An Empirical Comparison of Continuous Time Models of the Short Term Interest Rate

TURAN G. BALI

This article tests the performance of a wide variety of well-known continuous time models—with particular emphasis on the Black, Derman, and Toy (1990; henceforth BDT) term structure model—in capturing the stochastic behavior of the short term interest rate volatility. Many popular interest rate models are nested within a more flexible time-varying BDT framework that allows us to compare the models and find the proper specification of the dynamics of short rates. The empirical results indicate that the *equilibrium models* that do not allow the drift and diffusion parameters to vary over time and parameterize the volatility only as a function of interest rate levels overemphasize the sensitivity of volatility to the level of interest rate and fail to model adequately the serial correlation in conditional variances. On the other hand, the GARCH-based *arbitrage-free models* with time-dependent parameters in the drift and diffusion functions define the volatility only as a function of unexpected information shocks and fail to capture adequately the relationship between interest rate levels and volatility. This study shows that the most successful models in capturing the dynamics of short term interest rates are those that introduce time-dependent parameters to the short rate process and define the conditional volatility as a function of both the interest rate levels and the last period's unexpected news. © 1999 John Wiley & Sons, Inc. Jrl Fut Mark 19: 777–797, 1999

The author would like to thank the anonymous referee and Robert Webb for their helpful comments and suggestions.
Correspondence author, Department of Economics, Queens College, City University of New York, 65-30 Kissena Boulevard, Flushing, NY 11367-1597.

■ *Turan G. Bali is an Assistant Professor of Economics in the Department of Economics at Queens College, City University of New York in Flushing, New York.*

The Journal of Futures Markets, Vol. 19, No. 7, 777–797 (1999)

CCC 0270-7314/99/070777-21

INTRODUCTION

The specification of the short-term interest rate volatility is fundamental to much of theoretical and empirical finance. Yet no consensus has emerged on the choice of the parametric diffusion functions of the short rate process. This article tests the performance of a wide variety of well-known continuous time models—with particular emphasis on the Black, Derman, and Toy (1990; henceforth BDT) term structure model—in capturing the stochastic behavior of the short term interest rate volatility. A new class of discrete time interest rate models is introduced within the BDT framework to determine which model best captures the dynamics of short rates.

The correct specification of the dynamics of short rates are important for several reasons. The short term risk-free interest rate and its estimated volatility, driving the changes in the entire term structure, are the fundamental variables in most of the existing continuous time interest rate models. This makes the choice of a volatility estimation model for short term interest rates crucial to pricing bonds, pricing interest rate sensitive derivative securities, and hedging interest rate risk. For example, to price fixed income securities as well as interest rate options and futures correctly, we need to model both the instantaneous and time-series properties of interest rate volatility. This follows because both the current level of volatility and the stochastic properties of volatility affect dynamic hedge ratios and the distribution of future interest rate levels, which determine a derivative's price. The use of an incorrect volatility estimation model could lead to incorrect inferences, mishedged or unhedged risks, or pricing errors.

The econometric estimation of continuous time models of the short-term interest rate has been a relatively recent development in empirical finance (see, for example, Brown & Dybvig; 1986; Barone, Cuoco, & Zautzik, 1991; Chan et al., 1992; Chen & Scott, 1993; Gibbons & Ramaswamy, 1993; Pearson & Sun, 1994; Broze, Scaillet, & Zakoian, 1995; Nowman, 1997). In an important study, Chan, Karolyi, Longstaff, and Sanders (1992; hereafter CKLS) develop a general framework to estimate and compare a range of different single-factor term structure models for the short-term interest rate. A partial listing of these interest rate models includes those by Merton (1973), Cox (1975), Cox and Ross (1976), Vasicek (1977), Dothan (1978), Brennan and Schwartz (1980), Rendleman and Bartter (1980), Cox, Ingersoll, and Ross (1980, 1985), and Constantinides and Ingersoll (1984).

In this article, we introduce and test a new class of interest rate models that nests not only the aforementioned models but also the original and modified versions of the BDT (1990) term structure model. The BDT model that is extensively used by academics and Wall Street practitioners is based on common sense about the relation between today's long rates and future short rates. The current interest rates and their estimated volatilities are used in this model as input to determine its output; for example, the value of interest rate contingent claims.

We consider both *equilibrium* and *arbitrage-free* models of interest rate. In an equilibrium model the initial term structure is an output from the model, whereas in an arbitrage-free model the initial term structure is an input to the model. We test the predictive power of equilibrium and arbitrage-free models for both interest rate and squared interest rate changes using Theil Inequality Coefficient (TIC). The TIC values in this study illustrate that the primary distinguishing feature of the arbitrage-free models (in this study, the original BDT and its modified versions) with time-dependent parameters in the drift and diffusion is their higher ability to capture the time-varying volatility of the short term interest rate compared to the equilibrium models, such as those of Merton (1973), Vasicek (1977), Cox, Ingersoll, and Ross (1980, 1985; hereafter CIR), CKLS (1992), as well as others.

THE SHORT TERM INTEREST RATE MODELS

This article compares the relative performance of a wide variety of well-known models in capturing the stochastic behavior of the short rate. Our approach exploits the fact that many term structure models imply dynamics for the short term riskless rate that can be nested within the following continuous time differential equation:[1]

$$dr_t = (\alpha_{0,t} + \alpha_{1,t} r_t + \alpha_{2,t} r_t \ln r_t)\, dt + \Psi_t r_t^{\gamma_t} dW_t \tag{1}$$

where r_t is the interest rate level at time t and ln r_t is the natural logarithm of r_t. W_t is a standard Brownian motion so that dW_t is normally distrib-

[1]The discrete time approximation of (1) can be written as:

$$\Delta r_t = (\alpha_{0,t} + \alpha_{1,t}\, r_t + \alpha_{2,t}\, r_t \ln r_t)\, \Delta + \Psi_t r_t^{\gamma_t} \Delta W_t$$

$$\Delta W_t = \varepsilon_{t+\Delta}\sqrt{\Delta},\ \Psi^2_{t+\Delta} = \beta_{0,t} + \beta_{1,t}\varepsilon_t^2 + \beta_{2,t}\Psi_t^2,$$

where $\Delta r_t = r_{t+\Delta} - r_t$ is the change in interest rate level where Δ is the length of the time interval. $\varepsilon_{t+\Delta}$ is considered as an unexpected interest rate shock, which is a random drawing from a standardized normal distribution with a mean of zero and a standard deviation of one. W_t is a standard Brownian motion at time t so that ΔW_t is a normally distributed random variable with zero mean and variance of Δ. The time-varying scale factor $\Psi^2_{t+\Delta}$ is defined as a function of unexpected interest rate shocks. α_t's and β_t's are time-varying parameters that must be estimated.

uted with zero mean and variance of *dt*. As discussed in footnote 1, Ψ_t is the time-varying scale factor and it depends on the last period's unexpected news in the interest rate market. γ_t is the sensitivity parameter that allows the volatility to depend on the level of the interest rate. In this model, $\mu_t = (\alpha_{0,t} + \alpha_{1,t}r_t + \alpha_{2,t}r_t \ln r_t)$ is the drift (or the conditional mean), and $\sigma_t^2 = \Psi_t^2 r_t^{2\gamma_t}$ is the diffusion (or the conditional variance) of the interest rate change that is parameterized as a function of both the interest rate level and the unexpected interest rate shocks. α_t's are time-varying parameters that must be estimated.

The stochastic differential equation in (1) defines a broad class of interest rate processes that includes many well-known interest rate models. These models can be obtained from eq. (1) simply by imposing the appropriate restrictions on the parameters α's and γ. These short rate specifications with their corresponding parameter restrictions will be discussed later in this article.

To analyze the variety of models for the short rate and its volatility, CKLS (1992) uses the following continuous time short rate specification:

$$dr_t = (\alpha_0 + \alpha_1 r_t)dt + \Psi r_t^{\gamma} dW_t \tag{2}$$

where the parameters in the drift and diffusion functions are assumed to be independent of time. In this model, the conditional variance of interest rate changes $\sigma_t^2 = \Psi^2 r_t^{2\gamma}$ is defined only as a function of interest rate levels. The CKLS model does not allow the volatility parameter Ψ^2 to be a time-varying function of the information set as in eq. (1), ignoring the dependence of volatility on the unexpected information shocks. The sensitivity parameter γ is also assumed to be independent of time, implying that the sensitivity of volatility to interest rate levels in this model is assumed not to change over time, which is obviously impossible in dynamic, rapidly revolving financial markets.

In addition to *equilibrium models* such as Merton (1973), Vasicek (1977), CIR (1980, 1985), CKLS (1992), and many others, in which the drift and diffusion parameters are not allowed to vary over time, we also consider one of the most popular *arbitrage-free* term structure models developed by BDT (1990). The BDT model, like other arbitrage-free models, must ensure that model prices for zero-coupon bonds match market prices. In addition this model must ensure that the model's term structure of volatilities matches an empirical or given term structure of volatilities. To accomplish this goal, the BDT model allows volatility of the short rate to change over time.

The BDT model assumes that changes in short-term interest rates follow a lognormal probability distribution, which implies that the evolution of the short rate in continuous time can be described by:

$$d \ln r_t = (\alpha_{1,t} + \alpha_{2,t} \ln r_t)\, dt + \Psi_t dW_t. \qquad (3)$$

where the mean-reversion parameter, $\alpha_{2,t} = -\Psi_t'/\Psi_t$, is a function of the short rate volatility, Ψt, and its derivative with respect to time, $\Psi_t' = d\Psi t/dt$. This is the assumption about future volatilities and about how the drift responds to the level of interest rates, which is required for the BDT model to match the market term structure of volatilities.[2]

In this model, the short term interest rate is assumed to be attracted to some long run average level over time, which is known as *mean reversion*. Whereas the BDT displays mean reversion, that property comes about through the term structure of volatilities. This means that the extent to which the drift depends on the level of rates depends completely on the short term rate volatility process. The danger of this process is that the model might have to portray both mean reversion and future short rate volatilities inaccurately in order to replicate the desired term structure of volatilities. Of course, these inaccuracies will adversely affect contingent claims pricing. Therefore it might be desirable to model mean reversion independently of the volatility process. The next model is one example of how this can be accomplished.

Applying Ito's lemma to the evolution of the log-interest rate given in eq. (3), the uncertainty of the short rate can be characterized as

$$dr_t = r_t \left[\alpha_{1,t} + \alpha_{2,t} \ln r_t + \frac{1}{2} \Psi_t^2\right] dt + \Psi_t r_t \, dW_t \qquad (4)$$

where the functions "$r_t [\alpha_{1,t} + \alpha_{2,t} \ln r_t + 1/2\Psi_t^2]$" and "$\Psi_t^2 r_t^2$" are, respectively, the drift (or instantaneous mean) and diffusion (or instantaneous variance) of the short rate process. The scale factor Ψ_t is again parameterized as a function of the unexpected information shocks. In this model, the parameters in both the drift and diffusion are *time-dependent*, and the volatility of the interest rate change is defined as a function of both the interest rate level and the last period's unexpected news. Because the mean-reversion parameter $\alpha_{2,t}$ in this model is not defined as a function of the short rate volatility (that is, there is no required relationship between drift and diffusion functions) as in the original BDT

[2]See Appendix.

model in eq. (3), we refer to the specification in eq. (4) as the Unrestricted-BDT model.

THE ECONOMETRIC APPROACH

In this section, we describe the econometric approach used in estimating the parameters of the interest rate models and in examining their explanatory power for the dynamic behavior of short term interest rate volatility. To illustrate the approach clearly, we first focus on the econometric estimation of the unrestricted process given in eq. (1). The same approach can then be employed for the nested models after imposing the appropriate restrictions.

Following Brennan and Schwartz (1982), Marsh and Rosenfeld (1983), Dietrich-Campbell and Schwartz (1986), CKLS (1992), Nowman (1997), and others, we estimate the parameters of the continuous time interest rate model in eq.(1) using a discrete time econometric specification:

$$\begin{aligned} r_t - r_{t-1} &= \alpha_{0,t} + \alpha_{1,t} r_{t-1} + \alpha_{2,t} r_{t-1} \ln r_{t-1} + \varepsilon_t \\ E(\varepsilon_t | \mathcal{F}_{t-1}) &= 0,\ E(\varepsilon_t^2 | \mathcal{F}_{t-1}) \equiv \sigma_t^2 = \Psi_t^2 r_{t-1}^{2\gamma_t} \\ \Psi_t^2 &= \beta_{0,t} + \beta_{1,t} \varepsilon_{t-1}^2 + \beta_{2,t} \Psi_{t-1}^2 \end{aligned} \tag{5}$$

where $\mathcal{F}_{t-1}$ is the information set at time $t - 1$ so that $E(\ldots|\mathcal{F}_{t-1})$ represents conditioning on the information available at time $t - 1$, $\alpha_{0,t} + \alpha_{1,t}\, r_{t-1} + \alpha_{2,t}\, r_{t-1} \ln r_{t-1}$ is the time-varying conditional mean and $\Psi_t^2 r_{t-1}^{2\gamma_t}$ is the time-varying conditional variance of unexpected interest rate changes. In this model, large interest rate shocks (as measured by large residuals ε) cause increases in rate volatility through their effect on Ψ_t^2. The volatility parameter γ_t allows the volatility of interest rates to depend on the level of the interest rate. At higher γs, the volatility is more sensitive to interest rate levels. Because the specification in eq. (5) nests all the models considered in this article we refer to this model as the general parametric model (GPM, hereafter). The nested models with their corresponding parameter restrictions are listed in Table I.

The GPM model parameterizes the current volatility of interest rate changes as a function of both interest rate levels and unexpected information shocks. The time-varying scale factor (denoted by Ψ_t^2) in this model is a function of information flows. Specifically, it is higher when the shocks ε are larger in absolute magnitude.

TABLE I

Parameter Restrictions Imposed on GPM by Alternative Models of Short Term Interest Rate

Model	$\mu(r)$	$\sigma^2(r)$	*Parameter Restrictions*
Merton (1973)	α_0	β_0	$\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = \gamma = 0$
Vasicek (1977)	$\alpha_0 + \alpha_1 r_{t-1}$	β_0	$\alpha_2 = \beta_1 = \beta_2 = \gamma = 0$
Cox (1975)[a]	$\alpha_1 r_{t-1}$	$\beta_0 r_{t-1}^{2\gamma}$	$\alpha_0 = \alpha_2 = \beta_1 = \beta_2 = 0$
Cox-Ross (1976)			
Dothan (1978)	0	$\beta_0 r_{t-1}^{2}$	$\alpha_0 = \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0, \gamma = 1$
Cox-Ingersoll-Ross (1980)[b]	0	$\beta_0 r_{t-1}^{3}$	$\alpha_0 = \alpha_1 = \alpha_2 = \beta_1 = \beta_2 = 0, \gamma = 1.5$
Constantinides-Ingersoll (1984)			
Rendleman-Bartter (1980)[c]	$\alpha_1 r_{t-1}$	$\beta_0 r_{t-1}^2$	$\alpha_0 = \alpha_2 = \beta_1 = \beta_2 = 0, \gamma = 1$
Brennan-Schwartz (1980)	$\alpha_0 + \alpha_1 r_{t-1}$	$\beta_0 r_{t-1}^2$	$\alpha_2 = \beta_1 = \beta_2 = 0, \gamma = 1$
Courdaton (1982)			
Cox-Ingersoll-Ross (1985)[d]	$\alpha_0 + \alpha_1 r_{t-1}$	$\beta_0 r_{t-1}$	$\alpha_2 = \beta_1 = \beta_2 = 0, \gamma = 0.5$
Brown-Dyvig (1986)			
Gibbons-Ramaswamy (1993)			
CKLS (1992)	$\alpha_0 + \alpha_1 r_{t-1}$	$\beta_0 r_{t-1}^{2\gamma}$	$\alpha_2 = \beta_1 = \beta_2 = 0$
Unrestricted BDT	$\alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1}$	$\Psi_t^2 r_{t-1}^2$	$\alpha_0 = 0, \gamma = 1$
BDT-LEVEL	$\alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1}$	$\beta_0 r_{t-1}^2$	$\alpha_0 = \beta_1 = \beta_2 = 0, \gamma = 1$
BDT-GARCH	$\alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1}$	$\beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2$	$\alpha_0 = \gamma = 0$

Notes: The alternative term structure models for the short term interest rate are obtained from imposing the appropriate parameter restrictions on the general parametric model (GPM)

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \Psi_t^2 r_{t-1}^{2\gamma}, \ \Psi_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \Psi_{t-1}^2.$$

The Original-BDT model is not nested within the GPM model.

[a]This is the constant elasticity of variance (CEV) process introduced by Cox (1975) and by Cox-Ross (1976).

[b]This is the Cox-Ingersoll-Ross (1980) variable-rate (CIR VR) model

[c]This is the familiar geometric Brownian motion (GBM) process.

[d]This is the Cox-Ingersoll-Ross square root (CIR SR) single-factor general equilibrium term structure model.

This model has several interesting features. The specification in eq. (5) reduces to the Unrestricted-BDT model when $\alpha_0 = 0$ and $\gamma = 1$. A more restricted version of the Unrestricted-BDT model can be obtained when $\alpha_0 = \beta_1 = \beta_2 = 0$ and $\gamma = 1$. In this case, the time variation in Ψ_t^2 disappears and the conditional heteroscedasticity enters solely through the squared level of the interest rate. We therefore refer to this empirical model as the BDT-LEVEL model:

$$r_t - r_{t-1} = \alpha_{1,t}\, r_{t-1} + \alpha_{2,t}\, r_{t-1} \ln r_{t-1} + \varepsilon_t$$

$$E(\varepsilon_t \,|\mathcal{F}_{t-1}) = 0, \ E(\varepsilon_t^2|\mathcal{F}_{t-1}) \equiv \sigma_t^2 = \beta_{0,t} r_{t-1}^2 \tag{6}$$

Perhaps the strongest criticism of the BDT-LEVEL model is that it restricts volatility to be a function of interest rate levels only, and not of the news arrival process. Therefore, the specification in eq. (6) tends to overemphasize the sensitivity of volatility to interest rate levels and fails to model adequately the serial correlation in conditional variances.

An alternative model that addresses this is the GARCH model, in which the current volatility is a function of the last period's unexpected news.[3] Specifically, when $\alpha_0 = \gamma = 0$ then the level effect disappears, and we end up with the BDT-GARCH model:

$$r_t - r_{t-1} = \alpha_{1,t}\, r_{t-1} + \alpha_{2,t} r_{t-1} \ln r_{t-1} + \varepsilon_t$$
$$E(\varepsilon_t|\mathcal{I}_{t-1}) = 0,\ E(\varepsilon_t^2|\mathcal{I}_{t-1}) \equiv \sigma_t^2 = \beta_{0,t} + \beta_{1,t}\varepsilon_{t-1}^2 + \beta_{2,t}\sigma_{t-1}^2. \quad (7)$$

The main problem of the BDT-GARCH model is that this serial-correlation-based model fails to capture adequately the relationship between interest rate levels and volatility. In general, there are other problems of the GARCH models. For example, most empirical applications to interest rate data find that $\beta_1 + \beta_2 \cong 1$, which implies that volatility shocks persist forever.[4] The other problem is that GARCH models permit negative interest rates.

CKLS (1992) uses the standard linear mean-reverting drift $\mu_t = \alpha_0 + \alpha_1\, r_{t-1}$, and the constant elasticity of variance $\sigma_t^2 = \Psi^2 r_{t-1}^{2\gamma}$ for the diffusion function. CKLS estimates the *time-independent* parameters of the continuous time model given in eq. (2) using a discrete time econometric specification:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1\, r_{t-1} + \varepsilon_t$$
$$E(\varepsilon_t\, |\mathcal{I}_{t-1}) = 0,\ E(\varepsilon_t^2|\mathcal{I}_{t-1}) \equiv \Psi^2 r_{t-1}^{2\gamma} \quad (8)$$

where the autoregressive conditional heteroscedasticity depends only on the squared level of the last period's interest rate. CKLS finds that the best models for capturing the dynamics of short rates allow volatility of unexpected interest rate changes to depend positively on the rate level (that is, $\gamma > 0$). CKLS estimated γ to be about 1.50 and demonstrated that models with $\gamma \geq 1$ outperformed those with $\gamma < 1$. The strongest criticism of the CKLS approximation is that it fails to model adequately

[3]See Engle, Lilien, and Robins (1987); Cecchetti, Cumbi, and Figlewski (1988); and Engle, Ng, and Rothschild (1992), among many others, for applications of GARCH models to interest rate data.
[4]Note that if $\beta_1 + \beta_2 = 1$ the unconditional variance $\sigma^2 = \beta_0/(1 - \beta_1 - \beta_2)$ is undefined. In this case, we have a nonstationary GARCH model called the *Integrated* GARCH (IGARCH) model introduced by Engle and Bollerslev (1986).

the serial correlation in conditional variances because it ignores the dependence of volatility on unexpected interest rate shocks.

The econometric modeling of the original BDT (1990) model is quite different from its modified versions due to its crucial assumption about future volatilities discussed in the Appendix. The time-varying parameters of the Original-BDT model are estimated from the following discrete time system of equations:[5]

$$r_t - r_{t-1} = \alpha^*_{1,t} r_{t-1} + \alpha_{2,t}\, r_{t-1} \ln r_{t-1} + \varepsilon_t$$

$$E(\varepsilon_t\,|\mathcal{F}_{t-1}) = 0,\ E(\varepsilon_t^2|\mathcal{F}_{t-1})' \equiv \sigma_t^2 = \Psi_t^2 r_{t-1}^2 \qquad (5)$$

where $\Psi_t^2 = \Psi_{t-1}^2 e^{2\delta_{2,t}}$ is the conditional variance of the *log-interest rate* change obtained from the discrete time econometric specification of the Original-BDT model given in eq. (3):

$$\ln r_t - \ln r_{t-1} = \delta_{1,t} + \delta_{2,t} \ln r_{t-1} + \eta_t.$$

$$E(\eta_t|\mathcal{F}_{t-1}) = 0,\ E(\eta_t^2|\mathcal{F}) \equiv \Psi_t^2 = \Psi_{t-1}^2 e^{2\delta_{2,t}}. \qquad (10)$$

where the time-varying conditional variance of the log-interest rate change depends on the mean-reversion parameter $\delta_{2,t}$. The econometric specification of the log-interest rate in eq. (10) are estimated with the ordinary least squares (OLS) and then the estimate of $\delta_{2,t}$ together with Ψ_{t-1}^2 is used to compute Ψ_t^2. Finally the scale factor Ψ_t^2 is utilized to find the volatility of the interest rate change in eq. (9) satisfying the critical assumption of the Original-BDT model.

THE DATA

We use the extended version of the United States Treasury bill one-month yield data utilized in CKLS (1992). This data set is originally constructed by Fama (1984) and then updated by the Center for Research in Security Prices (CRSP). The data for our study are monthly and cover the period from June 1964 to December 1996, giving a total of 390 observations.

[5]The econometric model presented in eq. (9) is obtained by applying Ito's lemma to the continuous time evolution of the log-interest rate in eq. (3): $d\ ln\ r_t = (\alpha_{1,t} + \alpha_{2,t}\ ln\ r_t)\, dt + \Psi_t\, dW_t$or by Ito's lemma:

$$dr_t = r_t\,[\alpha_{1,t} + \alpha_{2,t}\ ln\ r_t + 1/2\Psi_t^2]\, dt + \Psi_t\, r_t\, dW_t. \qquad 4$$

The discrete-time econometric specification of eq. (4) is written as:

$$r_t - r_{t-1} = \alpha^*_{1,t} r_{t-1} + \alpha_{2,t}\, r_{t-1}\ ln\ r_{t-1} + \varepsilon_t \qquad (9)$$

$$E(\varepsilon_t|\mathcal{F}_{t-1}) = 0,\ E(\varepsilon_t^2|\mathcal{F}_{t-1}) \equiv \sigma_t^2 = \Psi_t^2 r_{t-1}^2 \text{ where } \alpha_{1,t} = \alpha^*_{1,t} - 1/2\Psi_t^2.$$

TABLE II

Summary Statistics

Variables	*N*	*Mean*	*Standard Deviation*	*ρ1*	*ρ2*	*ρ3*	*ρ4*	*ρ5*	*ρ6*	*ADF*
r_t	390	0.06272	0.02598	0.956	0.921	0.880	0.848	0.827	0.806	−2.7285
$r_t - r_{t-1}$	389	0.00004	0.00757	−0.108	0.066	−0.101	−0.125	−0.008	−0.026	−13.832

Notes: Means, standard deviations, and autocorrelations of monthly Treasury bill yields and yield changes are computed from June 1964 to December 1996. The variable r_t denotes the annualized yield on Treasury bills maturing in one 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. ADF denotes the Augmented Dickey-Fuller unit root statistic with a 1%, (5%), [10%] critical value of −3.4460, (−2.8692), [−2.5708].

The one-month yields expressed in annualized form are based on the average of bid and ask prices for Treasury bills.

Table II displays the means, standard deviations, and first six autocorrelations of the one-month yield and the monthly changes in the one-month yield. The augmented Dickey-Fuller (ADF) statistics are reported for the presence of a unit root. The unconditional average level of the one-month yield is 6.272% with a standard deviation of 2.598%, whereas the monthly yield change averaged 0.00411% with a standard deviation of 0.757%. The autocorrelations for the level decay slowly and those of the first differences are generally small and are not systematically positive or negative. The ADF statistics reject the null hypothesis of a unit root in the interest rate level at the 10% level of significance but fail to reject at the 5%. The ADF statistics for the interest rate change reject the presence of a unit root at the 1% level of significance.

THE EMPIRICAL RESULTS

In this section we present the maximum likelihood estimation results from estimating the time-varying parameters of the drift and diffusion functions of the interest rate models. We also show the estimated parameters for the well known existing models of the short-term interest rate obtained from imposing the appropriate restrictions on the GPM model. The explanatory power of the nested models is compared to that of the GPM and the different versions of the BDT model using the maximized log-likelihood functions. We also compare the models in terms of their predictive power for the future level and variance of interest rate changes.

Estimation Results and Model Comparisons

The procedure used in this study for measuring the predictive power of the interest rate models closely follows that of earlier studies (for example, CKLS, 1992; Nowman, 1997), although the present study differs from its predecessors in an important respect. The former studies compute the level and variance of interest rate changes using a *fixed* sample period assuming that the parameters in the drift and diffusion are independent of time. Whereas this study uses a rolling regression procedure and generates one-step-ahead forecasts by estimating the *time-varying* parameters of the drift and diffusion functions.

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 389 observations is split into two parts: The first 307 observations (from June 1964 to December 1989) are used for estimation of the parameters of the model, and then a one-step-ahead forecast is calculated.[6] The sample is then 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 is made. This "recursive" modeling and forecasting procedure is repeated until a forecast for observation 389 has been made using data available at time 388. Computation of forecasts employing a rolling window of data ensures that the forecasts are made using models for which parameters have been estimated utilizing all the information available at that time, while not incorporating old data.

Table III presents the whole-sample (June 1964–December 1996) parameter estimates, asymptotic t-statistics, maximized log-likelihoods for the GPM and the BDT models as well as, for each of the nested models, the likelihood ratio tests comparing the nested models with the general parametric model introduced in this article. As described in CKLS (1992), a large number of models (presented in Table I) are nested within the CKLS model. Because this model is considered as a nested model in this study and it covers most of the *subnested* models (see, for example, Merton, 1973; Vasicek, 1977; Dothan, 1978; Brennan-Schwartz, 1980; CIR, 1980, 1985), we will compare the relative performance of these subnested models with our more general specifications *only* using the likelihood ratio test. Their predictive power for the level and variance of interest rate changes will not be tested because it has already been done by CKLS (1992). The reader may wish to consult Table III for the parameter estimates of these models that are strongly rejected against the GPM model.

[6]CKLS (1992) used the same T-bill yield data but for the period June 1964 to December 1989. To compare our maximum likelihood estimates with their GMM results and estimate the time-varying parameters of the models for the period January 1990—December 1996, we choose to employ the same sample (from June 1964 to December 1989) for the in-sample estimation of the models.

TABLE III

GPM Estimates and Comparisons with the Nested Models

Model	α_0	α_1	α_2	β_0	β_1	β_2	γ	*Log Likelihood*	*LR*	$\chi^2_{(0.01)}$	*df*
GPM	0.0007	−0.0562	−0.0168	0.0018	30.403	0.3920	0.8975	1845.8660	—	—	—
	(0.2318)	(−0.5744)	(−0.3223)	(1.6042)	(1.9283)	(3.6703)	(9.6887)				
Unrestricted BDT	0.0	−0.0753	−0.0281	0.0036	48.279	0.3768	1.0	1845.3079	1.116	9.21	2
		(−2.2915)	(−2.4837)	(4.6943)	(3.5501)	(3.3349)					
CKLS	0.0015	−0.0216	0.0	0.0502	0.0	0.0	1.2998	1825.6710	40.390	11.34	3
	(2.2655)	(−1.5121)		(2.4938)			(18.556)				
CEV	0.0	0.0085	0.0	0.0512	0.0	0.0	1.3010	1823.1474	45.437	13.28	4
		(1.6849)		(2.4996)			(18.538)				
BDT GARCH	0.0	−0.1314	−0.0437	0.000001	0.2510	0.7500	0.0	1822.0180	47.696	9.21	2
		(−4.8467)	(−4.4117)	(3.0840)	(6.5632)	(32.595)					
BDT LEVEL	0.0	−0.0825	−0.0311	0.0094	0.0	0.0	1.0	1820.0108	51.710	13.28	4
		(−2.5312)	(−2.6372)	(19.162)							
Brennan Schwartz	0.0016	−0.0242	0.0	0.0094	0.0	0.0	1.0	1819.6699	52.392	13.28	4
	(2.2686)	(−1.8583)		(19.318)							
CIR VR	0.0	0.0	0.0	0.1629	0.0	0.0	1.5	1818.3535	55.025	16.81	6
				(19.478)							
GBM	0.0	0.0058	0.0	0.0095	0.0	0.0	1.0	1817.0550	57.622	15.09	5
		(1.1819)		(19.778)							
Dothan	0.0	0.0	0.0	0.0096	0.0	0.0	1.0	1816.3577	59.017	16.81	6
				(19.778)							
CIR SR	0.0020	−0.0317	0.0	0.0007	0.0	0.0	0.5	1781.8101	128.112	13.28	4
	(2.2950)	(−2.7496)		(23.543)							
Vasicek	0.0028	−0.0436	0.0	0.00006	0.0	0.0	0.0	1709.6213	272.485	13.28	4
	(2.3604)	(−3.7762)		(30.591)							
Merton	0.0001	0.0	0.0	0.00005	0.0	0.0	0.0	1704.3630	283.006	15.09	5
	(0.2961)			(37.814)							

Notes: This table displays the maximum likelihood estimates of alternative models of the short term interest rate r_t (annualized one-month U.S. Treasury bill yield) from June 1964 to December 1996 (390 observations). The parameters estimates with asymptotic *t*-statistics in parentheses are presented for each model. The maximized log-likelihoods for the unrestricted model and for each of the nested and subnested models are shown to compare the explanatory power of the nested and subnested models with the GPM model. Likelihood ratio tests evaluate restrictions imposed by the nested and subnested models against the GPM model. The likelihood ratio (LR) test statistics with associated degrees of freedom (df) and the associated chi-squared critical values $\chi^2_{(0.01)}$ at the 1% level of significance are reported. The parameters are estimated from the following discrete time system of equations:

$$r_t = r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \Psi_t^2 r_{t-1}^{2\gamma}, \ \Psi_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \Psi_{t-1}^2.$$

Based on maximized log-likelihood values, the GPM (log L = 1845.8660) and the Unrestricted-BDT (log L = 1845.3079) models perform the best compared to the CKLS and the other versions of the BDT model (BDT-LEVEL, BDT-GARCH, Original-BDT). The relative performance of the BDT-GARCH (log L = 1822.0180) and the BDT-LEVEL (log L= 1820.0108), models turn out to be the same and both of them are found to be slightly inferior to the CKLS (log L = 1825.6710) model because it does not restrict the sensitivity parameter γ to 0 (as in BDT-GARCH) or to 1.0 (as in BDT-LEVEL).

An important property of this ranking is that the models (GPM and Unrestricted-BDT) that parameterize the volatility of interest rate changes as a function of both the interest rate level and the unexpected information shocks best capture the dynamics of short term interest rates. According to the likelihood ratio (LR) test statistics demonstrated in Table III, the nested (CKLS, BDT-GARCH, BDT-LEVEL) as well as the subnested models which specify the linear drift with or without mean-reversion and/or parameterize the volatility only as a function of interest rate levels are *strongly* rejected.

The empirical results suggest that although the models (CKLS and BDT-LEVEL) that define the diffusion only as a function of interest rate levels are rejected the relation between interest rate volatility and the level of r is the important feature of any interest rate model discussed in this article. As presented in Table III, the unconstrained estimate of γ is approximately between 0.90 (in GPM) and 1.30 (in CKLS) and highly significant. In other words, when the sensitivity of volatility to interest rate levels depends on information flows (because $\Psi_t^2 = \beta_0 + \beta_1\varepsilon_{t-1}^2 + \beta_2\Psi_{t-1}^2$), as in the case of GPM, the parameter γ becomes smaller. In contrast, when the volatility is defined only as a function of the level of interest rate as in the case of CKLS and BDT-LEVEL, the models tend to overemphasize the sensitivity of volatility to interest rate levels.

Table IV shows the in-sample (June 1964–December 1989) parameter estimates, asymptotic t-statistics, maximized log-likelihoods for all models considered in this article, and the likelihood ratio test statistics comparing the nested and subnested models with the GPM model. The statistical results, based on the likelihood ratio tests displayed in Table IV, imply that both the equilibrium and the modified versions of the BDT model—except the Unrestricted-BDT model—are strongly rejected.[7]

[7]Because the Original-BDT model is not nested within the GPM model, we prefer not to present the in-sample parameter estimates of this model in Table IV. In addition, we cannot estimate the parameters of this model for the whole-sample because it requires rolling regression procedure, which cannot be performed with a *fixed* sample period. The in-sample parameter estimates with asymptotic

TABLE IV

In-Sample GPM Estimates and Comparisons with the Nested Models

Model	α_0	α_1	α_2	β_0	β_1	β_2	γ	*Log Likelihood*	*LR*	$\chi^2_{(0.01)}$	*df*
GPM	0.0001	−0.0659	−0.0199	0.0015	27.238	0.5019	0.9062	1427.7174	—	—	—
	(0.2367)	(−0.5383)	(−0.2936)	(1.3986)	(1.7030)	(4.3252)	(8.6605)				
Unrestricted BDT	0.0	−0.0900	−0.0347	0.0027	41.228	0.4941	1.0	1427.4373	0.5602	9.21	2
		(−2.3507)	(−2.5596)	(3.4880)	(3.4706)	(4.1116)					
BDT GARCH	0.0	−0.1431	−0.0493	0.00001	0.2882	0.7363	0.0	1412.5055	30.424	9.21	2
		(−4.6752)	(−4.3624)	(2.7307)	(5.8362)	(26.259)					
CKLS	0.0020	−0.0272	0.0	0.0701	0.0	0.0	1.3648	1411.1482	33.138	11.34	3
	(2.3143)	(−1.5988)		(2.2550)			(17.397)				
CEV	0.0	0.0100	0.0	0.0706	0.0	0.0	1.3632	1408.6281	38.179	13.28	4
		(1.7121)		(2.3053)			(17.803)				
BDT LEVEL	0.0	−0.0937	−0.0362	0.0097	0.0	0.0	1.0	1404.9402	45.554	13.28	4
		(−2.4679)	(−2.5548)	(16.860)							
Brennan Schwartz	0.0021	−0.0295	0.0	0.0097	0.0	0.0	1.0	1404.6781	46.079	13.28	4
	(2.2698)	(−1.9202)		(16.957)							
CIR VR	0.0	0.0	0.0	0.1537	0.0	0.0	1.5	1405.6118	44.211	16.81	6
				(18.068)							
GBM	0.0	0.0067	0.0	0.0098	0.0	0.0	1.0	1402.1256	51.184	15.09	5
		(1.1742)		(17.649)							
Dothan	0.0	0.0	0.0	0.0099	0.0	0.0	1.0	1401.4374	52.560	16.81	6
				(17.647)							
CIR SR	0.0026	−0.0369	0.0	0.0007	0.0	0.0	0.5	1374.2303	106.97	13.28	4
	(2.2008)	(−2.6123)		(20.341)							
Vasicek	0.0034	−0.0497	0.0	0.00007	0.0	0.0	0.0	1319.9922	215.45	13.28	4
	(2.2101)	(−3.3779)		(25.470)							
Merton	0.0001	0.0	0.0	0.00007	0.0	0.0	0.0	1315.9519	223.53	15.09	5
	(0.1513)			(28.038)							

Notes: This table displays the in-sample maximum likelihood estimates of alternative models of the short term interest rate r_t (annualized one-month U.S. Treasury bill yield) from June 1964 to December 1989 (306 observations). The parameters estimates with asymptotic t-statistics in parentheses are presented for each model. The maximized log-likelihoods for the unrestricted model and for each of the nested and subnested models are shown to compare the explanatory power of the nested and subnested models with the GPM model. Likelihood ratio tests evaluate restrictions imposed by the nested and subnested models against the GPM model. The likelihood ratio (LR) test statistics with associated degrees of freedom (df) and the associated Chi-Squared critical values $\chi^2_{(0.01)}$ at the 1% level of significance are reported. The parameters are estimated from the following discrete time system of equations:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1} \ln r_{t-1} + \varepsilon_t,$$

$$\sigma_t^2 = \Psi_t^2 r_{t-1}^{2\gamma}, \; \Psi_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \Psi_{t-1}^2.$$

Using the same data set for the period June 1964–December 1989, CKLS (1992) estimates the time-independent parameters of equilibrium models employing the Generalized Method of Moments (GMM) technique of Hansen (1982). Nowman (1997) estimates the same equilibrium models using the Gaussian estimation methods developed by Bergstrom (1983). Both CKLS (1992) and Nowman (1997) assume that the models parameters do not vary over time. Instead of assuming that the parameters are independent of time, using rolling regression procedure, we estimate the time-varying parameters employing the maximum likelihood estimation (MLE) technique with the Berndt-Hall-Hall-Hausman algorithm.[8] Our empirical results for the same nested models turn out to be exactly the same as Nowman's (1997) and very close to those obtained by CKLS (1992).

An Alternative Measure of Model Performance

In order to measure further the relative performance of the nested models against the more general parametric specifications introduced in this article, we test their predictive power for interest rate changes. Furthermore, we test their explanatory power for squared interest rate changes, which provide a simple ex post measure of interest rate volatility. This is done by first computing the time series of conditional mean and conditional variance of the monthly yield changes for each model using the fitted values of the various econometric models introduced in this article. We then define $(r_t - r_{t-1})$ as the interest rate change and $(r_t - r_{t-1})^2$ as the ex post measure of the variance of interest rate changes. Finally, we compute the Theil Inequality Coefficient for the conditional expected yield changes and the conditional variance.

For forecast evaluation, most financial researchers generally use Root Mean Square Error (RMSE) or Mean Absolute Error (MAE). We prefer to employ Theil Inequality Coefficient (TIC) because RMSE and MAE depend on the scale of the dependent variable whereas TIC is scale invariant. TIC always lies between zero and one, where zero indicates a

t-statistics in parentheses and maximized log-likelihood values are presented below:

$$r_t - r_{t-1} = \underset{(-3.7933)}{-0.1369}\, r_{t-1} - \underset{(-3.1876)}{0.0529}\, r_{t-1}\, ln\, r_{t-1} + \varepsilon_t \quad \log L = 1321.2258$$

$$ln\, r_t - ln\, r_{t-1} = \underset{(-2.4733)}{-0.0076} - \underset{(-2.8481)}{0.0421}\, ln\, r_{t-1} + \eta_t \quad \log L = 1214.1843$$

[8] Broze, Scaillet, and Zakoian (1995) prove in their proposition 3.3 that the GMM estimator of the CKLS model is not well behaved if $\gamma > 1$. Because we obtain estimates of γ that exceed 1.0 in some of our models, we prefer to use maximum likelihood estimation.

TABLE V

An Alternative Measure of Model Performance

$N = 389$	*GPM*	*Unrestricted BDT*	*BDT GARCH*	*CKLS*	*BDT LEVEL*	*Original[a] BDT*
TIC_{mean}	0.7197	0.8057	0.8877	0.9231	0.8955	—
$TIC_{variance}$	0.5585	0.5544	0.5894	0.6349	0.7107	—
N = 84	GPM	Unrestricted BDT	BDT GARCH	CKLS	BDT LEVEL	Original BDT
TIC_{mean}	0.7616	0.8512	0.8566	0.9002	0.8980	0.9144
$TIC_{variance}$	0.6232	0.6253	0.6413	0.6686	0.6558	0.6722

Notes: This table demonstrates the Theil Inequality Coefficient for the forecasted conditional mean denoted by TIC_{mean} and for the forecasted conditional variance denoted by $TIC_{variance}$. 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 predict the future level and variance of interest rate changes:

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}y_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i^2}}.$$

where y_t and $\hat{y}_t$ denote the actual and forecasted values of $(r_t - r_{t-1})$ for TIC_{mean} and $(r_t - r_{t-1})^2$ for $TIC_{variance}$. The first three rows show the TIC values for the whole-sample estimation results for the period June 1964-December 1996 (N = 389). The last three rows present the TIC values for the one-period-ahead forecasts obtained from the rolling regression procedure for the period January 1990-December 1996 (N = 84).

[a]Since the Original-BDT model requires rolling regression procedure to find the time-varying conditional mean and conditional variance series, TIC values for this model cannot be calculated with the whole-sample estimation.

perfect fit.[9] The TIC values in this study provide information about how well each model is able to forecast the future level and variance of interest rate changes.

The forecast results are presented in Table V. There is no TIC values for the Original-BDT model for the whole-sample estimation because the Original-BDT model requires rolling regression procedure to find the conditional mean and variance series. To compare the explanatory power of the models for forecasting the time-varying conditional mean and conditional variance, the TIC values for the one-period-ahead forecasts from rolling regression are shown in Table V.

TIC_{mean} and $TIC_{variance}$ describe the fit of the various models for the actual yield changes and for the ex post measure of the variance of interest rate changes, respectively. According to the values of TIC_{mean} shown in

[9]The Theil Inequality Coefficient (TIC) is defined as follows:

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\gamma_{;i} - \hat{\gamma})^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\gamma_i^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{\gamma}_i^2}}$$

where y_i and $\hat{\gamma}_i$ denote the actual and forecasted values of $(r_i - r_{i-1})$ or $(r_i - r_{i-1})^2$ in this study.

Table V, the GPM model that uses the most general drift and diffusion functions have more explanatory power (or less TIC_{mean}) for interest rate changes than the nested models. The performance of the BDT-GARCH model in terms of its explanatory power for the interest rate change is found to be superior to the BDT-LEVEL, the Original-BDT, and the CKLS models. In other words, the ranking of the models based on their predictive power for the interest rate change is not aligned to the ranking implied by the maximized log likelihood values.

However, this is not the case for the variance of interest rate changes. The ranking of the models based on their explanatory power for the variance of interest rate changes is closely aligned to the ranking implied by the maximized log likelihood values. $TIC_{variance}$ values for each model imply that the more flexible models (like GPM and Unrestricted-BDT) which allow the volatility to depend on both interest rate levels and unexpected interest rate shocks best forecast the volatility of interest rate changes. The models which parameterize the volatility either only as a function of interest rate levels (like CKLS and BDT-LEVEL) or only as a function of the news arrival process (like BDT-GARCH) appear similar in their forecasting ability that is found to be inferior to the explanatory power of the more general parametric models. The performance of the Original-BDT model for predicting the time-varying conditional variance of the interest rate change turns out to be similar to CKLS and the restricted versions of the BDT model (BDT-LEVEL and BDT-GARCH). Its performance is also found to be inferior to the GPM and the Unrestricted-BDT models.

These results suggest that pricing interest rate derivatives using the newly proposed volatility models should yield more accurate prices than using the existing models because both the GPM and the Unrestricted-BDT provide a more accurate representation of the underlying volatility structure than the more restricted models such as the Original-BDT, BDT-LEVEL and BDT-GARCH.

Bali and Karagozoglu (1999) demonstrate the effects of using different volatility estimators in the pricing of interest rate options within the BDT framework. Their empirical results, based on 4,228 estimated prices, indicate that the valuation of Eurodollar futures options is sensitive to the volatility model used. One of the econometric specifications they use to estimate the short rate volatility, which is very similar to the Unrestricted-BDT model, produces more accurate predictions of actual option prices than the Original-BDT, historical and moving average volatility estimators. For forecast evaluation, they use both MSE and TIC measures, and find that the mean square error of the estimated prices reduces from 0.57 to 0.21, whereas the TIC values decrease from 0.51 to 0.38 when the Original-BDT is generalized to the Unrestricted-BDT model.[10]

Bali (1999) modifies the BDT term structure model by developing a more flexible framework that uses non-recombining binomial trees to price interest rate contingent claims. The modified BDT uses the more flexible GPM model to estimate the volatility of interest rate changes, whereas the original BDT employs the more restricted volatility estimators. Bali compares the empirical performance of the original and modified BDT models in pricing Eurodollar futures options over the period March 1988 to June 1998.

To measure the relative performance of the BDT models in capturing the stochastic behavior of option prices, Bali computes the proportion of the total variation in the actual option prices that can be explained by the estimated values. He refers to this as the coefficient of determination denoted by R^2. According to his findings, the R^2 values increase from about 0.58 to 0.87 when the original BDT model is extended to the more general framework using the GPM volatility model.

CONCLUSION

In this article, a wide variety of well-known existing models of the short term interest rate is compared in order to determine which model best fits the one-month U.S. Treasury bill yield data. Many popular interest rate models are nested within a more flexible time-varying BDT framework that allows us to compare the models and find the proper specification of the short term interest rate volatility.

The empirical results for the one-month Treasury bills indicate that the *equilibrium models* that do not allow the drift and diffusion parameters to vary over time and parameterize the volatility of interest rate changes only as a function of interest rate levels overemphasize the sensitivity of volatility to the level of interest rate and fail to model adequately the serial correlation in conditional variances. On the other hand, the serial-correlation-based *arbitrage-free models* with time-dependent parameters in the drift and diffusion functions define the volatility only as a function of unexpected information shocks and fail to capture adequately the relationship between interest rate levels and volatility.

[10]Bali and Karagozoglu (1999) classify the estimation results according to factors affecting the option price, and then test whether the results from comparing the performance of the volatility models are sensitive to these factors. Their findings regarding the superior performance of the Unrestricted-BDT model in predicting the prices of Eurodollar futures options turn out to be robust across factors such as the type of option (call or put), trading volume, time-to-expiration, and moneyness.

This article shows that the most successful models in capturing the dynamics of short term interest rates are those that introduce time-dependent parameters to the short rate process and define the conditional volatility as a function of both the interest rate level and the last period's unexpected news.

APPENDIX

Financial researchers generally make use of a *recombining binomial tree* for the BDT model to price interest rate contingent claims. As mentioned earlier, this model allows the volatility of the short term interest rate to change over time. Time-varying volatilities create a bit of problem because they may lead to a non-recombining tree although the original BDT model is constructed based on a recombining tree. The binomial tree will recombine if the up-down state and down-up state are forced to generate the same rate. This condition requires that $\delta_{2,t} = \Psi_t'/\Psi_t$.

Proof: The discrete time version of (3) is written as follows (replacing α's with δ's):

$$\Delta \ln r_t = (\delta_{1,t} + \delta_{2,t} \ln r_t)\, \Delta + \Psi_t\, \Delta W_t. \tag{A1}$$

Eq. (A1) can be approximated via the binomial representation:

$$\ln r_{t+\Delta} - \ln r_t = \begin{Bmatrix} (\delta_{1,t} + \delta_{2,t} \ln r_t)\Delta + \Psi_t\sqrt{\Delta}; & \text{probability } 1/2 \\ (\delta_{1,t} + \delta_{2,t} \ln r_t)\Delta - \Psi_t\sqrt{\Delta}; & \text{probability } 1/2 \end{Bmatrix} \tag{A2}$$

In this model, $\delta_{2,t}$ is treated as a measure of the speed of mean reversion in log-rate levels. At time t, if the log-interest rate is $\ln r_t$ then one period later, using (A2), the log-rate will be either:

$$\ln r^u_{t+\Delta} - \ln r_t = (\delta_{1,t} + \delta_{2,t} \ln r_t)\, \Delta + \Psi_t\sqrt{\Delta} \text{ or}$$

$$\ln r^d_{t+\Delta} - \ln r_t = (\delta_{1,t} + \delta_{2,t} \ln r_t)\, \Delta - \Psi_t\sqrt{\Delta}$$

where "u" and "d" denote the "up" and "down" states, respectively. Subtracting the above two equations yields the *sensitivity formula*, which shows how the spread between the two possible log-interest rates at time $t + \Delta$ is determined by the volatility of log-interest rate changes:

$$(\ln r^u_{t+\Delta} - \ln r^d_{t+\Delta})/2 = \Psi_t\sqrt{\Delta}. \tag{A3}$$

Because the original BDT model employs the recombining binomial tree which requires that the up-down (denoted by "ud") and down-up (denoted by "du") must generate the same interest rate: that is, $\ln r^{ud}_{t+\Delta} = \ln r^{du}_{t+\Delta}$ must hold where

$$\ln r^{ud}_{t+2\Delta} - \ln r^{u}_{t+\Delta} = (\delta_{1,t+\Delta} + \delta_{2,t+\Delta} \ln r^{u}_{t+\Delta})\Delta - \Psi_{t+\Delta}\sqrt{\Delta}$$

$$\ln r^{du}_{t+2\Delta} - \ln r^{d}_{t+\Delta} = (\delta_{1,t+\Delta} + \delta_{2,t+\Delta} \ln r^{d}_{t+\Delta})\Delta + \Psi_{t+\Delta}\sqrt{\Delta}$$

The condition $\ln r^{ud}_{t+\Delta} = \ln r^{du}_{t+\Delta}$ requires that $\delta_{2,t+\Delta} = (\Psi_{t+\Delta} - \Psi_t)/(\Psi_t \cdot \Delta)$. The continuous time limit of this expression: $\lim_{\Delta\to 0} \delta_{2,t+\Delta} = ((\Psi_{t+\Delta} - \Psi_t)/\Delta)\ (1/\Psi_t)$ implies $\delta_{2,t} = (\Psi'_t/\Psi_t)$ where $\Psi'_t = d\Psi_t/dt$. The last expression can be rewritten as $\delta_{2,t} = (d \ln \Psi_t/dt)$ that yields in discrete time $\Psi_t = \Psi_{t-1}e^{\delta_{2,t}}$ or $\Psi^2_t = \Psi^2_{t-1}e^{2\delta_{2,t}}$.

BIBLIOGRAPHY

Bali, T. G. (1999). Another look at the BDT term structure model. Working paper, Queens College, City University of New York.

Bali, T. G., & Karagozoglu, A. K. (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.

Bergstrom, A. R. (1983). Gaussian estimation of structural parameters in higher-order continuous time dynamic models. Econometrica, 51, 117–152.

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.

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.

Brown, S. J., & Dybvig, P. H. (1986). The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates. Journal of Finance, 41, 617–630.

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.

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

Constantinides, G. M., & Ingersoll, J. E. (1984). Optimal bond trading with personal taxes. Journal of Financial Economics, 13, 299–335.
Courdaton, G. (1982). The pricing of options on default-free bonds. Journal of Financial and Quantitative Analysis, 17, 75–100.
Cox, J. C. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Working paper, Stanford University.
Cox, J. C., & Ross, S. (1976). The valuation of options for alternative stochastic processes. Journal of Financial Economics, 3, 145–166.
Cox, J. C., Ingersoll, J., & Ross, S. (1980). An analysis of variable rate loan contracts. Journal of Finance, 35, 389–403.
Cox, J. C., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. Econometrica, 53, 385–407.
Dietrich-Campbell, 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., & 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.
Fama, E. F. (1984). Term premiums in bond returns. Journal of Financial Economics, 13, 529–546.
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.
Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. Econometrica, 50, 1029–1054.
Marsh, T. A., & Rosenfeld, E. R. (1983). Stochastic processes for interest rates and equilibrium bond prices. Journal of Finance, 38, 635–646.
Merton, R. C. (1973). Theory of rational option pricing. Bell Journal of Economics and Management Science, 4, 141–183.
Nowman, K. B. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. Journal of Finance, 4, 1695–1706.
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.
Rendleman, R., & Bartter, B. (1980). The pricing of options on debt securities. Journal of Financial and Quantitative Analysis, 15, 11–24.
Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5, 177–188.