



Design and Estimation of Quadratic Term Structure Models ^{*}

MARKUS LEIPPOLD^{**}

Swiss Banking Institute, University of Zurich

LIUREN WU[‡]

Graduate School of Business, Fordham University

Abstract. We consider the design and estimation of quadratic term structure models. We start with a list of stylized facts on interest rates and interest rate derivatives, classified into three layers: (1) general statistical properties, (2) forecasting relations, and (3) conditional dynamics. We then investigate the implications of each layer of property on model design and strive to establish a mapping between evidence and model structures. We calibrate a two-factor model that approximates these three layers of properties well, and show that a flexible specification for the market price of risk is important in capturing the stylized evidence in forecasting relations while factor interactions are indispensable in generating the hump-shaped dynamics of bond yields.

Key words: quadratic model; term structure; positive interest rates; humps; expectation hypothesis; GMM.

JEL classification codes: G12, G13, E43.

1. Introduction

Term structure modeling has enjoyed rapid growth during the last decade. Among the vast number of different models, the affine class stands out as the most popular class due to its analytical tractability. Duffie and Kan (1996)'s characterization of the complete class has spurred a stream of studies on its empirical applications and model design. Examples include general econometric estimations by Chen and Scott (1993), Duffie and Singleton (1997), Dai and Singleton (2000), and Singleton

^{*} We thank Robert Bliss for providing the computer programs and data used to estimate forward rates and yields. For helpful comments and thoughtful suggestions, we thank Marco Avellaneda, David Backus, Geert Bekaert, Simon Benninga (the executive editor), Peter Carr, Massoud Heidari, the participants of the Western Finance Association Annual Meeting in Tucson, Arizona, 2001, and an anonymous referee for his insightful comments. All errors are solely our responsibility.

^{**} Correspondence Information: Plattenstr. 14, 8032 Zurich, Switzerland; Tel: (+41) 1-634-2951; Fax: (+41) 1-634-4903; E-mail: leippold@isb.unizh.ch.

[‡] Correspondence Information: 113 West 60th Street, New York, NY 10023; Tel: (212) 636-6117; Fax: (212)765-5573; E-mail: wu@fordham.edu.

(2001), applications to the predictability of interest rates by Frachot and Lesne (1994), Roberds and Whiteman (1999), Backus et al. (2001), Duffee (2002), and Dai and Singleton (2002a), and currency pricing by Backus et al. (2001). While these applications claim success in one or two dimensions, inherent tension exists when one tries to apply the model to a wider range of properties. Even more troublesome, however, is a seemingly irreconcilable tension between delivering a relatively good empirical performance and excluding negative interest rates. Indeed, all of the relatively “successful” model designs within the affine framework, in terms of empirical performance, imply positive probabilities of negative interest rates. Examples include Backus et al. (2001a), Backus et al. (2001b), Dai and Singleton (2000), Dai and Singleton (2002a), Duffee (2002), Duffie and Singleton (1997), and Singleton (2001).

Leippold and Wu (2002) identify and characterize an alternative class, the quadratic class of term structure models, where bond yields and forward rates are quadratic functions of the state vector. They illustrate that the quadratic class is comparable to the affine class for analytical tractability but is much more flexible for model design. In particular, positive interest rates can be guaranteed with little loss of generality or flexibility. Moreover, under the quadratic class, interest rates can be represented as quadratic forms of normal variates, for which moments and cross moments are known in closed form. The properties of the quadratic model can hence be readily analyzed. In this paper, we consider the model design and estimation problem within the quadratic framework.

Although examples of quadratic models have appeared in the literature since the late eighties,¹ empirical applications have at best been sporadic. The most systematic empirical study, and hence the most germane to our work, is by Ahn et al. (2002). They apply the efficient methods of moments (EMM) of Gallant and Tauchen (1996), calibrate the maximally flexible three-factor quadratic model and various restricted versions to the US Treasury data, and compare their performance with that of affine models. Our approach goes in the opposite direction and should be viewed as a complement to their work. We start with a list of what we view as the salient features of interest rates and attempt to find a parsimonious quadratic specification which accounts for them. What we gain are some helpful insights into the mapping between parameters and data. The approach also highlights the impact of different pieces of evidence on model structure.

We take a series of steps that we think serve to integrate evidence with theory. Using U.S. Treasury bond yields, we start by classifying the properties of interest rates into three categories: (1) general statistical properties, (2) forecasting relations, and (3) conditional dynamics. The most significant statistical properties, in our view, include an upward sloping mean yield curve, high (but different) persistence in yields of different maturities, and positive skewness in interest rate distributions. For forecasting relations, we examine the violations of the various

¹ Early examples include Longstaff (1989), Beaglehole and Tenney (1991, 1992), Elkaroui et al. (1992), Constantinides (1992), Jamshidian (1996), and Rogers (1997).

forms of the expectation hypotheses (EH). We find that underlying all interest rates is a common feature revealed by a single-period forward rate regression. That is, violations of the EH are mainly at the short end. On conditional dynamics, we find that the mean hump-shaped term structure for conditional variance not only shows up in the implied volatility quotes of interest rate derivatives, but also reveals itself vividly in the variance of multi-period changes in interest rates.

With facts in hand, we turn to the model. We analyze the properties of bond yields and forward rates under the quadratic class and examine the implications of the documented evidence on model design. The quadratic relation between interest rates and the state variables not only provides a convenient approach to guarantee positive interest rates, but also directly incorporates nonlinearity in interest rate dynamics. Furthermore, we find that the affine specification on market price of risk plays an important role in accounting for the violations in the expectation hypotheses while flexible interactions between state variables are indispensable in generating the observed hump-shaped dynamics. The quadratic class synthesizes positive interest rates, affine market price of risk, and factor interactions into one flexible framework.

In the third step, we transform the property analysis into moment conditions and calibrate a two-factor quadratic model by generalized methods of moments (GMM). While three or even more factors might be necessary to fully capture the interest rate dynamics [e.g., Litterman and Scheinkman (1991), Knez et al. (1994), and Heidari and Wu (2001)], a two-factor model is the minimum required to capture the stylized evidence listed above. The calibration exercise confirms with the property analysis on the relative contribution of each component of the model to different features of interest rates. We find that fitting the prominent hump shape observed in the annualized variance of multi-period interest rate differences asks for strong interaction between state variables; simultaneously fitting the slopes of the forward rate regressions and the mean yield curve, on the other hand, requires a significant input for the affine part of the market price of risk.

Compared to the efficient methods of moments applied in, for example, Ahn et al. (2002) and Dai and Singleton (2000), our moment choice is mainly motivated by economic and structural significance. Such an approach might potentially lose some efficiency in parameter estimates from a purely statistical perspective, but it makes apparent the inherent link between each piece of evidence and the necessary model structure to account for it and hence provides guidance for future model design.

The structure of the paper is as follows. The next section documents the evidence using 15 years of data on the U.S. government bond yields and forward rates. Section 2 lays out the framework for the quadratic class of term structure models. Section 3 analyzes the implication of quadratic models on the three layers of properties of bond yields and forward rates. Section 4 applies the generalized method of moments to estimate a two-factor quadratic model and tests various parameter restrictions. Section 5 concludes.

2. Evidence

We document the salient features of U.S. Treasury bond yields in three dimensions: (1) general statistical properties of bond yields, (2) forecasting relations, and (3) the conditional dynamics. We analyze 15 years of monthly data, from January 1985 to December 1999 (180 observations), with maturities from 1 month to 10 years. The yields are computed from U.S. treasury prices by the “smoothed Fama-Bliss” method using programs and data supplied by Robert Bliss. The estimation method is described in detail in Bliss (1997).

In discrete time notation, we denote the continuously compounded yield on an n -period bond at date t as y_t^n . It is defined as

$$y_t^n = -n^{-1} \ln P_t^n,$$

where P_t^n denotes the dollar price at date t of a claim to one dollar at $t + n$. Here the discrete period is in number of months, corresponding to the monthly data. Forward rates are defined by

$$f_t^n = \ln (P_t^n / P_t^{n+1}),$$

so that yields are averages of forward rates:

$$y_t^n = n^{-1} \sum_{i=0}^{n-1} f_t^i.$$

We use the one-month yield as a proxy for the instantaneous interest rate, or the short rate: $r_t = y_t^1 = f_t^0$.

Summary Statistics. Table I provides the summary statistics of the yields. The most significant features include:

1. The mean term structure is upward sloping. Figure 1 depicts mean term structure (the solid line) and the 5% and 95% percentiles (dashed lines).
2. Bond yields are highly persistent. The first-order autocorrelations for the bond yields range from 0.956 to 0.978.
3. The past 15 years have been uneventful for U.S. Treasuries. While the interest rates are slightly positively skewed, the tails are no thicker than that of a normal distribution.

Forecasting Relations. A long established fact about Treasury yields is that the current term structure contains information about future interest rate movements. While the expectation hypothesis (EH) has long been regarded as a poor approximation of the evidence, it presents useful forms for interest rate forecasting. An enormous body of research to this effect has been surveyed repeatedly, most recently by Bekaert et al. (1997), Campbell (1995), Campbell and Shiller (1991), Evans and Lewis (1994).

Table I. General properties of bond yields

Maturity	Mean	St Dev	Skewness	Kurtosis	Auto
1	5.348	1.492	0.288	2.450	0.967
3	5.570	1.525	0.235	2.432	0.978
6	5.787	1.536	0.178	2.410	0.977
9	5.949	1.539	0.165	2.379	0.975
12	6.081	1.537	0.178	2.358	0.972
24	6.448	1.513	0.298	2.388	0.963
36	6.696	1.491	0.403	2.527	0.959
48	6.884	1.474	0.465	2.687	0.957
60	7.031	1.460	0.500	2.833	0.956
84	7.246	1.434	0.540	3.064	0.956
120	7.450	1.400	0.581	3.304	0.957

The data are end-of-month estimates of continuously-compounded zero-coupon U.S. government bond yields expressed as annual percentages. They were supplied by Robert Bliss (“smoothed Fama-Bliss” method) and cover the period January 1985 to December 1999 (180 observations). Mean is the sample mean, St Dev the sample standard deviation, and Skewness is defined as the third central moment divided by the cube of the standard deviation, Kurtosis is the fourth central moment divided by the fourth power of the standard deviation, and Auto is the first-order autocorrelation. Our estimates replace population moments with sample moments.

While various forecasting relations have been formulated in the literature based on different forms of the EH, we focus on the simplest form proposed by Backus et al. (2001). It is a single period regression based on the martingale hypothesis on the forward rate:

$$f_{t+1}^{n-1} - r_t = a_n + g_n (f_t^n - r_t) + e_{t+1}^n. \quad (1)$$

Figure 2 plots the regression slope estimates over maturities. Expectation hypothesis implies that the regression should have a slope of one. Yet the deviation is obvious. Backus et al. (2001) show that measurement errors are too small to account for the deviation while small sample bias works in the opposite direction.

Conditional Dynamics. By conditional dynamics, we refer to the dynamics of the conditional volatilities of bond yields. It has been widely noted that the conditional volatility of interest rates has a “hump-shaped” mean term structure. The conditional volatility first increases with horizon, reaches a plateau, and then decreases.

The hump-shaped dynamics shows up in the data in a variety of ways. A relatively simple one comes from multiperiod differences: changes $y_t - y_{t-k}$ over

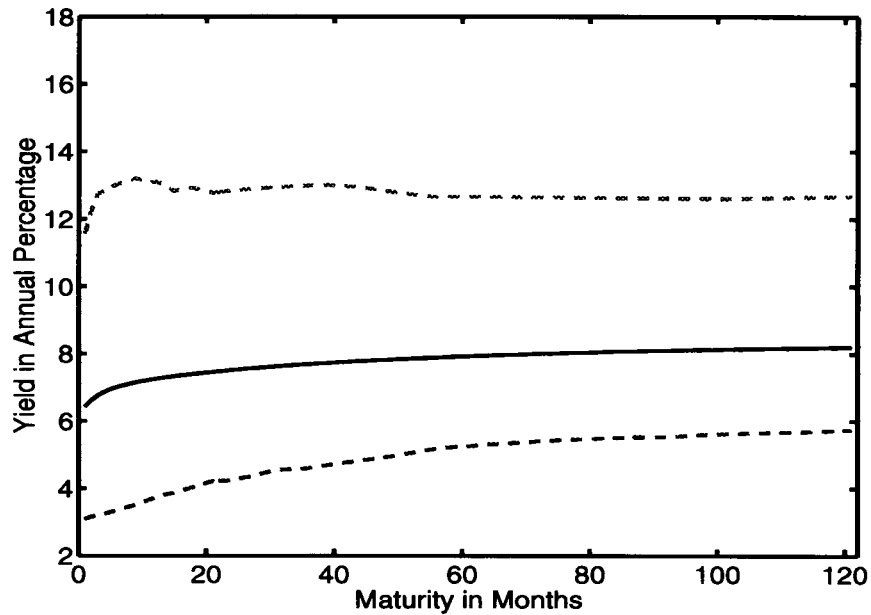


Figure 1. Term structure of U.S. Treasury yields. The gray lines in the background are term structures of the U.S. Treasury yields at each month from January 1985 to December 1999 (180 observations). The bold solid line represents the mean yield curve.

periods of length k . If y has hump-shaped dynamics, the variance of the multiperiod differences increases initially at a rate faster than k .² Equivalently, $\text{Var}(y_t - y_{t-k})/k$ is hump shaped. Figure 3 depicts $\text{Var}(y_t - y_{t-k})/k$ over k . The hump shape is very prominent, especially for bond yields of relatively short maturities.

However, the hump dynamics has caught practitioners' attention mainly from its appearance in the derivatives market. In particular, the Black implied volatilities of interest rate caps, floors, and swaptions on many currencies exhibit a hump-shaped mean term structure. As an example, Figure 4 depicts the mean term structure of the Black implied volatility quotes for interest rate caps on U.S. Dollar (left panel) and Deutsche Mark LIBOR (right panel), respectively. The hump shape is obvious.

² The unconditional variance of the difference captures the conditional variance of the level if y_t can be approximated as

$$y_t = \theta + \phi y_{t-1} + \sigma \varepsilon_t,$$

with $\phi = 1$. It can be used as an approximation for very persistent series such as interest rates when ϕ is less than but very close to one.

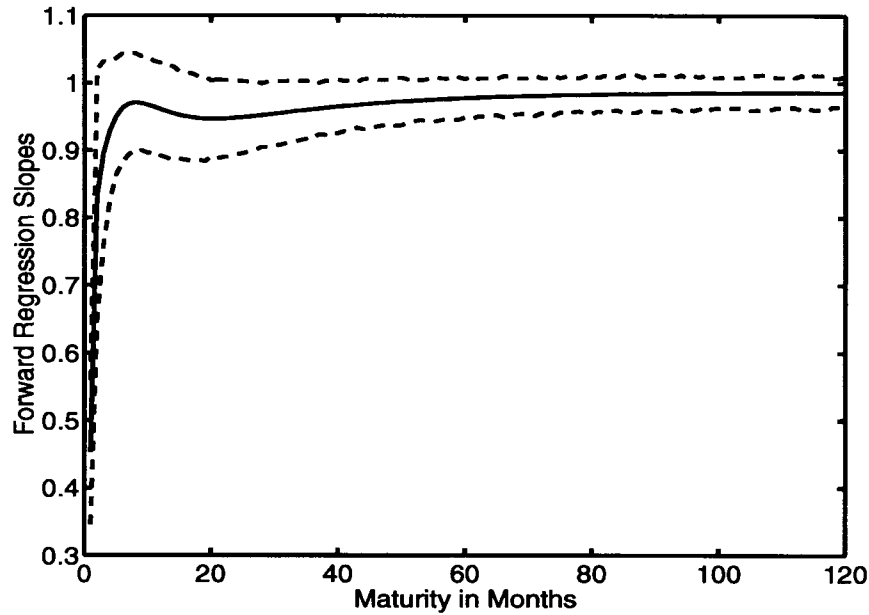


Figure 2. Term structure of forward regression slopes on U.S. Treasuries. The solid line is the slope estimates, g_n , of the following single period forward regression:

$$f_{t+1}^{n-1} - r_t = a_n + g_n(f_t^n - r_t) + e_{t+1}^n,$$

where f^n denotes the continuously compounded forward rate with maturity n and r denotes the one month rate. The dashed lines are the 95% confidence intervals constructed from the standard error estimate for each slope point estimate, with a normal distribution assumption. Standard errors are computed following Newey and West (1987) with six lags. The regressions are performed on monthly U.S. Treasury data from January 1985 to December 1999 (180 observations).

While both measures are approximations of the conditional dynamics,³ the similarities are suggestive. The fact that hump dynamics are observed across different currencies and from different measures implies that it is a robust feature of the interest rate data.

While the three dimensions do not exhaust the known properties of bond yields, they represent three layers of the data that a reasonable model should reproduce. The first layer summarizes the general statistical features of the time series and imposes minimal structure to the analysis. The second layer deals with predictability of interest rates and has far-reaching implications for interest rate forecasting. The third layer of property is even more subtle. It concerns the conditional dynamics

³ The two measures differ, among other things, in both units and measures. The variance of multiperiod changes approximates the conditional variance of continuously compounded spot rates under the objective measure while the implied volatility approximates the conditional volatility rate of the simply compounded LIBOR rate under the risk neutral measure.

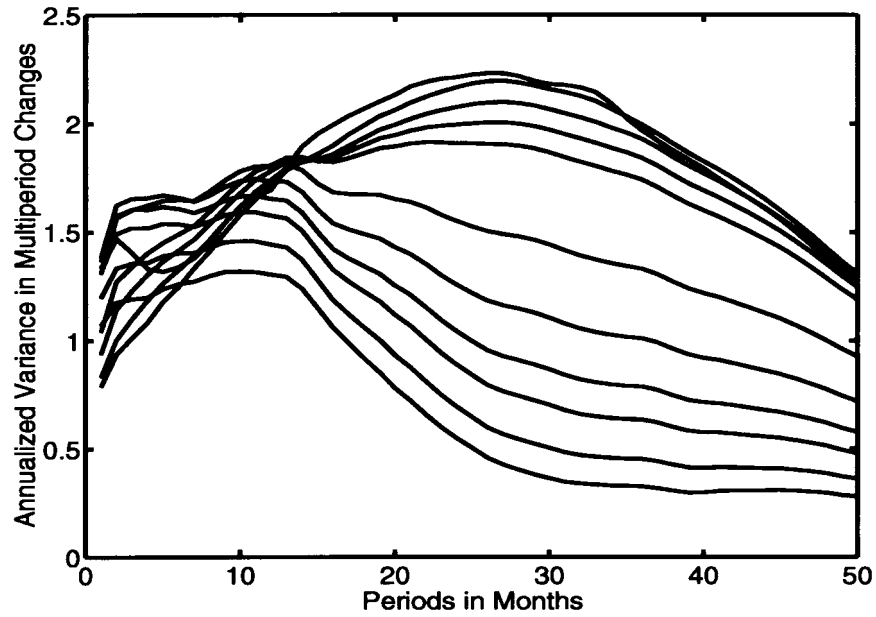


Figure 3. Hump-shaped conditional dynamics in U.S. treasuries. Lines are annualized variance estimates of multiperiod changes in yields (in annualized percentage), $\text{Var}(y_{t+k} - y_t)/k$. The maturities of the bond yields are, from top to bottom, 1, 3, 6, 9, 12, 24, 36, 48, 60, 84, and 120 months. The data are monthly from January 1985 to December 1999.

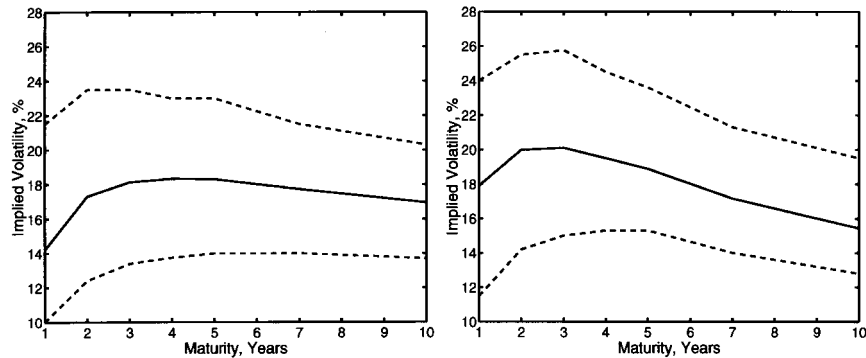


Figure 4. Hump-shaped conditional dynamics in interest rate caps. Solid lines are the mean term structure of at-the-money implied volatility quotes for interest rate caps for U.S. dollar (left) and Deutsche mark (right). Dashed lines depict the 95% and 5% quantiles. The data are from February 1st, 1995 to October 17th, 2000 (1490 observations), downloaded from Datastream.

of the second moments, which has strong implications for model design and even more so for applications in risk management and derivatives pricing.

3. Quadratic Models

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq \mathcal{T}}, \mathbb{P})$ be a stochastic basis. The filtration $\{\mathcal{F}_t\}_{0 \leq t \leq \mathcal{T}}$ satisfies the usual conditions of right-continuity and completeness. We fix a strictly positive horizon date $\mathcal{T} > 0$. The process W is a d -dimensional Wiener process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that the underlying filtration $\{\mathcal{F}_t\}_{0 \leq t \leq \mathcal{T}}$ coincides with the usual \mathbb{P} -augmentation of the natural filtration of W . We introduce uncertainty in our economy by assuming that all assets are functions of an underlying Markov process X valued in some open subset \mathcal{D} of \mathbb{R}^d . Suppose that for any time $t \in [0, \mathcal{T}]$ and time-of-maturity $T \in [t, \mathcal{T}]$, the market value at time t of a zero-coupon bond with time-to-maturity $\tau = T - t$ is fully characterized by $P(X_t, \tau)$. The discrete-time notation P_t^n then corresponds to $P(X_t, nh)$ with $h = 1/12$ denoting the monthly time interval.

DEFINITION 1. In the *quadratic class of term structure models*, the prices of zero-coupon bonds, $P(X_t, \tau)$, are exponential-quadratic functions of the Markov process X_t :

$$P(X_t, \tau) = \exp[-X_t^\top A(\tau) X_t - \mathbf{b}(\tau)^\top X_t - c(\tau)], \quad (2)$$

where $A(\tau)$ is a *nonsingular* $d \times d$ matrix, $\mathbf{b}(\tau)$ is a $d \times 1$ vector, and $c(\tau)$ is a scalar.

Leippold and Wu (2002) have identified the quadratic class in terms of the Markov process X_t , the instantaneous interest rate function $r(X_t)$, and the market price of risk $\gamma(X_t)$:

$$\begin{aligned} dX_t &= -\kappa X_t dt + dW_t; \\ r(X_t) &= X_t^\top A_r X_t + \mathbf{b}^\top X_t + c_r, \\ \gamma(X_t) &= A_\gamma X_t + \mathbf{b}_\gamma, \end{aligned} \quad (3)$$

where $\kappa, A_r, A_\gamma \in \mathbb{R}^{d \times d}$, $\mathbf{b}_r, \mathbf{b}_\gamma \in \mathbb{R}^d$, and $c_r, c_\gamma \in \mathbb{R}$. Due to the symmetric nature of the quadratic models, we assume A_r to be symmetric with no loss of generality. For the Markov process to be stationary, we also require that all the eigenvalues of κ be positive. We further normalize the diffusion of X_t to an identity matrix and its long run mean to zero.

As long as one is allowed to hold cash without cost, the instantaneous interest rate has to stay positive to guarantee no-arbitrage. A sufficient condition for the instantaneous interest rate to be positive is to restrict A_r to be positive definite and $c_r \geq \frac{1}{4} \mathbf{b}_r^\top A_r^{-1} \mathbf{b}_r$.

Straightforward application of Girsanov's theorem shows that, under the risk neutral measure \mathbb{P}^* , the drift of the Markov process remains affine with $\kappa^* = \kappa + A_\gamma$ and $\kappa^*\theta^* = -\mathbf{b}_\gamma$. Stationarity for the Markov process under the risk neutral measure requires that all the eigenvalues of $\kappa^* = \kappa + A_\gamma$ be also positive.

Under the above specification, the coefficients for the bond price are determined by the following ordinary differential equations (ODE):

$$\begin{aligned} A'(\tau) &= A_r - A(\tau)(\kappa + A_\gamma) - (\kappa + A_\gamma)^\top A(\tau) - 2A(\tau)^2; \\ \mathbf{b}'(\tau) &= \mathbf{b}_r - 2A(\tau)\mathbf{b}_\gamma - (\kappa + A_\gamma)^\top \mathbf{b}(\tau) - 2A(\tau)\mathbf{b}(\tau); \\ c'(\tau) &= c_r - \mathbf{b}(\tau)^\top \mathbf{b}_\gamma + trA(\tau) - \mathbf{b}(\tau)^\top \mathbf{b}(\tau)/2, \end{aligned} \quad (4)$$

subject to the boundary conditions: $A(0) = 0$, $\mathbf{b}(0) = 0$, and $c(0) = 0$. In calibration, corresponding to the monthly data we use, we adopt a discrete-time version of the model. In particular, we choose monthly frequency, use the one-month rate as a proxy for the short rate, and solve the ODEs numerically by the Euler's method.

Given bond prices, bond yields are obtained straightforwardly:

$$y(X_t, \tau) = -\frac{1}{\tau} \ln P(X_t, \tau) = \frac{1}{\tau} (X_t^\top A(\tau) X_t + \mathbf{b}(\tau)^\top X_t + c(\tau)). \quad (5)$$

The discrete-time notation for the monthly yield, y_t^n , corresponds to $y(X_t, nh)$, with $h = 1/12$. The monthly forward rate, f_t^n , denotes the time t forward rate between $t + nh$ and $t + (n + 1)h$:

$$f_t^n = \frac{1}{h} \ln P(X_t, nh) / P(X_t, (n + 1)h) = X_t^\top A_n^f X_t + X_t^\top \mathbf{b}_n^f + c_n^f, \quad (6)$$

with

$$\begin{aligned} A_n^f &= \frac{A((n + 1)h) - A(nh)}{h}, & \mathbf{b}_n^f &= \frac{\mathbf{b}((n + 1)h) - \mathbf{b}(nh)}{h}, \\ c_n^f &= \frac{c((n + 1)h) - c(nh)}{h}. \end{aligned}$$

The free parameters of the quadratic model include: $\Theta = (\kappa, A_\gamma, \mathbf{b}_\gamma, A_r, \mathbf{b}_r, c_r)$. Given these parameters and the current state vector X_t , bond prices, yields, and forward rates can be determined by (2), (5), and (6).

4. Property Analysis

In this section, we analyze the implications of quadratic models on the three layers of properties of interest rates. The analysis lays a foundation for moment choices in our generalized method of moments (GMM) estimation in the subsequent section. More importantly, we ask what requirements the stylized evidence imposes on model design.

4.1. STATISTICAL PROPERTIES

With the Markov process specified in (3), the state vector X is both conditionally and unconditionally normally distributed. In our discrete-time version with monthly intervals, let $\Phi = e^{-\kappa h}$ with $h = 1/12$ denote the monthly autocorrelation matrix of the state vector, we can then write the unconditional and conditional variance as

$$\text{vec}(V) = (I - \Phi \otimes \Phi)^{-1} \text{vec}(I)h; \quad V_k = \sum_{j=0}^{k-1} \Phi^j (\Phi^j)^\top,$$

where the subscript k denotes the conditional variance in k discrete periods (months). The k -period conditional mean is $\mu_{t,k} = \Phi^k X_t$ while the unconditional mean is zero.

Let q_t^n denote a generic quadratic form of normal variates:

$$q_t^n = X_t^\top A_n X_t + \mathbf{b}_n^\top X_t + c_n,$$

for any (A_n, \mathbf{b}_n, c_n) . q_t^n can be a bond yield, a forward rate, or the short rate. Fully utilizing the well-documented properties of quadratic forms of normal variates,⁴ we can derive the following properties for q_t^n :

$$\begin{aligned} \mathbb{E}[q_t^n] &= \text{tr}(A_n V) + c_n; \\ \text{Var}[q_t^n] &= 2\text{tr}((A_n V)^2) + \mathbf{b}_n^\top V \mathbf{b}_n; \\ \text{Cov}(q_{t+k}^n, q_t^n) &= 2\text{tr}((\Phi^k)^\top A_n \Phi^k V A_n V) + \mathbf{b}_n^\top \Phi^k V \mathbf{b}_n. \end{aligned}$$

The monthly k th-order autocorrelation, $\rho(k)$, of a quadratic form q_t^n is then given by

$$\rho(k) = \frac{2\text{tr}((\Phi^k)^\top A_n \Phi^k V A_n V) + \mathbf{b}_n^\top \Phi V \mathbf{b}_n}{2\text{tr}((A_n V)^2) + \mathbf{b}_n^\top V \mathbf{b}_n}.$$

In the case of a one-factor model, the autocorrelation function, $\rho(k)$, can be written as a weighted average of the autocorrelation and its square of the Markov process:

$$a(n)\phi^{2k} + b(n)\phi^k = \rho(k), \quad (7)$$

with $\phi \equiv \exp(-\kappa h)$ being the monthly autocorrelation of the Markov process X , and

$$a(n) = \frac{2(A_n V)^2}{2(A_n V)^2 + b_n^2 V}; \quad b(n) = \frac{b_n^2 V}{2(A_n V)^2 + b_n^2 V}.$$

⁴ See, for example, Holmquist (1996), Johnson and Kotz (1970), Kathri (1980), and Mathai and Provost (1992).

Therefore, in contrast to the AR(1) type behavior of one-factor affine models, bond yields under quadratic models enjoy a richer, nonlinear dynamics. In particular, the autocorrelation function of bond yields under a one-factor quadratic model can vary across maturities, in conformity with the data, while all one-factor affine models imply the *same* autocorrelation function across yields and forward rates of *all* maturities. Within the affine class, multiple factors are needed to generate the observed nonlinearities in the interest rate dynamics. In contrast, nonlinearity is intrinsically built into the quadratic model through the quadratic term.

4.2. FORECASTING RELATIONS

To derive the forward regression slope, g_n , in (1), we apply two important properties of quadratic forms of normal variates:

$$\begin{aligned} \text{Cov}(q_t^m, q_t^n) &= 2\text{tr}(A_m V A_n V) + \mathbf{b}_m^\top V \mathbf{b}_n; \\ \text{Cov}(q_{t+k}^m, q_t^n) &= 2\text{tr}\left((\Phi^k)^\top A_m \Phi^k V A_n V\right) + \mathbf{b}_m^\top \Phi^k V \mathbf{b}_n. \end{aligned}$$

The forward regression slope g_n can be written as

$$\begin{aligned} c_n &= \frac{\text{Cov}(f_{t+1}^{n-1} - r_t, f_t^n - r_t)}{\text{Var}(f_t^n - r_t)} \\ &= \frac{2\text{tr}\left[\left(\Phi^\top A_{n-1}^f \Phi - A_r\right) V A_{frn} V\right] + \mathbf{b}_{frn}^\top V \left(\Phi^\top \mathbf{b}_{n-1}^f - \mathbf{b}_r\right)}{2\text{tr}\left(A_{frn} V\right)^2 + \mathbf{b}_{frn}^\top V \mathbf{b}_{frn}}, \end{aligned} \quad (8)$$

where $A_{frn} = A_n^f - A_r$ and $\mathbf{b}_{frn} = \mathbf{b}_n^f - \mathbf{b}_r$. The relation is relatively opaque, but its convergence to the stationary state is not. Suppose indeed that a stationary state exists, as $n \rightarrow \infty$, $A_n^f \rightarrow 0$ and $\mathbf{b}_n^f \rightarrow 0$, the regression slope converges to one:

$$\lim_{n \rightarrow \infty} c_n = \frac{2\text{tr}(A_r V)^2 + \mathbf{b}_r^\top V \mathbf{b}_r}{2\text{tr}(A_r V)^2 + \mathbf{b}_r^\top V \mathbf{b}_r} = 1.$$

This result confirms with the analysis of Dybvig et al. (1996). As long as the interest rate processes are stationary, the variance of the forward rate falls with maturity. Therefore, for very long maturities, we are essentially regressing $-r$ on itself.

Backus et al. (2001) illustrate that intrinsic tension exists for a one-factor Cox et al. (1985) (CIR) model to simultaneously fit the mean yield curve and the regression slope. To make $c_1 < 1$, the market price risk needs to be greater than zero while an upward sloping mean yield curve requires it to be negative. To release the tension, they propose a negative CIR model, where the short rate is proportional to the negative of a CIR factor.

Duffee (2002) further illustrates that the inherent tension remains even for multi-factor CIR models when one tries to match the properties of the whole term

structure of excess returns. To release the tension, he proposes the application of Gaussian state variables with a flexible affine market price of risk specification. Dai and Singleton (2002a) incorporates such a specification to explain the EH violations. In particular, Dai and Singleton (2002a) show that such a specification also releases the tension identified by Backus et al. (2001) as \mathbf{b}_γ controls the shape of the mean yield curve (and hence should be negative) while the slope parameter A_γ controls the regression slope (and should be positive).

Note, however, that the one-factor affine example of Dai and Singleton (2002a) is merely a degenerating case of a one-factor quadratic model with $A_r = 0$. Under the quadratic class, we only use Gaussian state variables. Affine market price of risk is naturally incorporated into the framework. The incorporation of the quadratic term A_r further enriches the interest rate dynamics and can prevent the interest rate from being negative.

4.3. CONDITIONAL DYNAMICS

Conditional dynamics in general and conditional variance in particular have far-reaching implications in risk management and option pricing. A central feature of the conditional dynamics for bond yields, as we observed earlier, is that the conditional volatility or variance of bond yields has a hump-shaped mean term structure. Let $cv(k)_n = \mathbb{E}[Var_t[y_{t+k}^n]]$ denote the mean conditional variance of n -month yields with a conditional horizon of k periods. Let $av(k)_n = cv(k)_n/(kh)$ denotes the annualized mean conditional variance. The hump-shaped conditional dynamics implies that $av(k)$ increases with k at first, reaches a plateau, and then decreases as k further increases.

The conditional moments of quadratic forms of normal variates are given by

$$\begin{aligned}\mathbb{E}_t[q_{t+k}^n] &= \mu_{t,k}^\top A_n \mu_{t,k} + tr(A_n V_k) + \mathbf{b}_n^\top \mu_{t,k} + c_n; \\ Var_t[q_{t+k}^n] &= 2[tr((A_n V_k)^2) + 2\tilde{\mu}_{t,k}^\top A_n V_k A_n \tilde{\mu}_{t,k}],\end{aligned}$$

where $\tilde{\mu}_{t,k} = \mu_{t,k} + \frac{1}{2}A_n^{-1}\mathbf{b}_n$ and $\mu_{t,k} = \Phi^k X_t$. The mean term structure of the conditional variance under the quadratic class can be written as

$$cv(k)_n = 2tr[(A_n V_k)^2] + 2(\Phi^k)^\top A_n V_k A_n \Phi^k V + \mathbf{b}_n^\top V_k \mathbf{b}_n. \quad (9)$$

PROPOSITION 1. Strong interdependence between elements of the state vector is essential in generating a hump-shaped conditional dynamics. Neither one-factor nor independent multifactor quadratic models are capable of generating the hump.

Refer to Appendix A for the proof. Similar necessary conditions are also required for affine models. However, the correlation structures between multi-factor CIR models are restricted. For example, Dai and Singleton (2000) and Backus et al. (2001) both observe that the unconditional correlation between two square-root state variables can only be positive. Hence, while multi-factor CIR models in

theory can generate a hump shape, the hump is often not large enough to match the evidence. In contrast, quadratic models only incorporate Ornstein-Uhlenbeck processes as the driving Markov process, the correlation structure between state variables can be chosen freely. Fitting the hump shape hence becomes a relatively easy task.

From the time series data, the conditional dynamics can be approximately captured by the annualized variance of multiperiod changes. Under the quadratic model, the variance of k -period changes in yields or forward rates q_t^n is given by

$$\begin{aligned} v(k)_n &= \text{Var}(q_{t+k}^n - q_t^n) = 2 [\text{Var}(q_t^n) - \text{Cov}(q_{t+k}^n, q_t^n)] \\ &= 4tr \left(\left(A_n - (\Phi^k)^\top A_n \Phi^k \right) V A_n V \right) + 2\mathbf{b}_n^\top (I - \Phi^k) V \mathbf{b}_n. \end{aligned}$$

One can readily prove that, similarly, strong interactions between state variables are required for the annualized variance of multi-period changes to be hump-shaped.

Nevertheless, conditional dynamics implied from options prices and conditional dynamics inferred from the time series are dynamics under two different measures. The former is under the risk-neutral measure while the latter is under the objective measure. The correlation structure is hence captured by $\kappa^* = \kappa + A_\gamma$ in the former and by κ in the latter. Therefore, to simultaneously capture the observed conditional dynamics in both the time series and option prices, one also imposes constraints on the specification of the market price of risk.

In summary, the quadratic class of term structure models exhibits great potentials in simultaneously (1) matching the mean yield curve and the forecasting relations through the specification of the affine market price of risk $A_\gamma X_t + \mathbf{b}_\gamma$, and (2) generating the observed hump-shaped conditional dynamics by the flexible specification of the correlation structures between the state variables. Furthermore, the quadratic term enriches the dynamics of the interest rate by incorporating non-linearity between state variables and interest rates and also provides a consistent approach to guarantee positive interest rates. In the next section, we further illustrate these properties by calibrating the quadratic model to the U.S. Treasury data.

5. Calibration with GMM Estimation

This section corroborates the property analysis in the previous section with empirical estimation. For this purpose, we choose the simplest model within the quadratic class which can approximate the three layers of properties of interest rates. As demonstrated in Dai and Singleton (2002a), a one-factor quadratic model suffices in capturing both the mean yield curve and the EH regression slopes. To capture the hump dynamics, however, we need at least a two-factor model to incorporate non-trivial correlation structures between state variables. Moment conditions are chosen based on the property analysis in the previous section. We calibrate the model to the time series of U.S. Treasuries.

Table II. Summary of moment conditions for GMM estimation

Mean yield curve	$\mu_n = tr(A_n V) + c_n$
Autocorrelation	$\rho(k)_n = \frac{2tr\left((\Phi^k)^\top A_n \Phi^k V A_n V\right) + \mathbf{b}_n^\top \Phi^k V \mathbf{b}_n}{tr(2(A_n V)^2) + \mathbf{b}_n^\top V \mathbf{b}_n}$
Forward Regression	$g_n = \frac{2tr\left[(\Phi^\top A_{n-1}^f \Phi - A_r) V A_{f_{rn}} V\right] + \mathbf{b}_{f_{rn}}^\top V (\Phi \mathbf{b}_{n-1}^f - \mathbf{b}_r)}{2tr(A_{f_{rn}} V)^2 + \mathbf{b}_{f_{rn}}^\top V \mathbf{b}_{f_{rn}}}$
Hump dynamics	$v(k)_n = 4tr\left(\left(A_n - (\Phi^k)^\top A_n \Phi^k\right) V A_n V\right) + 2\mathbf{b}_n^\top (I - \Phi^k) V \mathbf{b}_n$

The left hand side of the equations denotes the moment conditions while the right hand side denotes them as functions of parameters of the quadratic term structure models. μ_n denotes the mean, $\rho(k)_n$ the k th order autocorrelation, and $v(k)_n$ the variance of the k -period changes, of the n -month bond yield. g_n denotes the forward regression slope. (A_n, \mathbf{b}_n, c_n) denote the coefficients of the n -month bond yield.

5.1. MOMENT CONDITIONS AND INFERENCE

The parameter set of the quadratic model is: $\Theta \equiv (\kappa, A_r, \mathbf{b}_r, c_r, A_\gamma, \mathbf{b}_\gamma)$. We choose 23 moment conditions for the GMM estimation. These are taken from

- Statistical properties:** Mean yields with maturities of 1, 6, 12, 60, and 120 months and the first order autocorrelation of the short rate (one-month yield).
- Forecasting relations:** One period forward regression slopes, g_n , with maturities $n = 1, 6, 12$.
- Conditional dynamics:** Variance of k -period (month) changes of 1-month and 6-month yields with $k = 6, 10, 14, 16, 18, 20, 24$.

Furthermore, given other parameters, we set c_r to match the mean short rate (μ_r), $c_r = \mu_r - tr(A_r V)$. The moment conditions and their values as functions of parameters are summarized in Table II.

Let $s_t^n = y_t^n - r_t$ denote the yield spread and $\Delta^{(k)} y_t^n$ the k -period change of the yield y_t^n . Let $\rho(x)$, $\mu(x)$, g_n , and $v(k)_n$ denote, respectively, the first-order autocorrelation of x , the mean of x , the slope of the forward regression with maturity n , and the variance of k -period changes in y_t^n , all implied from the model parameters and computed based on the formulae in Table II. Given these notations, the four types of moment conditions can be written as:

$$\begin{aligned}
 h_t^1(n) &= s_t^n - \rho(s^n) s_{t-1}^n - (1 - \rho(s^n)) \mu(s^n), & (n = 6, 12, 60, 120) \\
 h_t^2(n) &= \left[y_t^n - \rho(y^n) y_{t-1}^n - (1 - \rho(y^n)) \mu(y^n) \right] (y_{t-1}^n - \mu(y^n)), & (n = 1) \\
 h_t^3(n) &= \left[(f_t^{n-1} - r_{t-1}) - g_n (f_{t-1}^n - r_{t-1}) \right] (f_{t-1}^n - r_{t-1}), & (n = 1, 6, 12) \\
 h_t^4(n, k) &= v(k)_n - \left(\Delta^{(k)} y_t^n \right)^2, & (n = 1, 6; \\
 & & k = 4, 6, 10, 14, 18, 20, 24)
 \end{aligned}$$

where k denotes the order and n denotes the maturity. The forward rate and the short rate in h_t^3 are demeaned: $f_t^n = f_t^n - \bar{f}_t^n$.

We estimate one unrestricted version and three restricted versions of a two-factor quadratic model:

- *Model A* (unrestricted two-factor quadratic model);
- *Model B* (independent two-factor quadratic model): $\kappa(1, 2) = A_\gamma(1, 2) = A_r(1, 2) = 0$;
- *Model C* (two-factor quadratic model with constant market price of risk): $A_\gamma = 0$;
- *Model D* (two-factor Gaussian affine model): $A_r = 0$.

For identification reasons, we normalize κ and A_γ to be upper triangular and A_r symmetric. Model B disallows any interactions between the two state variables and hence serves as an over-identification test on how important such interactions are in capturing the properties of U.S. Treasuries. Model C, on the other hand, provides a test on the significance of the affine market price of risk A_γ . Finally, Model D tests the significance of the contribution of the quadratic term A_r .

The weighting matrix, W , for the GMM estimation is constructed according to Newey and West (1987) with demeaned moment conditions (Bekaert and Urias, 1996). We follow the approach of Andrews (1991) in setting the lag truncation parameter to 12 based on VAR(1) estimates on the moment conditions. Each model is estimated using the same moment conditions and the same weighting matrix: the Newey and West (1987) covariance matrix implied by estimates of Model A, which includes each of the other models as a special case and hence provides a common basis for comparison. The weighting matrix is approximately a fixed point for Model A: It both produces and is produced by the parameter estimates.

Excluding c_r as a free parameter and the mean short rate as a moment condition, we have 22 orthogonality conditions and 13 free parameters in the unrestricted model (Model A). Hence, the model is over-identified and has nine over-identification restrictions. Each of the three restricted models (B, C, and D) has three more constraints on parameters. The number of over-identification restrictions increases to 12.

Let T denote the number of observations and let $\mathbf{e}_T = [e^j(n, k)] \in \mathbb{R}^{22 \times 1}$, with

$$e^j(n, k) = \frac{1}{T} \sum_{t=1}^T h_t^j(n, k), \quad j = 1, 2, 3, 4,$$

denote the sample average of the orthogonality conditions. Let $\hat{\Theta}_T$ be the parameter estimates that minimize the objective function:

$$\mathbf{e}_T^\top W \mathbf{e}_T.$$

Asymptotically, under certain technical conditions (Hamilton, 1994), the estimator has a normal distribution:

$$\hat{\Theta}_T \sim N(\Theta, V/T), \quad \text{where } V = [D'WD]^{-1} D'WSWD [D'WD]^{-1}, \quad (10)$$

where S is the spectral density matrix estimated following Newey and West (1987) with 12 lags, W is the universal weighting matrix, and D is the Jacobian matrix, defined as

$$D = \left. \frac{\partial \mathbf{e}_T^\top}{\partial \Theta} \right|_{\Theta = \hat{\Theta}_T}.$$

For the unrestricted model (Model A), we have $W = S^{-1}$ and the covariance matrix is simplified to $V = [D'WD]^{-1}$.

Under certain technical conditions, the moment conditions also have an asymptotic normal distribution,

$$\mathbf{e}_T \sim N(0, M/T),$$

where

$$M = [I - D(D'WD)^{-1}D'W]S[I - D(D'WD)^{-1}D'W]^\top.$$

This matrix simplifies for $W = S^{-1}$, giving $M = S - DVD'$. When the model is over-identified, as is the case in this paper, a χ^2 test can be constructed for the over-identifying restrictions,

$$T\mathbf{e}_T^\top(M)^{-1}\mathbf{e}_T \sim \chi^2(r - a),$$

where $r = 22$ is number of orthogonality conditions and a is number of free parameters (13 for Model A and 10 for Models B, C, and D).

5.2. PERFORMANCE ANALYSIS

Table III reports the GMM estimates and standard errors of the four two-factor quadratic models. Also included are the J -statistics, p -values, and the probability of negative interest rates. Table IV reports the goodness-of-fit test on the 22 moment conditions under each model. Again, the t -statistics of the individual moment conditions are smallest overall under Model A.

Figure 5 depicts the superb performance of the unrestricted two-factor quadratic model in matching the three layers of the data: the mean yield curve, the forward regression slopes, and hump-shaped dynamics. The solid lines in the figure are implied by the model estimates in Table III. The dash-dotted lines correspond to the 10% and 90% quantiles. They are computed by the delta method, based on the covariance matrix of the parameter estimates and an asymptotic normal distribution assumption.

The three restricted versions test, respectively, the significance of three important parts of the quadratic model. All three restricted versions generate larger J -statistics than Model A.⁶ Furthermore, in Model A, the estimates of almost all

⁶ However, none of the models can be rejected under 10% confidence level.

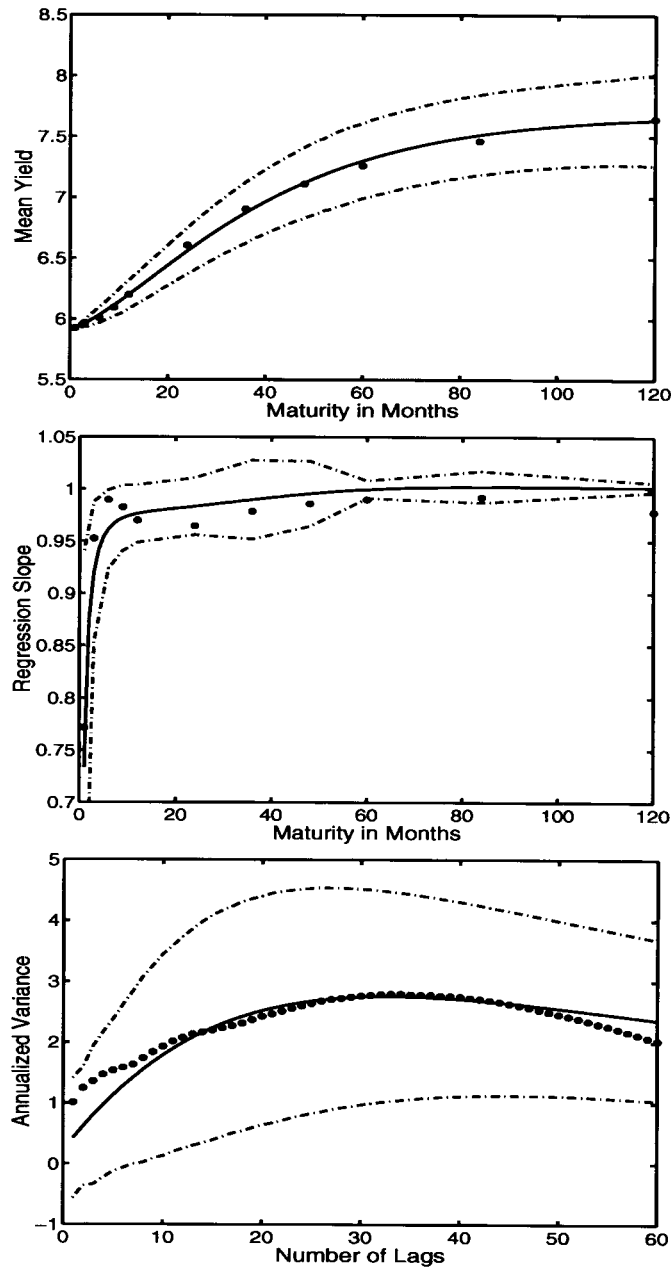


Figure 5. Performance of a two-factor quadratic model. Lines represent the performance of the unrestricted two-factor quadratic model (Model A) in matching the mean yield curve (top), the forward regression slope (middle), and the annualized variance of multiperiod changes of the six-month yield (bottom). Stars are data estimates. The solid lines correspond to the point estimates of Model A reported in Table III. The dashed lines correspond to the 10% and 90% quantiles, computed by the delta method based on the covariance matrix of the parameter estimates and an asymptotic normality assumption.

Table III. Estimates of quadratic models on the U.S. Treasuries

Parameter	A	B	C	D
$\kappa(1, 1)$	0.8291 (0.2462)	0.1815 (0.0533)	0.5521 (0.0622)	0.6234 (0.0207)
$\kappa(1, 2)$	-6.0332 (1.3188)	0	-24.7089 (4.0318)	-4.2288 (0.1254)
$\kappa(2, 2)$	0.6110 (0.2773)	14.1780 (0.6707)	1.0191 (0.1653)	0.7984 (0.3861)
$A_r(1, 1)$	0.0423 (0.0212)	0.2021 (0.0297)	0.0034 (0.0006)	0
$A_r(1, 2)$	-0.0102 (0.0029)	0	-0.0008 (0.0001)	0
$A_r(2, 2)$	0.1388 (0.0767)	1.5615 (0.4652)	0.0676 (0.0232)	0
$\mathbf{b}_r(1)$	0.0745 (0.0436)	0.8650 (0.1029)	-0.0137 (0.0085)	0.6085 (0.0157)
$\mathbf{b}_r(2)$	0.7176 (0.1583)	0.5887 (0.2316)	0.6832 (0.2133)	0.0885 (0.0036)
$A_\gamma(1, 1)$	-0.4494 (0.1589)	0.2027 (0.0386)	0	-0.5413 (0.0089)
$A_\gamma(1, 2)$	0.0333 (0.0303)	0	0	3.3835 (0.2266)
$A_\gamma(2, 2)$	2.4696 (0.5559)	-12.0450 (1.6271)	0	5.4851 (0.4210)
$\mathbf{b}_\gamma(1)$	-2.5700 (1.4833)	-1.9128 (0.3341)	-1.7748 (0.6666)	-0.6271 (1.1220)
$\mathbf{b}_\gamma(2)$	-3.1605 (1.2649)	-1.2224 (0.8372)	-2.4091 (0.4431)	-14.4570 (7.0320)
J -statistic	4.1808	15.2542	8.4497	8.9210
Deg of Fr, ν	9	12	12	12
p -value	89.91%	22.78%	74.91%	70.79%
$\Pr(r < 0)$	0	0	0	15.17%

Entries are GMM estimates (standard errors in parentheses) of the parameters of models: (A) an unrestricted two-factor quadratic model; (B) an independent two-factor quadratic model with $\kappa(1, 2) = A_\gamma(1, 2) = A_r(1, 2) = 0$; (C) a two-factor quadratic model with $A_\gamma = 0$; and (D) a two-factor Gaussian-affine model with $A_r = 0$. The moment conditions are (i) the mean yields with maturities of 6, 12, 60, and 120 months and the first order autocorrelation of the short rate, (ii) the forward regression slope with maturities 1, 6, and 12 months, and (iii) the variance of multi-period (k) changes of one and six-month yields with $k = 6, 10, 14, 16, 18, 20, 24$. Furthermore, c_r is used to match the mean short rate perfectly. J -statistics, degree of freedom, p -value, and the probability that the short rate becomes negative are also reported.

the parameters are significantly different from zero, implying that they all play important roles in delivering the observed superior performance. In what follows, in accordance with the three restricted versions of the quadratic model, we discuss the separate roles played by, respectively, (i) the off-diagonal terms of κ , A_γ , and A_r , (ii) the affine market price of risk term A_γ , and (iii) the quadratic term A_r in the short rate function.

Table IV. Goodness-of-fit tests of the moment conditions on U.S. Treasuries

Moments	Model A	Model B	Model C	Model D
$e^1(6)$	-0.5073	-0.6198	-0.4391	-0.4949
$e^1(12)$	-0.5305	-0.7566	-0.3704	-0.5428
$e^1(60)$	-0.7954	-0.8794	0.1771	-0.9119
$e^1(120)$	-0.9273	-0.7619	0.5515	-1.0168
$e^2(1)$	-0.8139	-0.3991	-0.8338	-0.9090
$e^3(1)$	-0.2315	1.1741	-0.2083	-1.3188
$e^3(6)$	-0.9562	0.0467	-0.6904	-0.8710
$e^3(12)$	-0.6262	-0.0794	-0.5174	-1.3231
$e^4(1, 6)$	1.0208	0.9671	0.7428	1.0306
$e^4(1, 10)$	0.5776	1.7283*	0.4067	0.6444
$e^4(1, 14)$	0.2841	1.9378*	0.2722	0.2740
$e^4(1, 16)$	0.0806	1.9288*	0.1109	0.0284
$e^4(1, 18)$	0.0567	1.9422*	0.1357	-0.0903
$e^4(1, 20)$	0.0617	1.9293*	0.1861	-0.1877
$e^4(1, 24)$	0.1047	1.8995*	0.2836	-0.3303
$e^4(6, 6)$	0.1675	0.4391	0.1620	0.7975
$e^4(6, 10)$	-0.1808	1.3905	-0.3018	0.3271
$e^4(6, 14)$	-0.0279	1.8350*	-0.1409	0.1523
$e^4(6, 16)$	0.0113	1.9507*	-0.0691	0.0652
$e^4(6, 18)$	0.0096	2.0324**	-0.0396	-0.0498
$e^4(6, 20)$	0.0405	2.0982**	0.0297	-0.1413
$e^4(6, 24)$	0.0972	2.0475**	0.1480	-0.2918

Entries are the t -statistics for each of the 16 moment conditions:

$$e^j(n, k) = \frac{1}{T} \sum h_t^j(n, k), \quad j = 1, 2, 3, 4,$$

where T is the number of observations, j refers to the four moment types, n is the maturity of the yield or forward rate and k is the order defined in each condition. The estimates, standard errors, and J -statistics of each of the four models are reported in Table III. * and ** indicate significance under a 10% and 5% test.

5.2.1. Interactions Between State Variables and the Hump Dynamics

Proposition 1 proves that interactions between the state variables play an indispensable role in generating the observed hump-shaped dynamics. These interactions are captured by the off-diagonal terms of κ and A_γ . Specifically, the off-diagonal terms of κ capture the interactions between the state variables under the objective measure, whereas the off-diagonal terms in A_γ govern the interactions of the state

variables under the risk-neutral measure. In addition, the off-diagonal terms in A_r play a more subtle role. While they do not incorporate direct interactions between state variables, these terms intertwine the impacts of different state variables on the behavior of interest rates.

In Model B, we restrict the off-diagonal terms of κ , A_γ , and A_r all to zero. As a result, the two state variables do not interact with each other and do not interact in their impacts on the behavior of interest rates. In absence of such interactions, Model B fails to generate the hump-shaped dynamics observed in the interest rate data. Of the four models tested, the J -statistic is largest for Model B. The goodness-of-fit tests on individual moment conditions, as reported in Table IV, reveal specifically its incapability of fitting the moment conditions related to the hump dynamics, i.e. $e^4(n, k)$. The t -tests on many of these moment conditions are large. The null hypothesis that these individual moment conditions are zero is rejected under either 5% or 10% level.

Figure 6 depicts the fit of the hump-dynamics by the four models with parameter estimates from Table III. While the unrestricted two-factor quadratic model (Model A) generates the best fit, the independent two-factor quadratic model (Model B) is simply incapable of generating a hump, further illustrating the importance of factor interactions in generating the hump dynamics. Dai and Singleton (2000) also find that the correlation structure between state variables plays an important role in improving the general empirical performance of affine models. The analysis here pinpoints a dimension of the data which specifically requires such interactions and therefore provides an explanation as to why the general performance is improved.

The real importance of factor interactions goes well beyond the hump dynamics, but the hump dynamics provides a unique perspective in identifying the significance of factor interactions and possesses potentially important implications for applications in risk management and derivatives pricing.

5.2.2. Time-Varying Risk Premium and Expectations Hypotheses

The importance of time-varying risk premium has long been recognized. In particular, one ultimately needs to resort to flexible forms on the specification of the risk premium in order to account for the well-documented puzzles or anomalies associated with various forms of expectations hypotheses (EH), not only in interest rates but also in currencies (Backus et al., 2000). The recent empirical work by Backus et al. (2001), Dai and Singleton (2002a), and Duffee (2002) are examples of continued efforts in searching for specifications of the pricing kernel that are flexible enough to account for not only the anomalies in EH but also other salient features of the data.

The quadratic model provides a natural framework in accommodating flexible specifications for the market price of risk. In particular, the affine form $\gamma(X_t) = A_\gamma X_t + \mathbf{b}_\gamma$ is in line with the recent findings of Dai and Singleton (2002a) and Duffee (2002). Our calibration result on Model C (with $A_\gamma = 0$) further illustrates the importance of such a flexible specification.

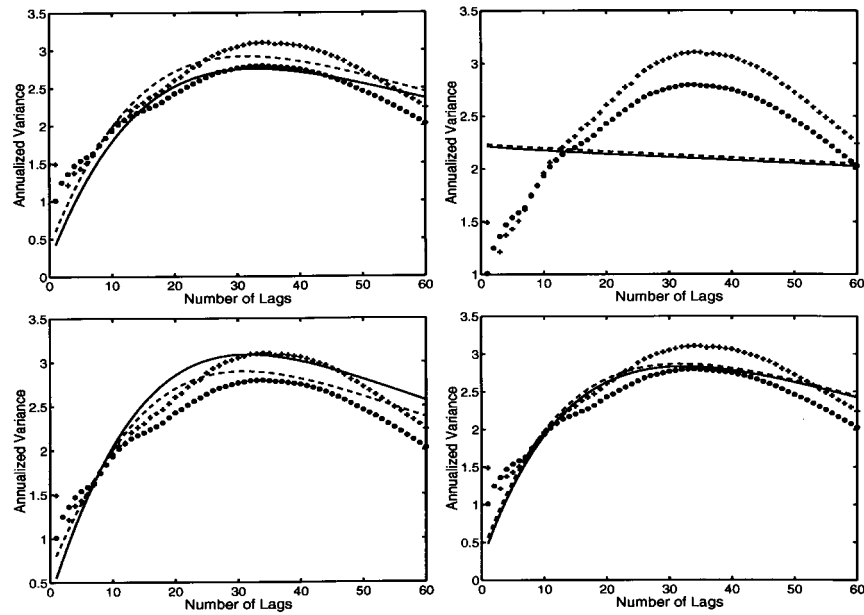


Figure 6. Fitting the hump dynamics for U.S. treasuries. The dashed and solid lines represent the performance of the four models: A (top left), B (top right), C (bottom left), and D (bottom right), in matching the annualized variance of multiperiod changes in one-month and six-month yields, respectively. Circles and stars are data estimates for the one-month and six-month yields, between January 1985 and December 1999. Parameter estimates of the models are reported in Table III.

With $A_\gamma = 0$, the market price of risk is reduced to a constant vector in Model C. Under the parameter estimates reported in Table III, Model C can fit the regression slope and the hump shape very well, but not the mean yield curve. The implied mean yield curve is much lower than observed in the data. This is illustrated in the top panel of Figure 7. Alternatively, we can find parameters for Model C that fit the mean yield curve and the hump dynamics well but miss badly on the regression slopes, as shown in the bottom panel of Figure 7. With the restriction of $A_\gamma = 0$ and positive interest rates, we are incapable of finding parameters that simultaneously fit both the mean yield curve and the regression slope. Such a tension is analogous to those observed in Backus et al. (2001). In contrast, with unrestricted A_γ , Models A, B, and C can all simultaneously fit the mean yield curve and regression slopes relatively well.

So far, we have confined ourselves to positive definite parameterizations for A_r (except for Model D, where $A_r = 0$) to guarantee positive interest rates. Inspired by the negative Cox et al. (1985) model of Backus et al. (2001), we experiment with A_r parameterizations which allow one of the eigenvalues to be negative. Interestingly, such a parameterization can fit both the mean yield curve and the regression slope

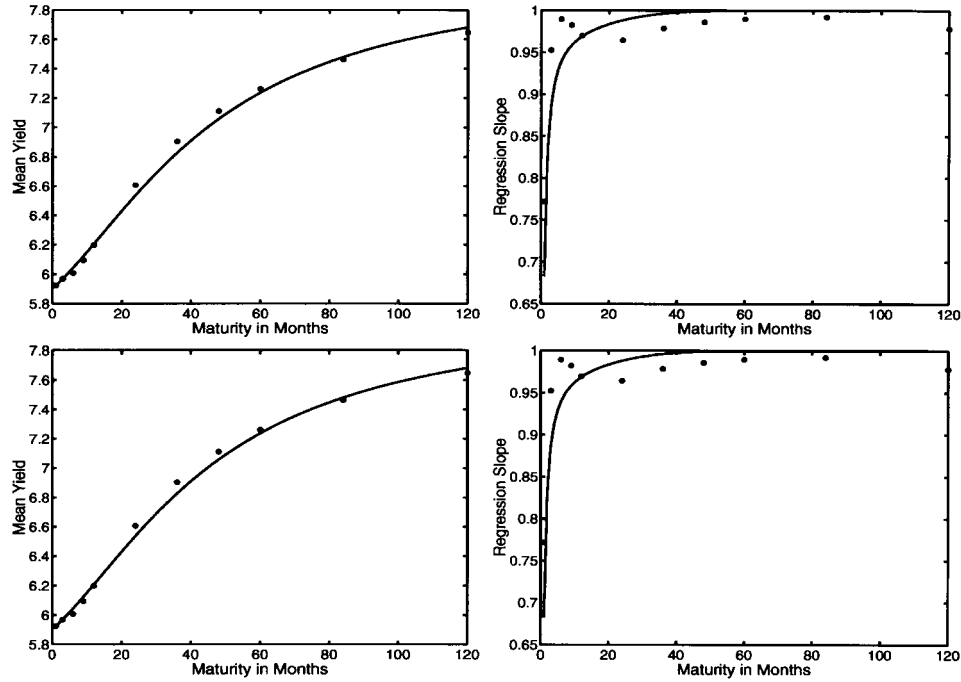


Figure 7. Performance of Model C with $A_\gamma = 0$. Lines represent the performance of Model C (with $A_\gamma = 0$) in matching the mean yield curve (left) and the forward regression slope (right). The top panels are computed from parameter estimates in Table III while the bottom panels are computed from the following parameter estimates:

$$\kappa = \begin{bmatrix} 0.4972 & -24.9860 \\ -24.9860 & 1.0231 \end{bmatrix}, A_r = \begin{bmatrix} 0.0029 & -0.0021 \\ -0.0021 & 0.1195 \end{bmatrix}, \mathbf{b}_r = \begin{bmatrix} 0.0053 \\ 0.7745 \end{bmatrix},$$

$$\mathbf{b}_\gamma = \begin{bmatrix} -2.7953 \\ -2.9148 \end{bmatrix}.$$

Stars are data estimates.

relatively well even with $A_\gamma = 0$. We, however, do not pursue further this route as it allows negative interest rates.

5.2.3. Positivity of Interest Rates

When we restrict the quadratic term A_r to be positive definite, we can rewrite the short rate function as

$$r(X_t) = \left(X_t + \frac{1}{2} A_r^{-1} \mathbf{b}_r \right)^\top A_r \left(X_t + \frac{1}{2} A_r^{-1} \mathbf{b}_r \right) + c_r - \frac{1}{4} \mathbf{b}_r^\top A_r^{-1} \mathbf{b}_r.$$

Obviously, as long as A_r is positive definite and $c_r \geq \frac{1}{4} \mathbf{b}_r^\top A_r^{-1} \mathbf{b}_r$, the short rate is guaranteed to be positive. The positivity of bond yields and forward rates can also be guaranteed by similar conditions. Parameter estimates for Models A, B, and C

in Table III all satisfy this positivity constraint for the short rate. But for Model D, as A_r is restricted to zero, positivity cannot possibly be satisfied because the model implies a normal distribution for the interest rate:

$$r \sim N(c_r, \mathbf{b}_r^\top V \mathbf{b}_r), \text{ where } V = \int_0^\infty e^{-s\kappa} e^{-s\kappa^\top} ds.$$

Note that V is the unconditional covariance matrix of the state vector X_t . Under Model D, the probability that the short rate becomes negative can be easily computed from this normality condition. Under the estimates in Table III, Model D implies a probability of 15.17% at which the short rate can be negative.

In addition to the unrealistic implication of normal interest rates, the fact that the Gaussian-affine model allows negative interest rates severely limits its application in practice as arbitrage opportunities can arise from such a model. Unfortunately, almost all the model designs under the affine framework which claim “success” in empirical performance allow negative interest rates. Examples include, but are not limited to, Backus et al. (2001a), Backus et al. (2001b), Dai and Singleton (2000), Dai and Singleton (2002a), Duffee (2002), Duffie and Singleton (1997), and Singleton (2001). In contrast, the unrestricted quadratic model (Model A) not only delivers superior empirical performance, but also guarantees that interest rates stay positive.

6. Concluding Remarks

We document the stylized evidence on U.S. Treasuries, study its implications on model design, and calibrate a two-factor quadratic term structure model that approximates the stylized evidence relatively well. Along the way, however, we ignore a number of issues that deserve comment. First, a two-factor model obviously does not capture all the variations in interest rates. Factor analysis by, for example, Litterman and Scheinkman (1991) claims the existence of at least three factors. Our choice of two factors represents the minimum number of factors required to capture the hump dynamics. Another issue is about the moment conditions for the GMM estimation, the selection of which is almost never unique. Ours is no exception. Admittedly, we could have used more “efficient” methods such as the maximum likelihood estimation, or the efficient methods of moments applied by, for example, Ahn et al. (2002) and Dai and Singleton (2000). Our choice of GMM and the particular moment conditions, on the other hand, focus on the economic interpretations and highlights the mapping between the stylized evidence and the model structure, which is the focus of our paper.

Recently, Heidari and Wu (2001) find that interest rate options seem to exhibit movements independent of the term structure of interest rates. Motivated by such evidence, a new wave of efforts, e.g., Dai and Singleton (2002b), Fan et al. (2001), Singleton and Umantsev (2001), have been devoted to search for term structure models which capture the behavior of both interest rates and interest rate options.

An important line of future research is to investigate the capability of quadratic models in simultaneously capturing the behaviors of both markets.

Appendix A. Proofs

PROOF OF PROPOSITION 1

Recall from (9) that the k -period conditional variance of a quadratic form $q(X_t) = X_t^\top A X_t + \mathbf{b}^\top X_t + c$ is given by

$$cv(k) \equiv \mathbb{E}[q(X_{t+k}) | \mathcal{F}_t] = 2tr \left[(AV_k)^2 + 2(\Phi^k)^\top AV_k A \Phi^k V \right] + \mathbf{b}^\top V_k \mathbf{b}. \quad (\text{A1})$$

For the term structure to exhibit a hump shape, we need the annualized variance $cv(k)/(kh)$ to increase initially with k , which implies that

$$cv(2) - 2cv(1) > 0.$$

In a one-factor case, from (A1), we have

$$\begin{aligned} cv(2) - 2cv(1) &= \frac{2A^2(1 + \phi^2)(1 + \phi^4)}{1 - \phi^2} + b^2(1 + \phi^2) - \frac{4A^2(1 + \phi^2)}{1 - \phi^2} - 2b^2 \\ &= -2A^2(1 + \phi^2)^2 - b^2(1 - \phi^2) < 0. \end{aligned} \quad (\text{A2})$$

as $\phi < 1$ for a stationary Markov process. Hence, the initial term structure is downward sloping. No hump can be generated from a one-factor model. The same is also true for completely independent multifactor models with A_r , κ , A_γ being diagonal matrices, in which case, the last line in (A2) represents the contribution from each diagonal element, which is negative as long as that element is stationary. $cv(2) - 2cv(1)$ can become positive only when the off-diagonal terms becomes dominating. Hence, strong interdependence between elements of the state vector is essential to generate hump dynamics.

The same argument also holds, almost in parallel, for the variance of multiperiod changes of $q(X_t)$,

$$\begin{aligned} v(k) &\equiv \text{Var}(q(X_{t+k}) - q(X_t)) \\ &= 4tr \left(\left(A - (\Phi^k)^\top A \Phi^k \right) V A V \right) + 2\mathbf{b}^\top (I - \Phi^k) V \mathbf{b}. \end{aligned} \quad (\text{A3})$$

For the term structure to exhibit a hump shape, we need the annualized variance to increase initially with periods k , or $v(2) - 2v(1) > 0$. Yet, for a one-factor model, we have again

$$v(2) - 2v(1) = -4(1 - \phi^2)^2 V^2 A_n^2 - 2(1 - \phi)^2 b_n^2 V < 0.$$

Hence, the initial term structure is also downward sloping. No hump can be generated from a one-factor model. And the same conclusion holds for completely independent multifactor models.

References

- Ahn, Dong-Hyun, Dittmar, Robert F., and Gallant, A. Ronald (2002), Quadratic term structure models: Theory and evidence, *Review of Financial Studies*, forthcoming.
- Andrews, Donald (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817–858.
- Backus, David, Foresi, Silverio, Mozumdar, Abon, and Wu, Liuren (2001a) Predictable changes in yields and forward rates, *Journal of Financial Economics* **59**, 281–311.
- Backus, David, Foresi, Silverio, and Telmer, Chris (2001b) Affine term structure models and the forward premium anomaly, *Journal of Finance* **56**, 279–304.
- Beaglehole, D. R. and Tenney, M. S. (1991) General solution of some interest rate-contingent claim pricing equations, *Journal of Fixed Income* **1**, 69–83.
- Beaglehole, D. R. and Tenney, M. S. (1992) A nonlinear equilibrium model of term structures of interest rates: Corrections and additions, *Journal of Financial Economics* **32**, 345–454.
- Bekaert, Geert and Urias, Michael (1996) Diversification, integration, and emerging market closed-end funds, *Journal of Finance* **51**, 835–869.
- Bekaert, Geert, Hodrick, Robert, and Marshall, David (1997) On biases in tests of the expectation hypothesis of the term structure of interest rates, *Journal of Financial Economics* **44**, 309–348.
- Bliss, Robert (1997) Testing term structure estimation methods, *Advances in Futures and Options Research* **9**, 191–231.
- Campbell, John (1995) Some lessons from the yield curve, *Journal of Economic Perspectives* **9**, 129–152.
- Campbell, John and Shiller, Robert (1991) Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* **58**, 495–514. Special Issue: The Econometrics of Financial markets.
- Chen, Ren-raw and Scott, Louis (1993) Maximum likelihood estimation of a multifactor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* **3**, 14–31.
- Constantinides, G. M. (1992) A theory of the nominal term structure of interest rates, *Review of Financial Studies* **5**, 531–552.
- Cox, John C., Ingersoll, Jonathan E., and Ross, Stephen R. (1985) A Theory of the term structure of interest rates, *Econometrica* **53**, 385–408.
- Dai, Qiang and Singleton, Kenneth (2000) Specification analysis of affine term structure models, *Journal of Finance* **55**, 1943–1978.
- Dai, Qiang and Singleton, Kenneth (2002a) Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* **63**, 415–441.
- Dai, Qiang and Singleton, Kenneth (2002b) Term structure dynamics in theory and reality, Working paper Graduate School of Business, Stanford University.
- Duffee, Gregory R. (2002) Term premia and interest rate forecasts in affine models, *Journal of Finance* **57**, 405–443.
- Duffie, Darrell and Kan, Rui (1996) A yield-factor model of interest rates, *Mathematical Finance* **6**, 379–406.
- Duffie, Darrel and Singleton, Kenneth (1997) An econometric model of the term structure of interest rate swap yields, *Journal of Finance* **52**, 1287–1322.
- Dybvig, Philip H., Ingersoll, Jonathan E., and Ross, Stephen A. (1996) Long forward and zero-coupon rates can never fall, *Journal of Business* **69**, 1–25.

- El Karoui, N., Myneni, R., and Viswanathan, R. (1992) Arbitrage Pricing and Hedging of Interest Rate Claims With State variables: I theory, working paper University of Paris.
- Evans, Martin and Lewis, Karen (1994) Do stationary risk premia explain it all? Evidence from the term structure, *Journal of Monetary Economics* **33**, 285–318.
- Fan, Rong, Gupta, Anurag, and Ritchken, Peter (2001) On pricing and hedging in the Swaption Market: How many factors, really?, manuscript Case Western Reserve University Ohio.
- Frachot, Antoine and Lesne, Jean-Philippe (1994) Expectation hypothesis and stochastic volatilities, manuscript Banque de France.
- Gallant, A. Ronald and Tauchen, George E. (1996) Which moments to match?, *Econometric Theory* **12**, 657–681.
- Hamilton, James D. (1994) *Time Series Analysis*, Princeton University Press Princeton, New Jersey.
- Heidari, Massoud and Wu, Liuren (2001) Are interest rate derivatives spanned by the term structure of interest rates?, manuscript Fordham University New York.
- Holmquist, Björn (1996) Expectations of products of quadratic forms in normal variates, *Stochastic Analysis and Applications* **14**, 149–164.
- Jamshidian, F. (1996) Bond, futures and option valuation in the quadratic interest rate model, *Applied Mathematical Finance* **3**, 93–115.
- Johnson, N. L. and Kotz, S. (1970) *Distributions in Statistics: Continuous Univariate Distributions* 2. Houghton Mifflin Co. Boston, chap. 29.
- Kathri, C. G. (1980) Quadratic forms in normal variables, in P. R. Krishnaiah, ed.: *Handbook of Statistics*, North-Holland Publisher Co., Amsterdam.
- Knez, Peter J., Litterman, Robert, and Scheinkman, Jose (1994) Explorations into factors explaining money market returns, *Journal of Finance* **49**, 1861–1882.
- Leippold, Markus and Wu, Liuren (2002) Asset pricing under the quadratic class, *Journal of Financial and Quantitative Analysis* **37**, 271–295.
- Litterman, Robert and Scheinkman, Jose (1991) Common factors affecting bond returns, *Journal of Fixed Income* **1**, 54–61.
- Longstaff, Francis A. (1989) A nonlinear general equilibrium model of the term structure of interest rates, *Journal of Financial Economics* **23**, 195–224.
- Mathai, A. M. and Provost, S. B. (1992) *Quadratic Forms in Random Variables*, Marcel Dekker, New York.
- Newey, Whitney K. and West, Kenneth (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.
- Roberds, William and Whiteman, Charles (1999) Endogenous term premia and anomalies in the term structure of interest rates: Explaining the predictability smile, *Journal of Monetary Economics* **44**, 555–580.
- Rogers, L. C. G. (1997) The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* **7**, 157–176.
- Singleton, Kenneth (2001) Estimation of affine Asset pricing models using the empirical characteristic function, *Journal of Econometrics* **102**, 111–141.
- Singleton, Kenneth and Umantsev, Len (2001) Pricing coupon-bond options and swaptions in affine term structure models, manuscript Stanford University Stanford, CA.

