Nonlinearity in the Term Structure*

Dong Heon Kim

Department of Economics, Korea University

Revised: December 2008

JEL *Classification*: E43, C13, G12

Keywords: Nonlinearity, Term structure, Affine model, Quadratic model, Flexible nonlinear model.

* Financial support from the Economic and Social Research Council (UK) under grant no. RES-000-22-0337 is gratefully acknowledged. This paper has benefited considerably from the comments and suggestions of James D. Hamilton. The author thanks Keith Blackburn, Ian Garrett, Stuart Hyde, Denise Osborn, Marianne Sensier and seminar participants in the University of Manchester, Korea University (Seoul, Korea), the Bank of Korea, and participants at the 2004 Econometic Society of European Meeting, and 2004 Far Eastern Meeting of Econometric Society for their constructive comments. The author is also grateful for the supports received from EEXCL in UCSD Economics Department and acknowledges financial support from the Korea University, the University of Manchester and the Manchester School Visiting grant. The author remains responsible for any errors or omissions.

Correspondence to: Dong Heon Kim, Department of Economics, Korea University, 5-1 Anam-dong, Seongbuk-Gu, Seoul, 136-701, South Korea, TEL: +82-2-3290-2226, E-mail: dongkim@korea.ac.kr

Nonlinearity in the Term Structure

Abstract

This paper investigates the nature of nonlinearities in the term structure using the flexible approach to nonlinear inference. The paper reports clear evidence of nonlinearity, in contrast to the affine term structure model and consistent with recent claims in the literature. We find that there is a threshold effect of volatility on the interest rate but this effect does not capture the entire nature of the nonlinearity. The quadratic term structure model recently proposed performs better for capturing the nonlinearity than the threshold model but the former model seems to miss some aspect of nonlinearity for short-term rates. However, our flexible nonlinear model which incorporates the threshold effect and the convexity of volatility into the quadratic model, generally performs well for all interest rates. The paper suggests that this model is a promising representation of nonlinearities and out-of-sample forecasts support the claim of nonlinearities.

The most commonly used term structure models are affine term structure models (ATSMs) in which the yield or log bond price is an affine function of the underlying state variables. Examples include Vasicek (1977), Cox, Ingersoll, and Ross (CIR; 1985), Brennan and Schwartz (1982), Schaefer and Schwartz (1984), Longstaff and Schwartz (1992), and De Jong (2000), among others. In a single-factor affine yield model, the short-term interest rate is the single factor driving movements in the term structure. Longstaff and Schwartz (1992) suggest that the two factors of the term structure model are the short-term interest rate and the instantaneous variance of changes in the short-term interest rate and these two factors summarize the state of the economy. De Jong (2000) provides evidence that a three-factor affine model with correlated factors is able to provide an adequate fit of the cross-section and the dynamics of the term structure. Generally, Duffie and Kan (1996) characterize the complete class of the affine model and provide necessary and sufficient conditions on the stochastic model for these affine representations. Their characterization has served as a general framework for more systematic study of model design. Dai and Singleton (2000) show how generalizations of earlier affine models lead to substantial improvements in their ability to account for dollar swap rates.[1]

However, despite a number of desirable features of these models of the affine class and their widespread use, the recent literature has challenged these models. In particular, empirical evidence on the diffusion process of the interest rate has been at odds with the affine class models. Longstaff (1989) derives an alternative closed-form general equilibrium model of the term structure within the CIR framework, in which discount bond yields are nonlinear functions of the risk-free interest rate, and shows that the double square root model that allows yield nonlinearity is more successful in capturing the level and variation of six- to twelve-month Treasury bill yields than the square root model during the 1964-1986 period.[2] Aït-Sahalia (1996a,b) shows, using a nonparametric approach, that there is evidence of nonlinearities in both the drift and diffusion functions. Using their semiparametric factor model of interest rates, Ghysel and Ng (1998) found that the empirical evidence does not support the restrictions imposed by affine models. Duffee (2002) shows that the standard class of affine models produces poor forecasts of future Treasury yields.

Ahn and Gao (1999) propose an alternative single-factor term structure model that is consistent with the dynamics of the interest rate process documented in the nonparametric literature. In this model, the specification for the diffusion is $r^{1.5}$ as suggested in empirical studies such as Chan et al. (1992) and Stanton (1997), and the drift is specified as a quadratic function instead of a linear one, so that it exhibits substantial nonlinear mean-reverting behaviour when the interest rate is above its long-run mean. They found that this model captures the nonlinearities in the drift and the diffusion and performs better than the affine model in explaining the stochastic process of the short rate. Ahn, Dittmar and Gallant (2002) point out that affine term structure models have a theoretical drawback that hampers their empirical performance and there is some omitted nonlinearity in these models. They propose a quadratic term structure model (QTSM) in which the yield on a bond is a quadratic function of underlying state variables and develop a comprehensive QTSM, which is maximally flexible and thus encompasses all features of the diverse models in the literature. They show that the QTSM outperforms the ATSMs in explaining historical bond price behaviour in the United States.

Boudoukh, Richardson, Stanton, and Whitelaw (1999) extend Stanton's (1997) nonparametric estimators for the underlying single-factor continuous-time process to a multivariate setting and provide for the non-parametric estimation of the drift and volatility functions of multivariate stochastic differential equations. Their approach is that rather than choosing the model parameterization, their processes are generated from the data using approximation methods for multifactor continuous-time Markow process. They present a general, nonlinear version of existing multifactor models, such as Longstaff and Schwartz (1992). They find that the volatility of interest rates is increasing in the level of interest rates only for sharply upward sloping term structures, and show, for the analysis of term premiums, that both the level of the short rate and the degree of interest volatility, and the underlying nonlinearities of their model, play an important role in fixed-income pricing.

Thus, recent claims for nonaffine family of term structure models cast doubt on the validity of ATSMs and whether discount bond yields are linear or nonlinear functions of the underlying state variables is an important question for evaluating term structure models.

Furthermore, if there is a nonlinear relationship between bond yields and state variables, what is the nature of the nonlinearity? From an unbounded universe of alternative nonlinear specifications, how does one decide which nonlinear specification is the right one to use?

This paper applies the flexible approach to nonlinear inference, recently developed by Hamilton (2001), to address these questions. This approach provides a valid test of the null hypothesis of linearity against a broad range of alternative nonlinear models, consistently estimates what the nonlinear function looks like, and makes a formal comparison of alternative nonlinear models. Hamilton (2003) and Kim, Osborn and Sensier (2005) show that this methodology is very useful for characterizing the nonlinear relationship between oil price changes and GDP growth and nonlinearities in the U.S. Federal Reserve System's monetary policy respectively.

Following Longstaff and Schwartz (1992), we consider the risk-free rate and its volatility as two state variables and develop a flexible two-factor model in which the yield is a unrestricted function of these two variables. While this approach is similar with Boudoukh et al. (1999) in terms of addressing general nonlinear term structure models with two factors–the instantaneous rate and its volatility, this paper provides a test statistic and examines nonlinear nature by using a flexible parametric framework for investigating nonlinear relations that combines the advantages of the parametric and nonparametric approaches.

The results of the linearity test against nonlinear alternatives suggest that there is clear evidence of nonlinearity, in contrast to the ATSMs and consistent with recent claims in the literature. While the relationship between the risk-free rate and the short-term rate seems to be linear, the dependence on the conditional variance of the risk-free rate seems to be nonlinear for all interest rates. We find there is a threshold effect of volatility on the interest rate, but this threshold effect seems not to capture the entire nature of nonlinearity in the term structure. More formal statistical comparison of the nonlinear dynamics implied by alternative specifications with what appears in the data from the flexible inference procedure used in this paper suggests that the QTSM performs better than the threshold model but the former seems to miss some aspect of nonlinearity for short-term rates. However, our flexible nonlinear model, which incorporates the threshold effect and the convexity of volatility into the quadratic term structure model, generally performs well for all interest rates and an out-of-sample forecasts support the claim of nonlinearities.

The plan of the paper is as follows. Section II reviews the QTSM and proposes a flexible two-factor model in which a discount bond yield is an unrestricted function of two state variables (factors). Section III briefly introduces the methodology applied in this paper. Empirical results are presented and suggested nonlinear specifications are evaluated in Section IV. Conclusions are offered in Section V.

1 A flexible two-factor term structure model

In this section, we review the QTSM developed by Ahn et al.(2002) and develop a flexible two-factor model for describing the bond yields. For this end, we follow Longstaff and Schwartz (1992) and thus assume that the short-term interest rate and its volatility are the most important factors and the bond yield can be described by these state variables. Thus, we consider the two-factor QTSM which is a two-state variables version of the N-factor QTSM of Ahn et al. (2002). Specifically, under a given complete probability space $(\Omega, \mathcal{F}, P)$ and the augmented filtration $F = \{\mathcal{F}_t : t \geq 0\}$, let $P(t, n)$ denote the price at time t of a zero-coupon bond maturing at time $t+n$, and let $M(t, t+1)$ denote a pricing kernel satisfying

$$P(t, n) = E_t^P[P(t+1, n-1)M(t, t+1)], \tag{1}$$

where E_t^P denotes expectation conditional on the information at time t under the physical probability measure P and $P(t+1, n-1)$ is the price of the $n-1$ period bond at time $t+1$. $M(t, t+n)$ is the stochastic discount factor, which discounts payoffs at time $t+n$ into time t value under the stochastic economy. Under the assumption of a complete market as in Harrison and Kreps (1979) and Harrison and Pliska (1981), there is a unique probability measure Q under which all money market scaled bond prices follow a martingale and we can write the stochastic discount factor, $M(t, t+n) = \frac{M(t+n)}{M(t)}$

$$M(t, t+n) = \frac{G(t)}{G(n)} N(t, t+n) = [\exp(-\int_t^n r_s ds)] N(t, t+n), \tag{2}$$

where $G(n)$ denotes a money market account, $G(t) = \exp(\int_0^t r(s)ds)$, r_s denotes the risk-free instantaneous rate at time s, and $N(t, t+n) = \frac{dQ(t,n)}{dP(t,n)}$ is called the Radon-Nikodym derivative. Ahn and Gao (1999) and Ahn et al. (2002) assume that the time-series process of the stochastic discount factor, $M(t)$, is represented as the stochastic differential equation (SDE)

$$\frac{dM(t)}{M(t)} = -r_t dt + \mathbf{1}_2' \mathbf{S}_t d\mathbf{W}_t, = -r_t dt + \mathbf{1}_2'[(\mathbf{a} + \mathbf{b}\mathbf{x}_t) \odot d\mathbf{W}_t] \tag{3}$$

where $\mathbf{S}_t$ is a 2×2 diagonal matrix with the ith diagonal element given by $[\mathbf{S}_t]_{ii} = a_i + \mathbf{b}_i'\mathbf{x}_t, \mathbf{b}_i = (b_{1i}, b_{2i})', \mathbf{a} = (a_1, a_2)', \mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2)'$, $\mathbf{1}_2$ is 2×1 ones vector, $\odot$ is an element by element multiplication, $\mathbf{x}_t = (x_{1t}, x_{2t})'$ is the 2×1 vector of two state variables, and $\mathbf{W}_t$ is a 2-dimensional vector of standard Wiener processes which are mutually independent. Equation (3) implies that $-r_t$ is the drift and the diffusion is an affine function of the state variables.

They also assume that the instantaneous short term rate is a quadratic function of the two state variables:

$$r_t = \delta_0 + \boldsymbol{\delta}_1'\mathbf{x}_t + \mathbf{x}_t{}'\boldsymbol{\Phi}\mathbf{x}_t, \tag{4}$$

where δ_0 is a constant, $\boldsymbol{\delta}_1$ is a 2×1 vector and $\boldsymbol{\Phi}$ is a 2×2 matrix of constants. To ensure the nonnegativity of the nominal interest rate, it is assumed that $\delta_0 - \frac{1}{4}\boldsymbol{\delta}_1'\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}_1 \geq 0$ and $\boldsymbol{\Phi}$ is a positive definite matrix. While the short rate is an affine function of a vector of state variables in the ATSM, it is a generalized positive semidefinite quadratic form in the equation (4). Note that the lower bound on the short rate is $\delta_0 - \frac{1}{4}\boldsymbol{\delta}_1'\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}_1$ when $\mathbf{x}_t = -\frac{1}{2}\boldsymbol{\Phi}^{-1}\boldsymbol{\delta}_1$.

Moreover, they assume that the dynamics of the state variables $\mathbf{x}_t$ are governed by

$$d\mathbf{x}_t = (\boldsymbol{\xi} + \boldsymbol{\theta}\mathbf{x}_t)dt + \boldsymbol{\Sigma} d\mathbf{Z}_t, \tag{5}$$

where $\boldsymbol{\xi}$ is an 2×1 constant vector, $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ are 2-dimensional square matrices, and $\mathbf{Z}_t$ is a 2-dimensional vector of standard Wiener processes that are mutually independent. It is further assumed that $\boldsymbol{\theta}$ is diagonalizable and has negative real components of eigenvalues to ensure the stationarity of the state variables. The correlation matrix between $d\mathbf{W}_t$ and $d\mathbf{Z}_t$, $Cov_t(d\mathbf{W}_t, d\mathbf{Z}_t)$, is denoted by $\boldsymbol{\Gamma}$, a 2×2 matrix of constants. Thus the time-series process of the state variables is represented as a Gaussian process with steady-state long-term means of $-\boldsymbol{\theta}^{-1}\boldsymbol{\xi}$, and covariance matrix $\boldsymbol{\Sigma\Sigma}'$.

Using the transitional densities for the state variables, Ito's lemma, and Girsanov Theorem, they drive following equation for the n-period bond yield, y_{nt}:

$$y_{nt} = \frac{1}{n}[-A(n) - \mathbf{B}(n)'\mathbf{x}_t - \mathbf{x}_t{}'\mathbf{C}(n)\mathbf{x}_t], \tag{6}$$

where, $A(n), \mathbf{B}(n)$ and $\mathbf{C}(n)$ with the initial conditions $A(0) = 0, \mathbf{B}(0) = \mathbf{0}_2$, and $\mathbf{C}(0) = \mathbf{0}_{2\times 2}$, satisfy the ordinary differential equations (ODEs)

$$\begin{aligned}
\frac{d\mathbf{C}(n)}{dn} &= 2\mathbf{C}(n)\boldsymbol{\Sigma\Sigma}'\mathbf{C}(n) + (\mathbf{C}(n)(\boldsymbol{\theta} - \boldsymbol{\kappa}_1) + (\boldsymbol{\theta} - \boldsymbol{\kappa}_1)'\mathbf{C}(n)) - \boldsymbol{\Phi}, \\
\frac{d\mathbf{B}(n)}{dn} &= 2\mathbf{C}(n)\boldsymbol{\Sigma\Sigma}'\mathbf{B}(n) + (\boldsymbol{\theta} - \boldsymbol{\kappa}_1)'\mathbf{B}(n) + 2\mathbf{C}(n)(\boldsymbol{\xi} - \boldsymbol{\kappa}_0) - \boldsymbol{\delta}_1, \\
\frac{dA(n)}{dn} &= tr[\boldsymbol{\Sigma\Sigma}'\mathbf{C}(n)] + \frac{1}{2}\mathbf{B}(n)'\boldsymbol{\Sigma\Sigma}'\mathbf{B}(n) + \mathbf{B}(n)'(\boldsymbol{\xi} - \boldsymbol{\kappa}_0) - \boldsymbol{\delta}_0,
\end{aligned}$$

$\boldsymbol{\kappa}_0 = -\boldsymbol{\Sigma\Gamma}\mathbf{a}$, and $\boldsymbol{\kappa}_1 = -\boldsymbol{\Sigma\Gamma}\mathbf{b}$. Equation (6) implies that the bond yield is a nonlinear (specifically, a quadratic) function of the state variables. This QTSM nests the affine-factor model when $\mathbf{C}(\tau) = \mathbf{0}_{2\times 2}$ and other nonlinear term structure models of Longstaff (1989), and Beaglehole and Tenny (1992).

Although the QTSM accommodates characteristics that can potentially overcome the shortcomings of the ATSM and has the potential to capture omitted nonlinearities, it is driven from the specific assumption that the instantaneous short term rate is a quadratic function of the two state variables. In reality, however, since there can be an unbounded universe of alternative nonlinear specifications, it is very hard to say which specification is the right one to use without examining the data. Hence, a more general approach is to leave the functional form unrestricted and to infer the functional form from the data. Thus, a more general equation for the n-period bond yield is:

$$y_{nt} = f(\mathbf{x}_t),$$

where $f(.)$ is unrestricted and the functional form is unknown. We seek the expectation of scalar y_{nt} conditional on the state-variable vector $\mathbf{x}_t$, $E(y_{nt}|\mathbf{x}_t) = \mu(\mathbf{x}_t)$. The regression of the n-period bond yield on the state variables with measurement error is

$$y_{nt} = \mu(\mathbf{x}_t) + \varepsilon_t, \tag{7}$$

where ε_t is i.i.d. with mean zero and independent of both $\mu(.)$ and $\mathbf{x}_\tau$ for $\tau = t, t-1, ..., 1$. If we identify the state variables, we can use the flexible nonlinear technique of Hamilton (2001) to estimate the bond yield regression (7) using maximum likelihood or Bayesian method. For example, under the model of Longstaff and Schwartz (1992), the level and the variance of risk-free rate are two the factors which drive the bond yields and $\mathbf{x}_t$ in equation (7) is the 2×1, vector $\begin{pmatrix} y_{1t} \\ v^2_{t+1|t} \end{pmatrix}$, where y_{1t} is the risk-free short term rate and $v^2_{t+1|t}$ is the conditional variance of $y_{1,t+1}$ given information at time t. In the following section, we describe the basic technique.

2 A flexible approach to nonlinear inference

Hamilton (2001) proposes a new framework that combines the advantages of non-parametric and parametric methods. While the procedure does not assume any specific parametric functional form for the conditional mean function, it has parameters to be estimated by maximum likelihood or Bayesian methods for the unknown conditional mean function and performs inference and hypothesis testing based on classical econometric theory. Consider a nonlinear regression model of the form

$$y_t = \mu(\mathbf{x}_t) + \varepsilon_t, \tag{8}$$

where y_t is a scalar dependent variable, $\mathbf{x}_t$ is a k-dimensional vector of explanatory variables, and ε_t is an error term with mean zero and is independent of $\mathbf{x}_t$ and of lagged values of y_{t-j} or $\mathbf{x}_{t-j}$. Since the form of the function $\mu(\cdot)$ is unknown, we seek to represent it using a flexible class. In our term structure application below, $y_t = y_{nt}$, $\mathbf{x}_t = (y_{1t}, v^2_{t+1|t})$ for the n-period bond yield. Following Hamilton (2001), we view $\mu(.)$ as the outcome of a random field. Specifically, the value of the function $\mu(\mathbf{x}_t)$ at $\mathbf{x}_t = \boldsymbol{\tau}$ is treated as being a Gaussian random variable with mean equal to the linear component $\alpha_0 + \boldsymbol{\alpha}'\boldsymbol{\tau}$ and variance λ^2, where $\alpha_0, \boldsymbol{\alpha}$, and λ are population parameters to be estimated. In the special case of $\lambda = 0$, then $\mu(\mathbf{x}_t)$ is fixed and (8) becomes the usual linear regression model. In general, the parameter λ measures the overall extent of nonlinearity.

The basic idea of the method is that nonlinearity implies the values for $\mu(\mathbf{x}_t)$ and $\mu(\mathbf{x}_s)$ will be positively correlated for periods t and s whenever the vectors $\mathbf{x}_t$ and $\mathbf{x}_s$ are close to each other. The key is then parameterizing this correlation based on the distance measure $h_{st} = (1/2)\left[\sum_{i=1}^{k} g_i^2 (x_{is} - x_{it})^2\right]^{1/2}$ where x_{it} denotes the ith element of the vector $\mathbf{x}_t$ and $g_1, g_2, ..., g_k$ are k additional parameters to be estimated. Hamilton proposes that $\mu(\mathbf{x}_s)$ should be uncorrelated with $\mu(\mathbf{x}_t)$ if $\mathbf{x}_s$ is sufficiently far away from $\mathbf{x}_t$. More precisely,

$$E\{[\mu(\mathbf{x}_s) - \alpha_0 - \boldsymbol{\alpha}'\mathbf{x}_s][\mu(\mathbf{x}_t) - \alpha_0 - \boldsymbol{\alpha}'\mathbf{x}_t]\} = 0 \quad \text{if } h_{st} > 1 \tag{9a}$$

However, when $0 \leq h_{st} \leq 1$, this correlation should increase as h_{st} decreases, with the correlation going to unity as h_{st} goes to zero. In our context where there are two nonlinear explanatory variables ($k = 2$), then the correlation is assumed to be given by

$$Corr(\mu(\mathbf{x}_s), \mu(\mathbf{x}_t)) = H_2(h_{st}) \;\; \text{if } 0 \leq h_{st} \leq 1 \tag{9b}$$

where

$$H_2(h_{st}) = 1 - (2/\pi)[h_{st}(1 - h_{st}^2)^{1/2} + \sin^{-1}(h_{st})]. \tag{10}$$

For the general specification and rationalization of this correlation, see Lemma 2.1 and Theorem 2.2 in Hamilton (2001). It should be emphasized that $H_k(.)$ does not assume any parametric form for the functional relation $\mu(.)$ itself, but rather it parameterizes the correlation between pairs of random outcomes $\mu(\mathbf{x}_s)$ and $\mu(\mathbf{x}_t)$. The coefficient g_i determines the extent to which variation in the i-th element of $\mathbf{x}_t$ contributes to nonlinear variation in $\mu(\mathbf{x}_t)$. For g_i small, the value of $\mu(\mathbf{x}_t)$ changes little when x_i changes, with $g_i = 0$ implying linearity of $\mu(\mathbf{x}_t)$ with respect to x_i.

Prior to estimation it is appropriate to determine whether nonlinearity exists by testing $H_0 : \lambda = 0$. As is usual in nonlinear modelling, certain parameters are unidentified under the null of linearity. In the present context, this applies to $g_1, g_2, ..., g_k$. For the purpose of the nonlinearity test, Hamilton suggests that the lack of identification can be avoided by setting $g_i = 2\left[k\left(T^{-1}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)^2\right)\right]^{-1/2}$, thereby scaling in terms of the individual sample standard deviations and the number of explanatory variables. Then, for T sample observations, the $(T \times T)$ matrix $\mathbf{H}$ of correlations can be formed, with the row s, column t element $H_k\{h_{st}\}$ given in (10) when $k = 2$ and $0 \leq h_{st} \leq 1$, or zero when $h_{st} > 1$. The Lagrange multiplier (LM) test of the null hypothesis can be obtained by using the residuals from an OLS linear regression of y_t on $(1, \mathbf{x}_t')'$. Denoting the OLS residual vector by $\widehat{\boldsymbol{\varepsilon}}$ and the OLS squared standard error as $\widetilde{\sigma}^2 = (T-k-1)^{-1}\widehat{\boldsymbol{\varepsilon}}'\widehat{\boldsymbol{\varepsilon}}$, and the $(T \times T)$ projection matrix $\mathbf{M} = \mathbf{I}_T - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ where $\mathbf{X}$ is a $(T \times (1+k))$ matrix whose tth row is given by $(1, \mathbf{x}_t')$ and $\mathbf{I}_T$ is the $(T \times T)$ identity matrix, the test statistic is

$$\nu^2 = \frac{[\widehat{\boldsymbol{\varepsilon}}'\mathbf{H}\widehat{\boldsymbol{\varepsilon}} - \widetilde{\sigma}^2 tr(\mathbf{MHM})]^2}{\widetilde{\sigma}^4(2tr\{[\mathbf{MHM} - (T-k-1)^{-1}\mathbf{M}tr(\mathbf{MHM})]^2\})}. \tag{11}$$

Under the null hypothesis of linearity, ν^2 has an asymptotic $\chi^2(1)$ distribution. Dahl's (2002) Monte Carlo investigations suggest that this test has good size and power properties against a variety of nonlinear alternatives.

In the presence of nonlinearity, Hamilton writes (8) as

$$y_t = \alpha_0 + \boldsymbol{\alpha}'\mathbf{x}_t + \lambda m(\mathbf{x}_t) + \varepsilon_t = \alpha_0 + \boldsymbol{\alpha}'\mathbf{x}_t + u_t, \tag{12}$$

where $m(.)$ is the realization of a scalar-valued Gaussian random field with mean zero, unit variance and covariance function given by (9a) and (9b). Assuming that the regression disturbance ε_t is i.i.d. $N(0, \sigma^2)$, the composite disturbance $u_t = \lambda m(\mathbf{x}_t) + \varepsilon_t$ is also Gaussian. With independence between $\mathbf{x}_t'$ and ε_t, this specification implies a GLS regression model of the form

$$\mathbf{y}|\mathbf{X} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{P}_0 + \sigma^2\mathbf{I}_T)$$

where $\mathbf{y} = (y_1, y_2, ..., y_T)'$, $\boldsymbol{\beta}$ is the $(1+k)$-dimensional vector $(\alpha_0, \boldsymbol{\alpha}')'$, and $\mathbf{P}_0$ is a $(T \times T)$ matrix whose row s, column t element is given by $\lambda^2 H_k(h_{st})\delta_{[h_{st}<1]}$ with h_{st} is defined above, and the function $H_k(.)$ is specified in (10) for the case $k = 2$.

In addition to the linear regression parameters $(\alpha_0, \boldsymbol{\alpha})$ and σ^2, parameters to be estimated are the variance of the nonlinear regression error, λ^2, which governs the overall importance of the nonlinear component, and the parameters $(g_1, g_2, ..., g_k)$ determining the variability of the nonlinear component with respect to each explanatory variable in $\mathbf{x}_t$. As the above discussion implies, estimation and inference can be achieved by a GLS Gaussian regression. However, Hamilton (2001) also describes the use of numerical Bayesian methods for the evaluation of the posterior distribution of any statistics of interest. The optimal inference of the value of the unobserved function $\mu(\mathbf{x}^*)$ at an arbitrary point $\mathbf{x}^*$ is given by

$$\widehat{\mu}(\mathbf{x}^*) = \alpha_0 + \alpha'\mathbf{x}^* + \mathbf{q}'(\mathbf{P}_0 + \sigma^2\mathbf{I}_T)^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \tag{13}$$

where the $(T \times 1)$ vector $\mathbf{q}$ has tth element $\lambda^2 H_k(h_t^*)\delta_{[h_t^*<1]}$ for $h_t^* = (1/2)\left[\sum_{i=1}^{k} g_i^2(x_{it} - x_i^*)^2\right]^{1/2}$, in which x_{it} denotes the ith element of $\mathbf{x}_t$ and x_i^* denotes the ith element of $\mathbf{x}^*$. Hamilton shows that $\widehat{\mu}(\mathbf{x}^*)$ converges to the true value $\mu(\mathbf{x}^*)$ for any $\mu(.)$ from a broad class of continuous functions. This permits the calculation of confidence intervals, using (13) along with its known standard error for each given parameter vector in conjunction with values of $\alpha_0, \boldsymbol{\alpha}, \sigma, \lambda$, and $\mathbf{g} = (g_1, g_2, ..., g_k)'$ generated from their posterior distributions, and examining the resulting distribution of inferences.

From a Monte Carlo investigation, Dahl (2002) shows that in many situations Hamilton's random field based estimator is substantially more accurate than the non-parametric spline smoother. He also finds that the procedure is useful in finite samples for characterizing a wide range of nonlinear time series models.

3 Empirical results

In this section, as in Longstaff and Schwartz (1992), we consider the risk-free rate and it's volatility as two factors. Following Hamilton's (2001) methodology described in the previous section, we estimate a flexible two-factor model:

$$y_{nt} = \mu(\mathbf{x}_t) + \varepsilon_t, \tag{14}$$

$$\mu(\mathbf{x}_t) = \alpha_0 + \boldsymbol{\alpha}_1'\mathbf{x}_t + \lambda m(\mathbf{g} \odot \mathbf{x}_t), \tag{15}$$

where $\mathbf{x}_t = (y_{1t}, v^2_{t+1|t})'$ is a 2×1 vector. We report estimates of equations (14) and (15) for various discount bond yields. Then, we evaluate alternative specifications of term structure models using a formal statistical basis for the comparison of the nonlinear dynamics. Finally, we consider the root-mean-squared-error (RMSE) for comparing the forecast performance of nonaffine models with that of the affine model.

3.1 Data

In order to investigate the implications of our flexible term structure model, we consider five zero-coupon bond yields; 3- and 6-month and 1-, 5- and 10-year (3M, 6M, 1Y, 5Y, 10Y respectively). These data are sampled at a monthly frequency and cover the period from February 1959 to December 1999. All data, except the 10-year Treasury bond, are taken from the Center for Research in Security Prices (CRSP). The yield data for the Treasury 10-year constant maturity bond were downloaded from the Federal Reserve Board of Governors. We assume that the 1-month interest rate is the risk-free rate, as usual in the literature, and consider its volatility as the conditional variance of the rate. To describe time variation in the volatility of interest rates, we use the generalized autoregressive conditional heteroskedasticity (GARCH) framework. Following Brenner, Harjes, and Kroner (1996) and Hamilton and Kim (2002), we model the conditional variance of the risk-free rate as a function of both the interest rate level and previous squared interest rate innovations:

$$y_{1t} = c + \phi y_{1,t-1} + \varepsilon_t, \varepsilon_t|\Omega_{t-1} \sim N(0, v^2_{t|t-1}), \tag{16}$$

$$v^2_{t|t-1} = \omega_0 + \omega_1\varepsilon^2_{t-1} + \omega_2 v^2_{t-1|t-2} + \omega_3 y_{1,t-1}. \tag{17}$$

Maximum likelihood estimates of equations (16) and (17) are as follows, with conventional standard errors in parentheses:

$$y_{1t} = \underset{(0.0367)}{0.1411} + \underset{(0.0085)}{0.9742} y_{1,t-1} + \varepsilon_t, \tag{18}$$

$$\upsilon^2_{t|t-1} = \underset{(0.0064)}{-0.0224} + \underset{(0.0477)}{0.3141} \varepsilon^2_{t-1} + \underset{(0.0444)}{0.6444} \upsilon^2_{t-1|t-2} + \underset{(0.0023)}{0.0093} y_{1,t-1}. \tag{19}$$

We then used the fitted values $\hat{\upsilon}^2_{t+1|t}$ for the conditional variance of the risk-free rate in regressions explaining the bond yields.

3.2 Estimates of a flexible two-factor model

Under the assumption that the data were generated from (14) and (15), we have the following regression:

$$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \hat{\upsilon}^2_{t+1|t} + \sigma[\zeta m(g_1 y_{1t}, g_2 \hat{\upsilon}^2_{t+1|t}) + \upsilon_t], \tag{20}$$

where $\hat{\upsilon}^2_{t+1|t}$ is the conditional variance of $y_{1,t+1}$ given information at time t. The innovation ε_t in (12) is written here as σ times υ_t, and the parameter λ in (12) is written as σ times ζ. Table 1 reports the test statistic ν^2 of the null hypothesis of linearity. The test statistics for 3M, 6M, 1Y, 5Y, and 10Y are 17.87, 33.88, 72.65, 53.42 and 26.94 respectively and a large value for a $\chi^2(1)$ variable implies overwhelming rejection of the null hypothesis that the relation between bond yields and the two factors, the risk-free rate and its volatility, is linear.

Insert Table 1

Bayesian posterior estimates and their standard errors for the flexible nonlinear alternative are reported in Table 2. The risk-free rate and its volatility exert positive effects on bond yields, though only the coefficients on the risk-free rate are statistically significant at the 1% level. When we consider a hypothesis of linearity for the risk-free rate and its volatility taken individually, we reject the null in relation to both variables for only 3-month rate. However, the t-statistic for $\zeta = 0$ in all five bond yields is highly significant and it is consistent with the results of the LM tests, implying that collectively the nonlinear component makes a highly significant contribution in all bond yields. Interestingly, as time to maturity increases, the value of estimated λ (σ times ζ) increases.

Insert Table 2

Given any particular values for the vectors $\mathbf{g}$, $\boldsymbol{\alpha}$, and the scalar λ, we can calculate the value of $H_2(.)$ as associated with any pair of observations on $\mathbf{x}_t$ and $\mathbf{x}_s$, the row t, column s element of the matrix $\mathbf{P}_0$ as $\lambda^2 H_2(.)$ and a value for (13) for any $\mathbf{x}^*$ of interest, which represents the econometrician's inference as to the value of the conditional mean $\mu(\mathbf{x}^*)$ when the explanatory variables take on the value represented by $\mathbf{x}^*$ and when the parameters are known to take on these specified values. By using values of $\mathbf{g}$ and other parameters from the posterior distribution whose mean and standard deviation are reported in Table 2, we generate a range of estimates of $\mu(\mathbf{x}^*)$, and the mean of this range then represents the econometrician's posterior inference as to the value of $\mu(\mathbf{x}^*)$.

To examine what the nonlinear function $\mu(.)$ looks like, we fixed the value of $\widehat{v}^2_{t+1|t}$ equal to its sample mean, and evaluated the Bayesian posterior expectation of (11) for various values of y_{1t}. Figure 1a - 1e plot the result as an function of y_{1t} along with 95% probability regions for 3M, 6M, 1Y, 5Y and 10Y bond yields respectively. While Figure 1a, 1b and 1c indicate that the relation between short-term bond yields (3M, 6M, 1Y) and risk-free rate is at least approximately linear, the relation for long-term bond yields (5Y, 10Y) seems to be nonlinear, as in Ahn and Gao (1999). This result suggests that the magnitude of the nonlinearity may increase as time to maturity increases.

Insert Figure 1a - 1e

Figure 2a - 2e answer the analogous question, fixing y_{1t} equal to its sample mean, $\overline{y}_{1t}$, and varying the value of $\widehat{v}^2_{t+1|t}$ for 3M, 6M, 1Y, 5Y, and 10Y bond yields respectively. All figures indicate that there is a threshold effect of volatility on the interest rates but the value of threshold looks different depending on the interest rate. In particular, the threshold effect is most marked in the 3M rate and it is around 1.2 of conditional variance. These results imply that while the effect of volatility on the bond yield is little at lower volatility values, relatively high volatility has a significant impact on the interest rates.

Insert Figure 2a - 2e

Another point to make is that all figures indicate that as volatility increases, the slope of the conditional expected function with respect to volatility is steeper, implying that the functional form of the bond yield is convex with respect to the volatility of risk-free rate. This convexity is compatible with the characteristic of the QTSM of Ahn *et al.* (2002). However, since the nonlinearity of the volatility component taken individually is not statistically significant in the Bayesian posterior estimates (Table 2), it is needed to examine this convexity in terms of the collective contribution of the nonlinear components.

Insert Figure 3a - 3e

To examine the interactive effect of nonlinear components, we consider contours. Figure 3a - 3e plot contours of the function $\widehat{E}[\mu(y_{1t}, \widehat{v}^2_{t+1|t})|\mathbf{Y}_T]$ for all five bond yields. For the 6M interest rate, there is little indication about nonlinearity resulting from the interaction. However, Figure 3d and 3e show that there is a significant relation between long-term bond yields and the multiplication of risk-free rate and its volatility, implying that the interaction between two components has an impact on the long-term bonds. For 3M and 1Y interest rate, there is an indication of interaction but not substantially. Furthermore, Figure 3a, 3c, 3d and 3e indicate that the value of threshold in volatility is an increasing function of the risk-free rate. To investigate this point, we calculated how the effect of volatility on bond yields are affected by different values of y_{1t}. Figure 4a - 4e compare the six functions $\widehat{\mu}(3, \widehat{v}^2_{t+1|t})$, $\widehat{\mu}(4, \widehat{v}^2_{t+1|t})$, $\widehat{\mu}(\overline{y}_{1t}, \widehat{v}^2_{t+1|t})$, $\widehat{\mu}(7, \widehat{v}^2_{t+1|t})$, $\widehat{\mu}(8, \widehat{v}^2_{t+1|t})$, and $\widehat{\mu}(9, \widehat{v}^2_{t+1|t})$, plotted as a function of $\widehat{v}^2_{t+1|t}$. All figures show that the value of the threshold in volatility increases as the risk-free rate rises though it is less clear in the 6-month rate. In particular, this phenomenon is significant for the 3-month rate and two long-term rates.

Insert Figure 4a - 4e

In sum, the flexible inference suggests three types of nonlinearity; threshold effect of volatility, interaction between the risk-free rate and its volatility, and convexity. The threshold effect of volatility on the bond yield seems to be an increasing function of the risk-free rate. The convexity with respect to volatility is compatible with the QTSM.

3.3 Alternative nonlinear specifications

Which form of nonlinear dynamics captures the exact nature of nonlinearity in the term structure? To address this question, we consider parametric models of nonlinear dynamics and estimate them. Then, we can use (11) as a specification test to see whether the nonlinearity has been successfully modeled. This procedure provides a formal statistical basis for comparing the nonlinear dynamics implied by alternative specifications with what appears in the data from the flexible inference. For example, the QTSM of (6) can be described as a linear regression model of the form

$$y_t = \alpha_0 + \boldsymbol{\alpha}'\mathbf{z}_t + \varepsilon_t, \tag{21}$$

with $\mathbf{z}_t = \left(y_{1t}, \widehat{v}^2_{t+1|t}, (y_{1t} * \widehat{v}^2_{t+1|t}), y^2_{1t}, (\widehat{v}^2_{t+1|t})^2\right)'$. Then, one can test directly whether such a specification of $\mathbf{z}_t$ adequately captures any nonlinearity that appears in the data by comparing (21) with the more general model

$$y_t = \alpha_0 + \boldsymbol{\alpha}'\mathbf{z}_t + \lambda m(\mathbf{x}_t) + \varepsilon_t,$$

for $\mathbf{x}_t = (y_{1t}, \widehat{v}^2_{t+1|t})'$ and $m(.)$ a realization of the random field whose correlations are characterized by (10). In what follows, we consider four alternative nonlinear specifications and Table 3 summarizes them.

Insert Table 3

Model A denotes the threshold model of volatility where the threshold value is an increasing function of the risk-free rate. The parameters c_0 and c_1 can be chosen from Figures 4a - 4e.[3] Model B indicates the interaction between risk-free rate and its volatility. Model C is the two-factor quadratic model of Ahn *et. al.* (2002). Since we found that the estimated coefficient on y_{1t}^2 is not statistically significant for all bond yields, we exclude y_{1t}^2 from Model C and consider the result as Model D. Finally, Model E incorporates the threshold effect and the convexity of variance on the interest rate into the QTSM and thus this model is based on our flexible inference. We consider that $\exp(\widehat{v}_{t+1|t}^2)$ catches the aspect of convexity in the model.

Insert Table 4

Table 4 reports the ν^2 statistics for five alternative nonlinear specifications and these can be compared with the corresponding values in Table 1. There is significant decrease in ν^2 statistics for Model A. We can not reject the null of linearity for 3M and 10Y at the 5% level whereas we reject the null for 6M, 1Y and 5Y interest rates, indicating that the threshold model does not capture the entire nature of nonlinearity in the term structure. The Model B is doing well for the long-term rates but does not capture the nonlinearity in the short-term rates. $\widehat{\alpha}_3$, the value in the estimated coefficient in the multiplication of risk-free rate and its volatility, is negative for all bond yields, indicating that the correlation between two factors has a negative impact on the interest rate. Dai and Singleton (2000), and Ahn *et. al.* (2002) state that the conditional correlation among the state variables plays an important role in explaining the dynamics of bond yields.

We cannot reject the null hypothesis that the general QTSM (Model C) accurately describes long-term bond yields relative to the ATSM and there is significant decrease in the ν^2 statistics for all three short-term bond yields. Nevertheless, we reject the QTSM at the 5% level for these short-term rates. When we exclude the squared risk-free rate, y_{1t}^2, from Model

C, the test statistics decreases significantly but we reject the null for two short-term rates, 6M and 1Y at the 5% level. Finally, in Model E, we cannot reject the null of linearity for all interest rates except 1Y in which the null is marginally rejected at the 5% level and the test statistics for 6M rate is relatively lower than in Model D though the statistic for 3M rate is slightly higher than in Model D. Thus, it would seem on the basis of these tests that the QTSM of Ahn *et al.* (2002) does a good job for describing the nonlinearity in the long-term rates but not in the short-term rates while the flexible nonlinear model is generally doing well for all interest rates. This result suggests that the incorporation of the threshold effect and the convexity of variance of the risk-free rate into the QTSM describes the nonlinearity in the term structure and thus is a promising representation of nonlinearities.

To further investigate if the implications of the estimates from the QTSM would turn out to look something like those of the flexible nonlinear model, we consider what Figure 4 would look like if calculated assuming that the estimated QTSM were true rather than assuming the flexible nonlinear model. For doing this, we estimated the QTSM and used estimated values of the parameters to calculate the expected values of the interest rates for the range of variance from zero to $\overline{\widehat{v}^2_{t+1|t}} + 2 * \sqrt{var(\widehat{v}^2_{t+1|t})}$ (the mean of variance plus 2 times standard deviation of the variance), when the risk-free rate has 6 values: 3%, 4%, 5.6% (mean), 7%, 8% and 9%.[4] Figure 5a - 5e plot the expected values of five interest rates against variance. Even though Figure 5s have some similar features with Figure 4s, particularly in the long-term rates, the expected interest rates based on the QTSM appear to be concave against the variance whereas those based on the flexible inference are convex (at least between 0 and 2.48 of the value of variance). These results imply that the QTSM seems to miss some aspect of the nonlinearity.

Insert Figure 5a - 5e

So far, our empirical results suggest that allowing for nonlinearities in the pricing of bond yields is quite important for describing the term structure. To examine how well nonlinear term structure models perform relatively to the ATSM, we consider the forecastability of three models; a two-factor affine model (ATSM), the quadratic term structure model (QTSM, Model D), and a flexible two-factor model (FNM, Model E). We estimate these two models for the sample from 1959:1 to 1995:12. Then, we use estimated coefficients to forecast five interest rates over the period 1996:1 - 1999:11 and then calculate the root-mean-squared-errors (RMSEs).

Insert Table 5

Table 5 reports the RMSEs in three models for all five bond yields. For all interest rates, the RMSEs in the ATSM are higher than both nonlinear models. In particular, the difference between the RMSE of the ATSM and that of nonlinear models increases as time to maturity increases, implying that the nonlinear model performs better relatively to the ATSM in the long-term rates than in the short-term rates. Thus, out-of-sample forecasts also support the claims of nonlinearities. The RMSEs of the short term rates in the QTSM are higher than those in the FNM while it is reverse in the long-term rates, confirming the results of Table 4 in which the QTSM is a good representation for the long-term rates but might miss some aspect of nonlinearity in the short-term rates.

4 Concluding remarks

The affine term structure models have received a lot of attraction from academic researchers as well as financial analysts due to analytical tractability and relatively easy implementation. However, these models have recently been challenged on the empirical performance. Recent empirical studies have shown that there are nonlinearities in the term structure and the nonlinear models capture the term structure dynamics considerably better than the affine models. From this point of view, whether the relation between bond yields and underlying state variables is linear is an important question for evaluating the term structure model. Furthermore, characterizing nature of the nonlinearity is a worthy task pursuing.

The contribution of this paper is to address this question using the framework of Hamilton (2001) that explicitly parameterizes the set of nonlinear relations in a flexible way and takes into account uncertainty about the functional form in conducting hypothesis tests. Since the risk-free rate and its volatility are the most important factors in modern financial markets, we considered these two factors as the underlying two state variables. We found that there is clear evidence of nonlinearities in the term structure, in contrast to the affine term structure model and consistent with recent claims in the literature.

When one looks at this nonlinear relation from a flexible, unrestricted framework, there seems to be three natures of nonlinearity: a threshold effect of volatility on bond yields, interaction between the risk-free rate and its volatility, and convexity. The threshold effect characterizes the nature of the nonlinearity but seems not to capture entire nature of nonlinearity. While the interaction captures the nonlinearity well for long-term rates, its performance is not good for short-term rates.

The quadratic term structure model specified by Ahn *et al.* (2002) does perform better than the threshold model but seems to miss some aspect of the nonlinearity for short-term rates. However, our flexible nonlinear model which incorporates the threshold effect and the convexity of volatility of the risk-free rate into the quadratic model, generally performs well for all interest rates and out-of-sample forecasts for comparing the performance of the nonlinear model with the affine model indicate that the nonlinear model are considerably better than the affine model. Hence, our results suggest that the flexible nonlinear model is a promising representation of nonlinearities and a better candidate for the hedging or pricing of interest rate contingent claims than the affine models.

Footnotes

1. For further studies on the affine term structure model, see Balduzzi et. al. (1996), Balduzzi et. al. (1998), Backus, Telmer and Wu (1999), Backus, Foresi, Mozumdar, and Wu (2001), Backus, Foresi, and Telmer (2001), Campbell, Lo, and MacKinlay (1997), Chen and Scott (1992, 1993), Dai and Singleton (2002), Duffie and Singleton (1999), Gong and Remolona (1997), Pearson and Sun (1994) and among others.

2. Beaglehole and Tenney (1992) point out that Longstaff's (1989) bond pricing equation is not the solution to the pricing problem because it is a failure to properly account for a boundary condition.

3. From the Figures 4a - 4e, we tried to find the initial threshold point and considered several values around it as a candidate of c_0. From this threshold point, we tried to draw the straight line and calculated the slope of the line. Then, we took the value as c_1. Thus, we have $c_0 = 0.6$ for $3M$, $c_0 = 0.6$ for $6M$, $c_0 = 0.8$ for $1Y$, $c_0 = 0.8$ for $5Y$, $c_0 = 0.8$ for $10Y$ and $c_1 = 0.207$ for all bond yields.

4. Figure 5a - 5e are based on the estimate of Model C. We also estimated Model D and found that the results are very similar. These results are available upon request.

References

Ahn, D-H, R. F. Dittmar, and A. R. Gallant. 2002. "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies* 15 No.1: 243-288.

Ahn, D-H, and B. Gao. 1999. "A parametric nonlinear model of term structure dynamics," *Review of Financial Studies* 12 No. 4: 721-762.

Backus, D., S. Foresi, A. Mozumdar, and L. Wu. 2001. "Predictable changes in yields and forward rates," *Journal of Financial Economics* 59: 281-311.

Backus, D., C. I. Telmer, and L. Wu. 1999. "Design and Estimation of Affine Yield Models," Working paper, http://bertha.gsia.cmu.edu

Backus, D., S. Foresi, and C. I. Telmer. 2001. "Affine term structure models and the forward premium anomaly," *Journal of Finance* LVI No.1: 279-304.

Balduzzi, P., S. R. Das, and S. Foresi. 1999. "The central tendency: a second factor in bond yields," *Review of Economics and Statistics* 80 No.1: 62-72.

Balduzzi, P., S. R. Das, S. Foresi, and R. K. Sundaram. 1996. "A simple approach to three factor affine term structure models, " *Journal of Fixed Income* 6: 43-53.

Beaglehole, D., and M. Tenney. 1992. "Corrections and additions to 'A nonlinear equilibrium model of the term structure of interest rate'," *Journal of Financial Economics* 32: 345-353.

Boudoukh, J., M. Richardson, R. Stanton, and R.F. Whitelaw. 1999. "A Multi-factor, Nonlinear, Continuous-Time Model of Interest Rate Volatility," Department of Finance Working paper FIN99-042, Stern School of Business, New York University.

Brennan, M. J., and E. S. Schwartz. 1982. "An equilibrium model of bond pricing and a test of market efficiency," *Journal of Financial and Quantitative Analysis* 17 No. 3: 301-329.

Brenner, R. J., R. H. Harjes, and K. F. Kroner. 1996. "Another Look at Models of the Short-term Interest Rate," *Journal of Financial and Quantitative Analysis*

31: 85-107.

Campbell, J. Y., A. W. Lo, and A. G. MacKinlay. 1997. *The Econometrics of Financial Markets*, Princeton: Princeton University Press.

Chan, K. C., A. Karolyi, F. A. Longstaff, and A. B. Sanders. 1992. "An empirical comparison of alternative models of the short-term interest rate," *Journal of Finance* XLVII No. 3: 1209-1227.

Chen, R-R., and L. Scott. 1993. "Maximum Likelihood Estimation for a multi-factor equilibrium model of the term structure of interest rates," *Journal of Fixed Income* 3: 14-31.

Chen, R-R., and L. Scott. 1992. "Pricing interest rate options in a two-factor Cox-Ingersoll-Ross model of the term structure," *Review of Financial Studies* 5: 613-636.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. 1985. "A theory of the term structure of interest rates," *Econometrica* 53: 385-407.

Dahl, C. M. 2002. "An investigation of tests for linearity and the accuracy of likelihood based inference using random fields," *Econometrics Journal* 5: 263-284.

Dai, Q., and K. J. Singleton. 2002. "Expectation puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics* 63: 415-441.

De Jong, F. 2000. "Time series and cross-section information in affine term-structure models," *Journal of Business and Economic Statistics* 18 No. 3: 300-314.

Duffee, G. R. 2002. "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance* LVII No. 1: 405-443.

Duffie, D., and K. J. Singleton. 1997. "An econometric model of the term structure of interest-rate swap yields," *Journal of Finance* LII No. 4: 1287-1321.

Duffie, D., and K. J. Singleton. 1999. "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies* 12 No. 4: 687-720.

Duffie, D., and R. Kan. 1996. "A yield-factor model of interest rates," *Mathematical Finance* 6 No.4: 379-406.

Ghysel, E., and S. Ng. 1998. "A semiparametric factor model of interest rates and tests of the affine term structure," *Review of Economics and Statistics* 80 No. 4: 535-548.

Gong, F., and E. M. Remolona. 1997. "Two factors along the yield curve," *The Manchester School Supplement*: 1-31.

Hamilton, J. D. 2001. "A parametric approach to flexible nonlinear inference," *Econometrica* 69 No. 3: 537-573.

Hamilton, J. D. 2003. "What is an oil shock?," *Journal of Econometrics* 113: 363 - 398.

Hamilton, J. D., and D. H. Kim. 2002. "A Reexamination of the Predictability of Economic Activity Using the Yield Spread," *Journal of Money, Credit, and Banking* 34 No. 2: 340-360.

Harrison, M., and D. Kreps. 1979. "Martingales and Arbitrage in Multiperiod Security Markets," *Journal of Economic Theory* 20: 381-408.

Harrison, M., and S. Pliska. 1981. "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Application* 11: 215-260.

Kim, D. H., D. R. Osborn, and M. Sensier. 2005. "Nonlinearity in the Fed's Monetary Policy Rule," *Journal of Applied Econometrics* 20: 621-639.

Longstaff, F. A. 1989. "A nonlinear general equilibrium model of the term structure of interest rates," *Journal of Financial Economics* 23: 195-224.

Longstaff, F. A., and E. S. Schwartz. 1992. "Interest rate volatility and the term structure: a two-factor general equilibrium model," *Journal of Finance* XLVII

No. 4: 1259-1282.

Pearson, N. D., and T-S. Sun. 1994. "Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model," *Journal of Finance* XLIX No. 4: 1279-1304.

Schaefer, S. M., and E. S. Schwartz. 1984. "A two-factor model of the term structure:an approximate analytical solution," *Journal of Financial and Quantitative Analysis* 19: 413-424.

Stanton, R. 1997. "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance* LII No. 5: 1973-2002.

Vasicek, O. A. 1977. "An equilibrium characterization of the term structure," *Journal of Financial Economics* 5: 177-188.

Table 1 The test of the null hypothesis that $\mu(\mathbf{X}_t) = \alpha_0 + \boldsymbol{\alpha}_1' \mathbf{X}_t$

maturity	ν^2 statistic	$p - value$
3M	17.874	2.36e-05
6M	33.883	5.85e-09
1Y	72.650	1.55e-17
5Y	53.421	2.69e-13
10Y	26.936	2.10e-07

Note: 3M, 6M, 1Y, 5Y and 10Y denote 3-month, 6-month, 1-year, 5-year, and 10-year bond yields respectively.

Table 2 Bayesian posterior estimates for the flexible nonlinear alternative

$$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \hat{v}^2_{t+1|t} + \sigma[\zeta m(g_1 y_{1t}, g_2 \hat{v}^2_{t+1|t}) + v_t].$$

maturity	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\sigma}$	$\hat{\zeta}$	$\hat{g}_1$	$\hat{g}_2$
3M	0.962	0.955	0.091	0.248	2.399	0.641	1.896
	(0.448)	(0.042)	(0.055)	(0.012)	(0.874)	(0.305)	(0.897)
6M	1.749	0.882	0.095	0.314	3.169	0.293	0.786
	(1.139)	(0.083)	(0.090)	(0.012)	(0.949)	(0.156)	(0.483)
1Y	3.039	0.769	0.052	0.428	3.193	0.201	0.767
	(0.943)	(0.086)	(0.161)	(0.020)	(0.893)	(0.125)	(0.563)
5Y	4.674	0.585	0.041	0.869	2.274	0.183	0.973
	(1.540)	(0.130)	(0.183)	(0.039)	(0.700)	(0.097)	(0.659)
10Y	5.019	0.584	0.098	1.015	1.980	0.206	1.252
	(1.628)	(0.137)	(0.193)	(0.041)	(0.673)	(0.117)	(0.836)

Note: The values in all parentheses are the standard errors of Bayesian posterior estimates with $N = 5,000$ Monte Carlo simulations.

Table 3 Alternative nonlinear specifications

model	specifications
A	$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2(\widehat{v}^2_{t+1\|t} - c_0 - c_1 y_{1t})\delta_{[\widehat{v}^2_{t+1\|t} > c_0 + c_1 y_{1t}]} + \varepsilon_t$
B	$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \widehat{v}^2_{t+1\|t} + \alpha_3(y_{1t} * \widehat{v}^2_{t+1\|t}) + \varepsilon_t$
C	$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \widehat{v}^2_{t+1\|t} + \alpha_3(y_{1t} * \widehat{v}^2_{t+1\|t}) + \alpha_4 (y_{1t})^2 + \alpha_5(\widehat{v}^2_{t+1\|t})^2 + \varepsilon_t$
D	$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \widehat{v}^2_{t+1\|t} + \alpha_3(y_{1t} * \widehat{v}^2_{t+1\|t}) + \alpha_5(\widehat{v}^2_{t+1\|t})^2 + \varepsilon_t$
E	$y_{nt} = \alpha_0 + \alpha_1 y_{1t} + \alpha_2 \widehat{v}^2_{t+1\|t} + \alpha_3 \delta_{[\widehat{v}^2_{t+1\|t} \geq c_2]} + \alpha_4(y_{1t} * \widehat{v}^2_{t+1\|t}) + \alpha_5(\widehat{v}^2_{t+1\|t})^2 + \alpha_6 \exp(\widehat{v}^2_{t+1\|t}) + \varepsilon_t$

Note: a. In model A, $\delta_{[\widehat{v}^2_{t+1|t} > c_0 + c_1 y_{1t}]} = 1$ if $\widehat{v}^2_{t+1|t} > c_0 + c_1 y_{1t}$ and 0 otherwise, where c_0 and c_1 are parameters.

b. In model E, $\delta_{[\widehat{v}^2_{t+1|t} \geq c_2]} = 1$ if $\widehat{v}^2_{t+1|t} \geq c_2$, where $c_2 = 1.2$ for all five yields and 0 otherwise.

Table 4. Tests of the linearity null hypothesis $\mu(\mathbf{x}_t) = \alpha_0 + \boldsymbol{\alpha}'\mathbf{x}_t$ for alternative nonlinear specifications in the term structure

maturity	model A		model B		model C		model D		model E	
	ν^2	$p-v.$	ν^2	$p-v.$	ν^2	$p-v.$	ν^2	$p-v.$	ν^2	$p-v.$
3M	1.962	0.161	8.962	0.003	5.297	0.021	2.484	0.115	3.466	0.063
6M	8.292	0.004	34.122	5.18e-09	6.481	0.011	4.975	0.026	3.190	0.074
1Y	8.074	0.005	13.783	0.0002	8.836	0.003	4.074	0.044	3.877	0.049
5Y	12.93	0.0003	0.991	0.320	0.253	0.615	0.033	0.857	0.042	0.837
10Y	3.100	0.078	1.628	0.202	0.621	0.431	0.236	0.627	0.1.48	0.701

Note. a. In model A, $c_0 = 0.6$ for $3M$, $c_0 = 0.6$ for $6M$, $c_0 = 0.8$ for $1Y$, $c_0 = 0.8$ for $5Y$, $c_0 = 0.8$ for $10Y$ and $c_1 = 0.207$ for all bond yields.

b. $p-v.$ denotes the $p-value$ of the test.

Table 5. The Forecastability of the affine model and the flexible nonlinear models

maturity	ATSM	QTSM	FNM
3M	0.2302	0.2202	0.2188
6M	0.3304	0.3083	0.3059
1Y	0.4151	0.3939	0.3910
5Y	0.7167	0.6776	0.6781
10Y	0.8115	0.7615	0.7668

Note. a. ATSM, QTSM, and FNM denote the affine term structure model, the quadratic term structure model and the flexible nonlinear model (Model E) of (4.7).

b. All figures are the root mean squared errors (RMSE) of multi-step-ahead out-of-sample forecasting from 1996:1 to 1999:12.

Figure 1a The effect of the risk-free rate on the 3-month rate

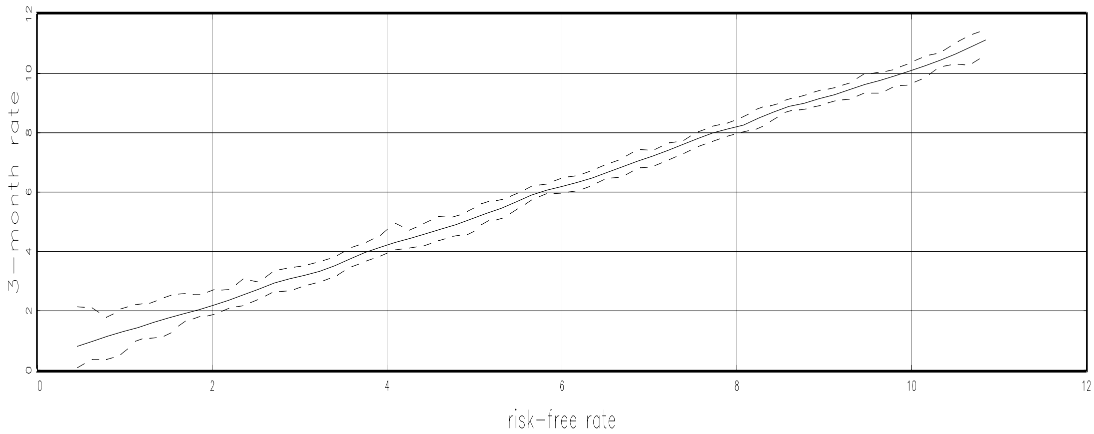


Figure 1b The effect of the risk-free rate on the 6-month rate

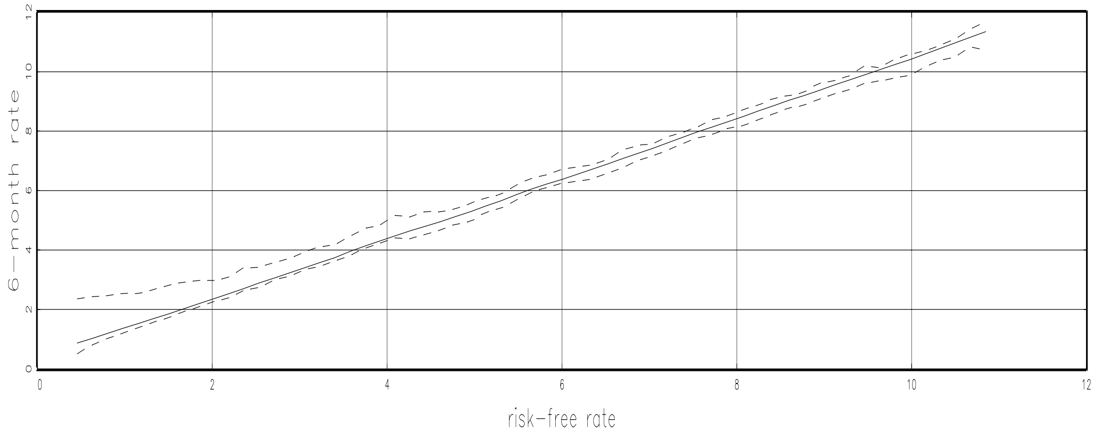


Figure 1c The effect of the risk-free rate on the 1-year rate

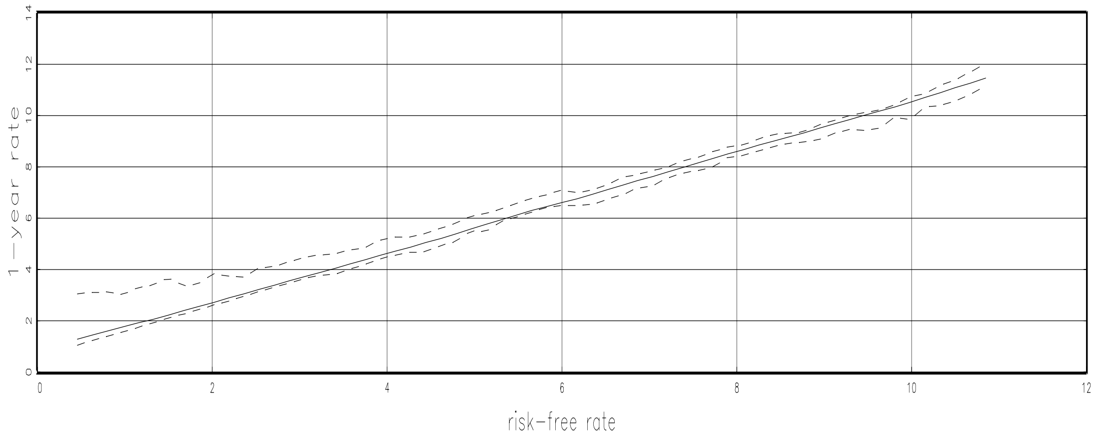


Figure 1d The effect of the risk-free rate on the 5-year rate

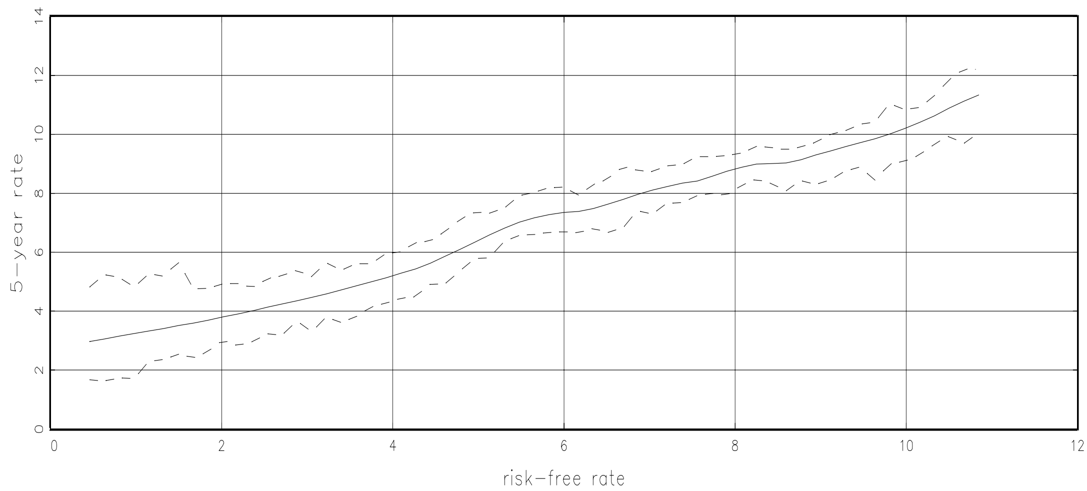


Figure 1e The effect of the risk-free rate on the 10-year rate

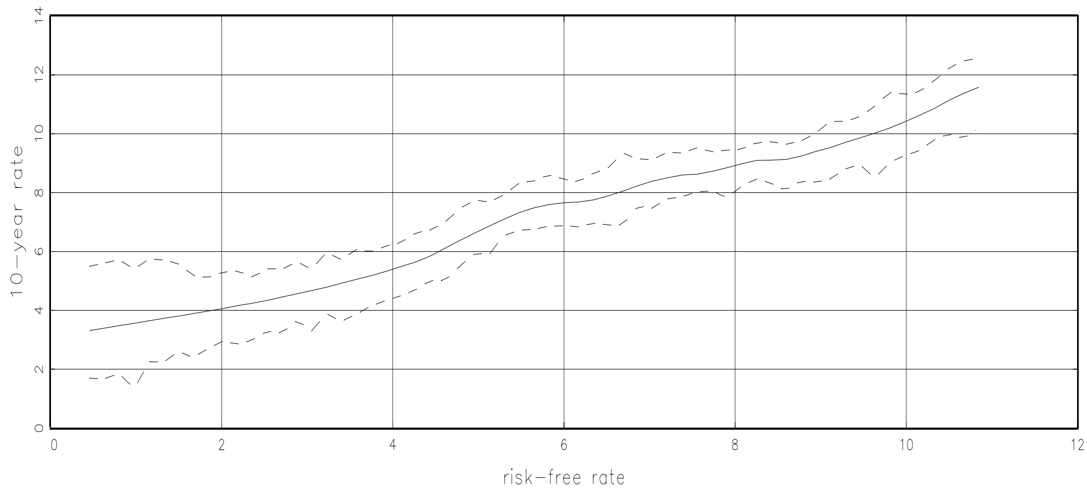


Figure 1a – 1e
The effect of the risk-free rate on the interest rates

Solid line plots the posterior expectation of the function $\alpha_0 + \boldsymbol{\alpha}_1^{'}\mathbf{x}_t + \lambda m(\mathbf{x}_t)$ evaluated at $\mathbf{x}_t = (x_1, \overline{\hat{v}^2_{t+1|t}})'$ as a function of x_1 where $\overline{\hat{v}^2_{t+1|t}} = T^{-1}\sum_{t=1}^{T}\hat{v}^2_{t+1|t}$ and where the expectation is with respect to the posterior distribution of $\alpha_0, \boldsymbol{\alpha}, \lambda,$ and $m(\mathbf{x}_t)$ conditional on observation of $\{y_t, \mathbf{x}_t\}_{t=1}^{T}$, with this posterior distribution estimated by Monte Carlo importance sampling with 5,000 simulations. Dashed lines give 95% probability regions.

Figure 2a The effect of volatility on the 3-month rate

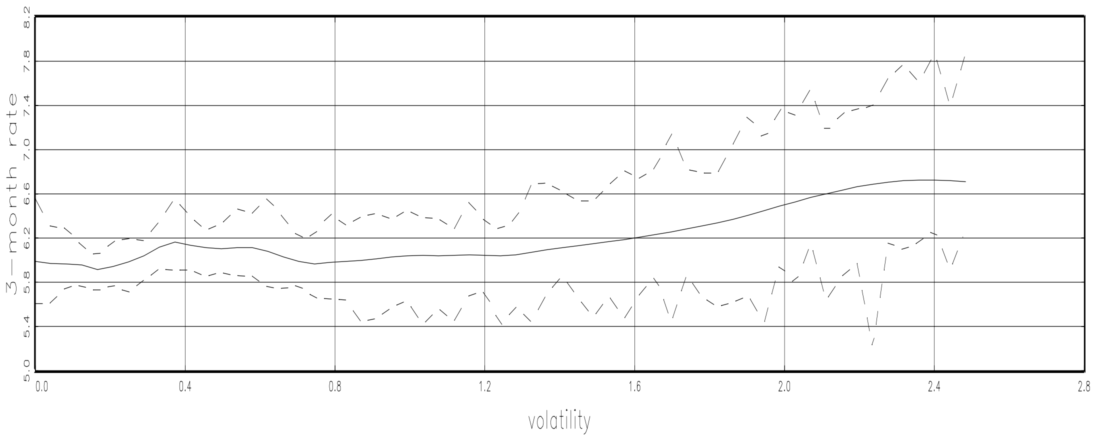


Figure 2b The effect of volatility on the 6-month rate

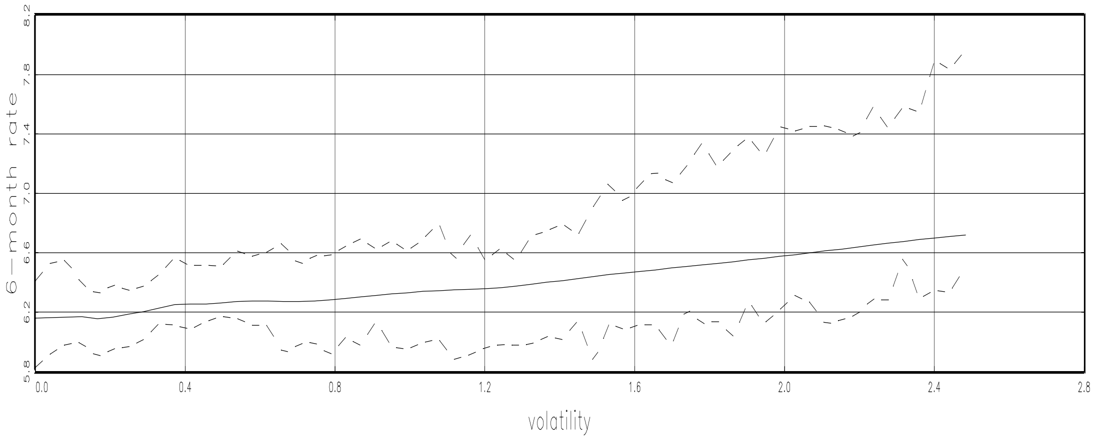
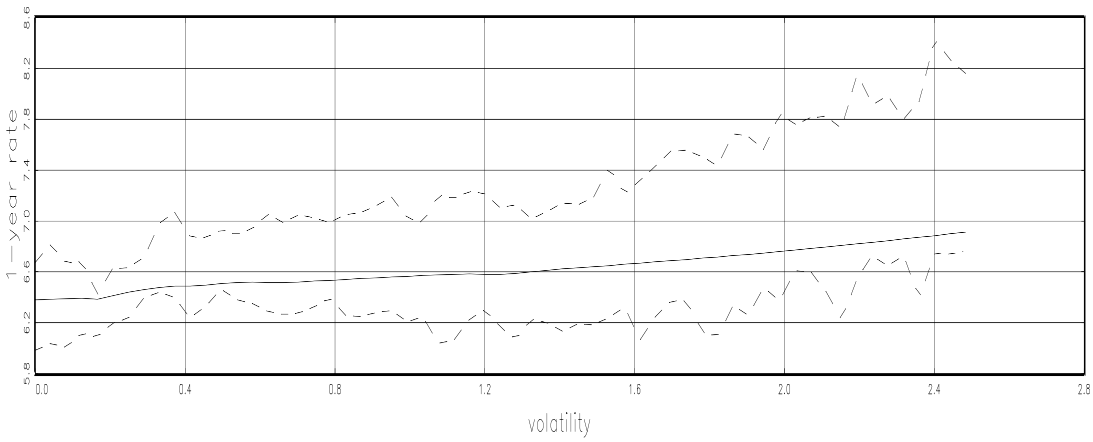
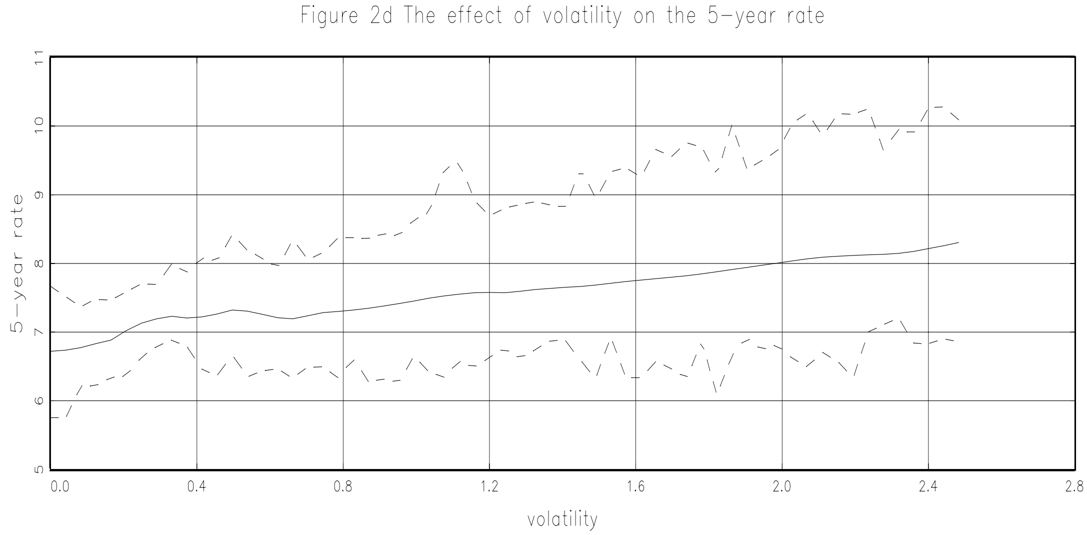


Figure 2c The effect of volatility on the 1-year rate





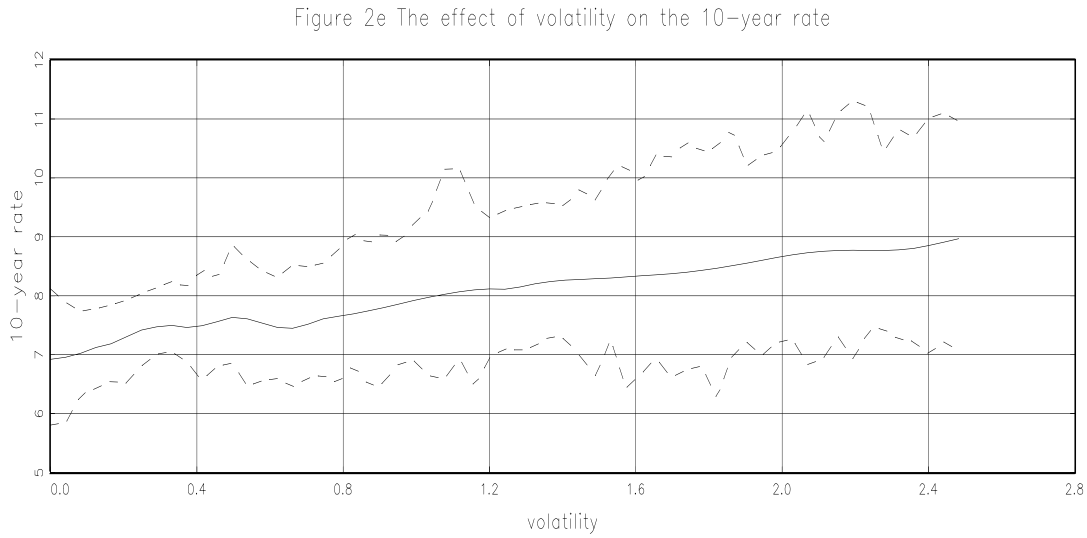


Figure 2a – 2e
The effect of volatility on the interest rates
Solid line plots the posterior expectation of the function $\alpha_0 + \boldsymbol{\alpha}_1^{'}\mathbf{x}_t + \lambda m(\mathbf{x}_t)$ evaluated at $\mathbf{x}_t = (\overline{y_{1t}}, x_2)'$ as a function of x_2. Dashed lines give 95% probability regions.

Figure 3a Contour of u(x) for 3-month rate

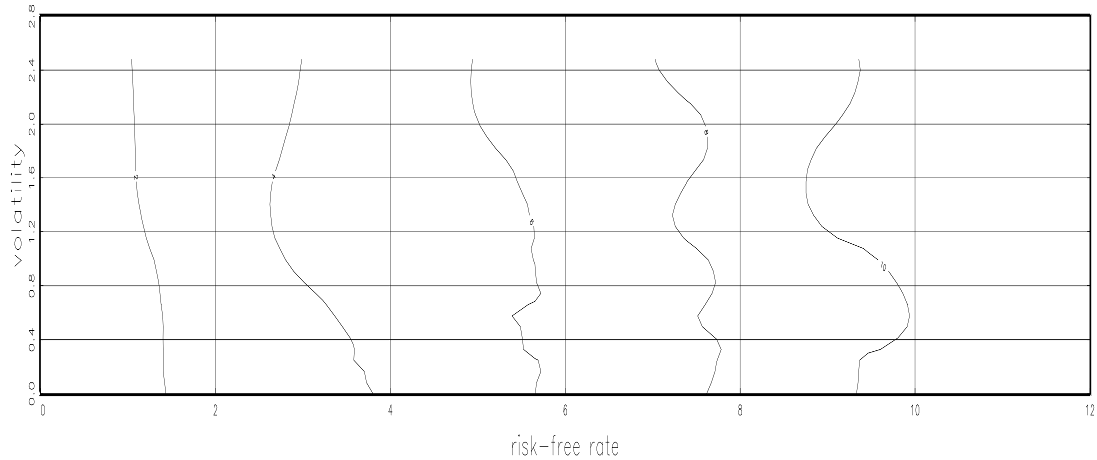


Figure 3b Contour of u(x) for 6-month rate

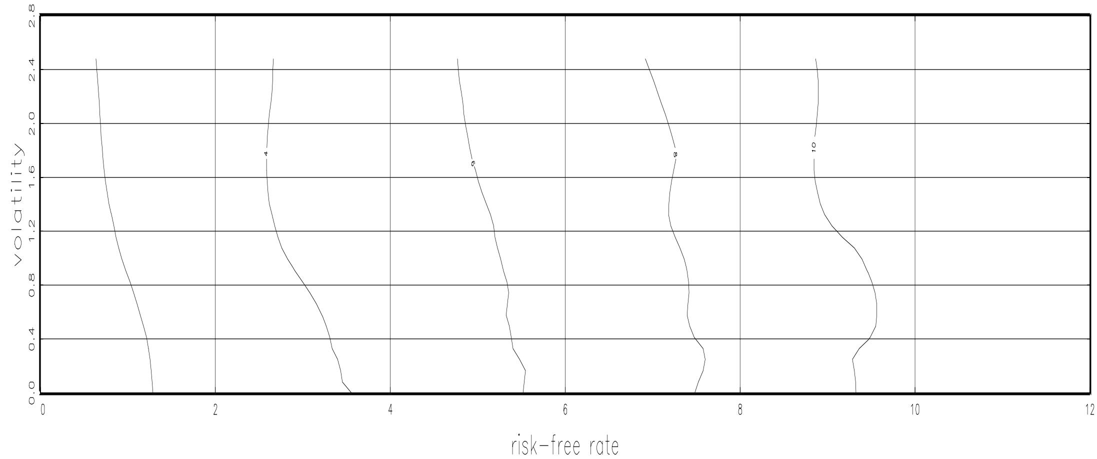
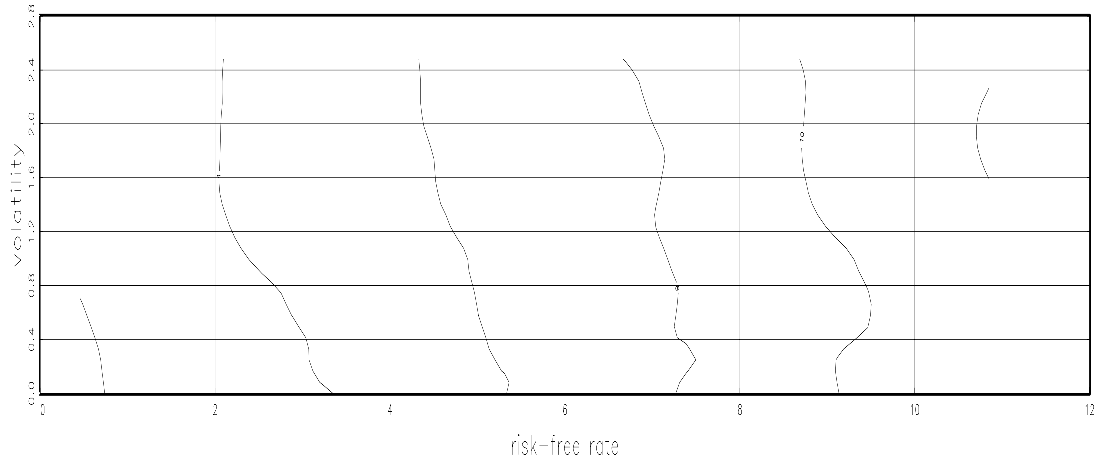
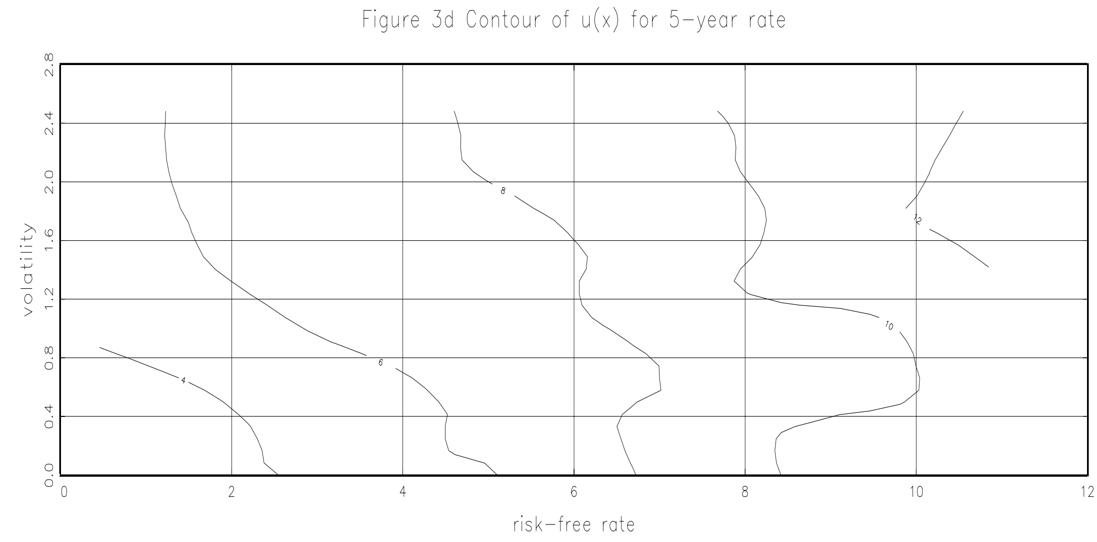


Figure 3c Contour of u(x) for 1-year rate





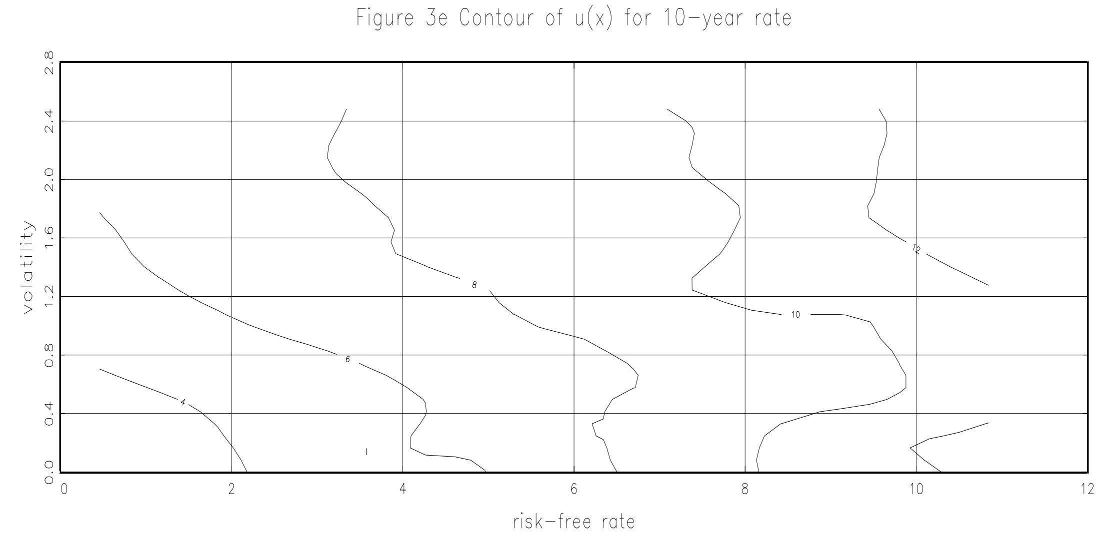


Figure 3a – 3e
The Contour of $\mu(\mathbf{x})$ for the interest rates

The figures are contour lines for estimated $\mu(\mathbf{x})$ function and plot combinations of x_1 and x_2.

Figure 4a. The effect of volatility on the 3-month rate: y1t = 3, 4, mean, 7, 8, 9

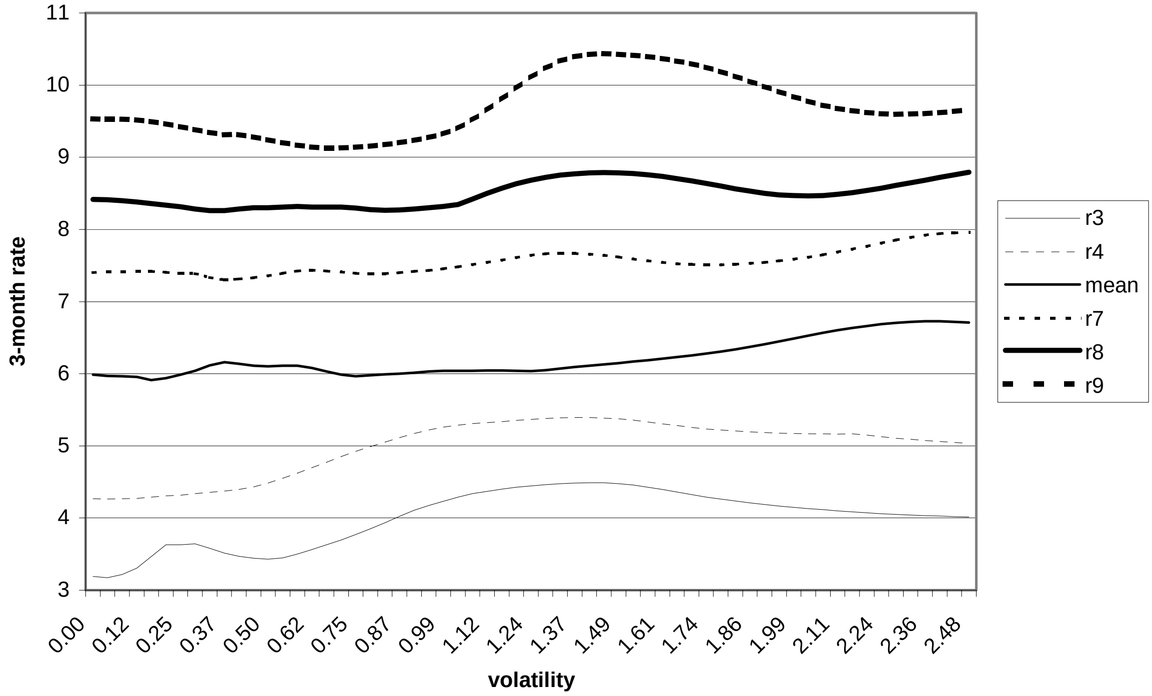


Figure 4b. The effect of volatility on 6-month rate: y1t = 3, 4, mean, 7, 8, 9

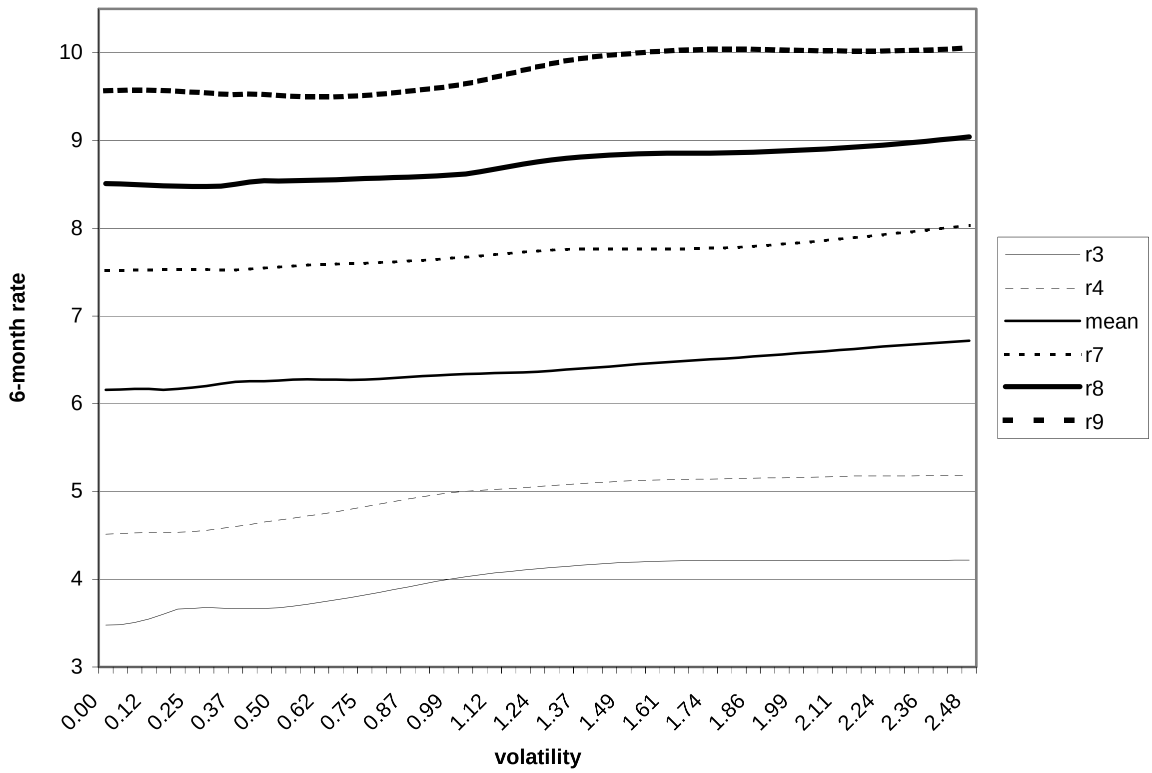


Figure 4c The effect of volatility on 1-year rate: y1t = 3, 4, mean, 7, 8, 9

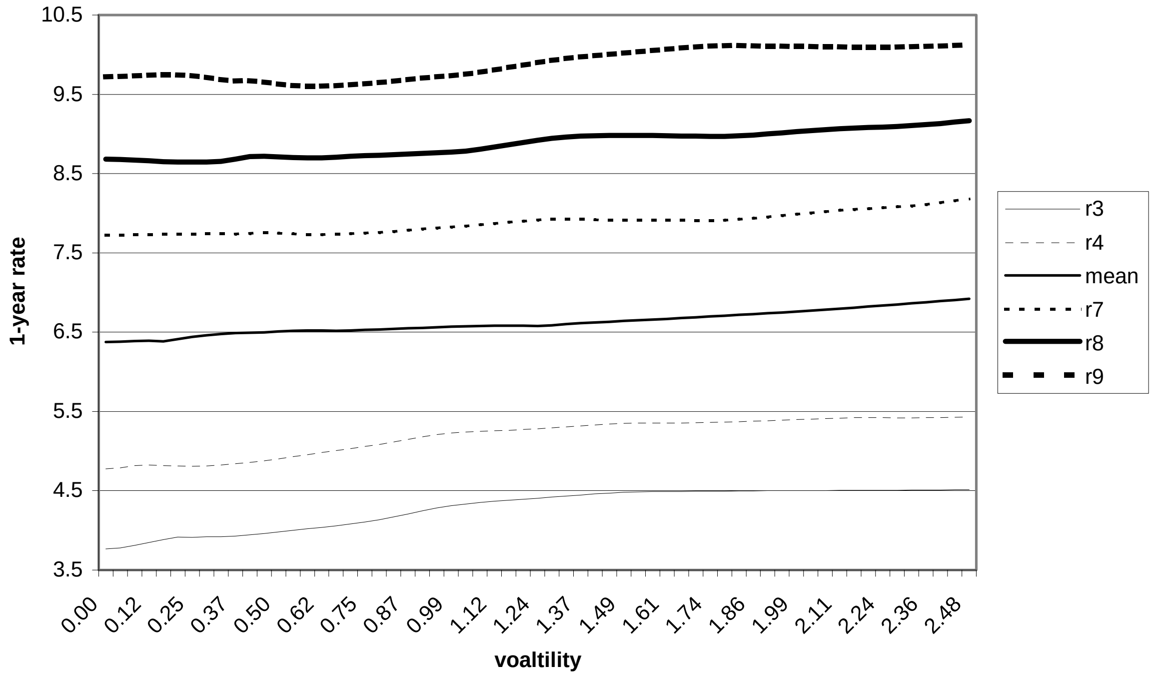


Figure 4d. The effect of volatility on 5-year rate: y1t = 3, 4, mean, 7, 8, 9

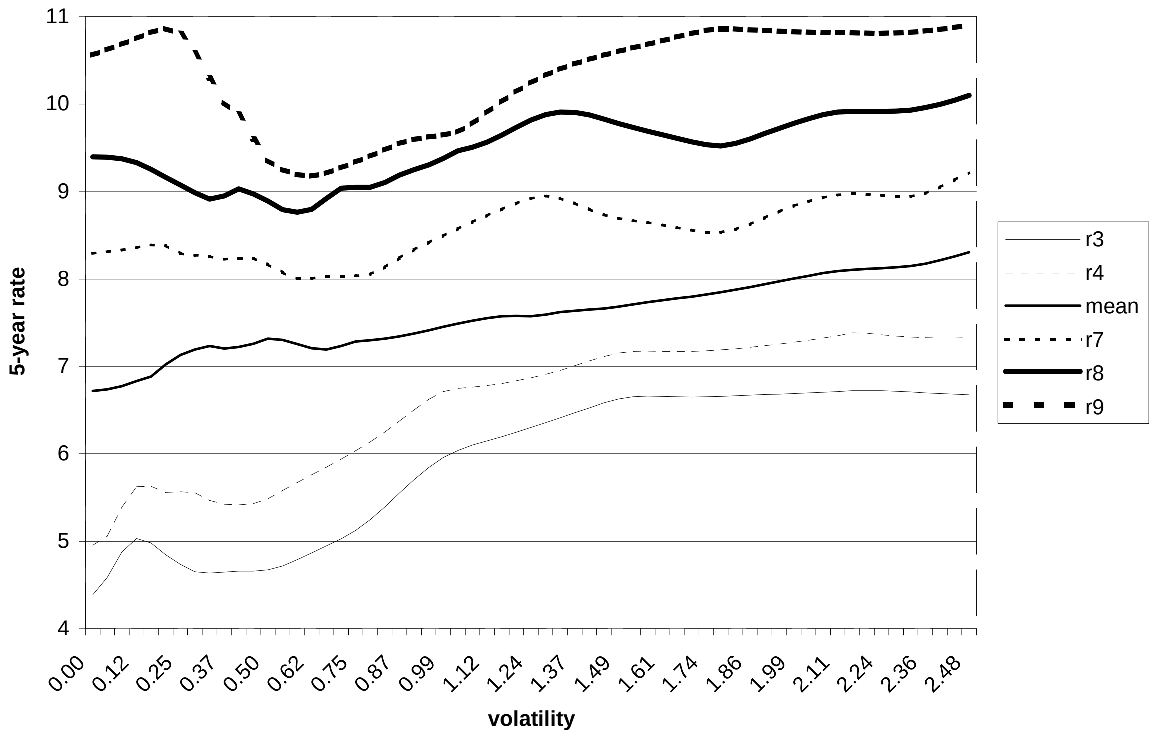


Figure 4e The effect of volatility on 10-year rate: y1t = 3, 4, mean, 7, 8, 9

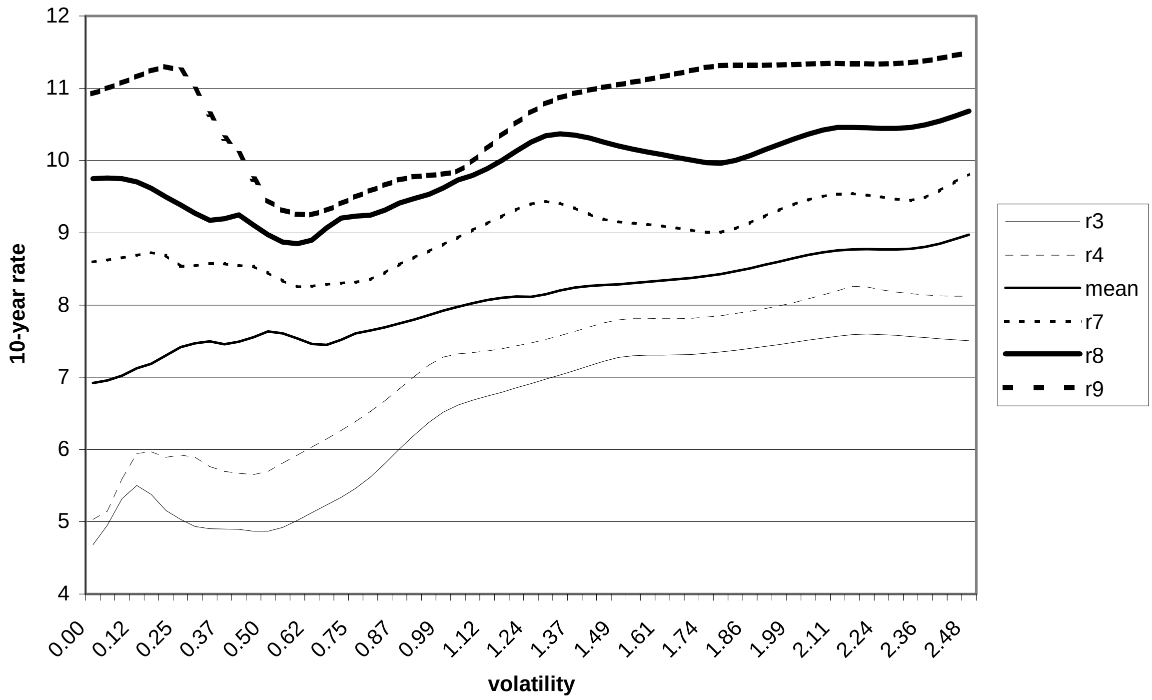


Figure 4a – 4e
The effect of volatility on the interest rates

Each line plots the posterior expectation of the function $\alpha_0 + \boldsymbol{\alpha}_1^{'}\mathbf{x}_t + \lambda m(\mathbf{x}_t)$ evaluated at $\mathbf{x}_t = (x_1, x_2)'$ as a function of x_2. For the thin-solid line, $x_1 = 3\%$, for the thin-dotted line, $x_1 = 4\%$, for the solid line, $x_1 = 5.65\%(mean)$, for the dotted line, $x_1 = 7\%$, for the thick-solid line, $x_1 = 8\%$, and for the thick-dotted line, $x_1 = 9\%$.

Figure 5a The effect of volatility on the 3-month rate in QTSM: y1t = 3, 4, mean, 7, 8, & 9%

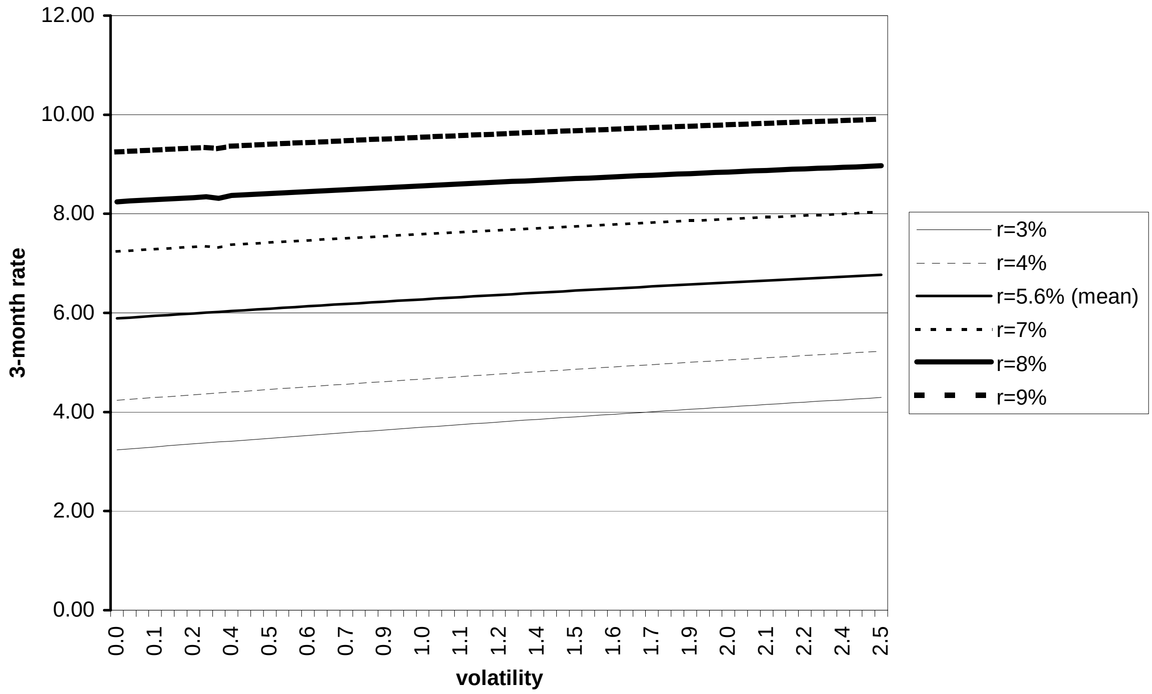


Figure 5b The effect of volatility on the 6-month rate in QTSM: y1t = 3, 4, mean, 7, 8 & 9%

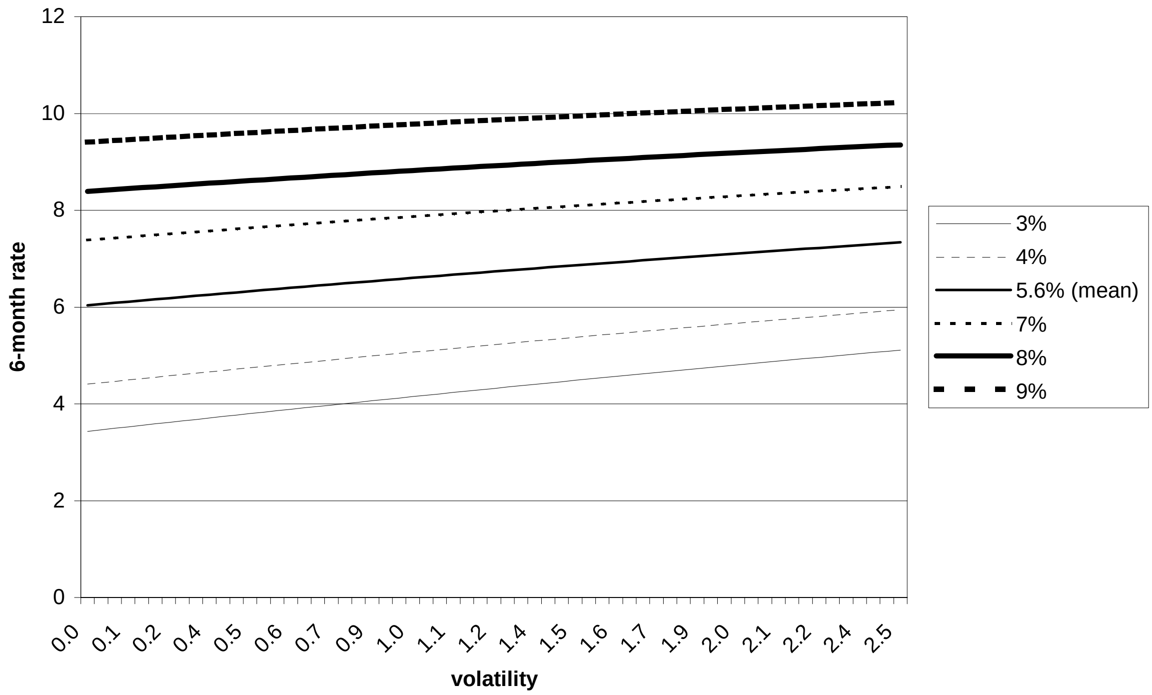


Figure 5c The effect of volatility on the 1-year rate in QTSM:
y1t = 3, 4, mean, 7, 8, & 9%

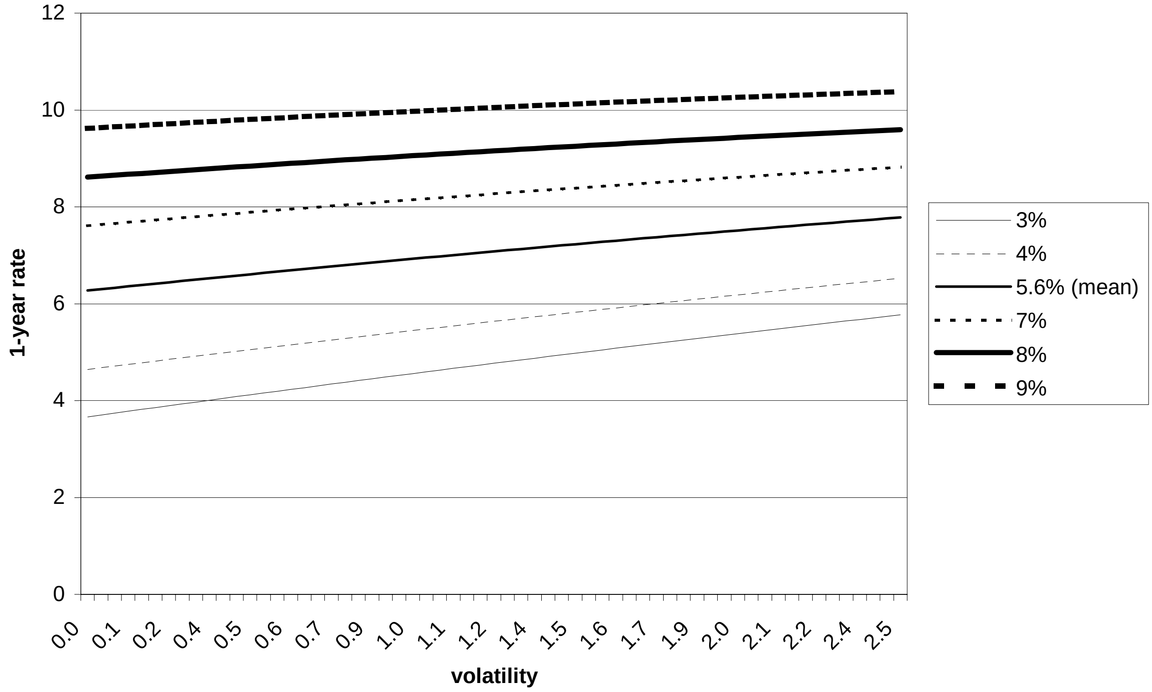


Figure 5d The effect of volatility on the 5-year rate in QTSM:
y1t = 3, 4, mean, 7, 8, & 9%

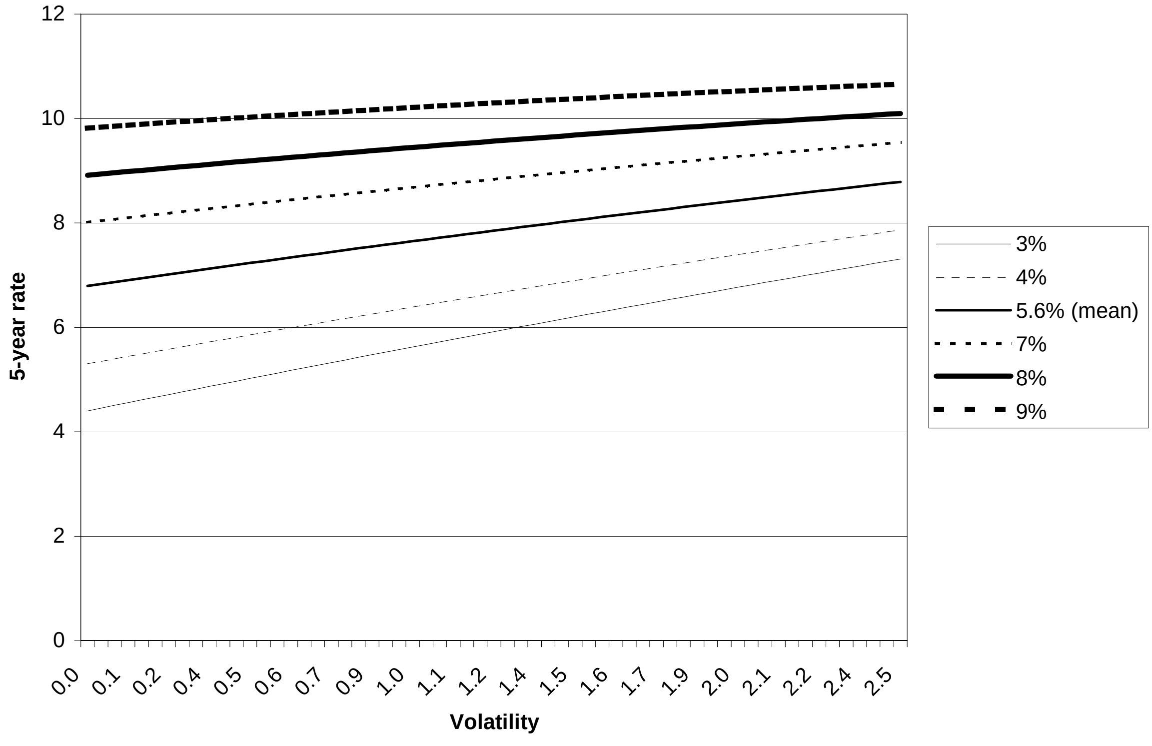


Figure 5e The effect of volatility on the 10-year rate in QTSM: y1t = 3, 4, mean, 7, 8, & 9%

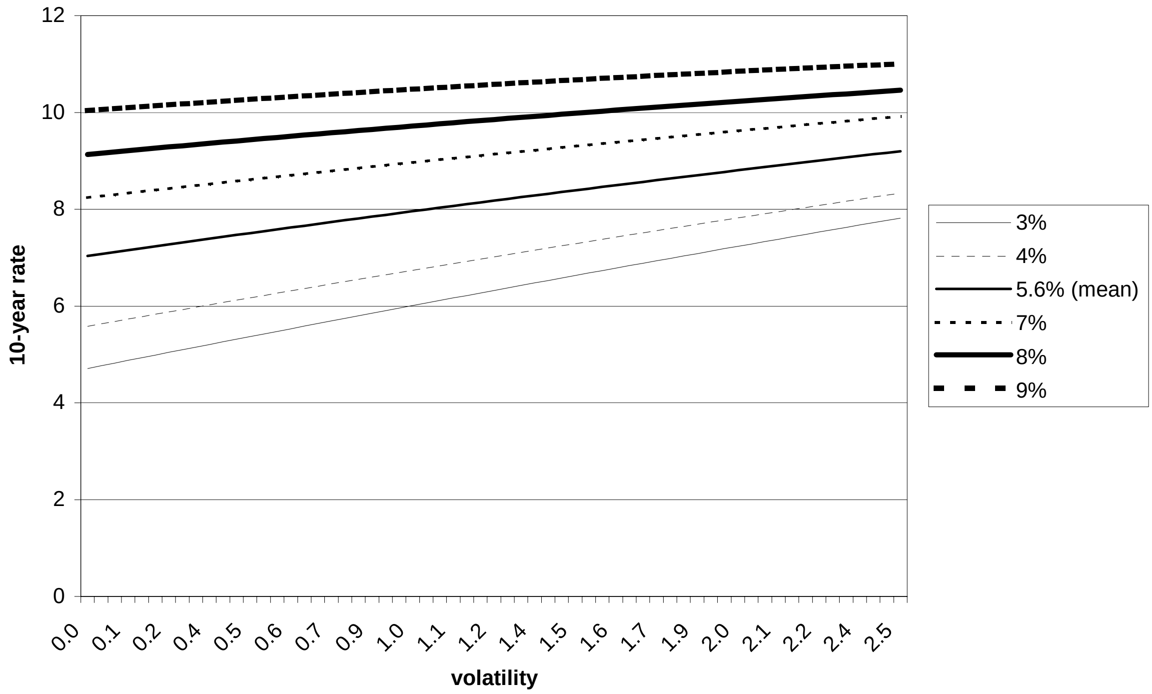


Figure 5a – 5e
The effect of volatility on the interest rates in QTSM
Figure 5's are based on the estimation of model C. They plot the expected values of the interest rates for the range of conditional variance from zero to $\overline{\hat{v}_{t+1|t}^2} + 2 \times \sqrt{\mathrm{var}(\hat{v}_{t+1|t}^2)}$ (the mean of variance plus 2 times standard deviation of the variance), when the risk-free rate has 6 values: 3% for the thin-solid line, 4% for the thin-dotted line, 5.6% (mean) for the solid line, 7% for the dotted line, 8% for the thick-solid line and 9% for the thick-dotted line.