Modeling the Conditional Mean and Variance of the Short Rate Using Diffusion, GARCH, and Moving Average Models

TURAN G. BALI

This article introduces a two-factor-discrete-time-stochastic-volatility model that allows for departures from linearity in the conditional mean and incorporates serially correlated unexpected news, asymmetry, and level effects into the definition of conditional volatility of the short rate. The new class of econometric specifications nests many popular existing symmetric and asymmetric GARCH as well as diffusion models of the short-term interest rate. This study attempts to determine the correct specification of conditional mean and variance of the short rate by developing a more general econometric framework that allows for nonlinear effects in the drift of the short rate, and that defines the conditional volatility as a nonlinear function of unexpected information shocks and interest rate levels. The existing and alternative models are compared in terms of their ability to capture the stochastic behavior of the short-term riskless rate. The

The author is very grateful to Robert I. Webb (the editor) and two anonymous referees of this journal for their extremely helpful comments and suggestions.
For correspondence, Turan G. Bali, Department of Economics & Finance, Baruch College, City University of New York, 17 Lexington Avenue, New York, New York 10010; e-mail: turanbali@worldnet.att.net.

Received June, 1999; Accepted January, 2000

■ *Turan G. Bali is an Assistant Professor of Finance in the Department of Economics and Finance at Baruch College, CUNY in New York, New York.*

The Journal of Futures Markets, Vol. 20, No. 8, 717–751 (2000)

empirical results indicate that the relative performance of the two-factor models in predicting the future level and variance of interest-rate changes is superior to the nested models. Jrl Fut Mark 20:717–751, 2000

INTRODUCTION

The short-term riskless rate of interest is fundamental to the classical approach to pricing fixed-income securities and interest-rate-sensitive claims. However, no consensus has emerged on the dynamics of either the level or the volatility of the short rate. In other words, the choice of the parametric conditional mean and conditional variance of interest rates often is arbitrary. The functional form of the conditional mean in most interest-rate models is specified to let next period's change in the short rate depend *linearly* on the current short level, although there is mounting evidence that the drift of the short-rate process is *not* linear.[1] The specification of the conditional variance, however, tends to vary widely. There are two popular classes of models for short-term interest-rate volatility in the finance literature. The first model belongs to the class of continuous-time or diffusion models in which the conditional volatility is defined only as a function of interest-rate levels.[2] The one-factor diffusion models tend to overemphasize the sensitivity of volatility to interest-rate levels and fail to model adequately the serial correlation in conditional variances. The second class of models, the generalized autoregressive conditional heteroskedasticity (GARCH), parameterizes volatility only as a function of unexpected information shocks to the interest-rate market.[3] The serial-correlation-based GARCH models fail to capture the relationship between interest-rate levels and volatility.

In order to overcome some of the empirical weaknesses of the existing models, this article proposes an alternative class of two-factor-discrete-time-stochastic-volatility models that not only allows for departures from linearity in the conditional mean, but also parameterizes the conditional volatility as a function of both the interest-rate level and the

[1]See, for example, Ait-Sahalia (1996a, 1996b), Pfann, Schotman, and Tschernig (1996), Conley, Hansen, Luttmer, and Scheinkman (1997), and Stanton (1997).

[2]For econometric estimation of continuous-time interest-rate models, see Marsh and Rosenfeld (1983), Barone, Cuoco, and Zautzik (1991), Chan et al. (1992), Chen and Scott (1993), Gibbons and Ramaswamy (1993), Pearson and Sun (1994), Melino (1994), Broze, Scaillet, and Zakoian (1995), and Bali (1999) among many others.

[3]For applications of GARCH and GARCH-type models to interest rate data, see Engle et al. (1987), Cecchetti, Cumbi, and Figlewski (1988), Engle et al. (1992), Cai (1994), Pfann et al. (1996), Bali and Karagozoglu (1999), among many others. See also Bollerslev, Chou, and Kroner (1992) for a comprehensive survey.

news-arrival process. The new class of models introduced in this article advances the current literature in a number of directions. First, the diffusion component of the short-rate process is defined flexibly enough to incorporate serially correlated unexpected news, asymmetry, and the interest-rate level into the conditional volatility of the short rate. Second, the drift component is specified flexibly enough to allow for nonlinear effects in the dynamics of the short rate. Third, the new class of econometric specifications nests many popular existing symmetric and asymmetric GARCH as well as diffusion models of the short-term interest rate. The nesting clearly shows the connection between the existing models and the new class of models. It also allows us to determine the relative performance of each model in predicting the future level and variance of interest-rate changes.

Econometric specifications that are more general are developed in this study to find the proper specification of the conditional mean and variance of the short-term interest rate. The nonlinear empirical generalizations are employed to test nonlinearity in the conditional mean, and serial correlation, asymmetry, and level effects in the conditional variance of the interest-rate change. In addition to symmetric and asymmetric GARCH and diffusion models, constantly weighted and exponentially weighted moving-average models are used to predict the future volatility of the short-term interest rate. The two-factor-stochastic-volatility models are compared with various existing GARCH, diffusion, and moving-average models in terms of the statistical fit of interest-rate data and the out-of-sample forecasting performance by using the Theil inequality coefficient and an alternative measure of model performance.

The empirical results indicate that the correct specification of the conditional mean is as important as that of the conditional variance in order to capture adequately the stochastic behavior of the short rate. In addition, the nested models strongly are rejected against the new class of models, which outperforms the existing models in forecasting the future level and variance of interest-rate changes. All these results point to the presence of nonlinearity in the conditional mean, and serial correlation, asymmetry, and level effects in the conditional variance of the short rate.

This article is organized as follows. In the next section, the existing models of the short-term interest rate will be reviewed. This review will be followed by the introduction of the new class of discrete-time interest-rate models, after which the data and estimation procedures will be described. Then the empirical results from comparing the models will be presented and a theoretical framework for pricing interest rate sensitive claims with the two-factor-stochastic-volatility models will be introduced. The paper closes with a conclusion.

EXISTING MODELS OF THE SHORT-TERM INTEREST RATE

Existing models of the short-term interest rate capture two well-known empirical attributes. First, the short rate is mean reverting. The simplest and most-common way of modeling mean reversion in the existing models is to define next period's interest-rate change as a function of the current short-rate level. Second, the unconditional distribution of interest-rate changes is leptokurtic. Engle (1982) showed that a possible cause of the leptokurtosis in the unconditional distribution is conditional heteroskedasticity, and two different ways of modeling this conditional heteroskedasticity are common in the literature. In various continuous-time or diffusion models, the conditional variance of interest-rate changes is a function of the level of the short rate. In the GARCH models, the conditional variance depends upon lagged-squared shocks to the short rate. The two-most-common classes of interest-rate models are reviewed below.

Diffusion Models

In continuous-time or diffusion models, the dynamics of the short rate is described by the Ito stochastic differential equation:

$$dr_t = \mu(r_t, t)\, dt + \sigma(r_t, t) dW_t \tag{1}$$

where r_t is the instantaneous spot rate, W_t is a standard Brownian motion, $\mu(r_t, t)$ and $\sigma(r_t, t)$, the drift and diffusion functions, depend on the short rate and possibly other state variables. Most empirical work allows for these functions to depend on the short-rate level only, i.e., $\mu(r_t, t) \equiv \mu(r_t)$ and $\sigma(r_t, t) \equiv \sigma(r_t)$. For example, imposing $\mu(r_t) = \alpha_0 + \alpha_1 r_t$, and $\sigma(r_t) = \sigma r_t^{1/2}$ and $\sigma(r_t) = \sigma$ results in the Cox, Ingersoll, and Ross (1985) and the Vasicek (1977) models, respectively.

To compare the performance of a wide variety of well-known models in capturing the stochastic behavior of the short-term interest rate, Chan, Karolyi, Longstaff, and Sanders (CKLS; 1992) used the standard linear mean-reverting drift and the constant elasticity of variance for the diffusion function:

$$dr_t = (\alpha_0 + \alpha_1 r_t)\, dt + \Psi r_t^{\phi}\, dW_t. \qquad (2)$$

Chan et al. (1992) noted that the stochastic differential equation given in eq. (2) encompassed a broad class of interest-rate processes, which included many well-known interest-rate models. These models can be obtained from imposing the appropriate parameter restrictions on α_0, α_1, and ϕ.[4]

Following Brennan and Schwartz (1982), Marsh and Rosenfeld (1983), Dietrich-Campbell and Schwartz (1986), Sanders and Unal (1988), and others, Chan et al. (1992) estimated the parameters of the continuous-time model given in eq. (2) using a discrete-time econometric specification:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \varepsilon_t \qquad (3a)$$

$$E(\varepsilon_t|\Omega_{t-1}) = 0,\ E(\varepsilon_t^2|\Omega_{t-1}) \equiv \sigma_t^2 = \Psi^2 r_{t-1}^{\phi} \qquad (3b)$$

where Ω_{t-1} is the information set available at time $t - 1$ so that $E(\ldots|\Omega_{t-1})$ represents the conditional expectation operator at time $t - 1$, and σ_t^2 is the conditional variance of unexpected interest-rate changes. In this model, autoregressive conditional heteroskedasticity enters solely through the level of the last period's interest rate. Rewriting the drift function $\mu(r_t) = \alpha_0 + \alpha_1 r_{t-1}$ as $\alpha_1(r_{t-1} - \alpha_0^*)$ reveals that α_1 can be viewed as a measure of the speed of mean reversion in interest-rate levels where α_0^* denotes the long-run average level of r. In the CKLS model, mean reversion can be captured by setting $\alpha_1 < 0$ and leptokurtosis by introducing conditional heteroskedasticity (setting $\phi > 0$). The volatility parameter Ψ^2 is simply a scale factor for the conditional variance of the interest-rate change. The parameter ϕ allows the volatility to depend on the level of the interest rate. At higher ϕ's, the volatility is more sensitive to interest-rate levels.

GARCH Models

There are some weaknesses of the CKLS model that encompass many popular diffusion models of the short-term interest rate. First, it does not

[4]For example, the constant elasticity of variance (CEV) process introduced by Cox (1975) and Cox and Ross (1976) is obtained form imposing the parameter restriction $\alpha_0 = 0$. The variable rate model of Cox et al. (1980) is nested within the CEV model by the parameter restriction $\alpha_1 = 0$ and $\phi = 3/2$. Setting $\phi = 0$ gives the Vasicek (1977) model and the Merton (1973) model is nested within the Vasicek model by the parameter restriction $\alpha_1 = 0$. Both of these models imply that the conditional volatility of changes in the riskless rate is constant. Setting $\phi = 1/2$ gives the Cox, Ingersoll, and Ross (1985) square root model. Imposing $\phi = 1$ yields the Brennan and Schwartz (1980) model, and the familiar geometric Brownian motion (GBM) process of Black and Scholes (1973) is nested within the Brennan and Schwartz model by the parameter restriction $\alpha_0 = 0$. In turn, the Dothan (1978) model is nested within the GBM model by the parameter restriction $\alpha_1 = 0$.

allow for nonlinear effects in the dynamics of the short rate because of the linear specification of the drift component. Second, the discrete-time approximation of the diffusion function given in eq. (3b) restricts the volatility to be a function of the interest-rate level only, and not of the news-arrival process. Therefore, the CKLS model fails to model adequately the serial correlation in conditional variances. An alternative model that addresses this in discrete time is the GARCH model, in which the current volatility is a function of the last period's unexpected news and the last period's volatility. A useful feature of the GARCH model is that it effectively can remove the excess kurtosis in interest-rate changes. The standard GARCH (1,1) model with linear conditional mean takes the following form:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \varepsilon_t \tag{4a}$$

$$E(\varepsilon_t|\Omega_{t-1}) = 0,\ E(\varepsilon_t^2|\Omega_{t-1}) \equiv \sigma_t^2 = \beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2\sigma_{t-1}^2 \tag{4b}$$

where $\beta_0 > 0$, $0 \le \beta_1 < 1$, $0 \le \beta_2 < 1$, and $\beta_1 + \beta_2 < 1$. In other words, in the GARCH (1,1) model, the effect of an interest-rate shock on current volatility declines geometrically over time. The introduction of conditional heteroskedasticity as a function of unexpected interest-rate shocks can capture leptokurtosis in the unconditional distribution.

In addition to the CKLS model, there are some weaknesses of the GARCH models as well. First, in direct contrast to much of the theoretical literature, GARCH models do not permit volatility to be a function of interest-rate levels. Second, most empirical applications to interest-rate data find that $\beta_1 + \beta_2 \cong 1$, which implies that the conditional-variance process is not covariance stationary.[5] In other words, current shocks affect volatility forecasts infinitely far into the future (or volatility shocks persist forever). Third, GARCH models permit negative interest rates.

Engle, Ng, and Rothschild (1992) found $\beta_1 + \beta_2 = 1.0096$ for a portfolio of U.S. Treasury securities. Koedijk, Nissen, Schotman, and Wolff (1997) reported $\beta_1 + \beta_2 = 1.10$ for one-month T-bills, Hong (1988) reported $\beta_1 + \beta_2 = 1.073$ for excess returns on three-month T-bills over one-month T-bills, and Engle, Lilien, and Robbins (1987) found similar results using an ARCH (12) model on quarterly data for excess-holding yield of six-month T-bills over three-month T-bills. When $\beta_1 + \beta_2 = 1$, the GARCH (1,1) volatility in equation (4b) can be rewritten as

[5]Bollerslev (1986) demonstrates that this is the case when $\beta_1 + \beta_2$ is greater than one. Nelson (1990) showed that in such cases shocks accumulate in the sense that $E(\sigma_{t+m}^2|\sigma_t^2) \to \infty$ as $m \to \infty$. Note, however, that this does not imply that $\sigma_{t+m}^2 \to \infty$ as $m \to \infty$.

$$\sigma_t^2 = \beta_0 + (1 - \beta_2)\varepsilon_{t-1}^2 + \beta_2\sigma_{t-1}^2. \tag{5}$$

Note that in this case, the unconditional variance is undefined. Indeed, the GARCH process in eq. (5) is a nonstationary GARCH called the integrated GARCH (IGARCH) model, originally introduced by Engle and Bollerslev (1986).

A stylized fact of financial time series is that the distribution of interest-rate changes can be skewed. For example, the short-term interest-rate change is skewed to the left, i.e., there are more negative than positive outlying observations.[6] Since, in a GARCH process given in eq. (4b), positive and negative interest-rate shocks of the same magnitude produce the same amount of volatility, the standard GARCH model cannot cope with such skewness. Hence, one can expect that forecasts and forecast-error variances from a symmetric GARCH model may be biased for skewed time series. Also, larger interest-rate shocks forecast more volatility at a rate proportional to the square of the size of the interest-rate shock. If a positive interest-rate shock causes more volatility than a negative interest-rate shock of the same size, the symmetric GARCH model underpredicts the amount of volatility following positive shocks and overpredicts the amount of volatility following negative shocks. Furthermore, if large interest- rate shocks cause more volatility than a quadratic function allows, then the symmetric GARCH model underpredicts volatility after a large interest-rate shock and overpredicts volatility after a small interest-rate shock.

Recently, a few modifications to the GARCH model, which explicitly take account of skewed distributions, have been proposed. In this article, we have considered two such modifications in which good news and bad news have different predictability for future volatility. The first model proposed to capture such asymmetric effects is the Asymmetric GARCH (AGARCH) model introduced by Engle (1990):

$$\sigma_t^2 = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2\sigma_{t-1}^2, \tag{6}$$

where $\beta_0 > 0$, $0 \leq \beta_1 < 1$, $0 \leq \beta_2 < 1$, $\beta_1 + \beta_2 < 1$, and importantly $\gamma > 0$ for the short-rate process. The asymmetry effect in the AGARCH model is captured by the volatility parameter γ. Since γ is assumed to be positive, an unexpected rise in the short rate (positive shock) increases predictable volatility more than an unexpected decrease of similar magnitude in the short rate (negative shock).

[6]See, for example, Broze et al. (1995), Gray (1996), and Koedijk et al. (1997).

The second model is the Quadratic GARCH (QGARCH) process, developed by Sentana (1995), which ensures positivity of the conditional variance. The QGARCH (1,1) model takes the following form:[7]

$$\sigma_t^2 = \beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2\sigma_{t-1}^2 + \gamma\, \varepsilon_{t-1}. \tag{7}$$

where $\beta_0 > 0$, $0 \leq \beta_1 < 1$, $0 \leq \beta_2 < 1$, $\beta_1 + \beta_2 < 1$, and $\gamma > 0$.[8] Similar to the AGARCH model, when $\gamma > 0$, positive interest-rate shocks (or $\varepsilon_{t-1} > 0$) will have a larger impact on σ_t^2 than negative interest-rate shocks (or $\varepsilon_{t-1} > 0$). The time-series properties of the QGARCH model are very similar to those of Bollerslev's (1986) GARCH model. However, the QGARCH model has the advantage that by adding a single parameter (in this case, γ), it can allow for both an asymmetric effect on the conditional variance and higher (unconditional) kurtosis, which goes in the right direction towards capturing some of the stylized facts characterizing many financial time series.[9]

ALTERNATIVE MODELS OF THE SHORT-TERM INTEREST RATE

In order to compare the relative performance of a variety of well-known models in capturing the stochastic behavior of the short rate, this section

[7]The QGARCH formulation can be interpreted as a second-order Taylor approximation to the unknown conditional variance function $\sigma^2(\quad)$, or, alternatively, as the quadratic projection of the square innovation on the information set. The main purpose of the QGARCH model is to use functional forms for $\sigma^2(\quad)$, which ensures positivity. To make the nature of the QGARCH model more clear, consider QARCH (2) process:

$$\sigma_t^2 = (\omega_0 + \omega_1\varepsilon_{t-1} + \omega_2\varepsilon_{t-2})^2 = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2\varepsilon_{t-2}^2 + \gamma_1\varepsilon_{t-1} + \gamma_2\varepsilon_{t-2} + \delta\, \varepsilon_{t-1}\varepsilon_{t-2}$$

where $\beta_0 = \omega_0^2$, $\beta_1 = \omega_1^2$, $\beta_2 = \omega_1^2$, $\gamma_1 = 2\omega_0\omega_1$, $\gamma_2 = 2\omega_0\omega_2$, $\delta = 2\omega_1\omega_2$. The generalized version of the QARCH (2) model, which is QGARCH (1,2), can be written as

$$\sigma_t^2 = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2\varepsilon_{t-2}^2 + \gamma_1\varepsilon_{t-1} + \gamma_2\varepsilon_{t-2} + \delta\varepsilon_{t-1}\varepsilon_{t-2} + \beta_3\sigma_{t-1}^2.$$

[8]In the QGARCH (1,1) case, $\sigma_t^2 = \beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2\, \sigma_{t-1}^2 + \gamma\varepsilon_{t-1}$, positivity of the variance is achieved if $\beta_1, \beta_2 \geq 0$ and $\gamma \leq 4\beta_0\beta_1$ (Sentana, 1995, p. 652). To impose these restrictions, one can estimate the likelihood function with $\sigma_t^2 = \rho_0^2 + \rho_1^2(\varepsilon_{t-1} - \mu)^2 + \rho_2^2\, \sigma_{t-1}^2$, so that $\beta_0 = \rho_0^2 + \rho_1^2\mu^2$, $\beta_1 = \rho_1^2$, $\beta_2 = \rho_2^2$ and $\gamma = -2\rho_1^2\,\mu$.

[9]In addition to AGARCH and QGARCH models, alternative asymmetric GARCH models that can generate skewed time-series patterns are the so-called Exponential GARCH models (EGARCH) proposed by Nelson (1991) and the Threshold GARCH model (TGARCH) introduced by Zakoian (1994) and Glosten, Jagannathan, and Runkle (1993). Although we have considered these models as possible useful candidates at an earlier stage of the study, they were found not to be very useful for repeated forecasting exercises. Given the latter purpose of repeatedly specifying and estimating GARCH models, an obvious requirement of these models is that the estimation method be reasonably simple and that parameter convergence occurs with reasonable quickness. Unfortunately, it has been our experience while running the estimation procedures that parameter estimation of the EGARCH and TGARCH models can be tedious. In view of these difficulties, we have restricted the specification of asymmetry in conditional variance to follow AGARCH and QGARCH processes.

introduces an alternative class of discrete-time interest-rate models. Our approach exploits the fact that many diffusion and GARCH models imply dynamics for the short-term riskless rate that can be nested within the following two-factor-continuous-time-stochastic-volatility model:

$$dr_t = (\alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \alpha_3/r_t)\, dt + \Psi_t r_t^{\phi}\, dW_{1,t} \tag{8a}$$

$$d\Psi_t^2 = (\kappa_0 + \kappa_1 \Psi_t^2)\, dt + \eta dW_{2,t} \tag{8b}$$

where $W_{1,t}$ and $W_{2,t}$ are independent standard Brownian motion processes so that $dW_{1,t}$ and $dW_{2,t}$ are normally distributed with zero mean and variance of dt. The specification of the short-rate process in eqs. (8a) and (8b) implies mean reversion of the interest-rate level, as well as mean reversion of the variance of interest-rate changes.[10]

We introduced a two-factor-interest-rate model as an extension of the one-factor diffusion models of Chan et al. (1992) and Ait-Sahalia (1996a) by incorporating a stochastic volatility factor in the short-rate process. The two-factor-discrete-time-stochastic-volatility model accommodates both the level and GARCH effects. The two factors of the model are the short-term interest rate and the instantaneous variance of changes in the short-term interest rate. It is known that the current level and volatility of interest-rate changes are two most important factors in explaining movements in the term structure of interest rates.

The diffusion model given in eq. (8) is a more-general version of the standard diffusion models. The model specifies a nonlinear-mean-reverting drift and allows the diffusion parameter Ψ_t to vary through time as new information arrives. To make the model flexible enough, and to encompass the popular symmetric and asymmetric GARCH models, we employed the discrete-time version of the model in our empirical analysis. The discrete-time approximation of eq. (8) can be written as:

$$r_{t+\Delta} - r_t = (\alpha_0 + \alpha_1 r_t + \alpha_2\, r_t^2 + \alpha_3/r_t)\, \Delta + \Psi_t r_t^{\phi} z_{1,t} \sqrt{\Delta} \tag{9a}$$

$$\Psi_{t+\Delta}^2 - \Psi_t^2 = (\kappa_0 + \kappa_1\, \Psi_t^2)\Delta + \eta\, z_{2,t} \sqrt{\Delta} \tag{9b}$$

where Δ is the length of the time interval, $z_{1,t}$ and $z_{2,t}$ are independent normal $N(0,1)$ variates so that $\Delta W_{1,t} = z_{1,t} \sqrt{\Delta}$ and $\Delta W_{2,t} = z_{2,t} \sqrt{\Delta}$ can be viewed as normally distributed random variables with zero mean and variance of Δ. Since $W_{1,t}$ and $W_{2,t}$ are assumed to be independent Wiener processes, there is no covariance between $\Delta W_{1,t}$ and $\Delta W_{2,t}$, i.e., cov(Δ

[10]The short rate process in eq. (8) exhibits nonlinear mean reversion if $\alpha_1 \geq 0$, $\alpha_2 \leq 0$, and $\alpha_3 \geq 0$. The diffusion process in eq. (8b) also implies mean reversion of the interest-rate volatility if $\kappa_1 < 0$.

$W_{1,t}, \Delta W_{2,t}) = 0$. The time-varying scale factor $\Psi^2_{t+\Delta}$ in the diffusion component of Δr_t is defined as a function of unexpected information shocks to the interest-rate market.

A New Class of Models

To test the presence of nonlinearity and asymmetry in the diffusion function, we modified eq. (9b) by incorporating nonlinear asymmetric dynamics along with the level effect into the short-rate process, and rewrote the two-factor-discrete-time-stochastic-volatility model as:

$$r_{t+\Delta} - r_t = (\alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \alpha_3/r_t)\,\Delta + \Psi_t z_t r_t^{\phi}\sqrt{\Delta} \quad (10a)$$

$$\Psi^2_{t+\Delta} = \beta_0 + \beta_1(\Psi_t z_t + \gamma)^2 + \beta_2 \Psi_t^2 \quad (10b)$$

$$\Psi^2_{t+\Delta} = \beta_0 + \beta_1 \Psi_t^2 z_t^2 + \beta_2 \Psi_t^2 + \gamma \Psi_t z_t \quad (10c)$$

Assuming that $\Delta = 1$, the discrete-time econometric specification of eqs. (10a), (10b), and (10c) can be written as

$$r_t - r_{t-1} = (\alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3/r_{t-1}) + r_{t-1}^{\phi}\sqrt{h_t}\, z_t \quad (11a)$$

$$\varepsilon_t = \sqrt{h_t}\, z_t \text{ and } z_t \overset{iid}{\sim} N(0,\ 1) \quad (11b)$$

$$h_t = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2\, h_{t-1} \quad (11c)$$

$$h_t = \beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2\, h_{t-1} + \gamma\, \varepsilon_{t-1} \quad (11d)$$

The parameterization of the diffusion function in this model incorporates serially correlated unexpected news, asymmetry, and level effects into the conditional volatility. The specification of h_t in eqs. (11c) and (11d) introduce, respectively, AGARCH and QGARCH processes to the conditional volatility. The model with the nonlinear drift, $\mu(r_t) = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3/r_{t-1}$, and the nonlinear diffusion function, $\sigma^2(r_t) = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 h_{t-1})\, r_{t-1}^{2\phi}$, is called the AGARCH-X model. The model with the same nonlinear drift, $\mu(r_t) = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3/r_{t-1}$, and the diffusion function, $\sigma^2(r_t) = (\beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2\, h_{t-1} + \gamma\varepsilon_{t-1})\, r_{t-1}^{2\phi}$, is referred to as the QGARCH-X model.

The econometric specification of the two-factor-discrete-time-interest-rate model presented in eq. (11) defines a broad class of interest-rate processes, which includes many well-known diffusion and GARCH models. This model encompasses the diffusion, symmetric, and asymmetric GARCH models considered in this article. These nested models can be obtained from eq. (11) by imposing the appropriate restrictions on the parameters α's, β's, ϕ, and γ. More specifically, the AGARCH-X (or QGARCH-X) model nests AGARCH (or QGARCH) model [if $\phi = 0$], symmetric GARCH model [if $\phi = \gamma = 0$], IGARCH model [if $\phi = \gamma = 0$ and $\beta_1 = (1 - \beta_2)$], and CKLS model [if $\gamma = \beta_1 = \beta_2 = 0$], all of which are assumed to have nonlinear conditional mean. Our nonlinear empirical generalization also encompasses the original CKLS model, which uses the standard linear mean-reverting drift [if $\alpha_2 = \alpha_3 = \gamma = \beta_1 = \beta_2 = 0$].

Moving Average Models

A moving average is an average taken over a rolling window of a fixed number of data points. Each time the window is rolled over, the estimate is updated by adding the observation from the preceding time period and dropping the oldest observation. Consequently, the sample size remains fixed. Moving averages have been a useful tool in financial forecasting for many years. However, recent advances in time-series analysis allow a more critical view of the efficacy of such methods and propose a more accurate representation of the underlying volatility. The following subsections explain the constantly weighted moving average (CWMA) and exponentially weighted moving average (EWMA) volatility.

CWMA Volatility

The simplest way to estimate the volatility of the interest-rate change is to find the usual estimate of the standard deviation. This volatility estimator, labeled as CWMA, is given by:

$$\sigma_t^2 = \frac{\sum_{i=1}^{n} (r_{t-i} - r_{t-i-1})^2}{n - 1}. \tag{12}$$

This method describes a simple structure on how volatility evolves through time and places constant weights, $1/(n - 1)$, on past observations. In this study, n is chosen to be 400 months. For example, the first 400 observations in the sample (from January 1957 through April 1990) are used to compute an estimate of the variance for the 401st month.[11]

[11]CWMA volatility employed in this study is different from the standard historical volatility in the sense that the traditional historical volatility places no structure on how volatility might evolve over time and is taking no account of dynamic properties of interest rates, i.e., it is a completely static model.

There are two major drawbacks of the CWMA-volatility-estimation method. First, if interest-rate volatility clusters, then more recent rates should be given more weight. Interest rates in recent periods provide more information about the current volatility than the ones in the distant past. Second, the choice of the length of the observation period, n, is somewhat arbitrary. A constant weight, $1/(n - 1)$, is given to observations that occur within the most recent n periods, while no weight is given to observations beyond this window. This procedure appears to be ad hoc at best.

EWMA Volatility

An EWMA volatility places more weight on more recent observations, and this weighting is done by using a *smoothing constant* (or a dampening parameter) λ, where $0 < \lambda < 1$. The larger the value of λ, the more weight is placed on past observations, and so the smoother the series becomes. An n-period EWMA volatility of the interest-rate change is defined as

$$\sigma_t^2 = \frac{(r_{t-1} - r_{t-2})^2 + \lambda(r_{t-2} - r_{t-3})^2 + \lambda^2(r_{t-3} - r_{t-4})^2 + \ldots + \lambda^{n-1}(r_{t-n} - r_{t-n-1})}{1 + \lambda + \lambda^2 + \ldots + \lambda^{n-1}} \tag{13}$$

Since the denominator of eq. (13) converges to $1/(1 - \lambda)$ as $n \to \infty$, an infinite EWMA can be rewritten as[12]

$$\sigma_t^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} (r_{t-i} - r_{t-i-1})^2. \tag{14}$$

Of course, the interest-rate series must be truncated since only a finite number of observations are available for estimation. In this study, following J. P. Morgan's RiskMetrics™, the dampening parameter λ is chosen to be 0.90 for the 400-month estimation period.

DATA AND ESTIMATION

The data used in this study are three-month U.S. Treasury bill rates obtained from the IMF International Financial Statistics. The data are

[12]The infinite EWMA model given in eq. (13) is almost identical to the IGARCH model presented in eq. (5). The similarity of these two models can be shown by repeated substitutions in eq. (5) by replacing β_2 with λ:

$$\begin{aligned}\sigma_t^2 &= \beta_0 + (1 - \lambda)\varepsilon_{t-1}^2 + \lambda(\beta_0 + (1 - \lambda)\varepsilon_{t-2}^2 + \lambda(\beta_0 + (1 - \lambda)\varepsilon_{t-3}^2 + \lambda(\ldots \\ &= \beta_0/(1 - \lambda) + (1 - \lambda)(\varepsilon_{t-1}^2 + \lambda\varepsilon_{t-2}^2 + \lambda^2\varepsilon_{t-3}^2 + \ldots).\end{aligned} \tag{5'}$$

When $\beta_0 = 0$, the representation in 5′ is equivalent to the EWMA model in eq. (13). While in RiskMetrics™ the parameter λ is estimated in a sub-optimal way (see Alexander, 1996, p.237), the estimation procedure for the IGARCH model results in optimal smoothing factor.

monthly and cover the period from January 1957 to December 1997, giving a total of 492 observations. The three-month Treasury bill rates expressed in annualized form are discount on new issues of three-month bills, and they are based on averages of end-of-the-month bid and ask prices. Figure 1 plots the data, illustrating the dramatic changes in interest rates that occurred during the Federal experiment and the OPEC oil crises. Figure 1, lower panel, contains plots of monthly changes in short-term interest rates. Again, the volatility associated with the Federal experiment is striking. Volatility also is noticeably higher than average during the 1973–1975 period (the time of the OPEC oil crisis) and immediately after the October 1987 stock-market crash.

Table I displays the means, standard deviations, and first six autocorrelations of the three-month yield and the monthly changes in the three-month yield. We also reported the skewness and excess kurtosis statistics for testing the distributional assumption of conditional normality. The unconditional average level of the three-month yield is 5.843% with a standard deviation of 2.761%, whereas the monthly yield change averaged 0.004% with a standard deviation of 0.522%. The autocorrelations for the level decay slowly, and those of the first differences are generally small and are not systematically positive or negative. The statistics given in Table I show that the monthly changes in the short-term interest rates are negatively skewed, although the skewness statistics are not large. The excess kurtosis values are larger than zero, implying that the distribution of interest-rate levels and changes has fat tails compared with the normal distribution.[13]

The discrete-time econometric specifications are estimated using the maximum-likelihood method with the Berndt–Hall–Hall–Hausman algorithm. The procedure used in this study for estimating the parameters of the interest rate models closely follows that of earlier studies, although the present study differs from its predecessors in an important respect. Former studies computed the level and variance of interest-rate changes if the parameters in the drift and diffusion were independent of time. Whereas this study, in addition to the entire sample estimation, uses a rolling-regression procedure and generates one-step-ahead forecasts by estimating the time-varying parameters of the drift and diffusion functions.

[13]The interest-rate change tends to exhibit non-normal unconditional sampling distributions in the form of skewness, but is more pronounced in the form of excess kurtosis. The conditional normality assumption in GARCH generates some degree of unconditional excess kurtosis, but typically less than adequate to account fully for the fat-tailed properties of the data. One solution to the kurtosis problem is the adoption of conditional distributions with fatter tails than the normal distribution. Other attempts to model the excess conditional kurtosis in the interest-rate change include the estimates of the AGARCH and the QGARCH models, as well as a more general version of these.

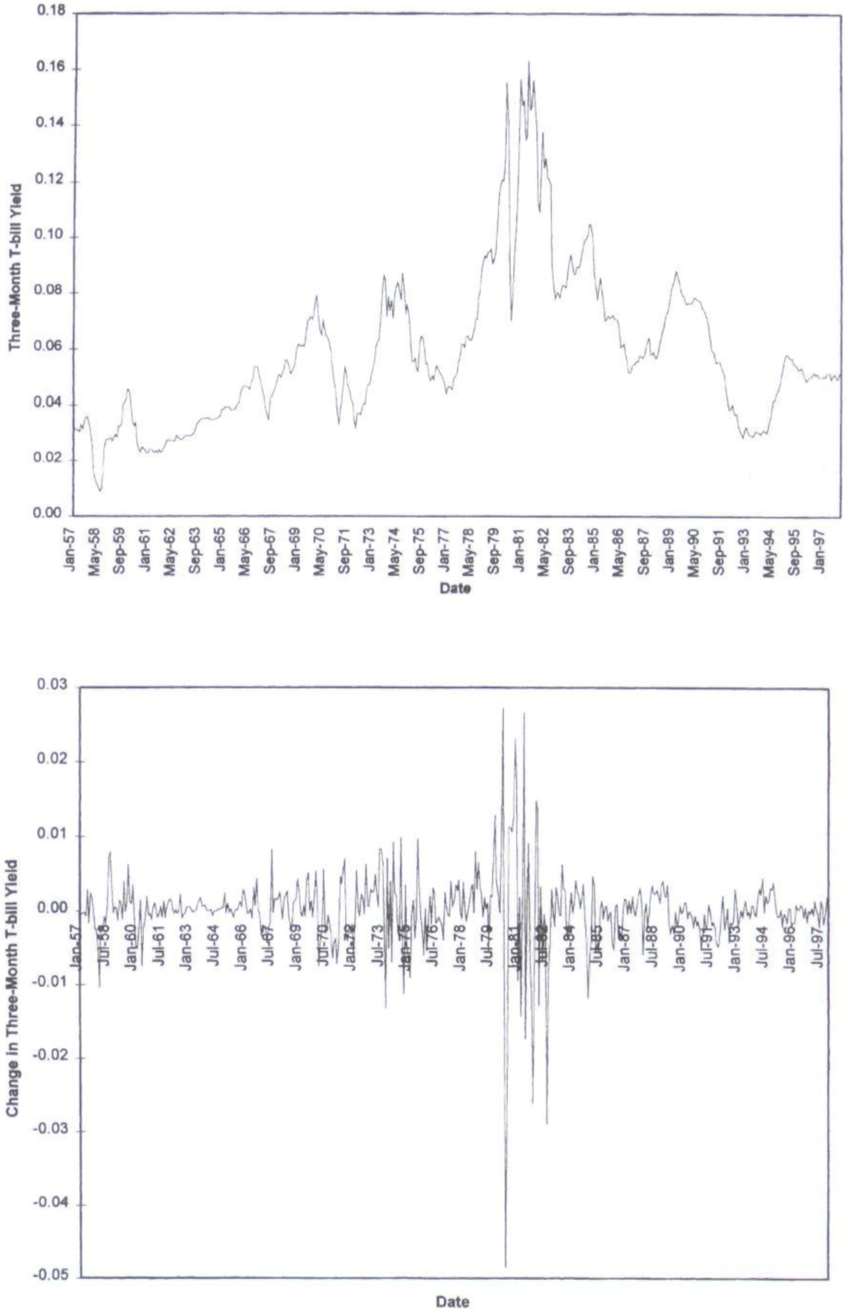

FIGURE 1

Three-Month Treasury Bill Yields.

The top panel contains a time-series plot of annualized three-month U.S. Treasury bill rates. The first differences of this series are shown in the bottom panel. The sample period is January 1957 to December 1997, a total of 492 observations.

TABLE I

Summary Statistics

Variables	*n*	*Mean*	*Standard Deviation*	*Skewness*	*Excess Kurtosis*	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6
r_t	492	0.05843	0.02761	1.2007	4.7440	0.981	0.952	0.927	0.906	0.885	0.864
$r_t - r_{t-1}$	491	0.00004	0.00522	−1.6489	24.742	0.277	−0.106	−0.094	−0.043	0.045	−0.188

Means, standard deviations, skewness, excess kurtosis, and autocorrelations of monthly Treasury bill yields and yield changes are computed from January 1957 to December 1997. The variable r_t denotes the annualized yield on Treasury bills maturing in three month and $r_t - r_{t-1}$ is the monthly yield change, ρ_j represents the autocorrelation coefficient of order j. n represents the number of observations used.

The construction of time-varying forecasts of the future level and variance of interest-rate changes can be explained as follows. The total sample of 492 observations is split into two parts: the first 400 observations (from January 1957 through April 1990) were used for estimation of the parameters of the model, and then a one-step-ahead forecast was calculated. The sample then was rolled forward by removing the first observation of the sample and adding one to the end, and another one-step-ahead forecast of the next month's conditional mean and variance was made. This "recursive" modeling and forecasting procedure was repeated until a forecast for observation 492 had been made using data available at time 491. Computation of forecasts using a rolling window of data ensures that the forecasts are made using models whose parameters have been estimated using all the information available at that time, while not incorporating old data.

EMPIRICAL RESULTS

This section first presents the maximum-likelihood-estimation results from estimating the conditional mean and variance of the interest-rate models and compares the existing models (diffusion and GARCH) with the two-factor-stochastic-volatility models (AGARCH-X and QGARCH-X) using the maximized log-likelihood functions. Second, the discrete-time econometric models are compared in terms of their ability to forecast the future volatility of the interest-rate change using the Theil inequality coefficient and an alternative measure of model performance. Third, the pricing implications of the stochastic volatility models are discussed.

Estimation Results and Model Comparisons

Tables II and III present, respectively, the in-sample and whole-sample parameter estimates, asymptotic t-statistics, maximized log-likelihoods

TABLE II

In-Sample Estimates of the Discrete-Time Interest-Rate Models

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	γ	ϕ	*Log Likelihood*	*LR*	*d.f.*
AGARCH-X	−0.0053	0.1510	−1.1968	0.00006	0.000001	1.0249	0.7045	0.00017	0.3893	2081.6036	—	—
Nonlinear	(−3.0459)	(3.6995)	(−4.1221)	(2.7897)	(1.5462)	(2.0618)	(23.324)	(1.9769)	(2.5656)			
QGARCH-X	−0.0053	0.1510	−1.1969	0.00006	0.000001	1.0250	0.7045	0.00036	0.3893	2081.6036	—	—
Nonlinear	(−3.0460)	(3.6998)	(−4.1225)	(2.7897)	(1.6354)	(2.0617)	(23.322)	(1.9533)	(2.5654)			
AGARCH	−0.0051	0.1497	−1.2272	0.00006	0.000001	0.3382	0.7125	0.00034	0.0	2079.1497	4.908	1
Nonlinear	(−2.0602)	(3.0435)	(−4.0491)	(1.6268)	(3.0891)	(7.3156)	(27.115)	(2.1633)				
QGARCH	−0.0051	0.1496	−1.2265	0.00006	0.000001	0.3382	0.7125	0.00023	0.0	2079.1497	4.908	1
Nonlinear	(−2.0598)	(3.0426)	(−4.0476)	(1.6269)	(3.3448)	(7.3170)	(27.121)	(2.1461)				
GARCH	−0.0053	0.1542	−1.2643	0.00006	0.000001	0.3444	0.7138	0.0	0.0	2077.6382	7.931	2
Nonlinear	(−2.0740)	(3.0682)	(−4.1048)	(1.5492)	(3.1192)	(7.0426)	(25.492)					
IGARCH	−0.0055	0.1611	−1.3286	0.00006	0.000001	0.2694	0.7306	0.0	0.0	2075.6453	11.917	3
Nonlinear	(−2.3660)	(3.4747)	(−4.7048)	(1.7779)	(3.4119)	(28.001)	(28.001)					
CKLS	−0.0040	0.1058	−0.7216	0.00005	0.0015	0.0	0.0	0.0	1.5209	1978.5396	206.1	3
Nonlinear	(−2.8530)	(2.8006)	(−3.0793)	(4.0643)	(7.8560)				(36.128)			
AGARCH-X	0.00067	−0.0073	0.0	0.0	0.000001	1.0164	0.7095	0.00015	0.3801	2077.4213	8.365	2
Linear	(2.8245)	(−1.3605)			(2.1387)	(3.0462)	(26.750)	(1.9592)	(3.5987)			
QGARCH-X	0.00067	−0.0073	0.0	0.0	0.000001	1.0164	0.7095	0.00033	0.3801	2077.4213	8.365	2
Linear	(2.8246)	(−1.3604)			(2.1385)	(3.0461)	(26.748)	(1.9662)	(3.5985)			
AGARCH	0.00084	−0.0113	0.0	0.0	0.000001	0.3431	0.7187	0.00034	0.0	2074.6677	13.872	3
Linear	(3.4349)	(−2.1292)			(3.1864)	(7.6921)	(29.926)	(2.1032)				
QGARCH	0.00084	−0.0113	0.0	0.0	0.000001	0.3431	0.7187	0.00023	0.0	2074.6677	13.872	3
Linear	(3.4348)	(−2.1291)			(3.3448)	(7.6925)	(29.926)	(2.0945)				
GARCH	0.00067	−0.0094	0.0	0.0	0.000001	0.3537	0.7167	0.0	0.0	2073.1681	16.871	4
Linear	(2.8510)	(−1.8176)			(3.4713)	(7.6413)	(28.513)					

TABLE II (Continued)

In-Sample Estimates of the Discrete-Time Interest-Rate Models

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	γ	ϕ	*Log Likelihood*	*LR*	*d.f.*
IGARCH	0.00076	−0.0119	0.0	0.0	0.000001	0.2601	0.7399	0.0	0.0	2070.1987	22.810	5
Linear	(3.1963)	(−2.5416)			(3.6243)	(33.512)	(33.512)					
CKLS	0.00082	−0.0115	0.0	0.0	0.0015	0.0	0.0	0.0	1.5080	1974.3605	214.5	5
Linear	(3.2198)	(−1.9535)			(8.7089)				(40.863)			

This table displays the in-sample-maximum-likelihood estimates of alternative models of the short-term interest rate r_t (annualized three-month U.S. Treasury bill rate) from January 1957 to April 1990 (400 observations). The parameter estimates with asymptotic t-statistics in parentheses are presented for each model. The maximized log-likelihoods for the interest-rate models (with linear and nonlinear conditional mean) are shown to compare the explanatory power of these models. Likelihood ratio tests evaluate restrictions imposed on the conditional mean and variance. Based on the likelihood ratio test, the LR statistics with associated degrees of freedom (d.f.) imply the presence of a nonlinearity in the conditional mean, asymmetry, and level effect in the conditional variance of the short rate. The parameters of the models are estimated from the following discrete-time econometric specifications:

$$r_t - r_{t-1} = (\alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3 / r_{t-1}) + r_{t-1}^{\phi} \sqrt{h_t}\, z_t,\ \varepsilon_t = \sqrt{h_t}\, z_t \text{ and } z_t \overset{iid}{\sim} N(0,1)$$

$$\sigma^2(r_t, t) = h_t\, r_{t-1}^{2\phi} \quad \text{where} \quad \begin{array}{ll} h_t = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 h_{t-1} & \text{(AGARCH-X)} \\ h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \gamma\, \varepsilon_{t-1} & \text{(QGARCH-X)} \end{array}$$

(1) AGARCH: $\sigma_t^2 = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 \sigma_{t-1}^2$

(2) QGARCH: $\sigma_t^2 = \beta_0 + \beta_1\, \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}$

(3) GARCH: $\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2$

(4) IGARCH: $\sigma_t^2 = \beta_0 + (1 - \beta_2)\varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2$

(5) CKLS: $\sigma_t^2 = \beta_0 r_{t-1}^{2\phi}$

TABLE III

Whole-Sample Estimates of the Discrete-Time Interest-Rate Models

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	γ	ϕ	Log Likelihood	LR	d.f.
AGARCH-X	−0.0032	0.0882	−0.7622	0.00005	0.000001	0.9763	0.7091	0.00009	0.3828	2623.9808	—	—
Nonlinear	(−1.9725)	(2.2472)	(−2.6752)	(2.1833)	(1.7953)	(2.3417)	(26.146)	(1.9499)	(2.8499)			
QGARCH-X	−0.0032	0.0882	−0.7622	0.00005	0.000001	0.9763	0.7091	0.00018	0.3828	2623.9808	—	—
Nonlinear	(−1.9725)	(2.2473)	(−2.6752)	(2.1834)	(1.8284)	(2.3416)	(26.147)	(1.9120)	(2.8500)			
AGARCH	−0.0032	0.0906	−0.8186	0.00004	0.000001	0.3288	0.7144	0.00019	0.0	2619.9932	7.975	1
Nonlinear	(−1.3893)	(1.9468)	(−2.7917)	(1.2958)	(3.6494)	(7.8894)	(28.247)	(1.9951)				
QGARCH	−0.0032	0.0905	−0.8179	0.00005	0.000001	0.3288	0.7144	0.00012	0.0	2619.9932	7.975	1
Nonlinear	(−1.3880)	(1.9450)	(−2.7892)	(1.2949)	(3.6953)	(7.8914)	(28.259)	(2.0156)				
GARCH	−0.0032	0.0906	−0.8239	0.00005	0.000001	0.3302	0.7161	0.0	0.0	2617.0867	13.788	2
Nonlinear	(−1.3689)	(1.9274)	(−2.7892)	(1.2330)	(3.6713)	(7.9084)	(28.490)					
IGARCH	−0.0034	0.0984	−0.8924	0.00004	0.000001	0.2707	0.7293	0.0	0.0	2614.9210	18.120	3
Nonlinear	(−1.5980)	(2.2672)	(−3.3210)	(1.4106)	(3.9776)	(30.749)	(30.749)					
CKLS	−0.0034	0.0855	−0.5984	0.00004	0.0014	0.0	0.0	0.0	1.5474	2482.1977	283.57	3
Nonlinear	(−2.8426)	(2.6807)	(−3.0052)	(4.3746)	(9.1969)				(43.837)			
AGARCH-X	0.00083	−0.0146	0.0	0.0	0.000001	0.9543	0.7106	0.00008	0.3682	2618.8669	10.228	2
Linear	(3.8282)	(−2.9897)			(2.1857)	(3.0873)	(28.062)	(1.9687)	(3.5668)			
QGARCH-X	0.00083	−0.0146	0.0	0.0	0.000001	0.9543	0.7106	0.00016	0.3683	2618.8669	10.228	2
Linear	(3.8282)	(−2.9897)			(2.2318)	(3.0873)	(28.062)	(1.9888)	(3.5669)			
AGARCH	0.00092	−0.0170	0.0	0.0	0.000001	0.3358	0.7158	0.00018	0.0	2615.7590	16.444	3
Linear	(4.3591)	(−3.7462)			(3.8140)	(8.2584)	(30.696)	(2.0966)				
QGARCH	0.00092	−0.0170	0.0	0.0	0.000001	0.3358	0.7158	0.00012	0.0	2615.7590	16.444	3
Linear	(4.3590)	(−3.7461)			(3.9073)	(8.2587)	(30.697)	(2.1011)				
GARCH	0.00085	−0.0165	0.0	0.0	0.000001	0.3382	0.7165	0.0	0.0	2613.0541	21.853	4
Linear	(3.9790)	(−3.5548)			(3.8973)	(8.2441)	(30.502)					

TABLE III (Continued)

Whole-Sample Estimates of the Discrete-Time Interest-Rate Models

Model	α_0	α_1	α_2	α_3	β_0	β_1	β_2	γ	ϕ	*Log Likelihood*	*LR*	*d.f.*
IGARCH	0.00090	−0.0177	0.0	0.0	0.000001	0.2659	0.7341	0.0	0.0	2611.8558	24.250	5
Linear	(4.2242)	(−4.1234)			(4.1161)	(34.860)	(34.860)					
CKLS	0.00081	−0.0133	0.0	0.0	0.0014	0.0	0.0	0.0	1.5325	2477.9589	292.04	5
Linear	(3.8064)	(−2.6407)			(10.175)				(49.233)			

This table displays the whole-sample maximum-likelihood estimates of alternative models of the short-term interest rate r_t (annualized three-month U.S. Treasury bill rate) from January 1957 to December 1997 (492 observations). The parameter estimates with asymptotic t-statistics in parentheses are presented for each model. The maximized log-likelihoods for the interest-rate models (with linear and nonlinear conditional mean) are shown to compare the explanatory power of these models. Likelihood ratio tests evaluate restrictions imposed on the conditional mean and variance. Based on the likelihood ratio test, the LR statistics with associated degrees of freedom (d.f.) imply the presence of nonlinearity in the conditional mean, asymmetry and level effect in the conditional variance of the short rate. The parameters of the models are estimated from the following econometric specifications:

$$r_t - r_{t-1} = (\alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + \alpha_3 / r_{t-1}) + r_{t-1}^{\phi} \sqrt{h_t}\, z_t,\ \varepsilon_t = \sqrt{h_t}\, z_t \text{ and } z_t \overset{iid}{\sim} N(0,1)$$

$$\sigma^2(r_t, t) = h_t r_{t-1}^{2\phi} \quad \text{where} \quad h_t = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 h_{t-1} \quad \text{(AGARCH-X)}$$

$$h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \gamma \varepsilon_{t-1} \quad \text{(QGARCH-X)}$$

(1) AGARCH: $\sigma_t^2 = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 \sigma_{t-1}^2$

(2) QGARCH: $\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 + \gamma \varepsilon_{t-1}$

(3) GARCH: $\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2$

(4) IGARCH: $\sigma_t^2 = \beta_0 + (1 - \beta_2)\varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2$

(5) CKLS: $\sigma_t^2 = \beta_0 r_{t-1}^{2\phi}$

for the unrestricted models (AGARCH-X and QGARCH-X) and for each of the nested models (AGARCH, QGARCH, GARCH, IGARCH, and CKLS), and the likelihood ratio tests comparing the nested models with the unrestricted models.[14] The coefficients of the variables in both the conditional mean and conditional variance of the interest-rate models usually are significant statistically at the 5-% level. Both the in-sample and whole-sample parameter estimates of the AGARCH-X and QGARCH-X models turn out to be almost the same. The drift parameters α_2 and α_3 in each econometric specification considered in the article are, respectively, negative and positive, as well as significant, suggesting the presence of mean reversion and nonlinearity in the conditional mean. The asymmetry parameter γ in the quadratic form in the AGARCH and AGARCH-X models, and the coefficient on the last period's unpredictable interest-rate change, ε_{t-1}, in the QGARCH and QGARCH-X models are significant and positive. The symmetric GARCH model has lower maximized-log-likelihood values than the asymmetric GARCH model, suggesting the existence of asymmetric volatility responses to shocks in the short-rate process. In addition, the sensitivity parameters, ϕ's, in the conditional variance of the AGARCH-X, QGARCH-X, and CKLS models also are significant and positive. All these results are consistent with the hypothesis that positive interest-rate shocks cause higher volatility than negative shocks, and volatility tends to be higher when the short rate is high.

Based on the maximized-log-likelihood values, the AGARCH-X and QGARCH-X models introduced in this study performed the best compared to the nested models.[15] According to the likelihood ratio test statistics presented in Tables II and III, the nested models are rejected strongly against the more-flexible AGARCH-X or QGARCH-X models for both the in-sample and whole-sample estimation periods. Among the nested models, the relative performance of the asymmetric GARCH models (AGARCH and QGARCH) in modeling the conditional distribution of short rates is found to be superior to the symmetric GARCH, IGARCH, and CKLS models. Overall, the nonlinear empirical generalizations outperform all other models in capturing the dynamic behavior of the three-month U.S. T-bill rates. All these results point to the presence of nonlin-

[14] As discussed in Existing Models of the Short-Term Interest Rate, the CKLS model encompasses many popular continuous-time models of the short-term interest rate. Since the CKLS model is nested within our more flexible models, and it covers most of the *subnested* models (e.g., Brennan & Schwartz, 1980; Cox, 1975; Cox & Ross, 1976; Cox et al., 1980, 1985; Dothan, 1978; Merton, 1973; Rendlemann & Bartter, 1980; Vasicek, 1977) we will not compare the relative performance of these subnested models with the newly proposed discrete-time interest-rate models.

[15] For the entire sample estimation, the maximized-log-likelihood values for the nonlinear mean AGARCH-X and QGARCH-X models are about 2623.98, whereas the maximized log likelihoods for the nonlinear mean GARCH and CKLS models are about 2617.09 and 2482.20, respectively.

earity in the conditional mean, and serial correlation, asymmetry, and level effects in the conditional variance of the short rate.

We found for the three-month Treasury-bill data that the volatility of the short-term interest rate is not as highly sensitive to the level of the short rate, as mentioned by Chan et al. (1992). In fact, the volatility of interest-rate changes was found to be more sensitive to unexpected news in the interest-rate market than to the level of the riskless rate. This can be understood simply by comparing the maximized-log-likelihood values of the CKLS and GARCH models. The empirical results also suggested that the diffusion models (like the CKLS model with linear or nonlinear drift) tend to overemphasize the sensitivity of volatility to interest-rate levels. For example, as presented in Tables II and III, the unconstrained estimate of the sensitivity parameter ϕ in the CKLS model is around 1.52 or 1.54 for the in-sample or whole-sample estimation period, whereas when the sensitivity of volatility to interest-rate levels depends on information flows (i.e., when $h_t = \beta_0 + \beta_1(\varepsilon_{t-1} + \gamma)^2 + \beta_2 h_{t-1}$ or $h_t = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 h_{t-1} + \gamma\varepsilon_{t-1}$) as in the case of the AGARCH-X or QGARCH-X models where the parameter ϕ reduces to 0.39 or 0.38 for the in-sample or whole-sample estimation period.

In order to determine the correct specification of both the drift and diffusion functions of the short-rate process, the nonlinearity in the conditional mean, asymmetry, and level effects in the conditional variance of interest-rate changes formally were tested for the in-sample and whole-sample estimation periods using the Likelihood Ratio (LR) test. According to the LR test statistics demonstrated in Table IV, the models with the standard linear mean-reverting drift were rejected strongly against those with the nonlinear conditional mean. In addition, the LR statistics indicated the strong presence of asymmetry and level effects in the conditional variance of the short rate since the symmetric GARCH models were rejected against the asymmetric GARCH models, and the models (AGARCH and QGARCH) that do not relate the volatility to interest-rate levels are rejected against the AGARCH-X and QGARCH-X models that incorporate the level effect into the diffusion function.

An Alternative Measure of Model Performance

In order to measure further the relative performance of the nested models against the more-general AGARCH-X and QGARCH-X models, we tested their predictive power for squared interest-rate changes, which provided a simple ex-post measure of interest-rate volatility. This was done by first computing the time-varying conditional variance of the monthly yield

TABLE IV

Likelihood Ratio Tests

Whole-Sample Estimates (n = 492)			*In-Sample Estimates (n = 400)*		
Nonlinearity *d.f. = 2*	*Level* *d.f. = 1*	*Asymmetry* *d.f. = 1*	*Nonlinearity* *d.f. = 2*	*Level* *d.f. = 1*	*Asymmetry* *d.f. = 1*
(LR = 10.228) AGARCH-X (linear vs nonlinear)	(LR = 7.975) AGARCH-X vs AGARCH (nonlinear)	(LR = 5.813) AGARCH vs GARCH (nonlinear)	(LR = 8.365) AGARCH-X (linear vs nonlinear)	(LR = 4.908) AGARCH-X vs AGARCH (nonlinear)	(LR = 4.023) AGARCH vs GARCH (nonlinear)
(LR = 10.228) QGARCH-X (linear vs nonlinear)	(LR = 7.975) QGARCH-X vs QGARCH (nonlinear)	(LR = 5.813) QGARCH vs GARCH (nonlinear)	(LR = 8.365) QGARCH-X (linear vs nonlinear)	(LR = 4.908) QGARCH-X vs QGARCH (nonlinear)	(LR = 4.023) QGARCH vs GARCH (nonlinear)
(LR = 8.468) AGARCH (linear vs nonlinear)	(LR = 6.216) AGARCH-X vs AGARCH (linear)	(LR = 3.910) AGARCH vs GARCH (linear)	(LR = 8.964) AGARCH (linear vs nonlinear)	(LR = 5.507) AGARCH-X vs AGARCH (linear)	(LR = 3.999) AGARCH vs GARCH (linear)
(LR = 8.468) QGARCH (linear vs nonlinear)	(LR = 6.216) QGARCH-X vs QGARCH (linear)	(LR = 3.910) QGARCH vs GARCH (linear)	(LR = 8.964) QGARCH (linear vs nonlinear)	(LR = 5.507) QGARCH-X vs QGARCH (linear)	(LR = 3.999) QGARCH vs GARCH (linear)
(LR = 6.565) GARCH (linear vs nonlinear)			(LR = 7.940) GARCH (linear vs nonlinear)		
(LR = 6.130) IGARCH (linear vs nonlinear)			(LR = 9.893) IGARCH (linear vs nonlinear)		
(LR = 8.478) CKLS (linear vs nonlinear)			(LR = 8.358) CKLS (linear vs nonlinear)		

This table displays the Likelihood Ratio (LR) test statistics for testing nonlinearity in the conditional mean, level and asymmetry in the conditional variance of the interest-rate change with the whole-sample and in-sample estimates. The critical values at the 5% level of significance are $\chi_{(1,0.05)} = 3.84$ and $\chi_{(2,0.05)} = 5.99$ with 1 and 2 degrees of freedom, respectively.

changes for each model using the fitted values of the discrete-time econometric specifications described previously. We then computed the proportion of the total variation in squared yield changes that can be explained by the estimated conditional variances. We refer to this as the coefficient of determination R^2. The R^2 values in this article provide information about how well each model is able to forecast the future volatility of the short-term interest-rate change.

The predictive power of the unrestricted and restricted models is evaluated by estimating a series of OLS regressions of the form:

$$\sigma_t^2 = a_0 + a_1\sigma_{t,f}^2 + \xi_t \tag{15}$$

where $\sigma_t^2 = (r_1 - r_{t-1})^2$ is the squared interest-rate changes at time t, $\sigma_{t,f}^2$ is the forecast of σ_t^2, and ξ_t is the forecast error. In addition to comparing the determination coefficients of the regressions for each interest-rate model, we also tested whether the estimates of a_0 and a_1 are equal to zero and one, respectively, in order to determine if the models' forecasts of conditional variances are "unbiased." The Wald test was used to determine whether $\sigma_{t,f}^2$ is an unbiased estimator of σ_t^2.

The results are presented in Table V. The AGARCH-X and QGARCH-X models (with linear and nonlinear drift) appear the same in their forecasting ability for the variance of interest-rate changes, although the predictive power of these models with linear conditional mean is found to be inferior to those with nonlinear drift. The coefficients of determination for the nonlinear mean AGARCH-X or QGARCH-X models were around 56 and 50% for the whole-sample and out-of-sample forecasting, whereas the corresponding R^2 values for the linear mean AGARCH-X or QGARCH-X models were about 52 and 49%. The explanatory power of the asymmetric and symmetric GARCH models turned out to be almost the same, although the forecasting ability of these restricted models was found to be slightly inferior to the AGARCH-X or QGARCH-X models in both the whole-sample and out-of-sample forecasting. The predictive power of the nonlinear mean IGARCH model, with R^2's about 53 and 46% for the whole-sample and out-of-sample forecasting, was found to be slightly inferior to the asymmetric and symmetric GARCH models.

Among the nested models, the predictive power of the CKLS model, which parameterizes the volatility only as a function of interest-rate levels, turned out to be the lowest, since it does not incorporate the effect of serially correlated unexpected news and asymmetry into the definition of volatility. The proportion of the total variation in volatility captured by the CKLS model with nonlinear drift is about 12 and 19% in the out-of-sample and in-sample forecasting, respectively. With the standard linear mean-reverting drift, these R^2 values were reduced to 11 and 16% for the CKLS model. Since there was no variation in the conditional volatility of the short rate for the moving average models, we could not calculate the coefficient of determination for the constantly weighted (CWMA) and exponentially weighted (EWMA) moving-average models. However, for out-of-sample forecasting, when the rolling-regression procedure was applied to the CWMA and EWMA models, the time-varying conditional variance of the interest-rate change and, in turn, the corresponding R^2 values were obtained. In out-of-sample forecasting, the relative performance of the EWMA model, with $R^2 = 23\%$, in forecasting the future volatility turns out to be considerably superior to the CWMA model, with

TABLE V

Alternative Measure of Model Performance for the Conditional Variance

	Whole-Sample Forecasting (n = 492)				*Out-of-Sample Forecasting (n = 92)*			
Model	a_0	a_1	R^2	*WALD*	a_0	a_1	R^2	*WALD*
AGARCH-X	−0.00001	0.8894	0.5559	16.686	−0.00001	1.2767	0.4952	8.2930
(nonlinear)	(−1.5498)	(24.739)			(−2.8360)	(9.3957)		
QGARCH-X	−0.00001	0.8894	0.5559	16.686	−0.00001	1.2768	0.4952	8.2930
(nonlinear)	(−1.5498)	(24.739)			(−2.8360)	(9.3957)		
AGARCH	−0.00001	1.0821	0.5542	5.4032	−0.00001	1.1815	0.4929	7.9928
(nonlinear)	(−1.9313)	(24.658)			(−2.6015)	(9.4054)		
QGARCH	−0.00001	1.0821	0.5542	5.4032	−0.00001	1.1816	0.4929	7.9919
(nonlinear)	(−1.9313)	(24.658)			(−2.6016)	(9.4052)		
GARCH	−0.00001	1.0924	0.5541	6.1026	−0.00001	1.1142	0.4833	8.5655
(nonlinear)	(−1.9526)	(24.652)			(−2.4116)	(9.2254)		
IGARCH	−0.00001	1.2869	0.5322	27.668	−0.00001	1.2921	0.4631	7.2324
(nonlinear)	(−1.9344)	(23.585)			(−2.6669)	8.8601		
CKLS	−0.00004	2.8851	0.1977	24.124	0.00003	0.3122	0.1212	59.032
(nonlinear)	(−3.3231)	(9.7896)			(3.1235)	(3.5231)		
AGARCH-X	−0.00001	1.1145	0.5217	6.9275	−0.00001	1.2755	0.4930	7.9987
(linear)	(−1.8965)	(23.095)			(−2.7909)	(9.3540)		
QGARCH-X	−0.00001	1.1146	0.5218	6.9417	−0.00001	1.2755	0.4930	7.9987
(linear)	(−1.8969)	(23.097)			(−2.7909)	(9.3540)		
AGARCH	−0.00001	1.1073	0.5213	6.3950	−0.00001	1.1188	0.4803	8.2085
(linear)	(−1.8869)	(23.077)			(−2.3887)	(9.1716)		
QGARCH	−0.00001	1.1073	0.5213	6.3950	−0.00001	1.1188	0.4803	8.2085
(linear)	(−1.8869)	(23.077)			(−2.3887)	(9.1716)		
GARCH	−0.00001	0.8798	0.5168	15.831	−0.00001	1.1638	0.4794	7.3492
(linear)	(−1.3054)	(22.870)			(−2.4499)	(9.1538)		
IGARCH	−0.00001	1.2696	0.5010	22.171	−0.00001	1.3354	0.4515	7.0317
(linear)	(−1.8458)	(22.158)			(−2.6513)	(8.6557)		
CKLS	−0.00003	2.9746	0.1678	28.023	0.00001	0.4221	0.1151	48.112
(linear)	(−3.2632)	(8.8563)			(2.9982)	(3.4213)		
EWMA	—	—	—	—	−0.00001	1.4927	0.2345	3.2223
					(−1.7561)	(5.2798)		
CWMA	—	—	—	—	−0.00013	3.9942	0.0356	2889.8
					(−1.7765)	(1.8234)		

The predictive power of the discrete-time interest-rate models for the ex-post measure of the variance of the interest rate change, $\sigma_t^2 = (r_t - r_{t-1})^2$, is tested by the following regression:

$$\sigma_t^2 = a_0 + a_1\,\sigma_{t,f}^2 + \xi_t$$

where σ_t^2 and $\sigma_{t,f}^2$ are the ex-post measure and forecasted values of the variance of $(r_t - r_{t-1})$, respectively. R^2 is the coefficient of determination from the regression. The *t*-statistics in parentheses reflect standard errors computed using the White (1980) correction for heteroskedasticity. WALD is the Wald test statistic for determining whether $\sigma_{t,f}^2$ is an unbiased estimator of σ_t^2. The critical values for the Wald test are $\chi^2_{(2,0.05)} = 5.99$ and $\chi^2_{(2,0.01)} = 9.21$.

$R^2 = 4\%$. The predictive power of the EWMA model also was found to be substantially higher than the explanatory power of the CKLS model.

The Wald test statistics shown in Table V imply that the estimated conditional variances, $\sigma_{t,f}^2$, obtained from the new class of models and the nested models usually were biased estimators of σ_t^2. In terms of producing "unbiased" forecasters of the actual variance of the interest-rate change, the performance of the AGARCH model with nonlinear mean was found to be superior to the other models. In particular, CKLS and moving-average models yield strongly biased estimators of σ_t^2.

As an alternative to R^2 measures, we also computed the Theil inequality coefficient (TIC) for the conditional variance of the short-term interest-rate change. The Theil inequality coefficient always lies between 0 and 1, where 0 indicates a perfect fit:

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma_i^2 - \sigma_{i,f}^2)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma_i^2)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma_{i,f}^2)^2}} \tag{16}$$

where σ_t^2 and $\sigma_{t,f}^2$ denote, respectively, the actual and forecasted values of squared interest-rate change. The *TIC* values describe the fit of the discrete-time models for the actual variance of the interest rate change.

The results for the diffusion, GARCH, and moving-average models are presented in Table VI. The *TIC* values shown in Table VI highlight the forecasting superiority of the new class of models (AGARCH-X and QGARCH-X) versus the GARCH and CKLS models. Among the nested models, the asymmetric GARCH models have slightly more explanatory power than the symmetric GARCH and IGARCH models. Both the asymmetric and symmetric GARCH models turn out to be much superior to the CKLS and moving-average models in capturing the stochastic behavior of interest rates. As presented in Tables V and VI, the ranking of the models based on their predictive power for the future volatility of interest-rate changes is aligned closely to the ranking implied by the maximized log-likelihood values of these models.

We also computed the *bias*, *variance*, and *covariance* proportions of the forecast errors for the discrete-time models discussed in the article.[16]

[16]The mean squared forecast error can be decomposed as

$$\sum (\sigma_{f,t} - \sigma_t)^2/n = (\bar{\sigma}_f - \bar{\sigma})^2 + (S_f - S)^2 + 2(1 - \rho)S_f S$$

where $\bar{\sigma}_f$, $\bar{\sigma}$, S_f, and S are the means and standard deviations of σ_f and σ, respectively, and ρ is the correlation coefficient between and σ_f and σ. The proportions then are defined as

$$\text{Bias prop.: } \frac{(\bar{\sigma}_f^2 - \bar{\sigma}^2)^2}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n}, \quad \text{Variance prop.: } \frac{(S_f - S)^2}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n},$$

$$\text{Covariance prop.: } \frac{2(1 - \rho)S_f\, S}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n}$$

TABLE VI

Theil Inequality Coefficient for Forecasting the Conditional Variance

	Whole-Sample Forecasting (n = 492)				*Out-of-Sample Forecasting (n = 92)*			
Model	*TIC*	*Bias Proportion*	*Variance Proportion*	*Covariance Proportion*	*TIC*	*Bias Proportion*	*Variance Proportion*	*Covariance Proportion*
AGARCH-X (nonlinear)	0.3554	0.0143	0.0571	0.9306	0.3375	0.0757	0.2535	0.6708
QGARCH-X (nonlinear)	0.3554	0.0143	0.0571	0.9306	0.3375	0.0757	0.2535	0.6708
AGARCH (nonlinear)	0.3832	0.0039	0.2164	0.7818	0.3393	0.0581	0.3021	0.6398
QGARCH (nonlinear)	0.3832	0.0039	0.2164	0.7818	0.3393	0.0581	0.3022	0.6397
GARCH (nonlinear)	0.3850	0.0035	0.2253	0.7732	0.3401	0.0446	0.3842	0.5817
IGARCH (nonlinear)	0.4326	0.0001	0.3804	0.6217	0.3602	0.0317	0.3912	0.5771
CKLS (nonlinear)	0.8060	0.0036	0.8407	0.1577	0.6107	0.6409	0.0001	0.3629
AGARCH-X (linear)	0.3746	0.0120	0.0672	0.9228	0.3389	0.0719	0.2591	0.6690
QGARCH-X (linear)	0.3746	0.0120	0.0672	0.9228	0.3389	0.0719	0.2591	0.6690
AGARCH (linear)	0.4057	0.0028	0.2501	0.7491	0.3423	0.0420	0.3816	0.5870
QGARCH (linear)	0.4057	0.0028	0.2501	0.7491	0.3423	0.0420	0.3816	0.5870
GARCH (linear)	0.4067	0.0026	0.2559	0.7435	0.3428	0.0565	0.2952	0.6483
IGARCH (linear)	0.4466	0.0001	0.3761	0.6259	0.3689	0.0226	0.4227	0.5547
CKLS (linear)	0.8071	0.0039	0.8409	0.1573	0.6138	0.6498	0.0001	0.3539
EWMA	—	—	—	—	0.5130	0.0015	0.4750	0.5234
CWMA	—	—	—	—	0.7657	0.9691	0.0028	0.0021

This table demonstrates the Theil Inequality Coefficients (TIC) for the discrete-time interest-rate models. TIC always lies between zero and one, where zero indicates a perfect fit. These TIC values provide information about how well each model is able to forecast the actual monthly change in T-bill rates.

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_t^2 - \sigma_{t,f}^2)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_t^2)^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\sigma_{t,f}^2)^2}}$$

where σ_t^2 and $\sigma_{t,f}^2$ and the ex-post measure and forecasted values of the variance of $(r_t - r_{t-1})$, respectively. The proportions are defined as

$$\text{Bias prop.: } \frac{(\sigma_f^{-2} - \sigma^{-2})^2}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n}, \quad \text{Variance prop.: } \frac{(S_f - S)^2}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n}, \quad \text{Covariance prop.: } \frac{2(1-\rho)S_fS}{\sum(\sigma_{t,f}^2 - \sigma_t^2)^2/n}$$

where σ^{-2}, σ_f^{-2}, S and S_f are the means and standard deviations of σ_t^2 and $\sigma_{t,f}^2$, respectively, and ρ is the correlation coefficient between σ_t^2 and $\sigma_{t,f}^2$. The bias, variance, and covariance proportions must add up to one. If the forecast is good, the bias and variance proportions should be small so that most of the bias should be concentrated on the covariance proportion.

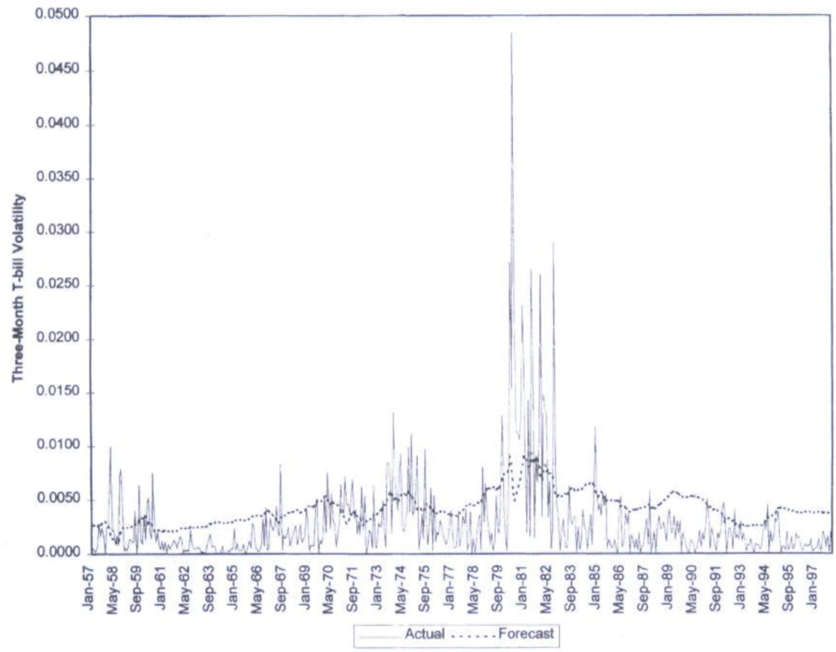

FIGURE 2

Forecasts of Three-Month Treasury-Bill Volatility Using the CKLS Model. This figure presents the time-series plot of the ex-post volatility (*solid line*), which is measured as the absolute value of the month-to-month change in the three-month Treasury bill yield, and the volatility forecast (*dashed line*), which is the square root of the conditional variance implied by the estimates of the CKLS model from January 1957 to December 1997.

As shown in Table VI, the bias proportion that tells us how far the mean of the forecast is from the mean of actual series is always much lower for the AGARCH-X or QGARCH-X models than for the nested and moving-average models. The variance proportion that measures how far the variation of the forecast is from the variation of the actual series is also much lower for the new class of models. In other words, the predictive power of our empirical generalizations is higher than the existing models since the bias and variance proportions are very small for the AGARCH-X or QGARCH-X models so that most of the bias is concentrated on the covariance proportion that measures the remaining unsystematic errors.

Further useful insights can be gained from Figures 2, 3, and 4, which present the time-series plot of the absolute value of the interest-rate

The bias, variance, and covariance proportions must add up to one. If the forecast is good, the bias and variance proportions should be small so that most of the bias should be concentrated on the covariance proportion.

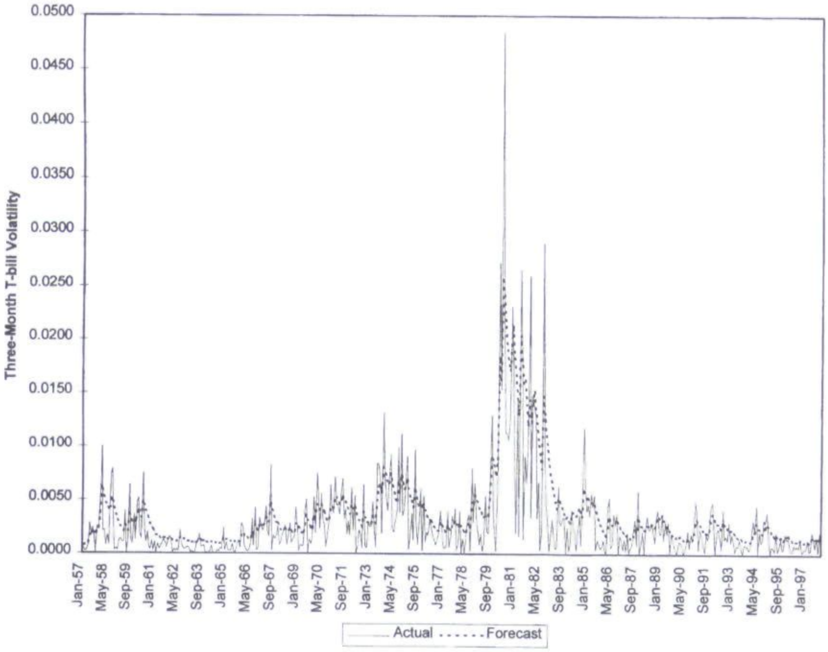

FIGURE 3

Forecasts of Three-Month Treasury-Bill Volatility Using the GARCH Model. This figure presents the time-series plot of the ex-post volatility (*solid line*), which is measured as the absolute value of the month-to-month change in the three-month Treasury bill yield, and the volatility forecast (*dashed line*), which is the square root of the conditional variance implied by the estimates of the GARCH model from January 1957 to December 1997.

change (i.e., the ex-post volatility) and the estimated conditional volatility obtained from the CKLS, GARCH, and AGARCH-X models, respectively. The ex-post volatility is measured as the absolute value of the month-to-month change in the three-month Treasury bill yield and the forecasted conditional volatility, $\sigma_{t,f}$, is the square root of the conditional variance implied by the estimates, $\sigma^2_{t,f}$, from the aforementioned models. To keep the graphs readable, we did not show all the volatility estimates in the same figure. The striking observation from these graphs is that in several periods, especially 1980 through 1983, the CKLS model in Figure 2 misrepresents realized volatility. In contrast, the GARCH model in Figure 3 tracks realized volatility much better during this period. The time-series plots of the estimated conditional volatilities along with the ex-post measure of the short-rate volatility depicted in Figure 4 suggest that the new class of models that explicitly allows for departures from linearity in the conditional mean, and incorporates serially correlated unexpected news,

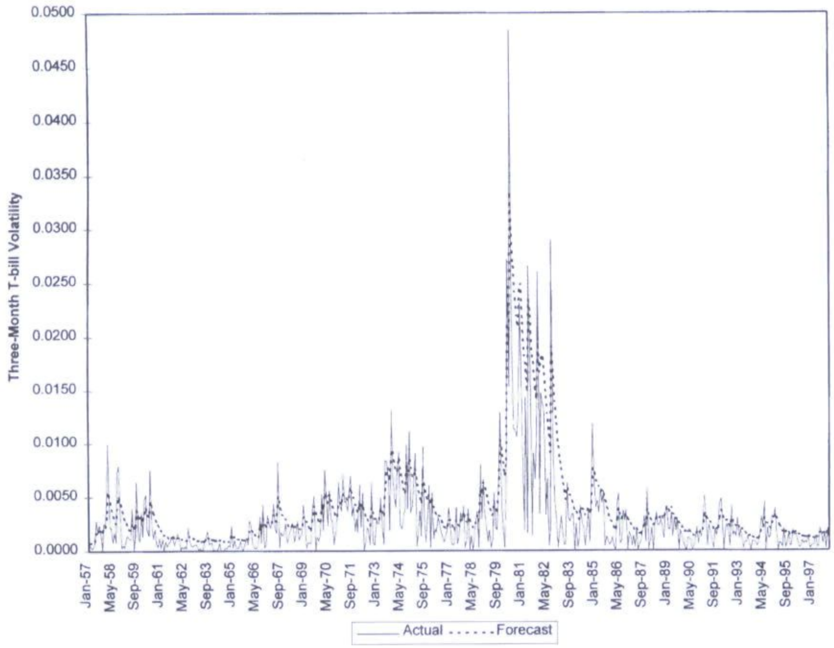

FIGURE 4

Forecasts of Three-Month Treasury-Bill Volatility Using the AGARCH-X Model. This figure presents the time-series plot of the ex-post volatility (*solid line*), which is measured as the absolute value of the month-to-month change in the three-month Treasury bill yield, and the volatility forecast (*dashed line*), which is the square root of the conditional variance implied by the estimates of the AGARCH-X model from January 1957 to December 1997.

asymmetry, and level effects into the diffusion function is superior to either of the baseline models (i.e., CKLS and GARCH).

Pricing Implications of the Stochastic Volatility Models

This section introduces a theoretical framework for the two-factor-stochastic-volatility model that values interest-rate-sensitive security. To explain how it works, we concentrated on valuing options on zero-coupon bonds. We discuss how the two-factor model uses current interest rates and their volatilities as input to determine its output, the value of interest-rate-sensitive options. This model, like other arbitrage-free interest-rate models, uses two sets of input: a *yield curve* (the yields for a set of zero-coupon Treasury securities) and a *volatility curve* (the estimated yield volatilities for these same securities). These two curves together are called the *term structure*. The model uses today's term structure to determine the expected means and standard deviations of future short rates.

For pricing implications, we used a restricted version of the two-factor-continuous-time model because the original model with nonlinear drift does not yield analytical solutions that can be used in the pricing of interest-rate derivatives. The two-factor-continuous-time model with the usual linear mean-reverting drift is given by eq. (17).

$$dr_t = (\alpha_0 + \alpha_1 r_t)\, dt + \Psi_t r_t^{\phi}\, dW_{1,t} \tag{17a}$$

$$d\Psi_t^2 = (\kappa_0 + \kappa_1\, \Psi_t^2)\, dt + \eta dW_{2,t}. \tag{17b}$$

The evolution of the short rate process presented in eqs. (17a) and (17b) can be approximated via the binomial representation:

$$r_{t+\Delta} - r_t = \begin{cases} (\alpha_0 + \alpha_1 r_t)\Delta + \Psi_t r_t^{\phi} \sqrt{\Delta}; & \text{probability } 1/2 \\ (\alpha_0 + \alpha_1 r_t)\Delta - \Psi_t r_t^{\phi} \sqrt{\Delta}; & \text{probability } 1/2 \end{cases} \tag{18a}$$

where

$$\Psi_{t+\Delta}^2 - \Psi_t^2 = \begin{cases} (\kappa_0 + \kappa_1\, \Psi_t^2)\Delta + \eta\sqrt{\Delta}; & \text{probability } 1/2 \\ (\kappa_0 + \kappa_1\, \Psi_t^2)\Delta - \eta\sqrt{\Delta}; & \text{probability } 1/2 \end{cases} \tag{18b}$$

At time t, if the interest-rate level is r_t, then one period later, using eq. (18a), the short rate will be either:

$$r_{t+\Delta}^{\text{up}} - r_t = (\alpha_0 + \alpha_1 r_t)\, \Delta + \Psi_t r_t^{\phi} \sqrt{\Delta} \tag{19a}$$

or

$$r_{t+\Delta}^{\text{down}} - r_t = (\alpha_0 + \alpha_1 r_t)\Delta - \Psi_t r_t^{\phi} \sqrt{\Delta} \tag{19b}$$

Subtracting eq. (19b) from eq. (19a) yields the *sensitivity formula,* which shows how the volatility of interest rate changes, $\sigma_t(r_t)$, in our stochastic-volatility model is determined by the spread between the two possible interest rates at time $t + \Delta$:

$$(r_{t+\Delta}^{\text{up}} - r_{t+\Delta}^{\text{down}})/2 = \Psi_t r_t^{\phi} \sqrt{\Delta} = \sigma_t(r_t). \tag{20}$$

In this model, following Black, Derman, and Toy (1990), we used the local-expectations hypothesis (LEH), which is consistent with no arbitrage. Applied to the valuation of interest-rate derivatives, the LEH yields the *valuation formula*:

$$P_t(T) = \frac{E_t[P_{t+\Delta}(T)]}{1 + r_t} = \frac{(1/2)P_{t+\Delta}(T)^{\text{up}} + (1/2)P_{t+\Delta}(T)^{\text{down}}}{1 + r_t} \tag{21}$$

where $P_t(T)$ is the price of a T-maturity claim at time t, which is assumed after Δ-period to move up to $P_{t+\Delta}(T)^{\text{up}}$ or down to $P_{t+\Delta}(T)^{\text{down}}$ with equal probability of 1/2. With the help of an iterative procedure, the valuation and sensitivity formulas are used to determine the evolution of the short-rate tree.

This model, like other arbitrage-free models, must ensure that model prices for interest-rate derivatives match market prices. In addition, the model's term structure of volatilities should match an empirical or given-term structure of volatilities. For example, Black et al. (1990) used a given-term structure based on the traditional historical volatilities that do not vary through time, although time-varying volatilities would match more closely the true underlying volatility structure. The use of time-varying volatilities leads, however, to a non-recombining tree. To illustrate this, consider the evolution of the short-rate process at time $t + 2\Delta$:

$$r^{uu}_{t+2\Delta} - r^{u}_{t+\Delta} = (\alpha_0 + \alpha_1 r^{u}_{t+\Delta})\Delta + \Psi^{u}_{t+\Delta}(r^{u}_{t+\Delta})^{\phi}\sqrt{\Delta} \tag{22a}$$

$$r^{ud}_{t+2\Delta} - r^{u}_{t+\Delta} = (\alpha_0 + \alpha_1 r^{u}_{t+\Delta})\Delta - \Psi^{u}_{t+\Delta}(r^{u}_{t+\Delta})^{\phi}\sqrt{\Delta} \tag{22b}$$

$$r^{du}_{t+2\Delta} - r^{d}_{t+\Delta} = (\alpha_0 + \alpha_1 r^{d}_{t+\Delta})\Delta + \Psi^{d}_{t+\Delta}(r^{d}_{t+\Delta})^{\phi}\sqrt{\Delta} \tag{22c}$$

$$r^{dd}_{t+2\Delta} - r^{d}_{t+\Delta} = (\alpha_0 + \alpha_1 r^{d}_{t+\Delta})\Delta - \Psi^{d}_{t+\Delta}(r^{d}_{t+\Delta})^{\phi}\sqrt{\Delta} \tag{22d}$$

Using eqs. (20), (22a), (22b), (22c), and (22d), one can show that

$$\left(\frac{r^{uu}_{t+2\Delta} - r^{ud}_{t+2\Delta}}{2}\right) = \sigma^{u}_{t+\Delta} \tag{23a}$$

$$\left(\frac{r^{du}_{t+2\Delta} - r^{dd}_{t+2\Delta}}{2}\right) = \sigma^{d}_{t+\Delta} \tag{23b}$$

$$\left(\frac{r^{ud}_{t+2\Delta} - r^{du}_{t+2\Delta}}{2}\right) = \sigma_t - \left(\frac{\sigma^{u}_{t+\Delta} + \sigma^{d}_{t+\Delta}}{2}\right) \tag{23c}$$

where $\sigma_t = \Psi_t r_t^{\phi}\sqrt{\Delta}$, $\sigma^{u}_{t+\Delta} = \Psi^{u}_{t+\Delta}(r^{u}_{t+\Delta})^{\phi}\sqrt{\Delta}$ and $\sigma^{d}_{t+\Delta} = \Psi^{d}_{t+\Delta}(r^{d}_{t+\Delta})^{\phi}\sqrt{\Delta}$. The mean reversion parameter, α_1, appears in eq. (23c). However, since it is a widely accepted idea that α_1 is a very small figure, we decided to eliminate it to simplify the framework.

An iterative procedure is used to find the expected future short rates and bond prices that will be consistent with the current market-term structure. Given a market-term structure and the resulting binomial tree of short rates, the model then can be used to price options on zero-coupon bonds. First, the zero-coupon bond prices at various times are found.

These prices are used to determine the option's value at expiration. Given the values of a call or put option at expiration, their possible values before expiration can be calculated as a present value of the future payoffs.

CONCLUSIONS

This article analyzes one potential source of misspecification of existing models of the short-term interest rate and introduces a new class of discrete-time econometric specifications that nests many existing interest-rate models as special cases. In existing continuous-time or time-series econometric models, the structural form of conditional means and variances is relatively inflexible in the sense that the existing models do not allow for departures from linearity in the conditional mean, and they do not parameterize the diffusion function flexible enough to incorporate serially correlated unexpected news, asymmetry, and level effects into the definition of conditional volatility. This study attempts to model the conditional distribution of interest rates by specifying a more general econometric framework that allows for nonlinear effects in the dynamics of the short rate and defines the conditional volatility as a nonlinear function of unexpected information shocks and interest-rate levels. The empirical results point to the presence of nonlinearity in the conditional mean, and serial correlation, asymmetry, and level effects in the conditional variance of the short rate. In addition, the relative performance of the new class of models in predicting the future level and variance of interest-rate changes has been found to be superior to the nested models.

BIBLIOGRAPHY

Ait-Sahalia, Y. (1996a). Testing continuous-time models of the spot interest rate. Review of Financial Studies, 9, 385–426.

Ait-Sahalia, Y. (1996b). Nonparametric pricing of interest rate derivatives. Econometrica, 64, 527–560.

Alexander, C. (1996). The handbook of risk management and analysis. Wiley: New York.

Bali, T. G. (1999). An empirical comparison of continuous time models of the short term interest rate. Journal of Futures Markets, 19, 777–798.

Bali, T. G., & Karagozoglu, A. (1999). Implementation of the BDT model with different volatility estimators: Applications to Eurodollar futures options. Journal of Fixed Income, 8, 24–34.

Barone, E., Cuoco, D., & Zautzik, E. (1991). Term structure estimation using the Cox, Ingersoll, and Ross model: The case of italian treasury bonds. Journal of Fixed Income, 1, 87–95.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81,637–654.

Black, F., Derman, E., & Toy, W. (1990). A one-factor model of interest rates and its applications to treasury bond options. Financial Analysts Journal, 46, 33–39.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. Journal of Econometrics, 31,307–327.

Bollerslev, T., Chou, R. T., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics, 52, 5–59.

Brennan, M. J., & Schwartz, E. S. (1980). Analyzing convertible bonds. Journal of Financial and Quantitative Analysis, 15, 907–929.

Brennan, M. J., & Schwartz, E. S. (1982). An equilibrium model of bond pricing and a test of market efficiency. Journal of Financial and Quantitative Analysis, 17, 75–100.

Broze, L., Scaillet, O., & Zakoian, J.-M. (1995). Testing for continuous-time models of the short-term interest rate. Journal of Empirical Finance 2, 199–223.

Cai, J. (1994). A Markov model of unconditional variance in ARCH. Journal of Business and Economic Statistics, 12, 309–316.

Cecchetti, S., Cumby, R., & Figlewski, S. (1988). Estimation of the optimal futures hedge. Review of Economics and Statistics, 70, 623–630.

Chan K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. Journal of Finance, 47, 1209–1227.

Chen, R.-R., & Scott, L. (1993). Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. Journal of Fixed Income, 3, 14–31.

Conley, T. G., Hansen, L. P., Luttmer, E. G. Z., and Scheinkman, J. A. (1997). Short-term interest rates as subordinated diffusions. Review of Financial Studies, 10, 525–577.

Cox, J. C. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Working paper, Stanford University.

Cox, J., J. Ingersoll, & Ross, S. (1980). An analysis of variable rate loan contracts. Journal of Finance, 35, 389–403.

Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. Econometrica, 53, 385–407.

Cox, J. C., & Ross, S. (1976). The valuation of options for alternative stochastic processes. Journal of Financial Economics, 3, 145–166.

Dietrich-Campbel, B., & Schwartz, E. S. (1986). Valuing debt options: Empirical evidence. Journal of Financial Economics, 16, 321–343.

Dothan, U. L. (1978). On the term structure of interest rates. Journal of Financial Economics, 7, 59–69.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance 7 of United Kingdom inflation. Econometrica, 50, 987–1007.

Engle, R. F. (1990). Discussion: Stock market volatility and the crash of '87. Review of Financial Studies, 3, 103–106.

Engle, R. F., & Bollerslev, T. (1986). Modeling the persistence of conditional variances. Econometric Reviews, 5, 1–50.

Engle, R. F., Lilien, D., & Robins, R. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model. Econometrica, 55, 391–407.

Engle, R. F., Ng, V., & Rothschild, M. (1992). Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills. Journal of Econometrics, 45, 213–238.

Gibbons, M. R., & Ramaswamy, K. (1993). A test of the Cox, Ingersoll, and Ross model of the term structure. Review of Financial Studies, 6, 619–658.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance, 48, 1779–1801.

Gray, S. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. Journal of Financial Economics, 42, 27–62.

Hong, C.-H. (1988). The integrated generalized autoregressive conditional heteroscedastic model: The process, estimation and Monte Carlo experiments. Working paper, University of California, San Diego, CA.

Koedijk, K. G., Nissen, F. G. J. A., Schotman, P. C., & Wolff, C. C. P. (1997). The dynamics of short-term interest rate volatility reconsidered. European Finance Review, 1, 105–130.

Marsh, T. A., & Rosenfeld, E. R. (1983). Stochastic processes for interest rates and equilibrium bond prices. Journal of Finance, 38, 635–646.

Melino, A., (1994). Estimation of continuous-time models in finance. In C. A. Sims (ed). Advances in econometrics: Sixth World Congress, Vol. II. Cambridge: Cambridge University Press.

Merton, R. C. (1973). Theory of rational option pricing. Bell Journal of Economics and Management Science, 4, 141–183.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH (1,1) model. Econometric Theory, 6, 318–344.

Nelson, D. B. (1991). Conditional heteroscedasticity in asset returns: A new approach. Econometrica, 59, 347–370.

Pearson, N. D., & Sun, T.-S. (1994). Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model. Journal of Finance, 49, 1279–1304.

Pfann, G. A., Schotman, P. C., & Tschernig, R. (1996). Nonlinear interest rate dynamics and implications for the term structure. Journal of Econometrics, 74, 149–176.

Rendleman, R., & Bartter, B. (1980). The pricing of options on debt securities. Journal of Financial and Quantitative Analysis, 15, 11–24.

Sanders, A. B., & Unal, H. (1988). On the intertemporal stability of the short term rate of interest. Journal of Financial and Quantitative Analysis, 23,417–423.

Sentana, E. (1995). Quadratic ARCH models. Review of Economic Studies, 62, 639–661.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. Journal of Finance, 52, 1973–2002.

Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5, 177–188.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica, 48, 817–838.

Zakoian, J.-M. (1994). Threshold heteroscedastic models. Journal of Economic Dynamics and Control, 18, 931–995.