A SIMPLE MODEL OF THE NOMINAL TERM STRUCTURE OF INTEREST RATES

BY YOUNGSOO CHOI AND TONY S. WIRJANTO

THIS VERSION: DECEMBER 2008

ABSTRACT. This paper presents a simple two-factor model of nominal term structure of interest rates, in which the log-price kernel has an autoregressive drift process and a nonlinear GARCH volatility process. With these two state-variable processes, closed-form solutions are derived for zero-coupon bond prices as well as yield to maturity for a given time to maturity.

1. INTRODUCTION

This paper presents a simple two-factor model of the term-structure of interest rates with the logarithm of the nominal discount factor (plus its long-term mean) and its conditional variance serving as state variables.[1] Motivated by prior research on interest-rate processes,[2] we model the dynamics of the state variables in the following way. The first state variable, which is the logarithm of the nominal discount factor, is specified to follow a first-order autoregressive, AR(1), process. The second state variable, which is the conditional variance of the first state variable, is modeled as a nonlinear asymmetric generalized autoregressive conditional heteroskedasticity (NGARCH) process. With the above two state-variable processes, closed-form solutions are derived for the zero-coupon bond price as well as the yield to maturity for a given time to maturity, with the resulting yield to maturity being affine in the two state variables. Equivalently, the yield is shown to be a function of the spread between the long-term and short-term rates, the difference between conditional variance and its long-term mean, as well as time to maturity. An alternative representation of the yield using another yield and its conditional volatility is also derived in this paper. Finally, a simple calibration exercise is performed to show that the proposed model is capable of producing a wide range

[1]As a two-factor model, it provides two state variables in squared-autoregressive-independent-variable nominal term structure (SAINTS) similar to the model proposed by Constantinides (1992).

[2]A partial list of the empirical studies includes Brennan and Schwartz (1982), Grossman, Melinio, and Shiller (1987), Longstaff (1989, 1990), and Chan, Karolyi, Longstaff, and Sanders (1992). See Chapman and Pearson (2001) for a critical survey on the empirical term-structure literature.

of shapes of the yield curves and its volatility curves can also assume a wide range of shapes other than that of the simple AR(1) model.

The remainder of the paper is organized as follows. In Section 2 we review the GARCH process, its continuous-time version, and a rearranged discrete-time version. The prices or yields of default-free bonds are derived and the functional dependence of the short-term rate and the yields on the logarithm of the discount factor are presented in Section 3. The importance of the selection of moments of both the volatility and yield curves is stressed by presenting typical patterns of the yield and volatility curves in Section 4. Section 5 contains concluding remarks. Finally detailed proofs for the prices and yields formula are relegated to the Appendix.

2. THE MODEL

Let $\ln m_t$ be the lograithm of the nominal discount factor at time t. Let $l_t = \ln m_t + \alpha$, where $-\alpha$ can be interpreted fruitfully as the long run mean for the short rate. Let σ_t^2 be the conditional variance of the logarithm of discount factor between t and $t + \Delta$, where Δ is the length of the equidistant time steps. This conditional variance is known from the information set at time t. Next let l_t and σ_t^2 be generated by the following processes over Δ

$$l_{t+\Delta} = (1 - \rho)l_t + \sigma_t v_{t+\Delta}, \qquad v_{t+\Delta} \overset{i.i.d}{\sim} N(0, 1) \tag{1}$$
$$\sigma_{t+\Delta}^2 = \beta_0 + \beta_1 \sigma_t^2 + \beta_2 \sigma_t^2 (v_{t+\Delta} - \gamma)^2, \tag{2}$$

where $v_{t+\Delta}$, conditional on information at time t, is a standard normal random variable and $i.i.d$ means "identically and independently distributed as".

The state variable l_t in (1) follows an AR(1) process, while the conditional variance σ_t^2 in (2) follows a nonlinear asymmetric GARCH (NGARCH) process, that has been studied by Engle and Ng (1993) and Duan (1995). Note that this model is quite similar to, but differs in some subtle ways from, the affine discrete-time GARCH models proposed by Heston and Nandi (2000). While the Heston-Nandi model is designed specifically to produce closed-form option prices, the specification in (1)-(2), like the Engle-Ng model, is designed foremost to provide a good empirical fit to the interest-rate data.

The variance process, $\sigma_{t+\Delta}^2$, and the logarithm of the nominal discount factor, $\ln m_{t+\Delta}$, are assumed to be correlated, such that

$$\text{Cov}_t\left(\sigma_{t+\Delta}^2, \ln m_{t+\Delta}\right) = -2\beta_2 \gamma \sigma_t^3. \tag{3}$$

Given $\beta_2 > 0$, the negative parameter γ captures the positive correlation between the discount factor and the volatility innovations. That is, γ controls the skewness or the asymmetry of the distribution of the discount factor. Furthermore, the third power term on σ_t accentuates the variation over time in the leverage effect. This feature of the model is intended to enhance the model's ability to fit the data to the extent that the leverage effect figures prominently in the term structure of interest rate. Thus our simple model accommodates two important stylized facts of the interest-rate data:

volatility clustering and leverage effect. We also note that when $\gamma = 0$, the model reduces to the GARCH model introduced by Bollerslev (1986).

Since $v_{t+\Delta}$ and $v_{t+\Delta}^2 - 1$ are uncorrelated by construction, the variance equation can be rearranged in the following form

$$\sigma_{t+\Delta}^2 - \sigma_t^2 = \beta_0 - \theta\sigma_t^2 - 2\beta_2\gamma\sigma_t^2 v_{t+\Delta} + \beta_2\sigma_t^2(v_{t+\Delta}^2 - 1),$$

where $\theta = 1 - \beta_1 - \beta_2(1 + \gamma^2)$ with $1 - \theta$ measuring the persistence of the variance. As the observation interval, Δ, shrinks to zero, a corresponding continuous-time system is obtained as:

$$\begin{aligned} dl_t &= -\rho\, l_t dt + \sigma_t dW_{1,t} \\ d\sigma_t^2 &= (\beta_0 - \theta\sigma_t^2)dt - 2\beta_2\gamma\sigma_t^2 dW_{1,t} + \beta_2\sigma_t^2 dW_{2,t}, \end{aligned} \tag{4}$$

where $(W_{1,t}, W_{2,t})$ is a bi-variate standard Brownian motion. From this continuous-time version, it is easy to see that the long-run variance of the logarithm of the discount factor is β_0/θ. That is, β_0/θ is the unconditional variance or, equivalently, the unconditional expectation of σ_t^2, which is $\mathrm{E}[\sigma_t^2] = \beta_0/\theta$. Also, we impose a condition for covariance stationarity such that $\mathrm{E}[\sigma_{t+\Delta}^2] = \mathrm{E}[\sigma_t^2]$ on the variance equation (2) via the following parameter restrictions:[3]

$$\beta_0 > 0 \qquad \text{and} \qquad \theta > 0, \ (\text{i.e., } \beta_1 + \beta_2(1 + \gamma^2) < 1).$$

We also impose another conditions for stationarity such that $\mathrm{E}[l_{t+\Delta}] = \mathrm{E}[l_t]$ on the mean equation (1) by requiring that the speed of the mean-reversion equation obeys the additional restriction that $0 < \rho < 2$.

The continuous-time model in matrix form can be expressed as

$$\begin{aligned} d\begin{pmatrix} l_t \\ \sigma_t^2 \end{pmatrix} &= \begin{pmatrix} -\rho\, l_t \\ \beta_0 - \theta\sigma_t^2 \end{pmatrix} dt + \begin{pmatrix} \sigma_t & 0 \\ -2\beta_2\gamma\sigma_t^2 & \beta_2\sigma_t^2 \end{pmatrix} \begin{pmatrix} dW_{1,t} \\ dW_{2,t} \end{pmatrix} \\ &= b\, dt + \begin{pmatrix} \sigma_{\cdot 1} & \sigma_{\cdot 2} \end{pmatrix} \begin{pmatrix} dW_{1,t} \\ dW_{2,t} \end{pmatrix} \end{aligned}$$

Next using the Itô-Taylor formula, the Euler-Maruyama approximation scheme of the continuous-time version of (4) can be written as

$$\begin{aligned} \begin{pmatrix} l_{t+\Delta} \\ \sigma_{t+\Delta}^2 \end{pmatrix} &= \begin{pmatrix} l_t \\ \sigma_t^2 \end{pmatrix} + b\Delta + \sum_{j=1}^{2} \sigma_{\cdot j}(W_{j,t+\Delta} - W_{j,t}) \\ &= \begin{pmatrix} l_t \\ \sigma_t^2 \end{pmatrix} + \begin{pmatrix} -\rho\, l_t \\ \beta_0 - \theta\sigma_t^2 \end{pmatrix} \Delta + \begin{pmatrix} \sigma_t \\ -2\beta_2\gamma\sigma_t^2 \end{pmatrix} \Delta_t W_1 + \begin{pmatrix} 0 \\ \beta_2\sigma_t^2 \end{pmatrix} \Delta_t W_2 \end{aligned}$$

---

[3]To ensure that the conditional variance is always positive further restrictions need to be imposed on β_1 and β_2 in (2). Alternatively, we can formulate the conditional variance by exponential GARCH (EGARCH) process (Nelson (1991)) instead of the NGARCH process (Equation (2)). The EGARCH process ensures positivity of the conditional variance and also allows for leverage effects and fat tails. However we choose the NGARCH process in this paper because it has been shown by the existing studies to be able to improve the fit of the model substantially better than the GARCH process.

where $\Delta_t W_j = W_{j,t+\Delta} - W_{j,t}$ is an independent normal distribution with zero mean and variance Δ, i.e., $N(0,\Delta)$. Thus, another way of writing the equations in (1) and (2) is

$$
\begin{aligned}
l_{t+\Delta} &= \rho_\Delta l_t + \sigma_t z_{t+\Delta}, \qquad z_{t+\Delta} \overset{i.i.d}{\sim} N(0,\Delta) \qquad (5)\\
\sigma^2_{t+\Delta} &= \beta_0 \Delta + \delta \sigma^2_t - 2\beta_2 \gamma \sigma^2_t z_{t+\Delta} + \beta_2 \sigma^2_t z^2_{t+\Delta}, \qquad (6)
\end{aligned}
$$

where $\rho_\Delta = 1 - \rho\Delta$ and $\delta = 1 - (\theta + \beta_2)\Delta = 1 - (1 - \beta_1 - \beta_2\gamma^2)\Delta$.

3. Zero-Coupon Bond Pricing and Yield to Maturity

It is well-known that the absence of arbitrage opportunities is characterized by the existence of an equivalent martingale measure $\mathbb{Q}$, so that the time-t price of a default-free, zero-coupon bond maturing at time $t + T$, $P_{t,T}$, is given by

$$
P_{t,T} = \mathrm{E}^{\mathbb{Q}}_t \left[\exp\left(\int_t^{t+T} \ln m_s \, ds \right) \right]. \qquad (7)
$$

By [i] partitioning the time interval $[t, t+T]$ into subintervals of equal size; [ii] utilizing the Euler-Maruyama approximation scheme in (5) and (6) and the tree property of conditional expectations; [iii] employing the trapezoidal rule to approximate the definite integral in the exponential function; and finally [iv] letting the subinterval size shrink to zero, we arrive at an analytical approximation formula for the nominal price at time t of the default-free, zero-coupon bond maturing at time $t + T$. The result is stated in the following theorem.[4]

Theorem 1. *If the yield factors follow the discrete stochastic differential equations in (5) and (6), the nominal price at time t of a default-free, zero-coupon bond maturing at time $t + T$, $P_{t,T}$, is given by*

$$
\ln(P_{t,T}) = -\alpha T + \frac{1 - e^{-\rho T}}{\rho} l_t + \frac{\beta_0}{2\theta\rho^2} f(T) + \frac{\sigma_t^2 - \beta_0/\theta}{2\rho^2} g(T), \qquad (8)
$$

where

$$
\begin{aligned}
f(T) &= f(T;\rho) = T - 2\frac{1 - e^{-\rho T}}{\rho} + \frac{1 - e^{-2\rho T}}{2\rho}\\
g(T) &= g(T;\rho,\theta) = \frac{1 - e^{-\theta T}}{\theta} - 2\frac{e^{-\theta T} - e^{-\rho T}}{\rho - \theta} + \frac{e^{-\theta T} - e^{-2\rho T}}{2\rho - \theta},
\end{aligned} \qquad (9)
$$

where $\theta = 1 - \beta_1 - \beta_2(1 + \gamma^2)$. Furthermore, $\mathrm{E}^{\mathbb{Q}}_t[\int_t^{t+T} \ln m_s \, ds]$ is given by the first two terms of $\ln P(t,T)$, $-\alpha T + \frac{1 - e^{-\rho T}}{\rho} l_t$.

In a simple case of constant conditional variance, $\sigma^2 = \beta_0/\theta$, its corresponding nominal price, $P_{t,T}$, has an exact analytical formula of the following form:

$$
\ln(P_{t,T}) = -\alpha T + \frac{1 - e^{-\rho T}}{\rho} l_t + \frac{\sigma^2}{2\rho^2} f(T).
$$

---

[4]This is a generalization of the result obtained in Choi and Wirjanto (2007) for the general one-factor model.

Proof: See Appendix.

Note that the nominal yield to maturity is defined as

$$y_t^{(T)} = -T^{-1}\ln(P_{t,T}).$$

From Theorem 1, the yield to maturity can be written in terms of the state variables, l_t and σ_t^2, as:[5]

$$y_t^{(T)} = \alpha - \frac{1 - e^{-\rho T}}{\rho T}l_t - \frac{\beta_0}{2\theta\rho^2 T}f(T) - \frac{\sigma_t^2 - \beta_0/\theta}{2\rho^2 T}g(T), \qquad (10)$$

where $f(T)$ and $g(T)$ are defined in (9).

As the time to maturity, T, tends to zero, the nominal (instantaneous) short rate reduces to the following expression

$$r_t = \lim_{T\to 0} y_t^{(T)} = \alpha - l_t = -\ln m_t, \qquad (11)$$

where the second equality is obtained by applying the L'Hospital's rule into functions such as $f(T)/T$ and $g(T)/T$. Thus, it follows as a simple computation using (4) that the dynamics for the nominal short rate can be written as a system of stochastic processes

$$\begin{aligned}
dr_t &= -dl_t = \rho(\alpha - r_t)dt - \sigma_t dW_{1,t} \\
d\sigma_t^2 &= (\beta_0 - \theta\sigma_t^2)dt - 2\beta_2\gamma\sigma_t^2 dW_{1,t} + \beta_2\sigma_t^2 dW_{2,t}.
\end{aligned} \qquad (12)$$

which essentially captures the tendency for the interest rate to revert to its long-run equilibrium level α at a speed ρ.

In a simple case of constant conditional variance, $\sigma^2 = \beta_0/\theta$, it reduces to the well-known Vasicek (1977) model and if the nominal price, $P_{t,T-t}$, is represented by $A(t,T)e^{-r_t B(t,T)}$, where $\ln(A(t,T)) = -\alpha(T-t) + \alpha(1 - e^{-\rho(T-t)})/\rho + \sigma^2 f(T-t)/(2\rho^2)$ and $B(t,T) = (1 - e^{-\rho(T-t)})/\rho$, then $A(t,T)$ and $B(t,T)$ satisfy the following system of differential equations

$$\begin{cases}
\dfrac{\partial A}{\partial t} - \rho\alpha AB + \dfrac{1}{2}\sigma^2 AB^2 = 0 \\
\dfrac{\partial B}{\partial t} - \rho B + 1 = 0.
\end{cases}$$

Similarly, as the time to maturity tends to infinity, the nominal long-term rate is defined as

$$y_t^{(\infty)} = \lim_{T\to\infty} y_t^{(T)} = \alpha - \frac{\beta_0}{2\theta\rho^2}. \qquad (13)$$

Note that it does not depend on the nominal short rate r_t. Thus, combining the short-term rate with the long-term rate, the nominal yield to maturity can be rearranged into the following form

$$y_t^{(T)} = y_t^{(\infty)} - \frac{1 - e^{-\rho T}}{\rho T}(y_t^{(\infty)} - r_t) + \frac{g(T)}{2\rho^2 T}(\beta_0/\theta - \sigma_t^2) + \frac{\beta_0}{4\theta\rho^3 T}(1 - e^{-\rho T})^2, \qquad (14)$$

---

[5]This shows that the yield to maturity is an affine function of the two state variables (logarithm of discount factor and conditional variance) defined in (1) and (2), in contrast to quadratic models (such as Ahn, Dittmar and Gallant (2002)) in which the yield is a quadratic function of the state variables.

which implies that the yield to maturity is obtained by adjusting the long-term rate by the spread between the long-term and short-term rates, the *difference between the current and long-run variances* (Note: this is a new feature resulting from the GARCH effect on the variance equation), and the time to maturity.

Since the state variable, l_t or r_t,[6] is unobservable, when we estimate the model parameters or calibrate the model, the shortest yield, for example, $y_t^{(\Delta)}$ for a Δ time-period, is an alternative state variable instead of l_t or r_t. To this end, we state the following proposition:

Proposition 1. *The mean of the logarithm of nominal discount factor, l_t, can be expressed in terms of the one-period ahead state variables, $y_{t-\Delta}^{(\Delta)}$ and $\sigma_{t-\Delta}$, as*

$$l_t = \sqrt{\Delta}\sigma_{t-\Delta}\psi_t - \frac{\rho\Delta(1-\rho\Delta)}{1-e^{-\rho\Delta}}\left(y_{t-\Delta}^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] + \frac{\sigma_{t-\Delta}^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)\right) \quad (15)$$

where ψ_t is a standard normal distribution.

Proof: See Appendix.

Substituting r_t with $\alpha - l_t$ in the equation (15), we have another form of (14) which is stated with its variance and kurtosis in the following theorem.

Theorem 2. *If the yield factors follow the discrete stochastic differential equations in (5) and (6), the yield to maturity, $y_t^{(T)}$, can be written in terms of the shortest yield, $y_{t-\Delta}^{(\Delta)}$, and the volatility, $\sigma_{t-\Delta}$, as*

$$\begin{aligned}
y_t^{(T)} &= \frac{\Delta(1-\rho\Delta)}{T}\frac{1-e^{-\rho T}}{1-e^{-\rho\Delta}}\left(y_{t-\Delta}^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] + \frac{\sigma_{t-\Delta}^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)\right) \\
&\quad + y_t^{(\infty)} - \frac{1-e^{-\rho T}}{\rho T}\sigma_{t-\Delta}\sqrt{\Delta}\psi_t + \frac{g(T)}{2\rho^2 T}(\beta_0/\theta - \sigma_t^2) + \frac{\beta_0}{2\theta\rho^2}\left(1 - \frac{f(T)}{T}\right),
\end{aligned}$$

(16)

where $\sigma_t^2 = \beta_0\Delta + \delta\sigma_{t-\Delta}^2 - 2\beta_2\gamma\sigma_{t-\Delta}^2\sqrt{\Delta}\psi_t + \beta_2\sigma_{t-\Delta}^2\Delta\psi_t^2$ and $\delta = 1 - (1-\beta_1-\beta_2\gamma^2)\Delta$ and its variance and kurtosis per Δ time period are

$$\mathrm{Var}_{t-\Delta}\big(y_t^{(T)}\big) = \mathrm{E}_{t-\Delta}[\,(u_t)^2\,] = C^2 + 2D^2$$

$$K\big(y_t^{(T)}\big) = \frac{\mathrm{E}_{t-\Delta}\big[(u_t)^4\big]}{\big(\mathrm{Var}_{t-\Delta}(y_t^{(T)})\big)^2} = \frac{3C^4 + 2D^4}{(C^2+2D^2)^2} < 3,$$

where $u_t = y_t^{(T)} - \mathrm{E}_{t-\Delta}[\,y_t^{(T)}\,]$ and

$$C = \frac{\sigma_{t-\Delta}\sqrt{\Delta}}{\rho T}\big(\beta_2\gamma g(T)\sigma_{t-\Delta}/\rho - (1-e^{-\rho T})\big)$$

$$D = -\frac{g(T)}{2\rho^2 T}\beta_2\sigma_{t-\Delta}^2\Delta.$$

Proof: See Appendix.

---

[6]From (11) and (13), it clearly does not matter whether we choose l_t or r_t.

Note that Theorem 2 implies that the excess kurtosis of nominal yields to maturity is negative when the logarithm of the nominal discount factor is specified as an AR(1)-NGARCH(1,1) process.[7]

Before proceeding to the next section, we make a few more remarks about the features of our proposed model vis-a-vis those of the existing models in the literature. First, for a fixed maturity time T, the zero-coupon bond price at time t is given by

$$P_{t,T-t} = F(l_t, \sigma_t^2, t) = \exp\big(A(T-t) + B_l(T-t)l_t + B_\sigma(T-t)\sigma_t^2\big), \qquad (17)$$

where

$$A(T-t) \;=\; -\alpha(T-t) + \frac{\beta_0}{2\theta\rho^2}\big[f(T-t) - g(T-t)\big]$$

$$B_l(T-t) \;=\; \frac{1 - e^{-\rho(T-t)}}{\rho}$$

$$B_\sigma(T-t) \;=\; \frac{g(T-t)}{2\rho^2}$$

Then by the Itô formula, the zero-coupon bond-price process can be expressed as

$$dP_{t,T-t} = \mathcal{D}F(l_t, \sigma_t^2, t)dt + \big(B_l(T-t)\; B_\sigma(T-t)\big)\begin{pmatrix} \sigma_t & 0 \\ -2\beta_2\gamma\sigma_t^2 & \beta_2\sigma_t^2 \end{pmatrix}\begin{pmatrix} dW_{1,t} \\ dW_{2,t} \end{pmatrix}, \quad (18)$$

where

$$\mathcal{D}F(l_t, \sigma_t^2, t) = F\Big(-A'(\tau) - B_l'(\tau)l_t - B_\sigma'(\tau)\sigma_t^2 - \rho B_l(\tau)l_t + B_\sigma(\tau)(\beta_0 - \theta\sigma_t^2)$$
$$+ \frac{1}{2}B_l^2(\tau)\sigma_t^2 - 2\beta_2\gamma B_l(\tau)B_\sigma(\tau)\sigma_t^3 + \frac{1}{2}\beta_2^2(1+4\gamma^2)B_\sigma^2(\tau)\sigma_t^4\Big) \quad (19)$$
$$= F\Big(\alpha - l_t - 2\beta_2\gamma B_l(\tau)B_\sigma(\tau)\sigma_t^3 + \frac{1}{2}\beta_2^2(1+4\gamma^2)B_\sigma^2(\tau)\sigma_t^4\Big)$$

with $\tau = T - t$. This result shows that our model nests the quadratic term structure model (QTSM) for $\gamma = 0$. This QTM is proposed by Constantinides (1992) and further studied by Ahn, Dittmar, and Gallant (2002). More importantly, equation (17) shows that our model is *exponential-affine* in the sense of Duffie and Kan (1996); however it is not affine, since equation (19) shows that equation (3.8) in Duffie and Kan (1996) is not satisfied.

Second, the instantaneous, arithmetic expected bond return for a fixed maturity time T with $\tau = T - t$ is defined as $R(t,\tau) = \mathrm{E}_t[dP_{t,\tau}]/P_{t,\tau}dt$. Then the arithmetic term premium, defined as $R(t,\tau) - r_t$, is given by

$$R(t, T-t) - r_t = \beta_2 B_\sigma(\tau)\sigma_t^3\big[\frac{1}{2}\beta_2(1+4\gamma^2)B_\sigma(\tau)\sigma_t - 2\gamma B_l(\tau)\big] \qquad (20)$$

which is obtained by taking the conditional expectation of (19) and using the result that $r_t = \alpha - l_t$. This result shows that, for a given maturity bond, the term premium can switch signs, depending on the state variable and the conditional variance σ_t^2.

---

[7]This should be interpreted as a short-term prediction only as a negative excess kurtosis of the nominal yields to maturity in the long run is hard to reconcile with the empirical data of the interest rates.

Third, given the set up of our proposed model, it is almost too tempting to argue that since equation (4) assumes that the drift of the short rate depends on the short rate only and the drift of the variance depends on the variance only, and also since the variance of the short rate is the only state variable, driving the variance of the variance, equation (4) represents a restricted version of the well-known two-factor model of Longstaff and Schwartz (1992) where both the short rate and the state variable V_t appears in the two drift terms and the two diffusion terms. However a closer examination reveals that our model is not nested within the Longstaff-Schwartz model. That is, although equation (21) in the Longstaff-Schwartz model bears similarity to our equation (10), there are also important differences between our proposed model and the Longstaff-Schwartz model. This is best seen by differentiating equation (14) with respect to r_t and use equation (12) for dr_t and the second equation in (14) for $d\sigma_t$ to obtain the sensitivity of the yield with respect to the interest rate

$$dy_t^T = f_1(T)(\alpha - r_t) - f_2(T)(\beta_0 - \theta\sigma_t^2)dt + [f_3(T)\sigma_t^2 - f_4(T)\sigma_t]dW_{1t} - f_5(T)\sigma_t^2 dW_{2t}$$

where $f_j(T)$ for $j = 1, ..., 5$ are coefficients which are functions of T; that is, although the drift term of our yield-to-maturity equation is the same as that in the Longstaff-Schwartz model, the diffusion terms have different specifications. In particular, the diffusion term in our model depends not only on σ_t as does the diffusion term in their model, but it also depends on σ_t^2. Moreover, the dynamics of the state variables, represented by equations (2) and (3) in Longstaff and Schwartz (1992) are clearly different from the dynamics characterizing our model in (4).

4. The yield and Volatility Curves

In this section we provide the typical patterns of the yield and volatility curves of the nominal yield-to-maturity of the AR(1) process (which is the Vasicek (1977) model) as well as the AR(1)-NGARCH(1,1) process with the time-to-maturity and show the importance of the moments of both the volatility and yield curves when we calibrate the model to match market data.

As we noted in (10) or (16), parameters affecting the shape of the yield curves are the current volatility, σ_t, the long-run variance, β_0/θ, the AR(1) coefficient, ρ, the reproduced parameter from a stationary restriction on the conditional variance equation, θ, the current short-term yield, $y_{t-\Delta}^{(T)}$, the long-run short-term yield, $\mathrm{E}[\, y^{(\Delta)}\,]$, and the long-term spread, α. Figures $1-4$ present the typical patterns of the yield and volatility curves of the AR(1) or AR(1)-NGARCH(1,1) processes and their related functions such as $f(T; \rho)$ and $g(T; \rho, \theta)$ defined in Theorem 1 and are calculated from the parameter values set in Table 1. For the random innovation at $t - \Delta$, we assume that $\psi_t \sim N(0, 1)$ have 7 possible values: $[-2, -1, -0.5, 0, 0.5, 1, 2]$.

The left (right) panels of Figures 1 and 3 present the yield curves for the AR(1) (AR(1)-NGARCH(1,1)) process and the left panels of Figures 2 and 4 depict the volatility curves of the AR(1) process, the AR(1)-NGARCH(1,1) process, and the market data for the yield volatility of On-the-Run Treasuries in 1987 presented in Fabozzi (1993) -

Parameters	ρ	θ	σ_t	β_0/θ	$y_{t-\Delta}^{(T)}$	$\mathrm{E}[\,y^{(\Delta)}\,]$	α	γ
1(a)	0.12	0.08	0.012	0.02	0.02	0.05	0.07	-0.56
1(c)	0.25	0.02	0.02	0.015	0.04	0.05	0.08	-2
1(e)	0.05	0.03	0.01	0.015	0.035	0.055	0.07	-0.6
3(a)	0.2	0.25	0.01	0.02	0.03	0.05	0.06	-0.4
3(c)	0.07	0.1	0.02	0.022	0.05	0.04	0.07	-0.5
3(e)	0.16	0.01	0.02	0.01	0.05	0.035	0.07	-0.6

TABLE 1. **Parameter Values for Figures 1– 4** This table provides parameter values for Figures 1–4. Since $\theta = 1 - \beta_1 - \beta_2(1 + \gamma^2)$ is a function of β_1 and β_2, given by γ, there are infinite pairs (β_1, β_2) satisfying θ; thus, we impose the restriction that $\beta_1 = (1 - \theta)/6$.

See in particular Exhibit 22-10 in Fabozzi (1993). Finally, the right panels of Figures 2 and 4 present the function, $f(T; \rho)$, which appears in both processes and the function, $g(T; \rho, \theta)$, which appears only in the NGARCH processes. These panels illustrate the effects of the magnitude of function values in $f(T)$ and $g(T)$ on the volatility curve.

Since the random part in (16) at time $t - \Delta$ can be written as

$$C\psi_t + D\psi_t^2,$$

where $C = \frac{\sigma_{t-\Delta}\sqrt{\Delta}}{\rho T}\big(\beta_2 \gamma g(T)\sigma_{t-\Delta}/\rho - (1 - e^{-\rho T})\big)$ and $D = -\frac{g(T)}{2\rho^2 T}\beta_2 \sigma_{t-\Delta}^2 \Delta$, the volatility of yield curve, which we call a volatility curve, is calculated as

$$\sqrt{\mathrm{Var}_{t-\Delta}\Big(\frac{y_t^{(T)} - y_{t-\Delta}^{(T)}}{y_{t-\Delta}^{(T)}}\Big)} = \frac{\sqrt{C^2 + 2D^2}}{y_{t-\Delta}^{(T)}}.$$

Figures 1a–1b, 3a–3b, and 2a and 4a illustrate that although the yield curves are uniformly upward-sloping, the volatility curves take different shapes depending on whether the corresponding model is an AR(1) process or an AR(1)-NGARCH(1,1) process. In addition they show that the yield curves retain their upward-sloping shape independent of the sign of the shock. However, Figures 1c–1d and 2c illustrate that the yield curve can change its shape when the shock is negative. Furthermore, Figures 1e–1f and 2e show that although the model parameter values are kept the same, except for the parameter value of γ, the yield curves take different shapes. The volatility curves display even more strikingly different shapes. In this regard, the AR(1)-NGARCH(1,1) process has a volatility curve, which has been shown empirically to have fitted the observed volatility curve very well. Thus, we have demonstrated that the NGARCH model for the conditional variance of interest rates is a necessary component of the model and it is important that we take into account the moments of the volatility and yield curves in calibrating the model to match the market data. Finally, Figures 3c–3d, 3e–3f, and 4c and 4e produce the results for the case of the downward-sloping yield curves.

5. Conclusions

This paper presented a simple model of the nominal term structure of interest rates. The proposed model was derived with specific considerations for data availability and model tractability as discussed so eloquently in Dai and Singleton (2003). To achieve this objective, we established a linkage between the discrete-time version of the model and its continuous-time counterparts. This was done in two steps. First, the nominal discount factor was selected as a state variable of the model. Second, the logarithm of the discount factor was specified as an AR(1) process with its conditional variance following an NGARCH process. This particular modeling strategy has several appeals: [i] the NGARCH process has been shown in the literature to have fitted market data empirically rather well; [ii] it allows mathematical tractability in deriving the formula for the prices or yields-to-maturity (yields) of default-free bonds; [iii] the functional dependence of the short-rate on this state variable was easily obtained from the established linkage between the discrete-time model and its continuous-time counterpart. In other words, the model tractability, which is obtained through the linkage between the discrete-time GARCH models and the bivariate diffusion processes as limiting cases, was exploited to show that the short-term rate is linearly dependent on the logarithm of this state variable in the limit; [iv] in a simple case of constant conditional variance, the discrete-time term-structure model reduces to the well-known Vasicek (1977) model; and finally [v] the comparison between the yield volatility of On-the-Run Treasuries with the volatility curve of our model indicates that it is important to take into account the moments of both the volatility and yield curves in calibrating the model for the purpose of matching with the real data.

The obvious next step is to actually establish the empirical advantages of the proposed model vis-a-vis the existing comparable models based on real data evaluation. To do so, we would first need to carefully discuss several critical statistical issues such as: what are the conditional densities of the state variables in this model? Which estimation technique is most appropriate to the model? What restrictions on the parameters of the model should be imposed to rule out arbitrage opportunities? What econometric issues may be involved when the state variables of the model are unobservable?, and so on. Moreover, an assessment of the empirical goodness-of-fit of the proposed model may also prove worthwhile. In particular, it is useful to analyze the models ability to capture the historical movements in the yields and volatilities for a full sample as well as for different subsamples, and perform across models comparisons. In this regard, since the proposed model in this paper is a two-factor model, good candidates for the model comparisons include a two-factor affine Gaussian model (with constant volatility), and Longstaff and Schwartz (1992, using interest rate and its volatility as state variables). Other multifactor models known to have empirical support, e.g. Ahn, Dittmar and Gallant (2002), and Dai and Singleton (2000), may also serve as good candidates for this model comparison exercise. These empirical exercises would, among other things, allow us [i] to better asses how well our model actually capture the historical movement in yields, [ii] whether it is able to produce different shapes of the yield curves, [iii] whether the model's parameters can easily be identified, and [iv] which parameters of the model actually govern the dynamics of the state variables. All of these empirical

questions need to be answered satisfactorily before the empirical appeal of the proposed model can be truly established.

Proof of Theorem 1:

By partitioning the interval $[t, t+T]$ into $t_0 = t < t_1 < \cdots < t_n = t + T$ with $t_j = t + j\Delta$ and $\Delta = \frac{T}{n}$, we compute the conditional expectation of $\exp\left(\int_t^{t+T} \ln m_s \, ds\right)$ under the equivalent martingale measure $\mathbb{Q}$. For simplicity, we denote l_{t_j}, σ_{t_j}, z_{t_j}, and $\mathrm{E}_{t_j}^{\mathbb{Q}}$ by l_j, σ_j, z_j, and E_j, respectively for $j = 0, 1, \ldots, n-1$ and assume that $t = 0$.

The trapezoidal rule is employed to approximate the definite integral in the exponent function of (7)), so that the integral part can be written in terms of l_t instead of $\ln m_t$

$$\int_t^{t+T} \ln m_s \, ds \approx -\alpha T + \Delta I_n,$$

where $I_n = l_0/2 + l_1 + \cdots + l_{n-1} + l_n/2$. It follows as an application of the tree property of conditional expectation that

$$\exp(\delta T) P_{t,T} \approx \mathrm{E}_0 \, \mathrm{E}_1 \cdots \mathrm{E}_{n-1} \left[e^{\Delta I_n} \right].$$

To compute the conditional expectation of $e^{\Delta I_n}$, it is necessary to represent l_j in terms of random variables $\{z_j\}_{j=1,\ldots,n}$ and obtain as an application of induction arguments on AR(1) process that

$$l_j = \rho_\Delta^j l_0 + \rho_\Delta^{j-1} \sigma_0 z_1 + \cdots + \rho_\Delta \sigma_{j-2} z_{j-1} + \sigma_{j-1} z_j$$

for $j = 1, \ldots, n$, where $\rho_\Delta = 1 - \rho\Delta$. A simple computation using the above formula yields the result that

$$I_n = a_{n+1} l_0 + \sum_{j=1}^n a_{n-j+1} \sigma_{j-1} z_j, \tag{21}$$

where $a_1 = 1/2$, $a_j = \frac{1 - \rho_\Delta^{j-1}}{1 - \rho_\Delta} + \frac{\rho_\Delta^{j-1}}{2}$ for $j = 2, \ldots, n$, and $a_{n+1} = \frac{1}{2} + \rho_\Delta \frac{1 - \rho_\Delta^{n-1}}{1 - \rho_\Delta} + \frac{\rho_\Delta^n}{2}$. To compute the conditional expectations, we stated a well-known result in the following lemma.

Lemma A1. *Suppose that ψ is a standard normal distribution, i.e., $\psi \sim N(0,1)$. Then we have*
$$\mathrm{E}[e^{b\sqrt{T}\psi}] = e^{\frac{1}{2}b^2 T}.$$
Furthermore, the moment generating function of $Q(\psi) = (\psi - w)^2$ is

$$\mathrm{E}[e^{vQ(w)}] = \exp\left(-\frac{w^2}{2}\right) \frac{1}{\sqrt{1 - 2v}} \exp\left(\frac{w^2}{2(1 - 2v)}\right) \tag{22}$$

First, applying the first equation of Lemma A1 into the case $a_1 \sigma_{n-1} z_n$ in I_n, we have

$$\mathrm{E}_{n-1}\left[e^{\Delta a_1 \sigma_{n-1} z_n} \right] = e^{\Delta^3 a_1^2 \sigma_{n-1}^2 / 2},$$

which implies that

$$\ln \mathrm{E}_{n-1}\left[e^{\Delta I_n} \right] = \Delta I_{n-1} + \Delta^3 a_1^2 \sigma_{n-1}^2 / 2$$

where

$$I_k = a_{n+1}l_0 + \sum_{j=1}^{k} a_{n-j+1}\sigma_{j-1}z_j$$

for $k = 1, 2, \ldots, n$. Using the recursive formula in (6) with $t = t_{n-2}$, the random part in $\ln \mathrm{E}_{n-1}[e^{\Delta I_n}]$ at time t_{n-2} is written as

$$I_{n-1,r} = \Delta \left(a_2\sigma_{n-2}z_{n-1} + \Delta^2 a_1^2\sigma_{n-1}^2/2\right)$$

$$= v_{n-1}(\psi_{n-1} - w_{n-1})^2 - v_{n-1}w_{n-1}^2 + \frac{\Delta^3}{2}a_1^2(\beta_0\Delta + \delta\sigma_{n-2}^2),$$

where $v_{n-1} = \beta_2\Delta^4 a_1^2\sigma_{n-2}^2/2$, $w_{n-1} = [\gamma - a_2/(\beta_2\Delta^2 a_1^2\sigma_{n-2})]/\sqrt{\Delta}$, and $\psi_{n-1} = z_{n-1}/\sqrt{\Delta}$ is a standard normal distribution.

Second, applying (18) into $v_{n-1}(\psi_{n-1} - w_{n-1})^2$ in $I_{n-1,r}$ with v_{n-1} and w_{n-1} and the tree property of conditional expectation, we have

$$\ln \mathrm{E}_{n-2}[e^{\Delta I_n}] = \Delta I_{n-2} + \frac{\Delta^3}{2}a_1^2(\beta_0\Delta + \delta\sigma_{n-2}^2)$$

$$+ \frac{w_{n-1}^2}{2}\left(1/(1 - 2v_{n-1}) - 1 - 2v_{n-1}\right) - \frac{1}{2}\ln(1 - 2v_{n-1}).$$

As a subinterval size, Δ, shrinks, v_{n-1} is sufficiently small. Thus we can approximate $\ln(1 - 2v_{n-1})$ by $-2v_{n-1}$ and $1/(1 - 2v_{n-1}) - 1 - 2v_{n-1} = 4v_{n-1}^2/(1 - 2v_{n-1})$ by $4v_{n-1}^2$, and then obtain

$$\ln \mathrm{E}_{n-2}[e^{\Delta I_n}] \approx \Delta I_{n-2} + \frac{\Delta^3}{2}a_1^2(\beta_0\Delta + \delta\sigma_{n-2}^2) + 2v_{n-1}^2 w_{n-1}^2 + v_{n-1}$$

$$= \Delta I_{n-2} + \frac{\Delta^4}{2}\beta_0 a_1^2 + \frac{\Delta^3}{2}a_1^2\sigma_{n-2}^2\left(\delta + \beta_2\Delta + (\beta_2\gamma a_1\sigma_{n-2}\Delta^2 - a_2/a_1)^2\right)$$

$$\approx \Delta I_{n-2} + \frac{\Delta^4}{2}\beta_0 a_1^2 + \frac{\Delta^3}{2}a_1^2\sigma_{n-2}^2\left(\delta + \beta_2\Delta + (a_2/a_1)^2\right)$$

$$= \Delta I_{n-2} + \frac{\Delta^4}{2}\beta_0 a_1^2 + \frac{\Delta^3}{2}\sigma_{n-2}^2 b_2,$$

where the second approximation comes from approximating $\beta_2\gamma a_1\sigma_{n-2}\Delta^2 - a_2/a_1$ by a_2/a_1, since $a_2 = 1 + \rho_\Delta/2$ and $a_1 = 1/2$ implies that $a_2/a_1 = O(1)$, and $b_2 = (\delta + \beta_2\Delta)b_1 + a_2^2$ with the initial value $b_1 = a_1^2$.

Similarly, using the recursive formula in (6) with $t = t_{n-3}$, the random part in $\ln \mathrm{E}_{n-2}[e^{\Delta I_n}]$ at time t_{n-3} is written as

$$I_{n-2,r} = \Delta \left(a_3\sigma_{n-3}z_{n-2} + \Delta^2 b_2\sigma_{n-2}^2/2\right)$$

$$= v_{n-2}(\psi_{n-2} - w_{n-2})^2 - v_{n-2}w_{n-2}^2 + \frac{\Delta^3}{2}b_2(\beta_0\Delta + \delta\sigma_{n-3}^2),$$

where $v_{n-2} = \beta_2\Delta^4 b_2\sigma_{n-3}^2/2$, $w_{n-2} = [\gamma - a_3/(\beta_2\Delta^2 b_2\sigma_{n-3})]/\sqrt{\Delta}$, and $\psi_{n-2} = z_{n-2}/\sqrt{\Delta}$ is a standard normal distribution. Thus, we have a similar computational structure as in the case of $t = t_{n-2}$. That is, we have

$$\ln \mathrm{E}_{n-3}[e^{\Delta I_n}] = \Delta I_{n-3} + \frac{\Delta^4}{2}\beta_0(b_1 + b_2) + \frac{\Delta^3}{2}\sigma_{n-3}^2 b_3,$$

where $b_j = (\delta + \beta_2 \Delta) b_{j-1} + a_j^2 = (1 - \theta\Delta) b_{j-1} + a_j^2$ with the initial value $b_1 = a_1^2$ for $j = 2, \ldots, n$.

Continuing this procedure, we have

$$\ln \mathrm{E}_1[\, e^{\Delta I_n}\,] = \Delta I_1 + \frac{\Delta^4}{2} \beta_0 (b_1 + b_2 + \cdots + b_{n-2}) + \frac{\Delta^3}{2} \sigma_1^2 b_{n-1}.$$

Finally,

$$\begin{aligned}
\ln \mathrm{E}_0[\, e^{\Delta I_n}\,] &= a_{n+1} l_0 \Delta + \frac{\Delta^4}{2} \beta_0 (b_1 + b_2 + \cdots + b_{n-2}) \\
&\quad + \ln \mathrm{E}_0[\, e^{\Delta(a_n \sigma_0 z_1 + \Delta^2 b_{n-1} \sigma_1^2 / 2)}\,] \\
&= a_{n+1} l_0 \Delta + \frac{\Delta^4}{2} \beta_0 (b_1 + b_2 + \cdots + b_{n-1}) + \frac{\Delta^3}{2} \sigma_0^2 b_n.
\end{aligned} \tag{23}$$

To obtain the convergence result as n approaches $+\infty$, we need several computational steps. Recall that the number e is defined as the limit of the sequence, i.e.,

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e,$$

which implies that for $\Delta = T/n$,

$$\lim_{n \to \infty} \left(1 - c\Delta\right)^n = e^{-cT}.$$

Applying this result to a sequence, we have the following lemma.

Lemma A2. *Given the sequence $\{a_j\}$, if a sequence $\{b_n\}$ is described by $b_1 = a_1^2$ and the recursive relationship $b_j = (1 - \theta\Delta) b_{j-1} + a_j^2$ for $j = 2, \ldots, n$, then b_j can be explicitly written as*

$$b_j = \sum_{i=1}^{j} a_i^2 (1 - \theta\Delta)^{j-i}.$$

Also, its partial sum is

$$\sum_{j=1}^{n-1} b_j = \left(\sum_{j=1}^{n} a_j^2 - b_n\right) / (\theta\Delta). \tag{24}$$

In particular, if $a_j = \frac{1}{\rho\Delta} + c_\rho \rho_\Delta^{j-1}$ with $c_\rho = \frac{1}{2} - \frac{1}{\rho\Delta}$ for $j = 1, 2, \ldots, n$,

$$\begin{aligned}
\Delta^3 b_n &= \frac{1 - (1 - \theta\Delta)^{n-1}}{\rho^2 \theta} + \left(\frac{\Delta^2}{4} - \frac{2 c_\rho \Delta}{\rho} - c_\rho^2 \Delta^2\right)(1 - \theta\Delta)^{n-1} \Delta \\
&\quad + \frac{2 c_\rho \Delta}{\rho} \frac{(1 - \theta\Delta)^n - (1 - \rho\Delta)^n}{\rho - \theta} + c_\rho^2 \Delta^2 \frac{(1 - \theta\Delta)^n - (1 - \rho\Delta)^{2n}}{2\rho - \theta - \rho^2 \Delta}
\end{aligned}$$

Thus, its limit, $\lim_{n \to \infty} \Delta^3 b_n$, is $g(T; \rho, \theta)/\rho^2$, where

$$g(T; \rho, \theta) = \frac{1 - e^{-\theta T}}{\theta} - 2\frac{e^{-\theta T} - e^{-\rho T}}{\rho - \theta} + \frac{e^{-\theta T} - e^{-2\rho T}}{2\rho - \theta}$$

Also, we have

$$\Delta^3 \sum_{j=1}^{n} a_j^2 = \frac{\Delta n}{\rho^2} + \frac{2 c_\rho \Delta}{\rho^2} (1 - \rho_\Delta^n) + c_\rho^2 \Delta^2 \frac{1 - \rho_\Delta^{2n}}{2\rho - \rho^2 \Delta}$$

and its limit, $\lim_{n\to\infty} \Delta^3 \sum_{j=1}^{n} a_j^2$, *is* $f(T;\rho)/\rho^2$, *where*

$$f(T;\rho) = T - 2\frac{1 - e^{-\rho T}}{\rho} + \frac{1 - e^{-2\rho T}}{2\rho}$$

where $c_\rho = 1/2 - 1/(\rho\Delta)$, *hence,* $\lim_{n\to\infty} \Delta c_\rho = -1/\rho$. *Finally,* $a_{n+1}\Delta = (\frac{1}{2} + \rho_\Delta \frac{1 - \rho_\Delta^{n-1}}{1 - \rho_\Delta} + \frac{\rho_\Delta^n}{2})\Delta$ *has the limit*

$$\frac{1 - e^{-\rho T}}{\rho}.$$

Substituting the relationship in (20) into (19) and rearranging terms give us the result that

$$\ln E_0[\, e^{\Delta I_n}\,] = a_{n+1}l_0\Delta + \frac{\beta_0}{2\theta}\Delta^3 \sum_{j=1}^{n} a_j^2 + \frac{\sigma_0^2 - \beta_0/\theta}{2}\Delta^3 b_n.$$

Thus, as the observation interval approaches zero, the desired result (8) is obtained as a simple application of Lemma A2.

Also, by the linear property of the expectation operator and the *i.i.d* property of $\{z_j\}$, it is easy to see that

$$E_0\Big[\int_t^{t+T} \ln m_s\, ds\,\Big] \approx -\alpha T + a_{n+1}l_0\Delta \to -\alpha T + \frac{1 - e^{-\rho T}}{\rho}l_0$$

In a simple case of constant conditional variance, $\sigma^2 = \beta_0/\sigma$, the recursive formula of l_j can be written as

$$l_j = \rho_\Delta^j l_0 + \rho_\Delta^{j-1}\sigma\sqrt{\Delta}\psi_1 + \cdots + \rho_\Delta\sigma\sqrt{\Delta}\psi_{j-1} + \sigma\sqrt{\Delta}\psi_j$$

for $j = 1, \ldots, n$, where $\rho_\Delta = 1 - \rho\Delta$ and $\{\psi_j = z_j/\Delta\}$ are i.i.d standard normals. Thus, we have a similar form for (17) as follows

$$I_n = a_{n+1}l_0 + \sum_{j=1}^{n} a_{n-j+1}\sigma\sqrt{\Delta}\psi_j.$$

Using the i.i.d property of $\{\psi_j\}$, it is easy to obtain a similar form for (19) as follows

$$\ln E_0[\, e^{\Delta I_n}\,] = a_{n+1}l_0\Delta + \frac{\Delta^3}{2}\sigma^2 \sum_{j=1}^{n} a_j^2.$$

As the observation interval approaches zero, the limit results in Lemma 2 provides us the desired result.

Proof of Proposition 1:

Recall that we have equation (10) with $T = \Delta$

$$y_t^{(\Delta)} = \alpha - \frac{1 - e^{-\rho\Delta}}{\rho\Delta}l_t - \frac{\beta_0}{2\theta\rho^2\Delta}f(\Delta) - \frac{\sigma_t^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)$$

Applying the stationary property such as $\mathrm{E}[\,l_t\,] = 0$ and $\mathrm{E}[\,\sigma_t^2\,] = \beta_2/\theta$, this equation can be rearranged as

$$y_t^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] = -\frac{1 - e^{-\rho\Delta}}{\rho\Delta}l_t - \frac{\sigma_t^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta) \tag{25}$$

Substituting l_t by (5) in terms of $l_{t-\Delta}$ and $\sigma_{t-\Delta}$ and replacing $l_{t-\Delta}$ by (21) in terms of $y_{t-\Delta}^{(\Delta)}$ and $\sigma_{t-\Delta}$, we have another form of (21) as follows

$$y_t^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] = (1 - \rho\Delta)\left[y_{t-\Delta}^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] + \frac{\sigma_{t-\Delta}^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)\right]$$
$$- \frac{1 - e^{-\rho\Delta}}{\rho\Delta}\sigma_{t-\Delta}\sqrt{\Delta}\psi_t - \frac{\sigma_t^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)$$

Subtracting the above equation from (21), we have

$$\frac{1 - e^{-\rho\Delta}}{\rho\Delta}l_t = \frac{1 - e^{-\rho\Delta}}{\rho\Delta}\sigma_{t-\Delta}\sqrt{\Delta}\psi_t - (1 - \rho\Delta)\left[y_{t-\Delta}^{(\Delta)} - \mathrm{E}[\,y^{(\Delta)}\,] + \frac{\sigma_{t-\Delta}^2 - \beta_0/\theta}{2\rho^2\Delta}g(\Delta)\right],$$

which gives us the desired result in (15).

Proof of Theorem 2:

Combining (11) with (12), we have

$$y_t^{(\infty)} - r_t = l_t - \frac{\beta_0}{2\theta\rho^2}.$$

Substituting l_t in this above equation by (15) and then plugging this result into (14) give us the desired result in (16).

The random part in (16) at time $t - \Delta$ can be written as

$$C\psi_t + D\psi_t^2,$$

where $C = \frac{\sigma_{t-\Delta}\sqrt{\Delta}}{\rho T}\big(\beta_2\gamma g(T)\sigma_{t-\Delta}/\rho - (1 - e^{-\rho T})\big)$ and $D = -\frac{g(T)}{2\rho^2 T}\beta_2\sigma_{t-\Delta}^2\Delta$. Thus, using the moments of the standard normal distribution ψ, that is, $\mathrm{E}[\,\psi^{2K}\,] = 1 \times 3 \times \cdots \times (2k - 1)$ for $k = 1, 2, \ldots$ and the fact that the odd moments are zero, we have

$$y_t^{(T)} - \mathrm{E}_{t-\Delta}[\,y_t^{(T)}\,] = C\psi_t + D(\psi_t^2 - 1).$$

A simple calculation provides us the desired result about the variance and kurtosis of the nominal yields to maturity.

References

[1] Ahn, D.-H., R.F. Dittmar, and A.R. Gallant, 2002, Quadratic Term Structure Models: Theory and Evidence, *Review of Financial Studies,* 15, 243-288.

[2] Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, 307-327

[3] Brennan, M. J. and E. S. Schwartz, 1982, An Equilibrium Model of Bond Pricing and a Test of Market Efficiency, *Journal of Financial and Quantitative Analysis* 17, 75-100

[4] Chan, K. C., G. Andrew Karolyi, Francis A. Longstaff, and Anthony B. Sanders, 1992, An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, Vol. XLVII, 1209-1227

[5] Chapman, D. and N. Pearson, 2001, What Can be Learned From Recent Advances in Estimating Models of the Term Structure ? *Financial Analysts Journal July/August,* 77-95

[6] Choi, Y. and T. S. Wirjanto, 2007, An Analytic Approximation Formula for Pricing Zero-Coupon Bonds, *Finance Research Letters*, Vol. 4, 116-126.

[7] Constantinides, G, 1992, A Theory of the Nominal Term Structure of Interest Rates, *Review of Financial Studies 5,* 531-552

[8] Dai, Q. and K. Singleton, 2000, Specification Analysis of Affine Term Structure Models, *Journal of Finance* LV , 1943-1978

[9] Dai, Q. and K. Singleton, 2003, Fixed Income Pricing *Handbook of Economics and Finance*, Chapter 20, C. Constantinides, M. Harris, and R. Stulz, eds. North Holland

[10] Duan, J., 1995, The GARCH Option Pricing Model, *Mathematical Finance* 5, 13-32

[11] Duffie D., and R. Kan, 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance* 6, 379-406

[12] Fabozzi J., *Fixed Income Mathematics*, Prentice Hall, New Jersey, 1993

[13] Grossman, S. J., A. Melino, and R. J. Shiller, 1987, Estimating the Continuous-time Consumption-based Asset Pricing Model, *Journal of Business and Economic Statistics* 5, 315-327

[14] Heston, S. and S. Nandi, 2000, A Closed-form GARCH Option Valuation Model, *Review of Financial Studies* 13, 585-625

[15] Longstaff, F. A., 1989, Temporal Aggregation and the Continuous-time Capital Asset Pricing Model, *Journal of Finance* 44, 871-887

[16] Longstaff, F. A., 1990, Time Varying Term Premia and Traditional Hypotheses about the Term Structure, *Journal of Finance* 45, 1307-1314

[17] Longstaff, F. A., and E. Schwartz, 1992, Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *Journal of Finance* 47, 1259-1282

[18] Longstaff, F., 1993, The Valuation of Option on Coupon Bonds, *Journal of Banking and Finance,* 27-42

[19] Nelson, D. B., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.

[20] Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177-188
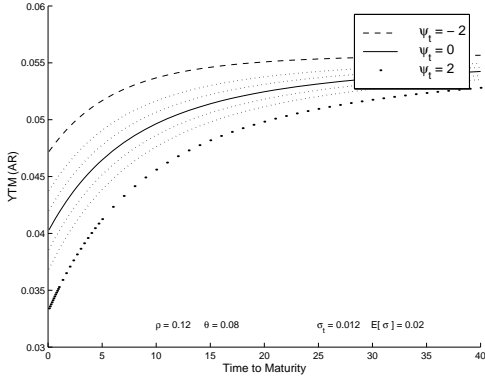
(a) $\rho = 0.12$, $\theta = 0.08$, $\sigma_t = 0.012$, $\beta_0/\theta = 0.02$
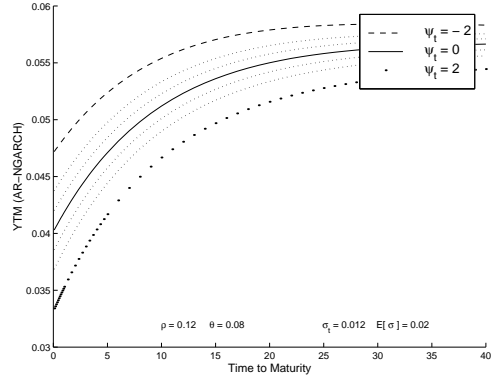
(b) $\rho = 0.12$, $\theta = 0.08$, $\sigma_t = 0.012$, $\beta_0/\theta = 0.02$
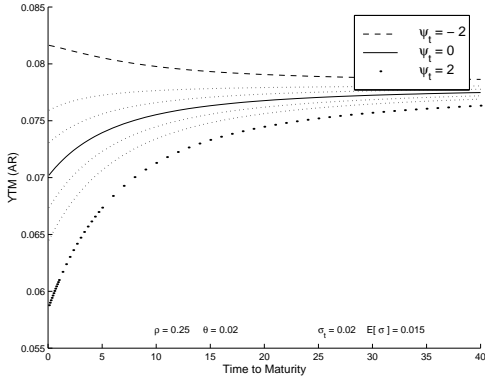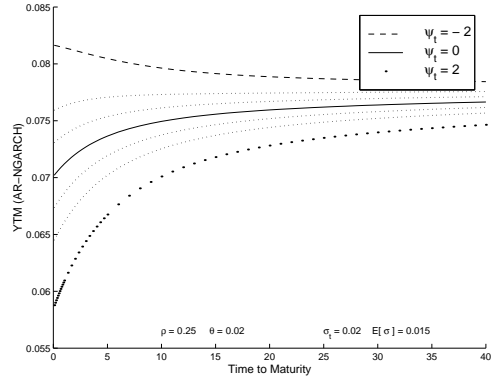
(c) $\rho = 0.25$, $\theta = 0.02$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.015$

(d) $\rho = 0.25$, $\theta = 0.02$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.015$
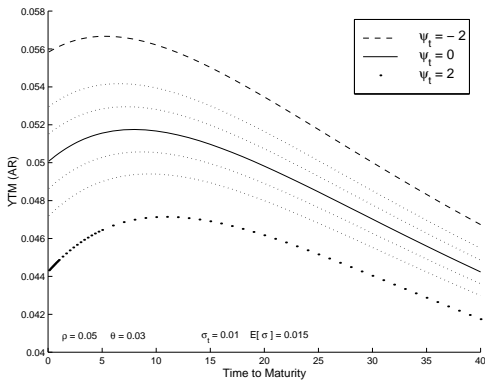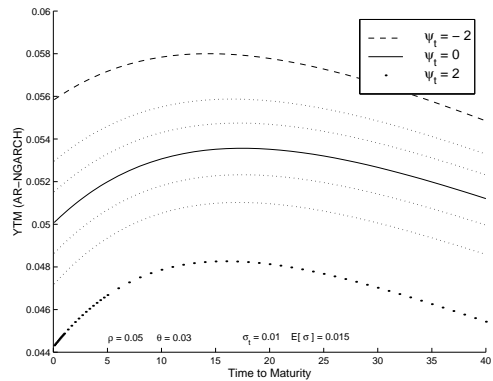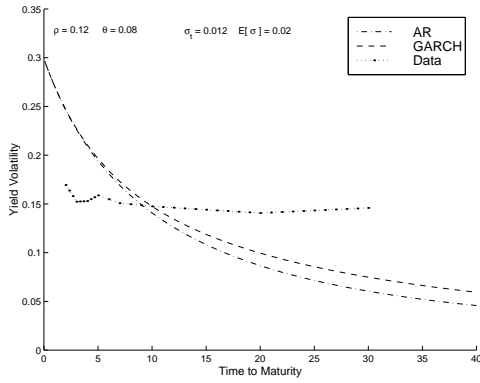
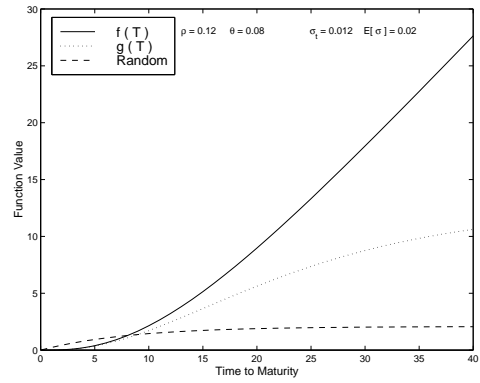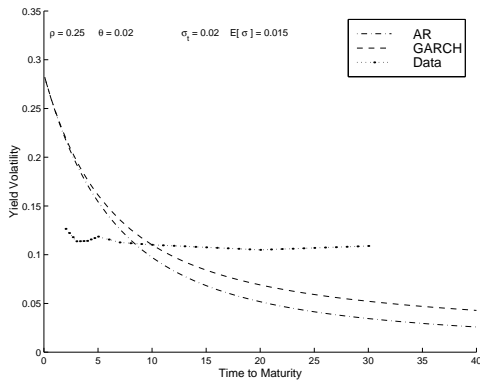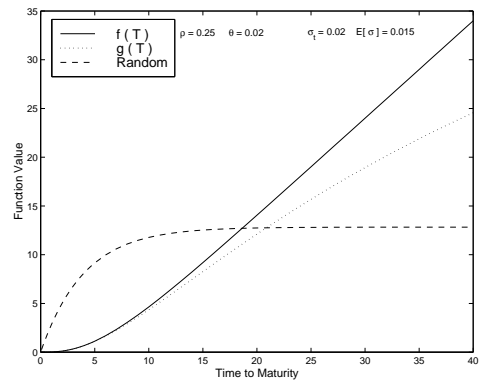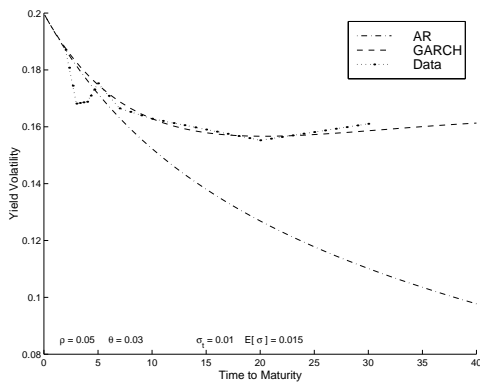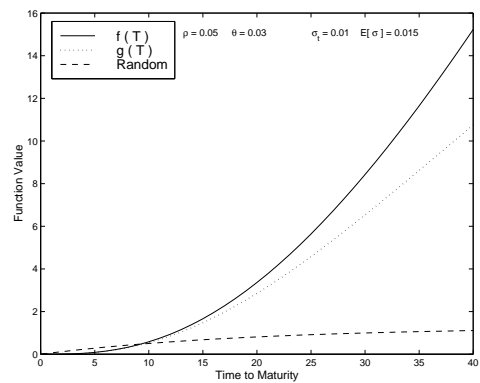(e) $\rho = 0.05$, $\theta = 0.03$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.015$

(f) $\rho = 0.05$, $\theta = 0.03$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.015$

FIGURE 1. Typical patterns of the yield curves in the case of the AR(1) process or AR (1)-NGARCH(1,1) process and the desired result (equation (8)) is obtained with the time to maturity.

(a) $\rho = 0.12$, $\theta = 0.08$, $\sigma_t = 0.012$, $\beta_0/\theta = 0.02$

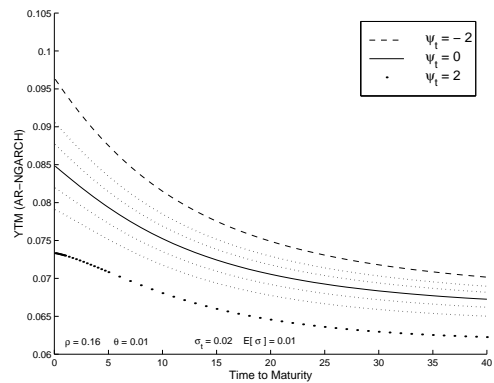(b) $\rho = 0.12$, $\theta = 0.08$, $\sigma_t = 0.012$, $\beta_0/\theta = 0.02$

(c) $\rho = 0.25$, $\theta = 0.02$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.015$

(d) $\rho = 0.25$, $\theta = 0.02$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.015$

(e) $\rho = 0.05$, $\theta = 0.03$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.015$

(f) $\rho = 0.05$, $\theta = 0.03$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.015$

FIGURE 2. Typical patterns of the volatility curves in the case of AR(1) process or AR(1)-NGARCH(1,1) process when function values appear in the yield curve and with the time to maturity.
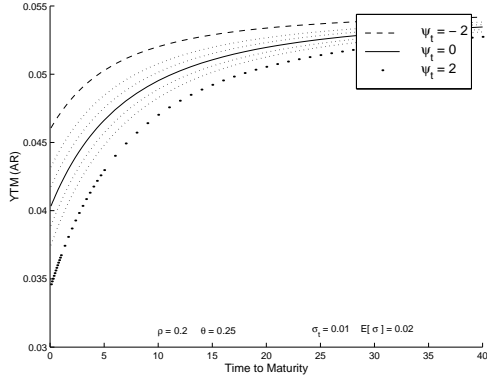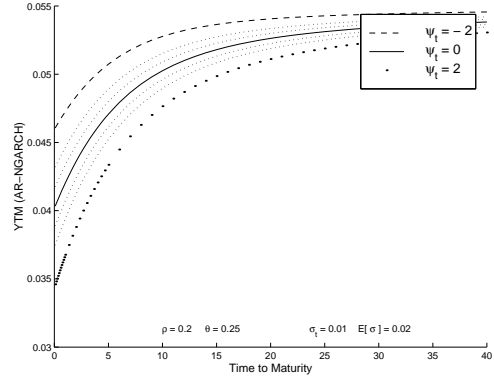
(a) $\rho = 0.20$, $\theta = 0.25$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.02$
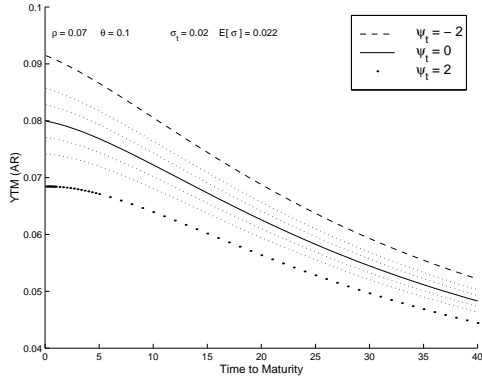
(b) $\rho = 0.20$, $\theta = 0.25$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.02$
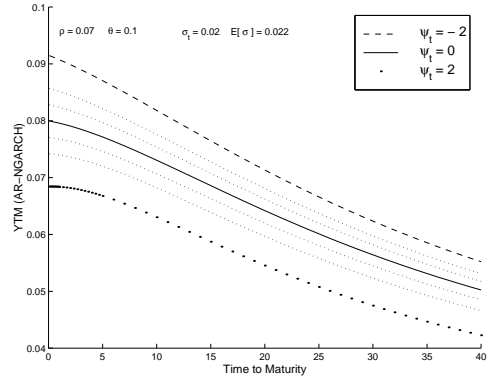
(c) $\rho = 0.07$, $\theta = 0.1$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.022$
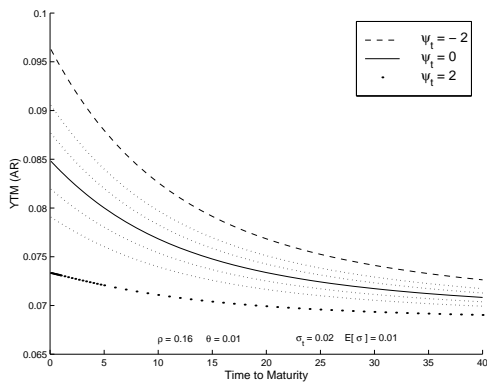
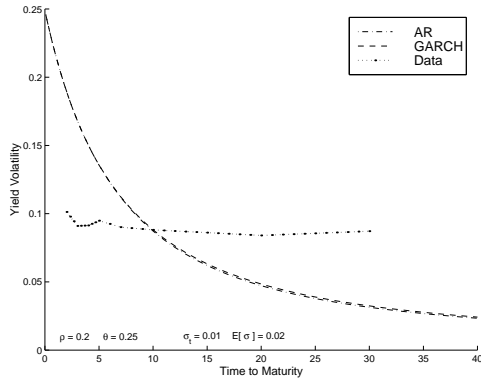(d) $\rho = 0.07$, $\theta = 0.1$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.022$

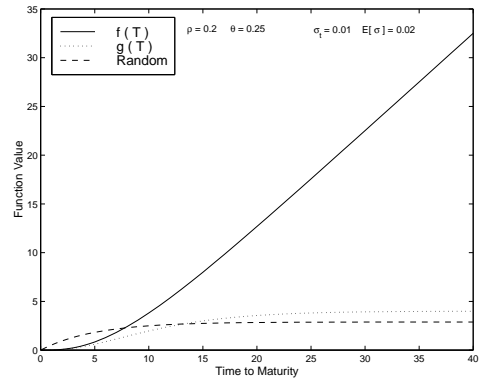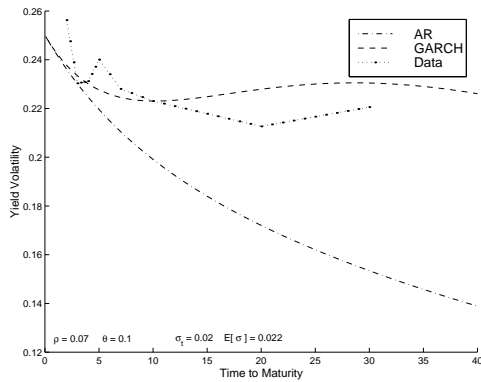(e) $\rho = 0.16$, $\theta = 0.01$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.01$

(f) $\rho = 0.16$, $\theta = 0.01$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.01$

FIGURE 3. Typical patterns for the yield curves in the case of AR(1) process or AR(1)-NGARCH(1,1) process and with the time to maturity.
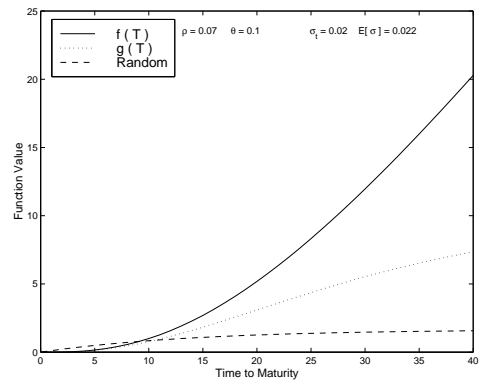
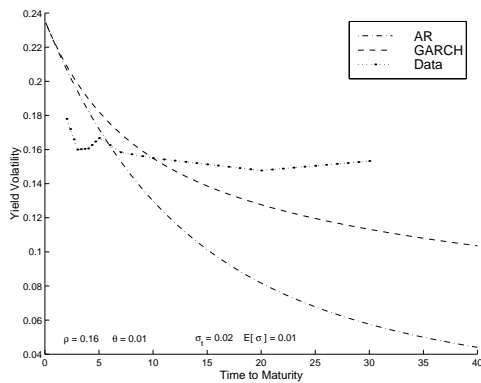(a) $\rho = 0.20$, $\theta = 0.25$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.02$

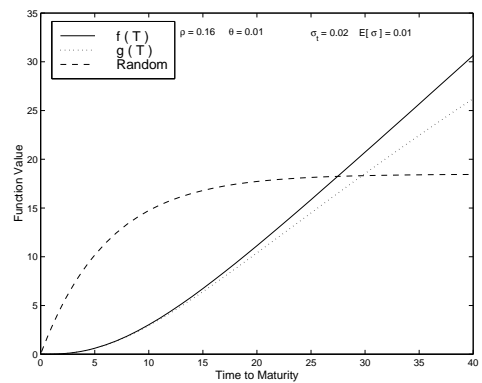(b) $\rho = 0.20$, $\theta = 0.25$, $\sigma_t = 0.01$, $\beta_0/\theta = 0.02$

(c) $\rho = 0.07$, $\theta = 0.1$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.022$

(d) $\rho = 0.07$, $\theta = 0.1$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.022$

(e) $\rho = 0.16$, $\theta = 0.01$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.01$

(f) $\rho = 0.16$, $\theta = 0.01$, $\sigma_t = 0.02$, $\beta_0/\theta = 0.01$

FIGURE 4. Typical patterns for the volatility curves in the case of AR(1) process or AR(1)-NGARCH(1,1) process where function values appear in the yield curve and with the time to maturity.