consider a riskless asset whose yield accrues entirely in the form of dividends. Since its price cannot vary with either x or t, it is characterized by

$$P_x = P_{xx} = P_t = 0$$

Equation (9) reduces to $0 = q - (f - gg')P$. The dividend yield q/P, or riskless interest rate, is thus

$$r = f - gg' \tag{10}$$

at each point in time. This is the instantaneous expected rate of return on real investment less its instantaneous variance (times the individual's index of relative risk aversion). The pricing relation for any other securities present may be expressed as a differential equation involving r by substituting rP for the last term in (9). By appropriate selection of the technology $\{f, g, a, s\}$, various equilibrium processes for r can be induced, and the corresponding risk adjustment gs' determined.

3 A Two Factor Interest Rate Process

The particular technology we assume in order to empirically specify the model is as follows. There are two state variables: $x \equiv (x_1, x_2)'$. Let the technology be

$$f(x) = (1 + G_1^2 + G_2^2)x_1 \qquad g(x) = (G_1 x_1^{1/2}, G_2 x_1^{1/2})$$

$$a(x) = \begin{pmatrix} \kappa_1 x_1 \ln(x_2/x_1) \\ \kappa_2 x_2 \ln(\gamma/x_2) \end{pmatrix} \quad s(x) = \begin{bmatrix} x_1\sigma_1 & 0 \\ \rho x_2 \sigma_2 & x_2\sigma_2(1-\rho^2)^{1/2} \end{bmatrix}$$

The parameters are all constants. Substituting these into (10) gives the equilibrium riskless rate

$$r = f - gg' = x_1 \tag{11}$$

Defining the new variables $\mu \equiv x_2$, $d\tilde{z}_1 \equiv dz_1$, $d\tilde{z}_2 \equiv \rho\, dz_1 + (1-\rho^2)^{1/2} dz_2$ and substituting the relevant functions into equation (2) determines the joint process followed by r, μ :

$$\begin{aligned} dr &= \kappa_1 r(\ln\mu - \ln r)\,dt + \sigma_1 r\, d\tilde{z}_1 \\ d\mu &= \kappa_2 \mu(\ln\gamma - \ln\mu)\,dt + \sigma_2\mu\, d\tilde{z}_2 \end{aligned} \tag{12}$$

where $d\tilde{z}_1, d\tilde{z}_2$ have variance per unit time of 1 and local correlation coefficient ρ. The riskless rate r tends toward a current target level μ , while μ in turn regresses toward some long run normal value γ, subject to stochastic shocks. The magnitude of these shocks are proportional to the current levels of r and μ respectively, insuring that interest rates remain positive.

The technology was chosen, of course, to induce this particular two factor interest rates process. It is a natural extension of the one factor 'regression towards normal value' models of Cox, Ingersoll and Ross (1978) or Vasicek (1977), and in the spirit of the two factor models of Brennan and Schwartz (1979,80). The reason for having the speed of adjustment depend on logarithms of current values will become apparent in the next section.

The corresponding risk adjustment term in the pricing equation (9) is

$$gs' = (\lambda_1 r^{3/2}, \lambda_2 \mu r^{1/2}) \tag{13}$$

in which $g_1 \equiv \sigma_1 G_1$ and $\lambda_2 \equiv \sigma_2[\rho G_1 + (1-\rho^2)^{1/2} G_2]$ are constants, and the substitutions $r \equiv x_1$, $\mu \equiv x_2$ have been made. These risk adjustments are proportional to the covariances of w with r and μ respectively. The partial differential equation satisfied by the pricing function $P(r, \mu, t)$ for securities with q and $P(., T)$ independent of w is thus

$$\begin{aligned} &\tfrac{1}{2} r^2 \sigma_1^2 P_{rr} + \rho r \mu \sigma_1 \sigma_2 P_{r\mu} + \tfrac{1}{2} \mu^2 \sigma_2^2 P_{\mu\mu} + [\kappa_1 r \ln(\mu/r) - \lambda_1 r^{3/2}] P_r \\ &+ [\kappa_2 \mu \ln(\gamma/\mu) - \lambda_2 \mu r^{1/2}] P_\mu + q + P_t - rP \ = \ 0 \end{aligned} \tag{14}$$

The eight pricing model parameters are $\Gamma \equiv \{\kappa_1, \kappa_2, \gamma; \sigma_1, \sigma_2, \rho; \lambda_1, \lambda_2\}$.

Theoretical prices for interest rate related securities are given by the solution to (14). Different securities are distinguished by their dividend rate q, terminal value $P(., T)$ and any other applicable boundary conditions. Our empirical concern is with the prices of bills, futures contracts on bills, and (European) options on bills. For notational convenience, let time now be measured backwards from the maturity date of an instrument. With t denoting time to maturity, (14) applies with the sign on P_t reversed. Using B rather than P when referring to prices of discount bonds, the relevant dividend rates and terminal values are as follows:[8]

Bills with maturity value 1:

$$q = 0, \qquad B(r, \mu, 0) = 1$$

Futures prices for M year bills:

$$q = rP, \qquad P(r, \mu, 0) = B(r, \mu, M)$$

Options on M year bills with exercise price X:

$$q = 0$$

calls $\quad P(r, \mu, 0) = \max\{0, B(r, \mu, M) - X\}$

puts $\quad P(r, \mu, 0) = \max\{0, X - B(r, \mu, M)\}$

4 Discussion

In this model fluctuations in security prices result from fluctuations in the expected return to current and future real investment. These expectations are in the minds of the investors. Their rationality depends on whether the assumed technology truly does describe the process of realized returns, and on whether individual perceptions of the process parameters are

[8] See Cox, Ingersoll and Ross (1981). By an arbitrage argument, they show that the equilibrium futures contract price (with daily settlement of gains and losses) equals the price P of an asset that pays a continuous dividend rP and whose value on delivery day equals the spot market price of the commodity on which the contract is written.

correct. But nothing so far requires rationality in that sense. The function $P(r, \mu, t)$ gives equilibrium prices that would prevail if investors held firm beliefs that the return on physical capital was as assumed and held point expectations about parameter values. The model is thus most appropriately viewed as a framework for describing *expectations* about interest rates in an internally consistent fashion. The issues of rationality of expectations (not assumed) and their consistency (assumed) should be sharply distinguished.

Before proceeding further, some comment should be made on the general equilibrium approach we have taken versus an arbitrage approach to the same pricing problem. Relation (14) can be obtained directly from (12) and (13) directly by arbitrage arguments alone (cf. Richard, 1978, or Brennan and Schwartz 1979). A straight arbitrage approach leaves unspecified the form of the risk adjustment terms, and a functional form would have to be arbitrarily selected for empirical work. As has been pointed out by Cox, Ingersoll and Ross (1985), however, reasonable forms for these terms can lead to an internally inconsistent model — one that does admit opportunities for riskless arbitrage. The problem is that there may be no economy (tastes and technology) that could both follow the claimed process for interest rates and support the assumed risk premiums. By starting from a general equilibrium model (despite the additional pain to the reader) we are assured that this type of inconsistency has not been introduced.

Strictly interpreted, arbitrage arguments also imply an exact raltionship between state variables and security prices. That is, they imply zero pricing residuals when comparing observed security prices with those predicted by the model because pricing residuals would imply arbitrage opportunities. There are many factors determining security prices which are not captured in theoretical models. There is a tradition in general equilibrium modelling of treating the resulting model as a stochastic relationship which only explains part of the data. If P denotes the theoretical price and $\hat{P}$ the observed security price, then the general equilibrium model is usually taken to imply that $\hat{P}_t = P_t + \epsilon_t$ where ϵ_t is a random variable. The theoretical model usually provides no guidance to the structure of ϵ_t. Its specification and analysis becomes a problem of econometrics. Problems in the estimation of model parameters are thus more readily formulated in the general equilibrium setting.

As is stands, we have derived a model of the term structure of real rather than nominal interest rates. However, one may interpret all real variables in the model of this section as nominal variables. The inflation rate becomes one component of the 'state of technology' x influencing the nominal return to investment in real capital. The influence of monetary policy would be subsumed in the overall process for x. A problem arises with utility being a function of nominal consumption. However the logarithmic utility function chosen is additively separable in prices and nominal consumption. As long as nominal money balances form a negligible component of aggregate wealth, it can be shown that the equilibrium process for nominal interest rates will take the form (12).

III EMPIRICAL IMPLEMENTATION

This section describes the estimation method in more detail. Three problems arise in trying to fit the theoretical model (12) and (14) to observed prices. First, (12) specifies the continuous time dynamics of r, μ; but only discrete data are available for estimation. Second, the predicted security prices conditional on r, μ given by (14) is not explicit. Finally, the state variable μ has no obvious empirical proxy — it is an unobservable, or 'latent', variable. These problems are treated respectively by solving for the relationship between the parameters of (12) and the parameters of the corresponding discrete time process, by numerically solving differential equation (14) for predicted security prices conditional on Γ, and by jointly estimating the sample path of μ along with Γ.[9] In addition, since all the parameters of (12) also enter (14), there may be a conflict between the values of Γ that give the smallest security pricing errors and the values that best describe the time series properties of r, μ. Attaching more weight to the former gives estimates of the market's apparent *ex ante* beliefs about the r, μ process; attaching more weight to the latter gives estimates of what would have been rational *ex post*.

1 Discrete Time Dynamics of the State Variables

Define the transformed variables $R \equiv \ln r$, $L \equiv \ln \mu$, $G \equiv \ln \gamma$. Applying Ito's lemma to (12), the transformed state variables follow the process

$$\begin{aligned} dR &= [\kappa_1(L - R) - \sigma_1^2/2]\, dt + \sigma_1 d\tilde{z}_1 \\ dL &= [\kappa_2(G - L) - \sigma_2^2/2]\, dt + \sigma_2 d\tilde{z}_2 \end{aligned} \tag{15}$$

That is, they follow a linear, constant coefficient, constant variance Ito process. In discrete time, the process followed by R and L can be shown to be (derived in the appendix; see Wymer 1972)

$$\begin{aligned} R_{t+1} &= (1 - k_1)R_t + \kappa_1\Big(\frac{k_1 - k_2}{\kappa_1 - \kappa_2}\Big)L_t \\ &\quad - \Big(\frac{\sigma_1^2 k_1}{2\kappa_1} + (\frac{\sigma_2^2}{2} - G\kappa_2)(\frac{k_2}{\kappa_2} - \frac{k_1 - k_2}{\kappa_1 - \kappa_2})\Big) + \psi_{1t} \\ L_{t+1} &= (1 - k_2)L_t + k_2 G - \frac{\sigma_2^2 k_2}{2\kappa_2} + \psi_{2t} \end{aligned} \tag{16}$$

[9]Brennan and Schwartz face the same estimation problem in their two factor model of the term structure. In their formulation, the two state variables are the risk free instantaneous rate and the unobserved consol rate. They use the long term bond rate as a proxy for the consol rate. That procedure anchors both ends of the yield curve so risk parameters only determine the curvature of the yield curve between these points. In our estimation the yield curve is only anchored at one end by the risk free rate. The value of μ at each point in time is then selected which, together with the model parameters, gives the best fit to the overall yield curve at that time point. Although our procedure is computationally more complex, it should lessen the measurement error bias that results from imperfect observation of the long term conceptual state variable.

where $k_1 \equiv 1 - \exp -\kappa_1$, $k_2 \equiv 1 - \exp -\kappa_2$. The stochastic terms ψ_{1t}, ψ_{2t} are serially uncorrelated and normally distributed, with zero means and approximate covariance matrix

$$\Omega_1 = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \tag{17}$$

Equations (16) and (17) give the distribution of one period ahead values of R and L in terms of their current values and the continuous time parameters. The reasons for selecting the technology of section II.3 should now be apparent. First, at least one state variable corresponds to an observable ($x_1 = r$). Second, the pair of state variables have an explicit distribution over discrete intervals (joint lognormal) with parameters readily related to those of the continuous time process.[10]

2 Predicted Security Prices

Although we cannot solve differential equation (14) explicitly for theoretical prices, it can be solved numerically for given values of Γ. A finite difference approach is used in which $P(r, \mu, t; \Gamma)$ is computed on a grid $[0, \bar{r}] \times [0, \bar{\mu}] \times [0, T]$ using centered differences as approximations to differentials, where $\bar{r}, \bar{\mu}$ denote the maximum values of r, μ considered. P at the grid points is calculated using the alternating direction implicit method of McKee and Mitchell (1970), used also by Courtadon (1982) and Schaefer and Schwartz (1983). For n, m grid points in the state variable directions, the solution at each time step involves solving n tridiagonal systems of m linear equations, plus m systems of n equations.

States $r = 0$ and $\mu = 0$ are absorbing barriers for process (12). One can show that $P(0, \mu, t) = P(r, 0, t) = 1$ for unit discount bonds. However $P_r, P_\mu \to -\infty$ as $r, \mu \to 0$, and imposition of these values renders the numerical solution unstable. Consequently all boundary grid values at each time step are generated by quadratic extrapolation of immediately interior values.[11]

[10] The joint lognormality of the state variables also makes computation of the distribution of future interest rates, and of the steady state distribution of rates, relatively straightforward for forecasting and other applications. The exact conditional covariance matrix, rather than the approximation of equation (17), must be used when seeking the conditional distribution of states more than a short time into the future.

[11] In essence we are imposing the condition $P_{rrr} = P_{\mu\mu\mu} = 0$ one and a half grid points in from the boundary. In the empirically relevant range (for the parameter estimates obtained), the resulting solution seemed quite insensitive changes in the size or mesh of the grid, transformation of the variables to $\ln r, \ln \mu$, etc.

First partial derivatives of security prices with respect to state variables, required in the estimation procedure and for computing hedge ratios, are computed as the derivatives of the quadratic Taylor series interpolation of the grid solution (cf. Isaacson and Keller, 1966). Derivatives with respect to model parameters, also required for the estimation procedure, are computed by incrementing the relevant parameter and resolving the system.

3 Joint Estimation of Γ and L

We now turn to the problem of jointly estimating the sample path of L (equivalently μ) and the parameters Γ. Part of the statistical model's structure was obtained above. From (16) we have

$$\psi_t = \begin{pmatrix} R_{t+1} \\ L_{t+1} \end{pmatrix} - \begin{pmatrix} F_1(R_t, L_t; \Gamma) \\ F_2(R_t, L_t; \Gamma) \end{pmatrix} \tag{18}$$

F_1 and F_2 are the right hand side expressions of (16), and $\psi_t \sim N(0, \Omega_1)$. At each sample point, we observe R_t and a vector of prices $\hat{P}_t = (\hat{P}_{it})$ of n interest rate related securities. Let $P_t = (P_i(R_t, L_t; \Gamma))$ denote the vector of corresponding theoretical prices obtained from the numerical solution of (14). There may be no L_t, Γ for which the theoretical and observed prices coincide exactly. We postulate pricing residuals

$$\epsilon_t = \hat{P}_t - P_t \tag{19}$$

with the following structure: for some scalar α,

$$\nu_t \equiv \epsilon_t - \alpha\epsilon_{t-1} \sim N(0, \Omega_2) \tag{20}$$

Thus the n pricing residuals at each point in time are permitted to have a single common level of autocorrelation and a constant covariance matrix.

We do not perform a full maximum likelihood estimation. However it is instructive to see what it would entail. The log likelihood function (times 2) of a sample of T observations is

$$-T\,\ln|\Omega_1| - \underbrace{\sum_t \psi_t'\Omega_1^{-1}\psi_t}_{SSR_1} - T\,\ln|\Omega_2| - \underbrace{\sum_t \nu_t'\Omega_2^{-1}\nu_t}_{SSR_2} + \text{const.} \tag{21}$$

Direct maximization of (21) with respect to all the unknown parameters would be difficult. Not only are there the usual problems associated with FIML estimation of a nonlinear system, but the elements of Ω_1 directly enter the theoretical price function P.[12]

Suppose that estimates of Ω_1 and Ω_2 were available. Maximum likelihood estimation of the remaining parameters would then reduce to the generalized least square estimator

$$\min_{\kappa,\lambda,L}\{\, SSR_1 + SSR_2 \,\} \tag{22}$$

This is in fact what we do. The solution to (22) is found by an iterative procedure, adjusting $\{\kappa, \lambda, L\}$ towards the values that minimize $SSR_1 + SSR_2$, with Ω_1 and γ derived from the L of the preceeding iteration. The method of Marquardt (1963) is used, which interpolates between a gradient and a Taylor series method of minimizing the SSR. An estimate for Ω_2

[12] If an estimator based on (21) was desired, the work of Dempster, Laird and Rubin (1977) suggests this could be accomplished by iteratively estimating $L|\Gamma$, then $\Gamma|L$, and repeating until convergence obtains. They term such a procedure an 'EM algorithm', and demonstrate that it converges to a joint maximum likelihood estimator under fairly general conditions.

was found by initially assuming a diagonal covariance matrix, solving (22), then using the resulting pricing residuals to form the estimate. An estimate of α was obtained in similar fashion.

Our reason for proceeding in this fashion is primarily computational. Since one endpoint and the general slope of the yield curve at each sample point are accomodated by R_t, L_t, the remaining parameters mainly influence its curvature properties. It seems unlikely that more than the four additional parameters κ, λ can be feasibly identified from these properties. Letting Ω_1 be determined solely by the time series characteristics of R and L resolves this indeterminacy with little effect on the maximized likelihood. As was later verified empirically, procedure did not cause parameter estimates to deviate significantly (in a statistical sense) from their maximum likelihood values.[13]

The theoretical model (12) and (14), or their empirical counterparts (16) and (19), in a sense predict two things: future values of the state R, L *and* security prices conditional on R, L. The estimates of κ from (22) also reflect these two things. The value for κ that minimizes SSR_1 by itself would be that which best describes the ex post time series properties of R, L. It would be the ex post rational belief about the process driving interest rate movements. The value of κ that minimizes SSR_2 by itself would be that which was most consistent with

[13] An example using the Cox, Ingersoll and Ross one factor term structure model will help clarify the estimation procedure. The interest rate process is $dr = \kappa(\mu - r)\,dt + \sigma r^{1/2}\,dz$ where μ is the assumed constant long run value of r. If we define $y_t = r_t^{1/2}$, Itos's lemma gives

$$dy = \frac{1}{2y}[\kappa(\mu - y^2) - \sigma^2/4]\,dt + \frac{1}{2}\sigma\,dz$$

Next expand $1/y$ in a Taylor series about the sample mean $\bar{y}$ to obtain $dy = (ay + b)\,dt + \frac{1}{2}\sigma\,dz$, whose solution is

$$y_t = \frac{b}{a}(e^a - 1) + e^a y_{t-1} + \epsilon_t$$

where $\epsilon_t \sim N(0, \sigma^2/4)$, $a = (\sigma^2/4 - \kappa\mu)/2\bar{y}^2$ and $b = (\kappa\mu - \sigma^2/4)/\bar{y}$.

Solution of the appropriate pricing equation gives prices for discount bonds maturing in τ periods of

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

where

$$\begin{aligned} A(\tau) &= \frac{2\gamma\, e^{(\kappa+\lambda+\gamma)\tau/2}}{(\kappa+\lambda+\gamma)(e\gamma\tau - 1) + 2\gamma}2\kappa\mu/\sigma^2 \\ B(\tau) &= \frac{2(e^{\gamma\tau} - 1)}{[(\kappa+\lambda+\gamma)(e^{\gamma\tau} - 1) + 2\gamma]} \\ \gamma &= \left[(\kappa+\lambda)^2 + 2\sigma^2\right]^{1/2} \end{aligned}$$

and λ is an additional parameter that specifies the market price of interest rate risk.

Estimating the model would require determining four parameters κ, μ, σ^2 and λ from time series data for r and discount bond prices. An examination of the equation for y_t indicates that only two parameters can be uniquely determined from the time series for r_t. The value of σ^2 can be determined from the autoregressive residuals, but only the product $\kappa\mu$ can be inferred from the intercept and slope parameters. The time series r_t, of course, provides no information on λ. Similarly, only three parameters can be identified from bond prices because only the combinations $\kappa\mu/\sigma^2$, $\kappa+\lambda$ and γ affect bond prices. As a result, no information is lost if σ^2 is estimated from the autoregressive residuals, λ is determined only from the bond prices, and κ and λ are jointly estimated from the data on r_t and bond prices.

The workings of the two factor model are not much different from this example. Equations (16) only provide information on the values of κ and σ^2 and not the values of λ. Similarly, the bond prices do not provide enough information to determine all the model parameters. Although there is not formally the lack of identification that occurs in the one factor model, numerical analysis indicated that a close to similar situation prevails in our two factor case. As a result, little or no information should be lost if equations (16) are used to determine Ω_1 and γ, with the remaining parameters estimated jointly.

the various cross-sections of security prices observed. It would be the value that the market appeared to believe in, given the structure of prices. The theoretical model assumes that expectations are consistent in the sense of section II.4. If expectations were also rational, then the two estimates of κ should coincide. The criterion function (22) imposes this restriction to improve efficiency. However this restriction can be lifted, two separate estimates of κ can be obtained (the 'ex post rational beliefs' and the 'market beliefs'), and the hypothesis of rational expectations tested.[14]

4 Data used in the estimation

The theoretical pricing relation (14) applies to *any* default free interest rate related claim. Hence the model parameters can be estimated from a time series $\{R_t, \hat{P}_t\}$, where the price data $\hat{P}$ is drawn from any market for interest rate securities.

The model was fitted to observations of US government Treasury Bill and Note prices in the cash market; then separately fitted to prices for futures contracts on 91 day Treasury Bills traded on the Chicago Mercantile Exchange. Weekly observations were obtained for 344 weeks from Jan. 5, 1978, to Aug. 9, 1984 (Thursday price quotes) for 4, 13, 26 and 52 week Bills, 3, 5 and 7 year Notes, and the four nearest to delivery T-Bill futures contracts. The Note yields were Federal Reserve Bank of NY estimates of the coupon rates for which Notes of those maturities would trade at par. That is, they represent FRBNY estimates of the yield curve rather than price quotes on actual Notes.[15]

For the riskless instantaneous interest rate, r_t, we used the overnight US Federal Funds rate minus the average differential between the reserve adjusted yield on 3 month bank certificates of deposit and 3 month Treasury Bills. Assuming that the cost of funds yield differential

[14]No properties of the resulting estimators have been demonstrated. The number of L_t's to be estimated grows in direct proportion to the length of the sample. There is thus an incidental parameter issue, raising the possibility of inconsistent structural parameter estimates. We do not feel this is a problem for several reasons. The model is similar to a 'random effects' model in that, although successive realizations of L_t are not independent, there exists a stationary distribution for the state variables with their degree of intertemporal dependence shrinking to zero through time. Moreover changes in R_t, L_t are independently and identically distributed after allowance for the deterministic component in their drift. Thus it may be possible to demonstrate consistency of the maximum likelihood estimator of the structural parameters based solely on (19) along the lines of Kiefer and Wolfowitz (1956). However we have not done this. Additionally, in our particular context, we make use of the dynamic relation (18) in the estimation procedure. This further identifies the parameters to be estimated beyond what is typically the case when incidental parameters are a problem.

[15]For a given maturity of bond, the height of the FRBNY estimated yield curve was interpreted as the coupon rate required on a bond trading at face value. I.e., a bond with that coupon rate was taken to have an observed price of 1.0. The theoretical prices of the discount bonds representing those semiannual coupon payments and the principal repayment were summed. The difference between this sum and 1.0 was used as the pricing residual. Although this procedure carries the disadvantage of not using actual bond quotes, it has the advantage of avoiding tax related coupon effects.

between unsecured bank liabilities and Treasury securities was the same for overnight funds as for 3 month funds, this is a proxy for an overnight Treasury rate.[16]

IV EMPIRICAL RESULTS

1 What are the questions?

The theoretical model of section II, and related arbitrage based models, take a unified view of the pricing of all interest rate claims. This is attractive for several reasons. It permits a parsimonious description of the world, allowing many markets to be treated as one. It explains relationships between yields on technically distinct securities. It provides a way of valuing customized contracts for which no established market exists. And it provides a basis for hedging the risk implicit in one contract or security with positions in other seemingly quite different securities. In this spirit, the fundamental hypothesis might be stated as follows:

On average, the prices of all interest rate securities are a common function of the current instantaneous interest rate r_t and one other common underlying factor μ_t.

By 'on average' is meant up to a pricing error structure of the form (20). By 'common function' is meant that the prices satisfy equation (14). And by 'common factor' is meant a variable that is perceived in a similar fashion by particpants in all markets. Nothing is said about the presence or absence of arbitrage opportunities, which would involve further hypotheses about the error structure, transaction costs, admissible strategies and so on. The hypothesis is not market efficiency, but rather the appropriateness of viewing these markets in a unified manner.

This fundamental hypothesis is of course a joint hypothesis about many things. A rejection of it could be interpreted as a rejection of the adequacy of a two factor description of the exogenous forces driving interest rates, of the correctness of specification (14), or of a multitude of other maintained assumptions underlying the empirical specification.

The issues that we do explore fall into two categories. Under the heading of quantitative issues, we ask: (a) Do the underlying parameters of the interest rate process appear stable over the sample period? (b) Is the theoretical distinction between forward and futures prices quantitatively important for empirical work? (c) For what length of observation interval would an approximate rather than exact discrete time analogue of the continuous model be adequate for estimation purposes? Under the heading of economic issues, we ask: (a) Are the parameters describing the joint movement of short term rates and the unobserved factor the same across the cash and futures markets? (b) Is the current state (L_t) perceived to be the same across the two markets? (c) Were market expectations about the process followed by

[16] Since this is a possibly noisy proxy for a hypothetical overnight Treasury rate, we explored using an exponentially smoothed moving average to reduce the effect of observation error. The best fit was obtained with smoothing parameter of 0.75. However the reduction in pricing residuals and effect on coefficient estimates were negligible, and the procedure was dropped.

the state variables *ex post* rational?

Issues that will be explored in a future paper include: (a) Market efficiency: are the pricing errors within and across markets such that 'profitable' trading rules can be devised? (b) Functional form: which of the various two-factor models proposed in the literature to date perform 'best'? (c) Financial management and intermediation: how well do hedging strategies based on alternative pricing models perform? (e.g., how well can interest rate options be synthesized with futures trading strategies?)

2 Estimation results

Table 1 gives parameter estimates obtained by fitting the model to the Bill and Note data, to the Bill data alone, and to the T-Bill futures contract prices alone. Table 2 gives some information on the size of the resulting pricing residuals.

The parameter values are based on a time unit of one week. The corresponding values when the time unit is one year (52 weeks), based on the Bills/Notes estimates (row 1 of Table 1) are

κ_1	12.72	(half-life .054 yrs.)	σ_1	.596
κ_2	.0718	(half-life 9.6 yrs.)	σ_2	.166
λ_1	-3.727		ρ	.236
λ_2	-.0808		γ	.0983

The very high value of κ_1 combines with the very high value of λ_1 to accomodate the very steep initial part of the yield curve.

The long half-life of deviations of of μ from γ relative to the sample period results in wide confidence bands for γ. Table 5, row 2, tests the constraint that this long run target for interest rates is .00289 (15.0%/yr.) as opposed to the estimated value of .00189 (9.83%/yr.), and cannot reject it. The Bill/Note pricing residuals are little affected since the market price of μ risk, λ_2, adjusts to compensate.

3 Quantitative issues

The sample period January 1978 – August 1984 was a time of considerable interest rate volatility and witnessed a variety of shapes to the yield curve. The state variables, expressed as annual yields, ranged from 5.43% to 16.98% for r_t, and from 6.64% to 14.38% for the estimated μ_t.

The stability of the model parameters was investigated by breaking the sample of Bill and Note prices into two halves. Row 1 of Table 5 details the relevant likelihood ratio test. Imposing identical parameter values over the two intervals raises the $SSSR$ by 0.45%, which rejects the hypothesis at the .95 confidence level but not at the .99 level.

Table 1: PARAMETER ESTIMATES

Prices fit	κ_1	κ_2	λ_1	λ_2	$SSSR$	σ_1	σ_2	ρ	γ
1. bills &	.2446	.00138	-.5169	-.0112	2730	.0826	.0230	.236	.00189
bonds	(.010)	(.00007)	(.020)	(.0007)					
2. bills	.4120	.01356	-.9817	.0093	1677	.0799	.0466	.358	.00188
only	(.027)	(.0011)	(.081)	(.008)					
3. futures	.2476	.00128	-.5281	-.0102	1717	.0856	.0445	.300	.00170
	(.002)	(.00003)	(.0036)	(.0005)					

Constrained Estimates

		$SSSR$	σ_1	σ_2	ρ	γ
4. bills	κ, λ, L same as (1)	25806				
only	κ, λ same as (1)	1915	.0786	.0425	.162	.00189
	unconstrained	1677	.0799	.0466	.358	.00188
5. futures	κ, λ, L same as (1)	8.02 E6				
	κ, λ same as (1)	1728	.0841	.0438	.310	.00173
	κ, λ, L same as (4b)	7.79 E6				
	unconstrained	1717	.0856	.0445	.300	.00170

$SSSR = SSR_1 + SSR_2$. Standard error of coefficient estimates in parentheses are asymptotic values based on the numerically computed partial derivatives of $SSSR$ on the last iteration of the minimization procedure. They are conditional upon fixed values of $\sigma_1, \sigma_2, \rho, \gamma$. A common value of Ω_1 and Ω_2 was used within (4) and within (5) to make $SSSR$ comparable for hypothesis testing. However Ω_2 differs between (4) and (5).

Turning next to the quantitative importance of the futures- forward price distinction, the theoretical difference between these prices arises from the fact that futures contract gains and losses are settled daily in cash between market participants, whereas all gains and losses are deferred to the delivery date for forward contracts. If the unanticipated components of future price and interest rate changes are negatively correlated, then individuals with long futures positions on average pay out cash when the opportunity cost of doing so is high, and receive cash when the interest that can be earned on it is low. Consequently equilibrium futures prices would be below forward prices.

For many markets, the correlation between interest rate and futures price movements might be weak. However there can be no such presumption for interest rate futures markets. To get some sense of the quantitative importance of the daily settlement effect, the theoretical

Table 2: STANDARD DEVIATION OF PRICING RESIDUALS

1. bills/bonds ($100 maturity value)

maturity:	4 wk.	13 wk.	26 wk.	52 wk.	3 yr.	5 yr.	7 yr.
	.038	.085	.130	.186	.181	.251	.453

2. bills only ($100 maturity value)

maturity:	4 wk.	13 wk.	26 wk.	52 wk.
	.030	.034	.045	.131

3. futures (contract on $100 maturity value 91 day bill)

delivery:	1-13 wk.	14-26 wk.	27-39 wk.	40-52 wk.
	.059	.017	.024	.042

Correlation Matrix of Price Residuals

bills/bonds

1.00	.44	.36	.30	−.13	−.44	−.52
	1.00	.87	.59	−.19	−.75	−.75
		1.00	.77	−.01	−.76	−.83
			1.00	.20	−.67	−.79
				1.00	−.04	−.34
					1.00	.59
						1.00

futures contracts

1.00	.23	−.97	−.89
	1.00	−.23	−.63
		1.00	.83
			1.00

Above are standard deviations of pricing errors after allowance for serial correlation. A common value of $\alpha = .806$ was used for all estimates. The standard deviation reported above is thus $\omega_{jj})^{1/2}$, where ω_{jj} is the j^{th} diagonal element of Ω_2. The standard deviations of the raw (uncorrected) pricing residuals ranged between two and four times larger, depending on the maturity of the security priced.

forward and futures prices of 13 week and 52 week Bills were computed for various current states, using the model parameters estimated from the Bill/Note data.

Table 3 presents the results, with prices converted to annualized yields to maturity. It is apparent that if interest rates are in a 'normal' range, and if the time to delivery is relatively short (a year or less), then the effect is negligible. The difference in implied yield is only 2 basis points for 13 week Bills deliverable in one year. However for more distant delivery dates and/or high levels of current interest rates (the stochastic components of state changes are proportional to the level of interest rates), then the effect is noticeable. For Bills deliverable in five years, with rates at the 20% level, the difference in implied yield is 192 basis points.

Finally, we explore the extent of bias that would be introduced by using approximate discrete time analogues for the interest rate process to estimate the continuous time parameters

Table 3: THEORETICAL FORWARD VERSUS FUTURES PRICES

	Current State							
	$r = \mu =$10%/yr.				$r = \mu =$20%/yr.			
Deliverable Bill:	13 week		52 week		13 week		52 week	
0 yr. forward	10.66		10.99		21.91		22.46	
1 yr. forward	11.23	(.02)	11.31	(.02)	22.46	(.08)	22.27	(.07)
3 yr. forward	11.62	(.16)	11.67	(.15)	21.36	(.79)	21.13	(.58)
5 yr. forward	11.84	(.41)	11.86	(.40)	20.09	(1.92)	19.85	(1.35)

The forward and futures prices have been expressed as continuously compounded annual yields to maturity on the delivered bill. The main entries in the table are the forward interest rates. The quantity in parentheses following each entry is the amount by which the theoretical futures market interest rate exceeds the corresponding forward rate. Theoretical forward and futures prices were all computed using parameter values from row 1 of Table 1.

entering the pricing relations. The production technology assumed in the theoretical model was chosen so that the exact discrete time process followed by the state variables could be obtained. However most specifications for the interest rate process that have been proposed cannot be similarly solved. The question arises whether it is reasonable to empirically implement these models by simply replacing infinitesimals in equations like (15) by differences, and ignoring the distinction between discrete and continuous time parameters.

This is an empirical issue in that as the interval between observations shrinks, the approximate discrete time parameters and exact parameters converge. The question is, how short an interval is short enough in a given context? To shed some light on this issue, we replaced dR and dL in (15) by differences $(R_{t+1} - R_t)$ and $(L_{t+1} - L_t)$, and used the average of L and R at times t and $t+1$ on the right hand side. This was then used in place of (16) in fitting the model to the Bill/Note data. With weekly observation intervals there was virtually no difference in the parameter estimates. The observation interval was then increased to a month and the experiment repeated.

Table 4 gives the results. As can be seen, the difference in parameter estimates is now noticeable. However the approximate model does not have much greater pricing residuals. This suggests that with even monthly observations, the use of approximate discretizations of continuous interest rate processes is adequate. Of greater concern is the disparity between estimates obtained from monthly observations and from weekly observations (row 3).

4 Market expectations and perceptions

For the same pricing equation (14) to apply for all interest rate claims, the same exogenous factors must drive the prices of distinct interest rate dependent securities, participants in these

Table 4: APPROXIMATE DISCRETE TIME MODEL ESTIMATES

	κ_1	κ_2	λ_1	λ_2	$SSSR$	σ_1	σ_2	ρ	γ
1. Exact discrete time model:									
	.5451	.00834	-.4787	-.0210	627	.1101	.0754	.391	.0081
2. Approximate discrete time model:									
	.4957	.00906	-.4840	-.0190	637	.1147	.0766	.394	.0081
3. Per month values implied by weekly parameter estimates:									
	.9784	.00552	-1.034	-.0224		.1652	.0460	.236	.0076

All estimates are based on monthly observations (4 weeks) of bill and bond prices. Row 3 gives the parameter estimates of row 1 of Table 1 converted to equivalent values for a monthly, rather than weekly, time unit. The value of the serial correlation coefficient for the monthly pricing residuals was $\alpha = .726$.

The parameter values of row 2 do not differ significantly from those of row 1. The appropriate test statistic, $n \ln(637/627)$, takes the value 3.63 and is asymptotically distributed as $\chi^2(4)$. The .90 and .95 significance level values for this statistic are 7.78 and 9.49 respectively.

markets must share common beliefs about the process governing their movement, and similar premiums for bearing equivalent risks must prevail across markets. In addition, the success of hedging strategies based on the theoretical price relations hinges on the rationality of these beliefs. In our context, we ask whether the parameters of the interest rate process are the same across the cash and futures market, whether the perception of the current state is the same in the two markets, and whether the cross equation restrictions implied by rationality can be rejected. The likelihood ratio test statistics for these hyotheses are provided in Table 5.

Row 3 tests the hypothesis that the process parameters and risk prices are the same in the T-Bill futures market as in the Bill/Note market. No joint fit to the Bill/Note/futures data combined was performed. Instead we test the restriction that values of κ and λ for the futures market equal the values obtained for the Bill/Note market (Table 1, row 1), without constraint on the values of L_t. The hypothesis of equal process parameters and risk prices across the two markets is not rejected at the .99 significance level.

Row 4 tests the hypothesis that, in addition, the perception of the current state (value of the unobserved second factor μ) is the same in the futures market as the cash market. Here the result is rather different. Imposing the values of μ estimated from the cash market increases the futures market $SSSR$ pricing residuals by a factor of over 4000, thoroughly rejecting the hypothesis. This suggests either that different factors influence these two markets, or 'news' is processed and incorporated into prices with a different lag by representative participants of

Table 5: STATISTICS FOR HYPOTHESIS TESTS

	Null hypothesis	$\frac{S_0}{S_1}$	n	q	$n\ln(\frac{S_0}{S_1})$	$\chi^2_{.95}(q)$	$\chi^2_{.99}(q)$
1.	κ, λ are the same for the 1st half of the sample period as the 2nd half	1.0045	2752	4	12.24	9.49	13.3
2.	$\gamma = .00289$ (versus unconstrained estimate of .00189)	1.0008	2752	1	2.07	3.84	6.63
3.	κ, λ for the futures market = values estimated for bill/bond market	1.0064	1720	4	10.98	9.49	13.3
4.	$\kappa, \lambda, \{\mu_t\}$ for the futures market = values estimated for bill/bond market	4641	1720	347	14521	390	408
5.	κ in bill/bond market = values estimated from time series on r_t, μ_t alone (with μ_t fixed at values estimated from bill/bond)	2.0659	2408	2	1747	5.99	9.21
6.	κ perceived by bill/bond market and that implied by r_t, μ_t time series are same (i.e., rational expectations)	1.0200	2752	2	54.37	5.99	9.21

S_0 and S_1 are the estimated sum squared residuals under the null and alternative (unconstrained) hypotheses respectively. The residuals were transformed by the appropriate covariance matrix first. n denotes the number of (independent) observations in the sample, and q the number of constraints imposed by the null hypothesis. The likelihood ratio test statistic in column 4 is asymptotically distributed as $\chi^2(q)$. The values corresponding to 95% and 99% significance level tests are given in the final two columns. The critical values in row 4 were computed using the normal approximation to the χ^2 distribution for large q.

the two markets.

The rationality of expectations was investigated using Bill/Note data only. The interest rate process parameters κ enter both equation (16), which describes the dynamic properties of the exogenous factors, and equation (19), which determines the cross-section structure of security prices. The value of κ that makes (16) fit best (minimizing SSR_1 alone) can be interpreted as the *ex post* rational belief about the interest rate process; while the value that makes (19) fit best (minimizing SSR_2 alone) can be interpreted as the market's belief about the interest rate process. The hypothesis of rational expectations is that these values are the same — a cross equation restriction on parameters.

The values of κ estimated using these different criteria are listed below. The values of μ_t for the time series estimate of κ were fixed at the values estimated for the Bill/Note market assuming rational expectations (i.e., Table 1, row 1).

	κ_1	κ_2
minimizingSSR_1	.1264	.01210
minimizingSSR_2	.2710	.00183
minimizing$SSSR$	.2446	.00138

Comparing the first two rows, the market appeared to overestimate the rate at which short term rates were being pulled toward their near term target μ, but underestimate the tendency for this near term target to be pulled back towards the long run value γ.

Row 6 of Table 5 formally tests the restriction that the values of κ are the same. The hypothesis is rejected at the .99 significance level. This suggests that expectations were not *ex post* rational. However imposing equality of the two values for κ only raises the combined $SSSR$ by 2%, indicating that the model's pricing residuals are only slightly affected by the misperception of the interest rate process. Put another way, the extent of irrationality may be statistically significant, but may not be quantitatively important for pricing and hedging purposes (see McClosky 1985).

Finally, row 5 of Table 5 details the outcome of taking a different route to testing for rational expectations — a two step procedure. If, instead of fitting (16) and (19) jointly subject to the constraint of a common κ, one estimates κ from the time series of the state variables (minimizing SSR_1), then imposes this value in the pricing equation (19), the sum squared pricing errors SSR_2 more than doubles, giving the appearance of quantitatively large deviations from rationality. The problem with performing the test this way is that it treats the value of κ as known rather than as an estimate, increasing the chance of erroneously rejecting the null hypothesis.

References

[1] Ayres, H. R., and Barry, J. Y., "The Equilibrium Yield Curve for Government Securities," *Financial Analyst's Journal*, May/June 1979, 31-39.

[2] Bard, Yonathan, *Nonlinear Parameter Estimation*, New York: Academic Press, 1974.

[3] Black, Fischer, "The Pricing of Commodity Contracts," *J. Financial Economics*, 3 (1976), 167-179.

[4] Brennan, Michael J., and Schwartz, Eduardo S., "Savings Bonds, Retractable Bonds and Callable Bonds," *J. Financial Economics*, 5 (1977), 67-88.

[5] Brennan, M.J., and Schwartz, E.S., "Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims: A Synthesis," *J.Financial and Quant. Analysis*, Sept.1978, 461-474

[6] Brennan, Michael J., and Schwartz, Eduardo S., "A Continuous Time Approach to the Pricing of Bonds," *J. of Banking and Finance*, 3 (1979), 133-155.

[7] Brennan, Michael J., and Schwartz, Eduardo S., *Savings Bonds: Theory and Empirical Evidence*, Monograph Series in Finance and Economics, 1979-4, Business School of New York University.

[8] Brennan, Michael J., and Schwartz, Eduardo S., "Conditional Predictions of Bond Prices and Returns," *Journal of Finance*, 35 (May 1980), 405-417.

[9] Brennan, Michael J., and Schwartz, Eduardo S., "Bond Pricing and Market Efficiency," *Financial Analyst's Journal*, Sept./Oct. 1982, 3-10.

[10] Brennan, Michael J., and Schwartz, Eduardo S., "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency," *J. Financial and Quantitative Analysis*, 17 (Sept. 1982),

[11] Brennan, Michael J., and Schwartz, Eduardo S., "Alternative Methods for Valuing Debt Options," UBC Working Paper 888, June 1982.

[12] Capozza, Dennis R., and Gau, George W., "The Pricing and Implementation of Mortgage Rate Insurance," mimeo, University of British Columbia, August 1983.

[13] Courtadon, Georges, "The Pricing of Options on Default-Free Bonds," *J. Financial and Quantitative Analysis*, 17 (March 1982), 75-100.

[14] Courtadon, Georges, "A More Accurate Finite Difference Approximation for the Valuation of Options," *J. Financial and Quantitative Analysis*, 17 (Dec. 1982), 697-703.

[15] Courtadon, George, "Interest Rate Uncertainty and the Pricing of American Call Options," mimeo, 1982.

[16] Cox, John C., Ingersoll, Jonathon E., and Ross, Stephen A., "A Theory of the Term Structure of Interest Rates," Research Paper 468, Stanford University, August 1978

[17] Cox, John C., Ingersoll, Jonathan E. Jr., and Ross, Stephen A., "Duration and the Measurement of Basis Risk," *Journal of Business*, 52 (1979), 51-61.

[18] Cox, J.C., Ingersoll, J.E., and Ross, S.A., "The Relation Between Forward Prices and Futures Prices," *J. Financial Economics*, 9 (Dec. 1981), 321-346.

[19] Cox, J.C., Ingersoll, J.E., and Ross, S.A., "A Theory of the Term Stucture of Interest Rates," *Econometrica*, 53 (March 1985), 385-409.

[20] Dempster, A.P., N.M. Laird and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, Series B, 39 (1977), 1-22.

[21] Dothan, L. Uri, "On the Term Structure of Interest Rates," *J. Financial Economics*, 6 (1978), 59-69.

[22] Fleming, Wendell H., and Rishel, Raymond W., *Deterministic and Stochastic Optimal Control*, New York: Springer-Verlag, 1975.

[23] Gladwell, I., and Wait, R., *A Survey of Numerical Methods for Partial Differential Equations*, Oxford: Clarendon Press, 1979.

[24] Goldfeld, Stephen M., and Quandt, Richard E., *Nonlinear Methods in Econometrics*, Amsterdam: North-Holland, 1972.

[25] Ingersoll, Jonathon E., Jr., "Forward Rates and Expected Spot Rates: The Effects of Uncertainty," mimeo, Wells Fargo Conference, Sept. 1977.

[26] Isaacson, Eugene, and Keller, Herbert Bishop, *Analysis of Numerical Methods*, New York: John Wiley & Sons, 1966.

[27] Keifer, J., and J. Wolfowitz, "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters," *Annals of Mathematical Statistics*, 27 (1956), 887-906.

[28] Kreps, David M., "Asset Pricing Models," IMSSS Workshop, August 1982.

[29] Kushner, Harold J., *Stochastic Stability and Control*, New York: Academic Press, 1967.

[30] Marquardt, Donald W., "An Algorithm for the Least-Squares Estimation of Non-linear Parameters," *J. Society of Industrial and Applied Mathematics*, 11 (June 1963), 431-41.

[31] Marsh, Terry A., "Equilibrium Term Structure Models: Test Methodology," *Journal of Finance*, 35 (May 1980), 421-435.

[32] Marsh, Terry A., and Rosenfeld, Eric R., "Stochastic Processes for Interest Rates and Equilibrium Bond Prices," *Journal of Finance*, 38 (May 1983), 635-646.

[33] Marsh, Terry A., and Rosenfeld, Eric R., "Discrete Time Approximations to Continuous Time Models in Finance," mimeo, Sloan School of Management, MIT, 1983.

[34] McKee, S., and A. R. Mitchell, "Alternating Direction Methods for Parabolic Equations in Two Space Dimensions with a Mixed Derivative," *Computer Journal*, 13, February 1970, 81-86.

[35] Merton, Robert, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3 (1971), 373-413.

[36] Phillips, P.C.B., "The Structural Estimation of a Stochastic Differential Equation System," *Econometrica*, 6 (November 1972), 1021-1041

[37] Richard, Scott F., "An Arbitrage Model of the Term Structure of Interest Rates," *J. Financial Economics*, 6 (1978), 33-57.

[38] Schaefer, Stephen M., and Schwartz, Eduardo S., "A Two-Factor Model of the Term Structure: An Approximate Analytical Solution," mimeo, June 1983. (1983 EFA Meetings)

[39] Vasicek, Oldrich, "An Equilibrium Characterization of the Term Structure," *J. Financial Economics*, 5 (1977), 177-188.

[40] Wymer, C.R., "Econometric Estimation of Stochastic Differential Equation Systems," *Econometrica*, Vol. 40, No. 3 (May 1972), 565-578.

Solution for Exact Discrete Time Model

Let $y_1(t) \equiv R_t \equiv \ln r(t)$, $\mathrm{y}_2 \equiv L_t \equiv \ln \mu(t)$ and $G \equiv \ln \gamma$. The stochastic process (15) can be written in vector form as

$$dy = Ay\,dt + b\,dt + d\psi \tag{A.1}$$

where

$$\begin{aligned} y &\equiv \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} & A &\equiv \begin{bmatrix} -\kappa_1 & \kappa_1 \\ 0 & -\kappa_2 \end{bmatrix} \\ b &\equiv \begin{pmatrix} -\sigma_1^2/2 \\ G\kappa_2 - \sigma_2^2/2 \end{pmatrix} & d\psi &\equiv \begin{pmatrix} \sigma_1\,d\tilde{z}_1(t) \\ \sigma_2\,d\tilde{z}_2(t) \end{pmatrix} \end{aligned} \tag{A.2}$$

From Wymer (1972), the exact discrete time solution to the constant coefficient linear system (A.1) is

$$y(t+h) = e^{hA}y(t) + A^{-1}[e^{hA} - I]b + \eta(t) \tag{A.3}$$

The expression e^A is defined as Te^DT^{-1} where T is a matrix whose columns are the eigenvectors of A and e^D is a diagonal matrix with elements e^{c_i}, where the c_i's are the corresponding eigenvalues of A (Note that the eigenvectors of hA are the same as those of A and the corresponding eigenvalues are hc_i). The vector $\eta(t)$ is normally distributed with mean 0 and covariance matrix

$$\hat{\Omega}_1 = \int_0^h e^{sA}\Omega_1 e^{sA'}\,ds \cong \Omega_1 \quad \text{for small } h \tag{A.4}$$

where Ω_1 is the instantaneous covariance matrix of the stochastic terms in the continuous time process. The exact covariance matrix for larger h is derived below.

The eigenvalues and corresponding eigenvectors of A are

$$\begin{aligned} c_1 &= -\kappa_1 & c_2 &= -\kappa_2 \\ T_{\cdot 1} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} & T_{\cdot 2} &= \begin{pmatrix} \kappa_1 \\ \kappa_1 - \kappa_2 \end{pmatrix} \end{aligned} \tag{A.5}$$

Hence we obtain:

$$A^{-1} = \begin{bmatrix} -1/\kappa_1 & -1/\kappa_2 \\ 0 & -1/\kappa_2 \end{bmatrix}$$

$$e^{hA} = \begin{bmatrix} e^{-h\kappa_1} & \frac{\kappa_1(e^{-h\kappa_2} - e^{-h\kappa_1})}{\kappa_1 - \kappa_2} \\ 0 & e^{-h\kappa_2} \end{bmatrix}$$

and $A^{-1}[e^{hA} - I]b \equiv (x_1, x_2)'$ where

$$\begin{aligned} x_1 &= -\sigma_1^2\Delta_1/2\kappa_1 - \sigma_2^2\Delta_2/2\kappa_2 + \sigma_2^2(\Delta_1 - \Delta_2)/2(\kappa_1 - \kappa_2) \\ &\quad + G\Delta_2 - G\kappa_2(\Delta_1 - \Delta_2)/(\kappa_1 - \kappa_2) \\ x_2 &= (G\kappa_2 - \sigma_2^2/2)\Delta_2/\kappa_2 \end{aligned} \tag{A.6}$$

In the above, $\Delta_1 \equiv 1 - e^{-h\kappa_1}$ and $\Delta_2 \equiv 1 - e^{-h\kappa_2}$. Substitution of (A.6) into (A.3) yields the discrete time relation:

$$\begin{aligned} R_{t+h} &= (1-\Delta_1)R_t + \kappa_1\Big(\frac{\Delta_1 - \Delta_2}{\kappa_1 - \kappa_2}\Big)L_t \\ &\quad - \Big(\frac{\sigma_1^2\Delta_1}{2\kappa_1} + (\frac{\sigma_2^2}{2} - G\kappa_2)(\frac{\Delta_2}{\kappa_2} - \frac{\Delta_1 - \Delta_2}{\kappa_1 - \kappa_2})\Big) + \eta_{1t} \\ L_{t+h} &= (1-\Delta_2)L_t + \Delta_2 G - \frac{\sigma_2^2\Delta_2}{2\kappa_2} + \eta_{2t} \end{aligned} \tag{A.7}$$

The stochastic terms ψ_{1t}, ψ_{2t} are serially uncorrelated and normally distributed, with zero means and covariance matrix determined below.

To get the covariance matrix, multiply out the terms in (A.4) to get the integrand

$$\begin{bmatrix} \sigma_1^2\nu_1 + \rho\sigma_1\sigma_2\kappa_1\frac{\nu_2-\nu_1}{\kappa_1-\kappa_2} + \sigma_2^2\kappa_1^2(\frac{\nu_2-\nu_1}{\kappa_1-\kappa_2})^2 & \rho\sigma_1\sigma_2\nu_1\nu_2 + \sigma_2^2\kappa_1\nu_2\frac{\nu_2-\nu_1}{\kappa_1-\kappa_2} \\ \rho\sigma_1\sigma_2\nu_1\nu_2 + \sigma_2^2\kappa_1\nu_2\frac{\nu_2-\nu_1}{\kappa_1-\kappa_2} & \sigma_2^2\nu_2^2 \end{bmatrix} \tag{A.8}$$

in which $\nu_1 \equiv e^{-s\kappa_1}$ and $\nu_2 \equiv e^{-s\kappa_2}$. The integral (A.4) is then the matrix of integrals of the respective terms, which evaluate to:

$$\begin{aligned} \hat{\sigma}_{11} &= \frac{\sigma_1^2}{2\kappa_1}(1-\delta_1^2) + \frac{2\rho\sigma_1\sigma_2\kappa_1}{\kappa_1-\kappa_2}\left(\frac{1-\delta_1\delta_2}{\kappa_1+\kappa_2} - \frac{1-\delta_1^2}{2\kappa_1}\right) \\ &\quad + \frac{\sigma_2^2\kappa_1^2}{(\kappa_1-\kappa_2)^2}\left(\frac{1-\delta_1^2}{2\kappa_1} - \frac{2(1-\delta_1\delta_2)}{\kappa_1+\kappa_2} + \frac{1-\delta_2^2}{2\kappa_2}\right) \\ \hat{\sigma}_{12} &= \frac{\rho\sigma_1\sigma_2}{\kappa_1+\kappa_2}(1-\delta_1\delta_2) + \frac{\sigma_2^2\kappa_1}{\kappa_1-\kappa_2}\left(\frac{1-\delta_2^2}{2\kappa_2} - \frac{1-\delta_1\delta_2}{\kappa_1+\kappa_2}\right) \\ \hat{\sigma}_{22} &= \frac{\sigma_2^2}{2\kappa_2}(1-\delta_2^2) \end{aligned} \tag{A.9}$$

in which $\delta_1 \equiv e^{-h\kappa_1}$ and $\delta_2 \equiv e^{-h\kappa_2}$.

The steady state distribution of R, L — obtained by letting $h \to \infty$ in (A.7) and (A.9) — is normal with mean

$$\begin{pmatrix} \bar{R} \\ \bar{L} \end{pmatrix} = \begin{pmatrix} G - \frac{\sigma_1^2}{2\kappa_1} - \frac{\sigma_2^2}{2\kappa_2} \\ G - \frac{\sigma_2^2}{2\kappa_2} \end{pmatrix}$$

and covariance matrix

$$\begin{bmatrix} \frac{\sigma_1^2}{2\kappa_1} + \frac{\sigma_2}{2\kappa_2(\kappa_1+\kappa_2)}(\sigma_2\kappa_1 + 2\rho\sigma_1\kappa_2) & \frac{\sigma_2}{2\kappa_2(\kappa_1+\kappa_2)}(\sigma_2\kappa_1 + 2\rho\sigma_1\kappa_2) \\ \frac{\sigma_2}{2\kappa_2(\kappa_1+\kappa_2)}(\sigma_2\kappa_1 + 2\rho\sigma_1\kappa_2) & \frac{\sigma_2^2}{2\kappa_2} \end{bmatrix}$$