Kalman Filtering of Generalized Vasicek Term Structure Models

Simon H. Babbs; K. Ben Nowman

Kalman Filtering of Generalized Vasicek Term Structure Models

Simon H. Babbs and K. Ben Nowman*

Abstract

We present a subclass of Langetieg's (1980) linear Gaussian models of the term structure. The bond price is derived in terms of a finite set of state variables with correlated innovations. The subclass contains a reformulation of the double-decay model of Beaglehole and Tenney (1991), enabling us to clarify interpretation of their parameters. We apply Kalman filtering to a state space formulation of the model, allowing measurement errors in the data. One-, two-, and three-factor models are estimated on U.S. data from 1987–1996 and the results indicate the subclass of models can fit the U.S. term structure.

I. Introduction

Since the mid-1980s, there has been a plethora of "arbitrage-based" models of the term structure of interest rates. Such models take the observed current term structure as given, and seek to price interest rate derivative securities by arbitrage alone, based on assumptions concerning the future dynamics of the term structure. Arbitrage-based models are able to leave arbitrary the market prices of risk while taking any initial term structure as given. The drifts of the state variables, under the objective probabilities, depend on the market prices of risk, and are therefore left totally unspecified. Clearly, these features make econometric investigation of arbitrage-based models problematic, whether to estimate their parameters or even to examine their plausibility. Moreover, in seeking to mount an econometric investigation of the realism of a model, it would be natural to enquire whether the initial term structure might plausibly be explained by the model rather than

*Babbs, First National Bank of Chicago, 1 Triton Square, London NW1 3FN, U.K., and University of Warwick; Nowman, City University Business School, Frobisher Crescent, London EC2Y 8HB, U.K.

treating the term structure at an arbitrarily chosen start date as a pre-specified functional parameter.

One obvious way to address this situation is by econometric analysis of "equilibrium-based analogues" of arbitrage-based models. We use the term equilibrium-based to identify versions of the models in which market prices of risk are (directly or indirectly) given, and the instantaneous spot interest rate depends explicitly upon the state variables, thus determining the initial term structure endogenously within the model. We define the term analogues to mean that the equilibrium-based version can be described in terms of a set of state variables whose risk-adjusted probability law is identical with those of the state variables of the arbitrage-based version. In addition, the two versions possess an identical equation to link the initial term structure to the term structure at a subsequent date via the evolution of the state variables over the intervening period.

In this paper, we consider a subclass of the general linear Gaussian model of the term structure, which goes back to Langetieg (1980). This class of models assumes that the absolute volatility of rates is independent of the level of rates, contrary to Chan, Karolyi, Longstaff, and Sanders (1992) who found a significant dependence for the U.S. with absolute volatility proportional to r^γ with $\gamma = 1.5$ using monthly data during 1964–1989. Their sample period included the exceptional 1979–1982 period (c.f., Nowman (1997), who also estimated the general equation in Chan et al., for both the U.S. and U.K.). Applying Nowman (1997) to U.S. monthly Eurocurrency data during the sample period April 1987–December 1996 to be used in our empirical work reported below gives an estimate of $\gamma = 0.287$ (c.f., Nowman (1998)). This provides tentative encouragement for trying the independence assumption; indeed our empirical results below indicate the linear Gaussian model provides a good description of the U.S. curve. Awareness that the absolute volatility of interest rates has exhibited limited dependence upon the level of rates since the mid-1980s has increased practitioner interest in Gaussian models.

Models in this class have been investigated by, for example, Langetieg (1980), Hull and White (1990), Beaglehole and Tenney (1991), and Babbs (1990), (1993). The subclass we consider here is distinguished by the characteristic that the drift of each state variable depends on no other state variable.[1] As detailed consideration of Langetieg (1980) readily makes clear, this drift restriction makes the model particularly analytically tractable,[2] even if the coefficients are time varying and innovations in the state variables are correlated,[3] in that pure discount bond prices can be related to the state variables by a closed-form formula. Such a formula is a considerable computational benefit possessed by our subclass of models.

---

[1] Strictly speaking, the assumption of deterministic market prices of risk restricts us further to what Langetieg termed the "multivariate elastic random walk model."

[2] This case corresponds to that in which the matrix B in Langetieg (1980) is diagonal. His matrix Ψ, introduced on p. 78 and which plays a major part in the subsequent analysis, is then also diagonal, with the jth entry satisfying $\Psi_j(v - t) = B_j(v)\Psi_j(v - t)$, where B_j signifies the jth entry on the diagonal of B.

[3] This contrasts with multi-factor CIR models for which there is no closed-form formula for the bond price unless innovations in the state variables are uncorrelated. In practice, we feel that this confers additional flexibility on Gaussian models.

We confirm that the arbitrage-based version of models in this subclass is precisely the multifactor "Generalized Vasicek" family discussed by Babbs (1993).

Superficially, it appears that the Generalized Vasicek subclass involves a further substantive restriction, namely that each state variable is mean reverting to zero rather than to an arbitrary function of time, and that the weight of each state variable in determining r is minus unity rather than an arbitrary function of time. It turns out however, that, except in very restricted cases, all linear Gaussian models can be recast in our form, so long as the state variables are unobservable (rather than, for instance, identified exactly[4] with spot rates of various maturities as in Duffie and Kan (1996)). This has the advantage of reducing the number of coefficient functions in the model by twice the number of state variables.

It appears that our formulation excludes the "double decay" model of Beaglehole and Tenney (1991) (and an equivalent model of Hull and White (1994)) in which the short rate reverts towards a level that itself follows a mean-reverting random walk about a constant long-run average. However, not only can those models be re-expressed as special cases of our model, but also our formulation may be preferable in that the common intuitive interpretation of the mean-reversion speeds turns out to be illusory!

Recent empirical testing of term structure models has concentrated on the dynamic implications of the models using time-series data. Recent examples include Chan, Karolyi, Longstaff, and Sanders (1992), Broze, Scaillet, and Zakoian (1995), Brenner, Harjes, and Kroner (1996), Nowman (1997), (1998), and Andersen and Lund (1997). An alternative approach has concentrated on the cross-sectional implications of term structure models: examples include Brown and Dybvig (1986), Brown and Schaefer (1994), and De Munnik and Schotman (1994). Both the above approaches suffer from the disadvantage that they do not use the full information available from the yield curve obtained over time and across maturities in the estimation procedure. Recently, approaches providing a solution to this have been put forward by Gibbons and Ramaswamy (1993), Chen and Scott (1993), and Pearson and Sun (1994). A drawback of Pearson and Sun (1994) is that they assumed the two data points on the yield curve are measured without error while Chen and Scott (1993) assumed some bond prices are observed without error (though allowing for measurement errors in others).

The application of Kalman filtering methods in the estimation of term structure models using cross-sectional/time-series data, has been investigated by Pennacchi (1991), Lund (1994), (1997), Chen and Scott (1995), Duan and Simonato (1995), Geyer and Pichler (1996), Ball and Torous (1996), and Jegadeesh and Pennacchi (1996) (see also Harvey (1989) for an extensive treatment of Kalman filtering in econometrics). The use of the state space model formulation of term structure models and the application of the Kalman filter has the advantage that it allows the underlying state variables to be handled correctly as unobservable variables compared to using a short-term rate as a proxy (e.g., Chan et al. (1992), Nowman (1997), (1998)). In this paper, we consider the application of Kalman filtering methods to one-, two-, and three-factor versions of our subclass of models using U.S. data.

---

[4]We view such an identification as undesirable for our purposes, since it would presuppose that the spot rate concerned could be observed without measurement error (see, for example, Lund (1994)).

The paper is organized as follows. In Section II, we present our subclass of models; present the general formula for the bond price; establish the analogue relationship between the equilibrium-based model and its arbitrage-based version; and discuss the relationship to the models of Langetieg (1980) and Beaglehole and Tenney (1991). The bond price for the general constant parameter case is then given in Section III. Section IV discusses the state space formulation of the model and the estimation of the parameters. The data and empirical analysis are presented in Section V. Some conclusions are offered in Section VI.

II. Subclass of Models

A possible description of the instantaneous spot interest rate, r, is

$$(1) \qquad r(t) \;=\; \mu(t) - \sum_{j=1}^{J} X_j(t),$$

where μ is some deterministic process (i.e., mostly a function of time), and $X_1(t)$, $\ldots, X_J(t)$ represent the current effect of J streams of economic "news" whose impact dies away exponentially,

$$(2) \qquad dX_j \;=\; -\xi_j X_j dt + c_j dW_j,$$

where each ξ_j and c_j are deterministic, and $W_1, \ldots, W_J$ are standard Brownian motions with deterministic instantaneous correlation processes, $\rho_{jk} : j, k = 1, \ldots, J$. Equivalently,[5]

$$(3) \qquad dX_j \;=\; -\xi_j X_j dt + \sum_{q=1}^{Q} \kappa_{jq} dZ_q, \quad (Q \le J),$$

where $Z_1, \ldots, Z_Q$ are independent standard Brownian motions, and

$$(4) \qquad \sum_{q=1}^{Q} \kappa_{jq} \kappa_{kq} \;=\; \rho_{jk} c_j c_k.$$

If we assume that the market price of risk, θ_q, attaching to each Z_q, is deterministic, the resulting[6] formula for the price $B(M, t)$ at time t of unit nominal of a pure discount bond maturing at time M is

$$(5) \qquad B(M, t) \;=\; \exp\left\{ - \int_t^M \mu(u) du - \sum_{q=1}^{Q} \int_t^M \theta_q(u) \sigma_q(M, u) \right.$$

[5]The possibility $Q < J$ is realized if the matrix of the instantaneous correlations ρ_{jk} is at all times of less than full rank. Babbs (1993) illustrates the potential usefulness of allowing $Q < J$ in the arbitrage-based version of Generalized Vasicek models.

[6]This formula can be derived in the manner of Cox, Ingersoll, and Ross (1985) by making assumptions on technology and preferences, and restricting information to that generated by the state variables. Alternatively, it can be derived under arbitrary information as part of an incomplete markets equilibrium, using the ideas in Babbs and Selby (1983); Babbs and Nowman (1997) provide details in an earlier version of this paper.

$$-\frac{1}{2}\sigma_q^2(M,u)du + \sum_{j=1}^{J}\frac{G_j(M) - G_j(t)}{G_t'(t)}X_j(t)\Bigg\},$$

(6a) $$\text{where}\quad G_j(t) = \int_0^t \exp\left\{-\int_0^u \xi_j(s)ds\right\}du$$

(6b) $$\text{and where}\quad \sigma_q(M,t) = \sum_{j=1}^{J}\frac{G_j(M) - G_j(t)}{G_j'(t)}\kappa_{jq}(t)$$

represents the component of the (proportional) volatility of $B(M,t)$ attributable to Z_q. Dividing (5) through by the expressions that equation gives for $B(M,0)$ and $B(t,0)$, and rearranging, we obtain

(7) $$B(M,t) = \frac{B(M,0)}{B(t,0)}\exp\left\{-\frac{1}{2}\sum_{q=1}^{Q}\int_0^t \sigma_q^2(M,u) - \sigma_q^2(t,u)du\right.$$
$$\left. + \sum_{j=1}^{J}\{G_j(M) - G_j(t)\}\,Y_j(t)\right\},$$

(8) $$\text{where}\quad Y_j(t) = \sum_{q=1}^{Q}\int_0^t \frac{\kappa_{jq}(u)}{G_j'(u)}dZ_q^*,$$
$$\text{and each}\quad dZ_q^* = \theta_q(u)dt + dZ_q$$

defines a driftless standard Brownian motion under risk-neutral probabilities. Equations (7) and (8) hold also under the arbitrage-based version of the Generalized Vasicek family of models, as discussed in Babbs (1993), being precisely equivalent to his equations (9)[7] and (7), respectively. This establishes the desired analogue relationship between the equilibrium-based model in this paper and the arbitrage-based version.

Our class of models is obviously a subclass of the general linear Gaussian model described by Langetieg (1980) in which the dynamics of the state variables can be expressed as

(9) $$dX_j = \left(a_j + \sum_{k=1}^{J}B_{jk}X_k\right)dt + \sum_{q=1}^{Q}\kappa_{jq}dZ_q,$$

---

[7]Equations (5a) and (9) in the published text of Babbs (1993) contained some typographical errors. His equation (9) should read,

$$B(M,t) = \frac{B(M,0)}{B(t,0)}\exp\left\{\sum_{j=1}^{J}\{G_j(M) - G_j(t)\}\left(Y_j(t) + \sum_{i=1}^{J}\sum_{q=1}^{Q}\int_0^t G_i(u)\lambda_{iq}(u)\lambda_{jq}(u)du\right)\right.$$
$$\left. -\frac{1}{2}\sum_{i=1}^{J}\sum_{j=1}^{J}\{G_i(M)G_j(M) - G_i(t)G_j(t)\}\sum_{q=1}^{Q}\int_0^t \lambda_{iq}(u)\lambda_{jq}(u)du\right\}.$$

and the short rate is a general linear combination of the state variables,

$$
(10) \qquad\qquad r \;=\; w_0 + \sum_{j=1}^{J} w_j X_j.
$$

Our subclass restricts the off-diagonal coefficient functions, B_{jk}, to zero. Given this substantive restriction, the remaining restrictions (dropping the levels coefficient functions, a_j, and setting the weighting functions, w_j, identically to unity) are apparent, not real, so long as the weighting functions, w_j, do not vanish. To see this, we define the alternative set of state variables,

$$
(11) \qquad \tilde{X}_j(t) \;\equiv\; -w_j(t) \left[X_j(t) - \int_0^t a_j(u) \exp\left\{ \int_u^t B_{jj}(s)ds \right\} du \right],
$$

and define

$$
(12) \qquad \mu(t) \;=\; w_0(t) + \sum_{j=1}^{J} w_j(t) \int_0^t a_j(u) \exp\left\{ \int_u^t B_{jj}(s)ds \right\} du,
$$

enabling our formulation to be recovered.[8] Thus, having opted to exclude off-diagonal entries from the matrix B, the set of remaining parameter functions can be reduced from the $3J + 1$ functions (excluding the diffusion coefficients), $w_0, w_1, \ldots, w_J, a_1, \ldots, a_J, B_{11}, \ldots, B_{JJ}$, to the set of $J+1$ functions, $\mu, \xi_1, \ldots, \xi_J$.

At first glance, our restriction that off-diagonal B_{jk} be zero appears to exclude the double-decay model of Beaglehole and Tenney (1991),

$$
(13a) \qquad\qquad dr \;=\; \xi_1(y - r)dt + \kappa_{11}dZ_1,
$$
$$
(13b) \qquad\qquad dy \;=\; \xi_2(m - y)dt + \kappa_{21}dZ_1 + \kappa_{22}dZ_2,
$$

since identifying X_1 directly with r and X_2 with y leads to X_2 appearing in the drift coefficient of X_1. Fortunately, Beaglehole and Tenney's model can be written in our form,[9] i.e., in terms of (1) and (3). Moreover, given constant mean-reversion speeds, if m is a constant, then μ is also constant at precisely the same level. We note in passing that the intuitive interpretation of ξ_1 as the speed of mean reversion of the short rate, r, and ξ_2 as the speed of mean reversion of the stochastic level, y, towards which r is tending, is built on sand since, assuming that y is unobservable, (13a)–(13b) are observationally equivalent to a model in which ξ_1, ξ_2 take the same values, but have their roles interchanged![10]

---

[8]For a detailed derivation, see Babbs and Nowman (1997).

[9]Except in the case where $\xi_2 - \xi_1$ vanishes. We shall neglect the case $\xi_1 = 0$, which, as casual inspection of (13a) reveals, corresponds to the degenerate and non-stationary model in which r follows a driftless random walk. Babbs and Nowman (1997) provide a detailed proof.

[10]Except under the constant parameters non-stationary case where ξ_2 vanishes, $\xi_1 \neq 0$. The equivalent model is obtained by defining the new unobservable level,

$$
y^* \;=\; \left(1 - \frac{\xi_1}{\xi_2}\right) r + \frac{\xi_1}{\xi_2} y,
$$

and rewriting the model in the form,

$$
dr \;=\; \xi_2(y^* - r)dt + \kappa_{11}dZ_1,
$$

III.   The Constant Parameter Case

In the case where the mean-reversion level μ, the mean-reversion speeds ξ_j, the diffusion coefficients κ_{jq}, and the market price of risk processes θ_q are all constant, the key pricing formula (6) for a pure discount bond evaluates to

$$(14) \qquad B(M,t) \;=\; \exp\left\{-\tau\left[R(\infty) - w(\tau) - \sum_{j=1}^{J} H(\xi_j \tau) X_j(t)\right]\right\},$$

$$(15a) \qquad \text{with} \quad R(\infty) \;=\; \mu + \sum_{q=1}^{Q} \theta_q \sum_{j=1}^{J} \frac{\kappa_{jq}}{\xi_j} - \frac{1}{2}\sum_{q=1}^{Q}\left(\sum_{j=1}^{J}\frac{\kappa_{jq}}{\xi_j}\right)^2,$$

$$(15b) \qquad w(\tau) \;=\; \sum_{j=1}^{J} H(\xi_j \tau)\left[\sum_{q=1}^{Q}\theta_q\frac{\kappa_{jq}}{\xi_j} - \sum_{q=1}^{Q}\sum_{i=1}^{J}\frac{\kappa_{iq}\kappa_{jq}}{\xi_i\xi_j}\right]$$

$$+ \frac{1}{2}\sum_{i=1}^{J}\sum_{j=1}^{J} H((\xi_i + \xi_j)\tau)\sum_{q=1}^{Q}\frac{\kappa_{iq}\kappa_{jq}}{\xi_i\xi_j},$$

$$(15c) \qquad\qquad \text{where} \quad \tau \;\equiv\; M - t,$$

$$(15d) \qquad\qquad \text{and} \quad H(x) \;=\; \frac{1 - e^{-x}}{x}.$$

Note from the form of (14) that the spot rate from t to M depends upon calendar time only through the state variables, $X_1(t), X_2(t)$, being otherwise a function of residual term to maturity, $M - t$.

IV.   The State Space Model and the Kalman Filter

In this section, we are concerned with one-, two-, and three-factor versions of the constant parameter case of our subclass of models. We derive the state space model formulation of the term structure model and present the Kalman filter algorithm. This is used in the evaluation of the exact likelihood function of the observed interest rates and the computation of the unobserved state variables and parameters of the model. The theoretical yield curve is given by

$$(16) \qquad R(t+\tau, t) \;\equiv\; -\frac{\log B(t+\tau, t)}{\tau} \;=\; A_0(\tau) - A_1(\tau)'X(t),$$

where $A_0(\tau) = R(\infty) - w(\tau)$ and $A_1(\tau) = H(\xi_j \tau)$ is a $J \times 1$ vector (where the superscript denotes transpose). The scalar $A_0(\tau)$ and the vector $A_1(\tau)$ are functions

---

$$dy^* \;=\; \xi_1(m - y^*)dt + \kappa_{21}^* dZ_1 + \kappa_{22}^* dZ_2,$$

$$\text{where} \quad \kappa_{21}^* \;=\; \left(1 - \frac{\xi_1}{\xi_2}\right)\kappa_{11} + \frac{\xi_1}{\xi_2}\kappa_{21} \;; \quad \kappa_{22}^* \;=\; \frac{\xi_1}{\xi_2}\kappa_{22}.$$

of the time to maturity τ and the parameters of the model. We have N observed interest rates at time t_k for $k = 1, 2, \ldots, n$, which are denoted by $R_k = (R_{1k}, \ldots, R_{Nk})$, where $R_{ik} = -\log B(t_k + \tau_i, t_k)/\tau_i$.

We assume that measurement errors in interest rates are additive and normally distributed. The measurement equation is then given by

$$(17) \qquad R_k \;=\; d(\psi) + Z(\psi)X_k + \epsilon_k \,; \quad \epsilon_k \sim N(0, H(\psi)),$$

where ψ contains the unknown parameters of the model including the parameters from the distribution of the measurement errors. The ith row of the matrices $d(N \times 1)$ and $Z(N \times J)$ are given by $A_0(\tau_i ; \psi)$ and $-A_1(\tau_i ; \psi)'$, respectively. The error terms ϵ_k are measurement errors to allow for noise in the sampling process of the data. The variance-covariance matrix of the measurement errors can take various forms. Typically, it is assumed in empirical work that either $H = h^*I$ or we have maturity-specific variances $H = h_1, \ldots, h_N$ along the diagonal. The first assumption has the advantage of reducing the computational burden in the Kalman filter (see below). But as Geyer and Pichler (1996) point out, having maturity-specific variances takes into consideration that trading activity is going to vary across maturities and, therefore, the bid-ask spread will differ across maturities.

The transition equation is the exact discrete-time distribution of the state variables obtained from the solution of (2) (see Bergstrom (1984) and Lund (1994)) and is a var(1) model,

$$(18) \qquad\qquad X_k \;=\; \Phi(\psi)X_{k-1} + \eta_k,$$

where $\Phi(\psi) = e^{-\xi_J(t_k - t_{k-1})}$. The error term η_k is normally distributed with $E[\eta_k] = 0$ and $\mathrm{cov}[\eta_k] = V(\psi)$, where V is given in Bergstrom (1984), Theorem 3, and Lund (1994). The measurement and transition equations represent the state space formulation of our model. We now present the Kalman filter algorithm and the exact likelihood function.

Let $\widehat{X}_{k/k-1}$ and $\widehat{X}_k$ denote the optimal estimator (in a mean square error sense, MSE) of the unknown state vector X_k, based on the available information (i.e., the observed interest rates) up to time t_{k-1} and t_k, respectively. The optimal estimator is the conditional mean of X_k in both cases, denoted $E_{k-1}[\cdot]$ and $E_k[\cdot]$, respectively. The prediction step is given by

$$(19) \qquad\qquad \widehat{X}_{k/k-1} \;=\; E_{k-1}(X_k) \;=\; \Phi\widehat{X}_{k-1},$$

with mean square error (MSE) matrix

$$(20) \qquad \Sigma_{k/k-1} \;=\; E_{k-1}\left[\left(X_k - \widehat{X}_{k/k-1}\right)\left(X_k - \widehat{X}_{k/k-1}\right)' \right]$$
$$\qquad\qquad\quad =\; \Phi\Sigma_{k-1}\Phi' + V.$$

In the update step, the additional information given by R_k is used to obtain a more precise estimator of X_k,

$$(21) \qquad\qquad \widehat{X}_k \;=\; E_k(X_k) \;=\; \widehat{X}_{k/k-1} + \Sigma_{k/k-1}Z'F_k^{-1}v_k,$$

$$(22) \qquad \Sigma_k = E_k \left[\left(X_k - \widehat{X}_k \right) \left(X_k - \widehat{X}_k \right)' \right]$$

$$= \Sigma_{k/k-1} - \Sigma_{k/k-1} Z' F_k^{-1} Z \Sigma_{k/k-1}$$

$$= \left(\Sigma_{k/k-1}^{-1} + Z' H^{-1} Z \right)^{-1},$$

$$(23) \qquad \text{where} \quad v_k = R_k - \left(d + Z \widehat{X}_{k/k-1} \right),$$

$$(24) \qquad \text{and} \quad F_k = Z \Sigma_{k/k-1} Z' + H,$$

(c.f., Harvey (1989), Ch. 3). This new estimate of X_k is called the *filtered* estimate. The aim of the Kalman filter is to obtain information about X_k from the observed interest rates. The Kalman filter also has the advantage of being able to evaluate the likelihood function using the prediction error decomposition. The log-likelihood function is given by (apart from a constant)

$$(25) \qquad \log L(R_1, \ldots, R_n; \psi) = -\frac{1}{2} \sum_{k=1}^{n} \log |F_k| - \frac{1}{2} \sum_{k=1}^{n} v_k' F_k^{-1} v_k,$$

where v_k and F_k are given by equations (23) and (24). We can also use the formulas of Harvey (1989), p. 108, for computing the inverse and determinant of F_k given by

$$F_k^{-1} = H^{-1} - H^{-1} Z \left(\Sigma_{k/k-1}^{-1} + Z' H^{-1} Z \right)^{-1} Z' H^{-1},$$

$$|F_k| = |H| \cdot |\Sigma_{k/k-1}| \cdot \left| \Sigma_{k/k-1}^{-1} + Z' H^{-1} Z \right|.$$

In the case of $H = h^* I$, these formulas can be simplified and the computational burden reduced, which is especially important if the number of maturities used in the estimation is large. Finally, the Kalman filter recursions are started by setting the initial state vector X_0 and covariance matrix Σ_0 to their unconditional mean and covariance.

V.   Empirical Results

A.   Data Description

The data used in our empirical work consist of constructed zero coupon yields obtained from interbank interest rates. In particular, the raw data include money market rates with maturities including the overnight rate, one-, three-, and six-month rates, Euro dollar futures, and swap rates with two–five, seven, and 10 years to maturity obtained from Datastream. The interest rates are sampled daily from April 1987 to December 1996. However, we use weekly data on a Wednesday in the empirical analysis following Lund (1997) to avoid missing observations and week-day effects. We have a total of 507 weekly observation dates and at each date we have N-interest rates. The following maturities: three and six months, one, two, three, five, seven, and 10 years were chosen $(N = 8)$. Table 1 reports the summary statistics.

TABLE 1

Summary Statistics: U.S. Data

April 1987–Dec. 1996

$r(t)$	Mean	Stan. Dev.
3-Month	0.0617	0.0196
6-Month	0.0623	0.0192
1-Year	0.0645	0.0185
2-Year	0.0700	0.0171
3-Year	0.0725	0.0159
5-Year	0.0762	0.0142
7-Year	0.0787	0.0135
10-Year	0.0809	0.0125

B.   Estimation Results

The application of the Kalman filter to one-, two-, and three-factor models to U.S. data are now discussed. The estimation results are presented in Table 2, which contains the parameter estimates of μ, ξ_j, κ_{jq}, θ_q, and the estimated standard deviations of the measurement errors $(\sqrt{h_1}, \ldots, \sqrt{h_N})$. The table also contains the BIC Information Criterion and log-likelihood value.

In the one-, two-, and three-factor models, the mean reversions ξ_j and diffusion parameters c_j are significant and have plausible values. In the two- and three-factor models, the market prices of risk θ_q are not significant, whereas they are in the one-factor model. The long-run average rate μ is significant in the one- and two-factor models and has plausible values. The log-likelihood value increases strongly as the number of factors is increased. The standard deviations for the measurement errors naturally decrease as the number of factors is increased.

The estimated standard deviations for the measurement errors are much larger for the one-factor model. In particular, in the one-factor model, these standard deviations are 36 basis points for the three-month rate, 22 basis points for the six-month rate, four basis points for the one-year rate, 37 basis points for the two-year rate, 42 basis points for the three-year rate, 52 basis points for the five-year rate, 62 basis points for the seven-year rate, and 73 basis points for the 10-year rate. The standard deviations are much smaller in the multi-factor models. As we move from the two-factor model to the three-factor model, there is a smaller decrease in the standard deviations than when we moved from the one-factor model. In particular, in the two-factor model, there are 17 basis points for the three-month rate, four basis points for the six-month rate, 17 basis points for the one-year rate, 28 basis points for the two-year rate, 19 basis points for the three-year rate, nine basis points for the five-year rate, less than one basis point for the seven-year rate, and eight basis points for the 10-year rate. This compares to a range of less than one basis point to 23 basis points for the three-factor model in these maturities.

Our measurement errors compare very favorably to recent studies by, for example, Chen and Scott (1995) and Geyer and Pichler (1996), who both estimate the multi-factor CIR model on U.S. data. In particular, Chen and Scott (1995), Table 2, report, for weekly data over the period 1980–1988, measurement errors for the one-factor model of 40 basis points for the three-month rate, zero basis

TABLE 2

Estimates of Generalized Vasicek Models: 1987–1996

	One-Factor	Two-Factor	Three-Factor
ξ_1	0.1908	0.5529	0.6553
	(0.0017)	(0.0108)	(0.0452)
ξ_2		0.0652	0.0705
		(0.0120)	(0.0010)
ξ_3			0.0525
			(0.0094)
c_1	0.0132	0.0195	0.0214
	(0.0004)	(0.0023)	(0.0018)
c_2		0.0186	0.0189
		(0.0018)	(0.0007)
c_3			0.0163
			(0.0016)
ρ_{12}		−0.8360	−0.9394
		(0.0440)	(0.0094)
ρ_{13}			0.8753
			(0.1725)
ρ_{23}			−0.9200
			(0.0190)
μ	0.0594	0.0728	0.0701
	(0.0024)	(0.0168)	(1.8607)
θ_1	0.6483	−0.0849	0.1582
	0.0192)	(1.0361)	(14.1842)
θ_2		0.0963	0.0961
		(1.2321)	(70.5551)
θ_3			0.0173
			(0.7276)
$\sqrt{h_1}$	0.0036	0.0017	0.0010
	(0.0001)	(<0.0001)	(<0.0001)
$\sqrt{h_2}$	0.0022	0.0004	0.0003
	(<0.0001)	(0.0001)	(<0.0001)
$\sqrt{h_3}$	0.0004	0.0017	0.0015
	(0.0001)	(<0.0001)	(<0.0001)
$\sqrt{h_4}$	0.0037	0.0028	0.0023
	(<0.0001)	(<0.0001)	(0.0001)
$\sqrt{h_5}$	0.0042	0.0019	0.0016
	(<0.0001)	(<0.0001)	(<0.0001)
$\sqrt{h_6}$	0.0052	0.0009	0.0008
	(0.0001)	(<0.0001)	(<0.0001)
$\sqrt{h_7}$	0.0062	0.0001	0.0001
	(0.0001)	(0.0001)	(<0.0001)
$\sqrt{h_8}$	0.0073	0.0008	0.0006
	(0.0001)	(<0.0001)	(<0.0001)
$\log L$	20494	24397	30309
BIC	−40914	−48695	−60518

points for the six-month rate, 104 basis points for the five-year rate, and 122 basis points for a long-term bond rate. In the two-factor model they report 35 basis points for the three-month rate, zero basis points for the six-month rate, 33 basis points for the five-year rate, and seven basis points for the long-term bond rate.

The mean reversion parameters imply mean half-lives for the interest rate process (i.e., the expected time for the process to return halfway to its long-term mean, defined as $-\ln(0.5)/\xi_j$). In the two-factor model, the mean half-lives are 1.3 years for the first factor and 10.6 years for the second factor. In the three-factor model, the mean half-lives for the three factors are 1.1 years, 9.8 years, and 13.2 years. The correlation coefficient in the two-factor model is -84% and significant. Calibration of an arbitrage-free analogue of this model to market prices of interest rate caps and swaptions yielded similar large values for the correlation parameter. In the three-factor model, the correlations are -94% between factors one and two, 87% between factors one and three, and -92% between factors two and three. The log-likelihood values for the one-factor model are 20494, for the two-factor model 24397, and for the three-factor model 30309. Based on the BIC Information Criterion (BIC), we find that moving from a one-factor to the two-factor model, the BIC improves by 19%, and moving from the two-factor to the three-factor model, the BIC improves by 24%. Based on the BIC rule, the three-factor model outperforms the two-factor model. The likelihood ratio test of the one-factor vs. the two-factor model gives a value of 7806, and one can reject the null hypothesis of a one-factor model at the 5% significance level. The likelihood ratio test of the two-factor vs. the three-factor model gives a value of 11824, and one can reject the null hypothesis of a two-factor model at the 5% significance level.

We also look at the factor loadings for the two- and three-factor models as a function of maturity that should help determine the nature of the factors calculated by the Kalman filter. Litterman and Scheinkman (1991), using principal components analysis, investigated a number of U.S. yields and identified three factors that they interpreted as changes in level, steepness, and curvature. Factor loadings correspond to orthogonal Brownian motions whereas innovations in our state variables are correlated. In the two-factor model, we, therefore, choose to produce factor loadings by re-expressing W_1 and W_2 in terms of uncorrelated Brownian motions Y_1 and Y_2 in such a way that dY_2 does not impact the term structure at approximately a particular maturity τ^*. Litterman and Scheinkman (1991), Table 2, find that factor two has approximately zero impact on the term structure at the five-year maturity and this is imposed here for comparison with their graphs ($\tau^* = 5$). The factor loadings are given below for factor one, $\gamma_1(\tau)$, and factor two, $\gamma_2(\tau)$,

$$(26) \qquad \gamma_1(\tau) = \frac{H(\xi_1\tau)c_1\beta\alpha_{22}}{\alpha_{22}H_1c_1 - \alpha_{21}H_2c_2} + \frac{H(\xi_2\tau)c_2\beta\alpha_{21}}{\alpha_{21}H_2c_2 - \alpha_{22}H_1c_1},$$

$$(27) \qquad \gamma_2(\tau) = \frac{-H(\xi_1\tau)c_1H_2c_2}{\alpha_{22}H_1c_1 - \alpha_{21}H_2c_2} - \frac{H(\xi_2\tau)c_2H_1c_1}{\alpha_{21}H_2c_2 - \alpha_{22}H_1c_1},$$

where $\quad \beta = \sqrt{H_1^2c_1^2 + 2\rho H_1 H_2 c_1 c_2 + H_2^2 c_2^2}\,;$

$$H_j = H(\xi_j \tau^*) ;$$

$$\alpha_{21} = -\frac{H_1 c_1 \rho + H_2 c_2}{\beta \sqrt{1 - \rho^2}} ; \quad \alpha_{22} = -\alpha_{21} \frac{H_1 c_1 + H_2 c_2 \rho}{H_1 c_1 \rho + H_2 c_2} ; \quad \tau^* = 5.$$

Figure 1 plots the factor loadings of the two-factor model as a function of maturity. The first factor's impact on yield changes has an increasing positive effect on the maturities up to four years, then has an equal impact on the remaining maturities. We conclude, as Litterman and Scheinkman identified, that the first factor could represent a level factor. The second factor has a strong influence on short-term rates up to five years, lowers them, and then has a positive impact on longer maturities by raising them. We conclude that the second factor could represent a steepness factor as identified by Litterman and Scheinkman. Overall, the model generates factor loadings in line with their results. Using a similar approach for the three-factor model, we can obtain factor loadings such that the second factor loadings disappear at around five years, the third factor loading at around two years and at around 12 years, in keeping with the results obtained by Litterman and Scheinkman. Figure 2 plots the factor loadings for the three-factor model. Compared to the two-factor model, the first loading is monotonically declining rather than humped; the second loading is similar to that obtained in the two-factor model, and finally the third loading has a negligible effect.
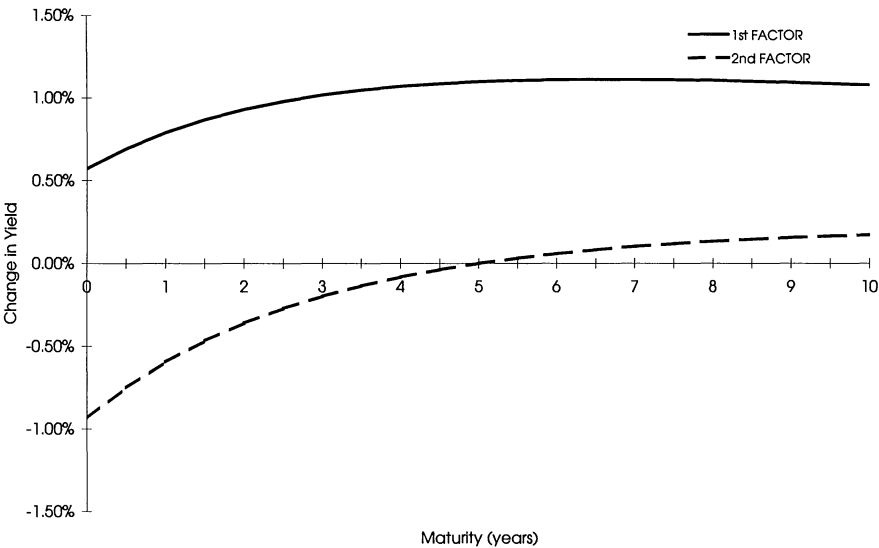
FIGURE 1

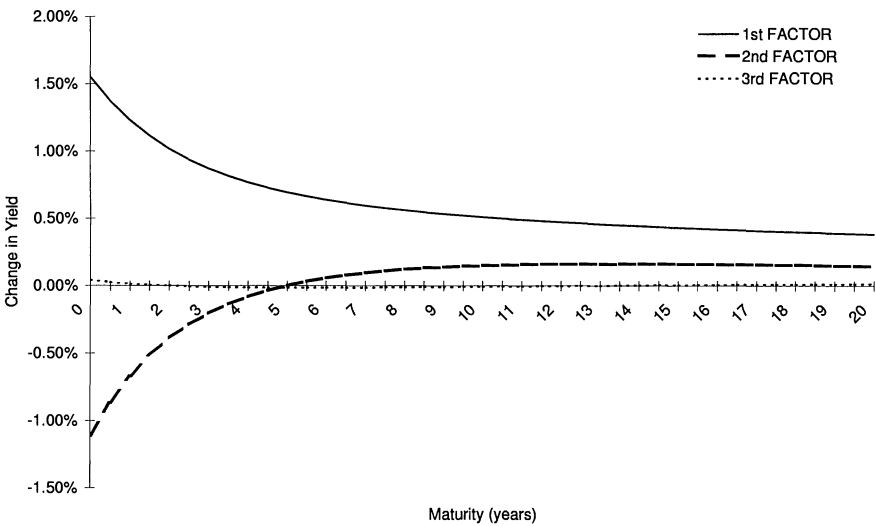Factor Loadings of the Two-Factor Model

FIGURE 2

Factor Loadings of the Three-Factor Model

VI.   Conclusions

In this paper, we have been concerned with a subclass of the general linear Gaussian model of Langetieg (1980). We have confirmed that the arbitrage-based version of models in this subclass is precisely the multifactor "Generalized Vasicek" family discussed by Babbs (1993). The subclass has the advantage of reducing the number of coefficient functions in the model by twice the number of state variables so long as the state variables are unobservable. We have shown also that the double-decay model of Beaglehole and Tenney (1991) can be re-expressed as a special case of our model, and that our formulation may be preferable since the obvious intuitive interpretation of the mean-reversion speeds is shown to be, in fact, illusory! For our empirical work, the model is expressed in a state space formulation that allows us to take into account both the cross-sectional and time-series restrictions on the data and that the observed yield curve contains measurement errors. Estimates are obtained for one-, two-, and three-factor models using U.S. data during the period 1987–1996. We find overall that, in formal statistical terms, the two-factor model is rejected with the three-factor model as the alternative hypothesis, but the measurement errors of the two-factor indicates that it frequently performs as well as the three-factor model.

References

Andersen, T. G., and J. Lund. "Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate." *Journal of Econometrics*, 77 (March 1997), 343–377.

Babbs, S. H. "Generalized Vasicek Models of the Term Structure." In *Applied Stochastic Models and Data Analysis: Proceedings of the Sixth International Symposium*, Vol. 1, J. Janssen and C. H. Skiadas, eds. Singapore: World Scientific (1993).

_____. "The Term Structure of Interest Rates: Stochastic Processes and Contingent Claims." Unpubl. Ph.D. Diss., Imperial College, Univ. of London (1990).

Babbs, S. H., and K. B. Nowman. "Kalman Filtering of Generalized Vasicek Term Structure Models." Financial Option Research Centre, Univ. of Warwick, Preprint 97/80 (1997).

Babbs, S. H., and M. J. P. Selby. "Pricing by Arbitrage in Incomplete Markets." Unpubl. Paper, London (1993).

Ball, C. A., and W. N. Torous. "Unit Roots and the Estimation of Interest Rate Dynamics." *Journal of Empirical Finance*, 3 (June 1996), 215–238.

Beaglehole, D. R., and M. S. Tenney. "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations." *Journal of Fixed Income*, 1 (Sept. 1991), 69–83.

Bergstrom, A. R. "Continuous Time Stochastic Models and Issues of Aggregation over Time." In *Handbook of Econometrics*, Vol. ll, Z. Griliches and M. D. Intriligator, eds. Amsterdam: Elsevier Science (1984).

Brenner, R. J.; R. H. Harjes; and K. F. Kroner. "Another Look at Models of the Short-Term Interest Rate." *Journal of Financial and Quantitative Analysis,* 31 (March 1996), 85–107.

Brown, S. J., and P. H. Dybvig. "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates." *Journal of Finance*, 41 (July 1986), 617–630.

Brown, R. H., and S. M. Schaefer. "The Term Structure of Real Interest Rates and the Cox, Ingersoll, and Ross Model." *Journal of Financial Economics*, 35 (Feb. 1994), 3–42.

Broze, L.; O. Scaillet; and J.-M. Zakoian. "Testing for Continuous Time Models of the Short-Term Interest Rate." *Journal of Empirical Finance*, 2 (Sept. 1995), 199–223.

Chan, K. C.; G. A. Karolyi; F. A. Longstaff; and A. B. Sanders. "Comparison of Models of the Short-Term Interest Rate." *Journal of Finance,* 47 (July 1992), 1209–1227.

Chen, R., and L. Scott. "Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates." *Journal of Fixed Income*, (Dec. 1993), 14–31.

_____. "Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests From a Kalman Filter Model." Unpubl. Paper, Rutgers Univ. (1995).

Cox, J. C.; J. E. Ingersoll; and S. A. Ross. "A Theory of the Term Structure of Interest Rates." *Econometrica*, 53 (1985), 385–407.

De Munnik, J., and P. Schotman. "Cross Sectional Versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market." *Journal of Banking and Finance*, 18 (Oct. 1994), 997–1025.

Duan, J. C., and J. G. Simonato. "Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter." Unpubl. Paper, Hong Kong Univ. of Science and Technology (1995).

Duffie, D., and R. Kan. "A Yield-Factor Model of Interest Rates." *Mathematical Finance*, 6 (Oct. 1996), 379–406.

Geyer, A. L. J., and S. Pichler. "A State-Space Approach to Estimate and Test Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure." Unpubl. Paper, Univ. of Economics, Vienna (1996).

Gibbons, M. R., and K. Ramaswamy. "A Test of the Cox, Ingersoll, and Ross Model of the Term Structure." *Review of Financial Studies*, 6 (1993), 619–658.

Harvey, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge, England: Cambridge Univ. Press (1989).

Hull, J., and A. White. "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models." *Journal of Derivatives*, 2 (1994), 37–48.

_____. "Pricing Interest-Rate Derivative Securities." *Review of Financial Studies*, 3 (1990), 573–592.

Jegadeesh, N., and G. G. Pennacchi. "The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures." *Journal of Money, Credit, and Banking,* 28 (Aug. 1996), 426–446.

Langetieg, T. C. "A Multivariate Model of the Term Structure." *Journal of Finance*, 35 (March 1980), 71–91.

Litterman, R., and J. Scheinkman. "Common Factors Affecting Bond Returns." *Journal of Fixed Income*, 1 (June 1991), 54–61.

Lund, J. "A Model for Studying the Effect of EMU on European Yield Curves." Unpubl. Paper, The Aarhus School of Business (1997).

_____. "Econometric Analysis of Continuous-Time Arbitrage-Free Models of the Term Structure of Interest Rates." The Aarhus School of Business (1994).

Nowman, K. B. "Continuous Time Short Rate Interest Rate Models." *Applied Financial Economics*, 8 (1998), 401–407.

_____. "Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates." *Journal of Finance*, 52 (Sept. 1997), 1695–1706.

Pearson, N. D., and T. S. Sun. "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll and Ross Model." *Journal of Finance*, 49 (Sept. 1994), 1279–1304.

Pennacchi, G. G. "Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data." *Review of Financial Studies*, 4 (1991), 53–86.