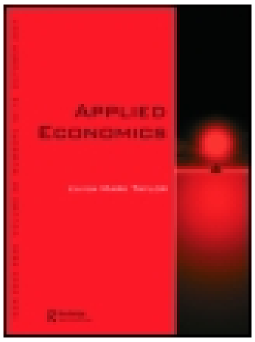

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: http://www.tandfonline.com/loi/raec20

A general model for short-term interest rates

Ching-Fan Chung & Mao-Wei Hung

Applied Economics, 2000, **32**, 111–121

A general model for short-term interest rates

CHING-FAN CHUNG* and MAO-WEI HUNG
The Institute of Economics, Academia Sinica, Taipei, Taiwan

A general one-factor model for short-term interest rates is proposed. Besides the long memory fractionally integrated mean process, the model also consists of a power function of the interest rate as well as the GARCH effect in the conditional variance. The estimation results show that, while there is no evidence for fractional integration in the mean beyond the well-known martingale property, both the power function of the interest rate and the GARCH effect (but not the ARCH effect) are highly significant in the formation of the conditional variance. Test results also confirm a structure change in October 1979 due to the shift in the Federal Reserve monetary policy.

I. INTRODUCTION

The behaviour of the short-term interest rate has been the focus of extensive studies in economics and finance. Many researchers have investigated the relationship between macroeconomic variables and the term structure of interest rates. For instance, Estrella and Hardouvelis (1991) examine the relationship between output and the term structure of interest rates. In finance, pricing interest rate options and designing hedging strategies all depend on the dynamic behaviour of the term structure of interest rate. A great number of models have been suggested in the literature to explain such dynamics. See, for example, Brennan and Schwartz (1977, 1979, 1980), Vasicek (1977), and Cox *et al.* (1981, 1985).

In a recent paper Chan *et al.* (1992) propose a continuous-time model for the short-term interest rate that nests most of the previous diffusion models. In their model both mean and conditional variance of the interest rate are functions of the level of interest rate. Their main empirical result is that incorporating a power function of interest rate into the conditional variance substantially improves the fit of the model.

More recently, along a different line of research Backus and Zin (1993) challenges the well-known belief that the short-term interest rate process is a martingale, suggested by, for example, Mankiw and Miron (1986) and Mishkin (1992). Backus and Zin demonstrate theoretically that long memory may exist in the short-term interest rates and find some empirical support for their argument. Such long memory in interest rates can have profound implications on asset pricing and thus deserves further investigation.

In this paper, we propose a generalized one-factor model for the short-term interest rate process. In the continuous-time version of the model we propose Fractional Brownian Motion, instead of the standard Brownian Motion, for the innovations to capture the potential long memory in the mean of the short-term interest rate. Besides this long memory mean process, our model also includes the GARCH effect and a power function of the interest rate in the conditional variance. We believe the GARCH effect is a natural extension of Chan *et al.*'s (1992) model of the conditional variance. With a long memory mean as well as the power function of the interest rate and the newly added GARCH effect in the conditional variance, our framework permits much greater flexibility than all previous models.

Our empirical results show that, in addition to the power function of the interest rate, the GARCH effect (but not ARCH effect) is significant in determining the conditional variance of the short-term interest rate. We also find that there is no empirical evidence for the long memory in the mean beyond the well-known martingale property. It is

* To whom all correspondence should be addressed; e-mail: chungcf@ccms.ntu.edu.tw

suggested that the long memory detected by Backus and Zin (1993) may result from an overly simplified constant conditional variance in their model.

Although it is widely believed that a structure break in the interest rate process occurs around October 1979 due to changes in Federal Reserve policy, Chan *et al.* (1992) were not able to find such a structure break in their empirical investigation. They attribute their results to the more flexible treatment of the conditional variance. However, being equipped with an even more general model, we obtain evidence contrary to Chan *et al.*'s finding of no structure break. Instead, our estimation confirms the earlier result that a structure break in the short-term interest rates did exist in October 1979. It seems that the model of Chan *et al.*, although more flexible than its predecessors, may still suffer from misspecification bias due to the omission of the GARCH effect. In terms of the testing of structural breaks, the incorporation of the additional GARCH effect appears to be critically important.

The remainder of the paper is organized as follows. The proposed one-factor model for short-term interest rate is described in Section II. The estimation method and the estimation results are presented in Section III and Section IV, respectively. The testing results for the structure break are given in Section V. Section VI discusses the estimation and testing results.

II. THE MODEL

Chan *et al.* (1992) combine several well-known continuous-time models for the riskless interest rate and propose the following specification:

$$\mathrm{d}r_t = (\alpha + \beta r_t)\,\mathrm{d}t + \sigma r_t^{\kappa/2}\,\mathrm{d}W_t \tag{1}$$

where r_t is the spot rate at time t and W_t is the standard Brownian motion. The key feature of their model is that the volatility of interest rate changes depends on a power function of the level of the interest rate. The discrete-time approximation to Equation 1 is

$$r_{t+1} - r_t = \alpha + \beta r_t + \sigma r_t^{\kappa/2}\varepsilon_{t+1} \tag{2}$$

where $E(\varepsilon_{t+1}) = 0$ and $E(\varepsilon_{t+1}^2) = 1$. Using the generalized method of moments (GMM) estimation, Chan *et al.* (1992) find that the term $r_t^{\kappa/2}$ plays a very important role in the formation of volatility. More specifically, they demonstrate that allowing κ to be a free parameter constitutes a significant improvement over those models with κ being fixed at a given value such as 0 (Vasicek, 1977), 1 (Cox *et al.*, 1985), or 2 (Brennan and Schwartz, 1977). In view of this interesting result it is natural to ask the following question: how is such a model of volatility related to the popular GARCH (generalized autoregressive conditionally heteroscedastic) specification? It is particularly interesting to know whether $r_t^{\kappa/2}$ is alternative or complementary to the GARCH effects. An obvious way to examine this issue is to consider a model which includes both $r_t^{\kappa/2}$ and the GARCH effects.

Once the GARCH effects are incorporated into the volatility process, the maximum likelihood estimation (MLE) seems to be more straightforward than the GMM procedure. So in this paper we employ the approximate MLE which is standard in the GARCH literature. We note one prerequisite for the MLE is the distributional assumption. Under an appropriate distributional assumption the MLE will be more efficient than its GMM counterpart. The price we pay for such efficiency gain is of course the potential biases in MLE when the assumed distribution is not supported by the data. However, recent literature (e.g. Bollersley and Wooldridge, 1992) indicates that the MLE under the normality assumption may still be robust to certain distributional misspecifications.

In this paper we also investigate the effects of the large kurtosis of the interest rate distribution through the *t*-distribution with its degree of freedom being treated as a free parameter. A small estimate of the degree of freedom implies a kurtosis that may be larger than the one from the normal distribution. Procedures based on *t*-distributions are common in the GARCH literature (Bollerslev, 1986, 1987).

In addition to the distributional assumption, an even more important condition for using MLE is that the data must be stationary. In this regard, we note it is widely believed that the short-term interest rate, especially before October 1979, was a martingale. See for example, Mankiw and Miron (1986) and Mishkin (1992). As a result, applying the OLS estimation (which is MLE under the normality assumption and the assumption that $\kappa = 0$) to Equation 2 where the dependent variable $r_{t+1} - r_t$ is martingale difference while the independent variable r_t is non-stationary, we will inevitably get a small estimate of the β coefficient.[1] These previous empirical results indicate that the usual stochastic differential Equation 1 may not be a suitable model for the short-term interest rate.[2]

[1] For example, using Mishkin's (1990, 1992) monthly data from February 1964 to October 1979 (see Section IV), we obtain the following OLS estiamtes (*t*-statistics in parentheses). On one-month interest rate, we have $\hat{\alpha} = 0.2428\ (1.8722)$, $\hat{\beta} = -0.0399\ (-1.7098)$, $\hat{\sigma}^2 = 0.502$, and $\mathrm{R}^2 = 0.016$. On two-month rate, we have $\hat{\alpha} = 0.2212\ (1.7717)$, $\hat{\beta} = -0.0347\ (-1.5900)$, $\hat{\sigma}^2 = 0.473$, and $\mathrm{R}^2 = 0.014$. On three-month rate, we have $\hat{\alpha} = 0.2227\ (1.7513)$, $\hat{\beta} = -0.0339\ (-1.5634)$, $\hat{\sigma}^2 = 0.473$, and $\mathrm{R}^2 = 0.013$. None of the β estimates is significant and R^2 are all small.

[2] Given the stochastic differential Equation 1, it is the difference $r_t - \int_0^t (\alpha + \beta r_s)\,\mathrm{d}s$, instead of the interest rate r_t itself, that is a martingale. See Øksendal (1992).

One way to modify model in Equation 1 is to replace the standard Brownian motion W_t by the increasingly more popular fractional Brownian motion so that more flexible non-integer orders of integration may be considered. The main difference between the standard Brownian motion and the fractional version is that the latter permits persistent autocorrelations (long memory). We thus propose the following two-equation continuous-time model for the short-term interest rate r_t:

$$dr_t = (\alpha + \beta r_t)dt + \sigma_t dW_t^d \tag{3}$$

$$d\sigma_t^2 = (\gamma r_t^\kappa - \tau\sigma_t^2)dt + \sqrt{2}\delta\sigma_t^2 dW_t \tag{4}$$

where W_t is a standard Brownian motion with a unit variance and W_t^d is a fractional Brownian motion that is independent of W_t and is defined by[3]

$$W_t^d(t) = \frac{1}{\Gamma(d+1)}\int_0^t (t-x)^d dW_t'(x)$$

Here, d is an arbitrary real number, W_t' is an ordinary Brownian motion which has a unit variance and is independent of W_t, and $\Gamma(\cdot)$ is the gamma function.

To specify the discrete-time counterparts of Equation 3, we note that Avram and Taqqu (1987, Theorem 2) have shown that if the stochastic process $\{Z_t\}$ follows a fractionally integrated process

$$(1-L)^d Z_t = \varepsilon_t \qquad t = 1,\ldots,T \tag{5}$$

for $d < 0.5$, where L is the usual lag operator and ε_t are independently and identically distributed (i.i.d.) with finite second moments, then

$$\frac{1}{\omega_T}\sum_{t=1}^{[Ts]} Z_t \xrightarrow{d} W_s^d \qquad \text{as } T \to \infty$$

where $\xrightarrow{d}$ denotes convergence in distribution, $\omega_T^2 \equiv \mathrm{Var}(\sum_{t=1}^T Z_t)$ and $[x]$ is the largest integer that is smaller than or equal to x. This is a functional central limit theorem for the fractionally integrated process (Equation 5).[4] It helps justify the following discrete-time approximation to Equation 3:

$$r_{t+1} - r_t = \alpha + \beta r_t + \sigma_{t+1}(1-L)^{-d}\varepsilon_{t+1} \tag{6}$$

where ε_t are assumed to be i.i.d. with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = 1$. Note that we also rely on the standard functional central limit theorem to justify the discrete-time model in Equation 2 as an approximation to the continuous-time model (Equation 1).

Simple algebra yields an alternative expression for Equation 6:

$$[1 - (\beta+1)L](1-L)^d(r_{t+1} - \alpha^*) = \sigma_{t+1}\varepsilon_{t+1} \tag{7}$$

where $\alpha^* = \alpha/\beta$. We note this model reduces to Equation 2 if $d = 0$. Since as mentioned earlier there is plenty of empirical evidence for the martingale property of the short-term interest rate, we believe it is the first-difference $\Delta r_{t+1} = (1-L)r_{t+1} = r_{t+1} - r_t$, instead of the level r_t, that is stationary. In terms of Δr_{t+1} the Equation 7 becomes

$$[1 - (\beta+1)L](1-L)^{d-1}\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1} \tag{8}$$

where the α^* term vanishes. It should be pointed out that the fractionally integrated model is particularly useful here since not only does it allow for a more flexible order of integration so that both orders d and $d-1$ are equally natural, it also nests the unit root (or the martingale) case with $d = 1$. Hence, with Equation 8 we are able to test the significance of unit root against the fractional alternatives.

The fractionally integrated process (Equation 5) was first introduced by Granger and Joyeux (1980), Granger (1980) and Hosking (1981). In contrast to the unit root process which exhibits perfect persistence with the impulse responses approaching 1, the fractionally integrated process with $d < 1$ is mean reverting and its impulse responses converge to 0 at a slow hyperbolic rate. We note that the impulse responses of the conventional ARMA model converge at fast exponential rates. While the ARMA model is said to have short memory, the fractionally integrated process has long memory. See Chung (1995) for a more complete analysis of the impulse responses of various time series models and Baillie (1996) for a survey of the applications of the long memory process in finance and economics. The specifications Equations 7 and 8 belong to the class of fractionally integrated autoregressive moving-average models, or the ARFIMA models, and can be denoted as ARFIMA$(1, d, 0)$ and ARFIMA$(1, d-1, 0)$, respectively. Finally, we note Backus and Zin (1993) also propose the ARFIMA models for the interest rates to explain certain properties of the term structure of interest rates. By fitting a number of ARFIMA models, Backkus and Zin suggest that fractional integration, although modest, exists in the interest rates.

Let's turn to the conditional variance process in Equation 4. To implement it empirically, we simply note Bollerslev *et al.* (1994) have shown its discrete-time approximation is

$$\sigma_{t+1}^2 = \gamma r_t^\kappa + \lambda\sigma_t^2 + \delta\varepsilon_t^2 \tag{9}$$

[3] The basic properties of the fractional Brownian motion W_t^d are as follows: (1) $E(W_t^d) = 0$. (2) Both W_t^d and $W_{\alpha t}^d/\alpha^{d+0.5}$ have the same finite-dimensional distributions for any $\alpha > 0$. (3) If $d \in (-0.5, 0.5)$, then W_t^d has stationary increments and $E|W_t^d - W_s^d|^2 = \sigma^2|t-s|^{2d+1}$. The standard Brownian motion is a special case of the fractional Brownian motion with $d = 0$. See Mandelbrot and Van Ness (1968) and Jonas (1983) for an overview of the fractional Brownian motion.

[4] A slightly more restricted version can also be found in Davydov (1970).

where $\lambda = 1 - \tau - \delta$. That is, the conditional variance σ_t^2 depends on a power function of r_t and follows a GARCH(1,1) process. Thus the empirical model we estimate consists of Equations 8 and 9. To guarantee that the model is stationary and the conditional variances are positive, we impose the following restrictions on the parameters: $d - 1 < 0.5$, $|\beta + 1| < 1$, $0 \leq \gamma$, $0 \leq \lambda < 1$, and $0 \leq \delta < 1$. Note that the model (2) of Chan *et al.* is a special case of Equations 7 and 9 with $d = 0$ and $\lambda = \delta = 0$, while Equation 7 is an alternative to Equation 8.

III. THE ESTIMATION METHOD

Given the proposed model Equations 8 and 9, the assumption that the innovations $\varepsilon_t, t = 1, \ldots, T$, are i.i.d. standard normal random variables, then the approximate maximum likelihood estimation maximizes the following function:

$$L(\alpha^*, \beta, d, \gamma, \kappa, \lambda, \delta) = -\frac{T-1}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T-1}\log\sigma_{t+1}^2 - \frac{1}{2}\sum_{t=1}^{T-1}\varepsilon_{t+1}^2 \qquad (10)$$

Apart from the Jacobian of the transformation from r_{t+1} to ε_{t+1} which is close to one and asymptotically negligible, Equation 10 is the same as the true log-likelihood function. If the initial observations $r_0, r_{-1}, r_{-2}, \ldots$ are assumed to be zero or, equivalently, $\varepsilon_0 = \varepsilon_{-1} = \varepsilon_{-2} = 0$, then maximizing Equation 10 is asymptotically equivalent to the MLE. As Bollerslev (1986) suggests, it is also straightforward to write down the approximate log-likelihood function when the innovations $\varepsilon, t = 1, \ldots, T$, are assumed to have the standardized t-distribution with the degree of freedom being treated as a free parameter. The asymptotic theory for the approximate MLE in the ARFIMA models is considered by Li and McLeod (1986) and Chung and Baillie (1993). The approximate MLE has been applied to inflation rate data by Baillie *et al.* (1996).

Given the model (Equation 8) for the first-differenced process Δr_{t+1}, we compute ε_{t+1} in the log-likelihood function as follows:

$$\varepsilon_{t+1}^2 = \frac{1}{\sigma_{t+1}^2}[1 - (\beta + 1)L](1 - L)^{d-1}\Delta r_{t+1} = \frac{1}{\sigma_{t+1}^2}[1 - (\beta + 1)L]\sum_{j=0}^{t-1}\pi_j L^j \Delta r_{t+1}$$

with $\pi_0 = 1$ and $\pi_j \equiv \Gamma(j - d + 1)/[\Gamma(1 - d)\Gamma(j + 1)]$, while σ_{t+1}^2 are computed recursively from Equation 9. Here we note the parameter α^* is not involved.[5]

IV. THE ESTIMATION RESULTS

We use Mishkin's (1990, 1992) monthly data on one-month, two-month, and three-month rates from February 1964 to December 1986, which are a well-known data set and largely overlap with those used by Chan *et al.* Hence, the findings reported in this paper can be readily compared with those earlier results.

We first fit the level data to model Equations 7 and 9 and find that, unsurprisingly, the fractional differencing parameter d is uniformly greater than 0.5 which renders the model non-stationary. We thus switch to the model Equations 8 and 9 for the first-differenced data. The estimation is conducted under both normal and t-distributions. In the latter case there is an additional parameter, the degree of freedom, that characterizes the kurtosis. The estimation results are presented in Table 1.

The estimates of both $d - 1$ and $\beta + 1$ are quite small and statistically insignificant. These findings strongly suggest that the short-term interest rates are martingales and thus verify what Mankiw and Miron (1986) and Mishkin (1992) have suggested. As a result, the best predictors of all the future short-term interest rates, conditional on all available information, are the current level: $E_t(r_{t+j}) = r_t$, for all $j = 1, 2, \ldots$ In contrast to such relatively simple dynamics in the level of the interest rate, the formation of volatility appears to be much more complex. While the estimates of the ARCH parameter δ are small, almost all the estimates of the GARCH parameter λ and the power parameter κ are highly significant. These results seem to imply that the GARCH effect, as opposed to the ARCH effect, is complementary to the term r_t^κ. that is, the GARCH effect and r_t^κ may represent different aspects of volatility formation, while the ARCH effect may be offset by the r_t^κ term.[6]

[5] If the estimation were based on the level data as in Equation 7, then the innovations ε_{t+1} would be

$$\varepsilon_{t+1}^2 = \frac{1}{\sigma_{t+1}^2}[1 - (\beta + 1)L](1 - L)^d(r_{t+1} - \alpha^*) = \frac{1}{\sigma_{t+1}^2}[1 - (\beta + 1)L]\sum_{j=0}^{t-1}\tilde{\pi}_j L^j(r_{t+1} - \alpha^*)$$

with $\tilde{\pi}_0 = 1$ and $\tilde{\pi}_j \equiv \Gamma(j - d)/[\Gamma(-d)\Gamma(j + 1)]$.

[6] Like the majority of the previous studies of GARCH models, our analysis concentrates on the GARCH(1, 1) model. We do try to fit other GARCH specifications of higher orders and then use the standard model selection criteria such as AIC and SIC to pick the best one. It is found that none of those GARCH models of higher orders survives this screening process.

Table 1. *Maximum likelihood estimates*

	$d-1$	$\beta+1$	γ	κ	λ	δ	d.f.
One-month interest rate							
Normal distribution	−0.0816	0.0612	0.7290	2.9775	0.3719	0.0787	
	(−0.9059)	(0.5193)	(3.6299)	(16.1586)	(2.2552)	(1.2734)	
t-distribution	−0.0912	0.0875	0.5759	3.1278	0.5449	0.0635	5.5013
	(−1.0352)	(0.7573)	(2.3993)	(10.1021)	(2.9982)	(0.7661)	(2.2467)
Two-month interest rate							
Normal distribution	−0.1134	0.2074	0.7696	2.7897	0.0000	0.1121	
	(−2.2093)	(1.5955)	(3.7134)	(15.0875)	(0.0000)	(1.9846)	
t-distribution	0.0114	0.0989	0.2931	0.9242	0.5651	0.1746	4.0439
	(0.1338)	(0.8750)	(2.3012)	(6.8739)	(4.2070)	(1.6276)	(2.5870)
Three-month interest rate							
Normal distribution	0.0267	0.0602	0.2473	3.2628	0.5392	0.2060	
	(0.3219)	(0.4415)	(4.1310)	(10.9247)	(6.1178)	(3.5113)	
t-distribution	0.0383	0.1179	0.2615	3.1514	0.6120	0.1834	3.4458
	(0.4386)	(1.0606)	(2.1384)	(6.7683)	(5.0113)	(1.6665)	(2.8764)

Notes: This table presents the approximate MLE of the model Equations 8 and 9:

$$[1-(\beta+1)L](1-L)^{d-1}\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$$

$$\sigma^2_{t+1} = \gamma r_t^{\kappa} + \lambda\sigma_t^2 + \delta\varepsilon_t^2$$

under the normal and t-distributional assumptions. The estimation is based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^{κ} term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^{κ}.

As to the degree of freedom of the t-distribution, we find its estimates range from 3.44 to 5.50 so that the corresponding kurtosis is larger than that of the normal distribution. One of the most striking results in Table 1 is that under the t-distribution assumption we obtain very similar estimation results for all three types of rates. We note the two most important parameters κ and λ have similar estimates around 3 and 0.6, respectively. For the one-month and three-month rates, the parameter estimates are also quite similar under the normality assumption. However, in the case of the two-month rate the different distributional assumptions appear to have a stronger effect on the estimates, especially those of the ARCH and GARCH parameters. Because obviously all three interest rate distributions have larger kurtosis, the estimation results based on the t-distribution seem to be more reliable.[7]

Following the estimation results in Table 1, especially those from the t-distribution, we consider a reduced specification where the values of the three insignificant parameters are restricted as follows: $d-1=\beta+1=\delta=0$. The restricted model consists of two equations:

$$\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1} \tag{11}$$

$$\sigma^2_{t+1} = \gamma r_t^{\kappa} + \lambda\sigma_t^2 \tag{12}$$

This is a preferred martingale model in which volatility contains the GARCH effect and the power term r_t^{κ}. The estimation results are presented in Table 2. There we find a highly significant effect of a power function of the interest rate and a moderate GARCH effect.

The main conclusion we draw from Tables 1 and 2 is that the level of the short-term interest rate follows a simple martingale process while volatility is determined by a more sophisticated dynamic mechanism.

We now show how misspecification in volatility might affect the estimation of the $d-1$ and $\beta+1$ parameters. In Table 3 we report the estimation results for a model where the GARCH effect is suppressed, while in Table 4 we further constrain the value of κ to be 1 so that we effectively have an Ornstein–Uhlenbeck model. These increasingly simplified specifications for volatility seem to move the estimates of $d-1$ and $\beta+1$ away from zero. In particular we find the estimate of $d-1$ for the one-month rate becomes a significant negative value. That is, if the volati-

[7] Even though theoretically the estimation under the normality assumption is more robust to the distributional assumptions, the parameter estimates for the two-month rate under the normality assumption appear quite different from the rest of Table 1. These aberrant results may indicate problems with the normality assumption.

Table 2. *Maximum likelihood estiamtes*

	γ	κ	λ	d.f.
One-month interest rate				
Normal distribution	0.7522 (4.2964)	3.1141 (19.9312)	0.4724 (3.7374)	
t-distribution	0.5979 (2.7092)	3.2324 (12.0498)	0.6112 (4.6698)	5.7232 (2.3281)
Two-month interest rate				
Normal distribution	0.7101 (4.7010)	2.8268 (17.2830)	0.2229 (1.3515)	
t-distribution	0.7125 (2.2687)	2.8185 (10.4621)	0.2826 (1.0383)	4.3351 (2.4575)
Three-month interest rate				
Normal distribution	0.5746 (4.4239)	2.8168 (18.7928)	0.3106 (1.9860)	
t-distribution	0.5466 (1.9486)	2.9432 (9.6628)	0.4972 (2.3945)	3.3082 (3.2029)

Notes: This table presents the approximate MLE of the model Equations 11 and 12:

$$\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$$

$$\sigma_{t+1}^2 = \gamma r_t^\kappa + \lambda\sigma_t^2$$

under the normal and t-distributional assumptions. The estimation is based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^κ term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^κ.

Table 3. *Maximum likelihood estimates*

	$d-1$	$\beta+1$	γ	κ	d.f.
One-month interest rate					
Normal distribution	−0.1500 (−1.7786)	0.1569 (1.6416)	1.4044 (8.8818)	3.0046 (18.4650)	
t-distribution	−0.1681 (−2.0320)	0.1718 (1.8057)	1.4688 (4.9357)	3.0508 (12.2612)	6.5717 (1.9495)
Two-month interest rate					
Normal distribution	−0.0885 (−1.1682)	0.1532 (1.9889)	0.9023 (10.2997)	2.7830 (17.7075)	
t-distribution	−0.0854 (−1.0972)	0.1944 (2.0817)	0.9652 (3.9445)	2.7718 (10.3324)	4.3448 (2.6582)
Three-month interest rate					
Normal distribution	−0.0824 (−1.1492)	0.1510 (2.1316)	0.8243 (10.7931)	2.7566 (19.3166)	
t-distribution	−0.0719 (−0.9167)	0.2192 (2.4173)	0.9753 (3.1869)	2.8541 (10.2758)	3.5367 (3.4541)

This table presents the approximate MLE of the following models:

$$[1-(\beta+1)L](1-L)^{d-1}\Delta r_{t+1} = \sqrt{\gamma} r_t^{\kappa/2}\varepsilon_{t+1}$$

under the normal and t-distributional assumptions. The estimation is based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^κ term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^κ.

Table 4. *Maximum likelihood estimates*

	$d-1$	$\beta+1$	γ	d.f.
One-month interest rate				
Normal distribution	−0.1826	0.1729	0.7398	
	(−0.5961)	(2.0796)	(20.0671)	
t-distribution	−0.1551	0.1499	0.9996	2.7846
	(−2.2185)	(1.8816)	(1.9496)	(4.1882)
Two-month interest rate				
Normal distribution	−0.1411	0.2186	0.5411	
	(−2.0028)	(3.0314)	(19.3219)	
t-distribution	−0.0492	0.1601	0.7936	2.6475
	(−0.6756)	(1.8573)	(1.6349)	(4.3363)
Three-month interest rate				
Normal distribution	−0.1377	0.2286	0.5111	
	(−1.8376)	(3.2065)	(19.0018)	
t-distribution	−0.0459	0.2029	0.7507	2.6066
	(−0.6006)	(2.3415)	(1.6443)	(4.6905)

Notes: This table presents the approximate MLE of the following model:

$$[1-(\beta+1)L](1-L)^{d-1}\Delta r_{t+1} = \sqrt{\gamma}\varepsilon_{t+1}$$

under the normal and t-distributional assumptions. The estimation is based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^{κ} term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^{κ}.

It should be noted that the t-distributions are not consistent with the well-known theory by Cox *et al.* (1985): the short-term interest rates which follow the present model should have non-central χ^2 distributions. While the normal distribution results can be considered large-sample approximations to the case of the non-central χ^2 distributions, the testing results for the t-distribution cases reported here are simply for the reference purpose only.

lity of this short-term interest rate is not properly specified, then its level can turn into a mean-reverting fractionally integrated process which is quite different from a martingale. Other than this long memory result, the significant $\beta+1$ estimates also indicate a form of short-run dynamics in the interest rate level which contradicts the martingale property. So from Tables 3 and 4 we conclude that the first and the second conditional moments of the short-term interest rates cannot be separately estimated. If the volatility process is misspecified, then the estimation of the first moments can also be spurious. We note our estimation results from Tables 1–4 not only refute Backus and Zin's (1993) theoretical proposition of the fractional integration in the interest rate but also explain why empirical support for their theory could have been 'found' when an overly simplified constant conditional variance is assumed.

The model specified in Table 3 is similar to the model (2) of Chan *et al.*, but with a more flexible order of integration. We note that, by restricting the order of integration and applying the GMM estimation to the one-month rate, Chan *et al.* obtained a significant estimate 2.9998 for κ and an insignificant but large estimate 0.4079 for $\beta+1$. These results are more or less compatible with our estimates on the one-month rate in Table 3.[8] Here, we should note the methodological difference between the GMM estimation and the MLE. The GMM estimation procedure does not require distributional assumption so it is robust to the distributional assumptions.[9] Also, whether data are stationary or not does not seem to matter in the GMM estimation so that Chan *et al.* do not consider the unit root problem in the interest rate data. In contrast, the MLE procedure requires us to make a great deal of effort to take each of these issues explicitly into consideration. Because of this, we believe the resulting MLE should achieve greater efficiency than the GMM method does, provided that the distributional assumptions, the t-

[8] We have tried to estimate the model of Chan *et al.* using the approximate MLE. That is, we estimate the model Equation 7 with $d = 0$ and no GARCH effects. The parameter estimates (t-statistics) for κ and $\beta+1$ are 1.0072 (66.4082) and 2.9987 (18.5858), respectively, under the normality assumption. The estimation results under the t-distribution are almost identical. Here, we note that the estimate of $\beta+1$ is larger than but insignificantly different from 1. So the estimated model essentially contains a unit root, which is not inconsistent with the martingale result we reported earlier.

[9] Chan *et al.* claim the GMM estimation is also robust to the GARCH effects without explanation. We will provide some evidence about how the GARCH effect might bias the GMM estimation and the ensuing hypothesis testing.

distribution in particular, are not too far away from the truth.

V. STRUCTURAL CHANGES IN OCTOBER 1979

It is widely believed that the short-term interest rate experiences a structural change in October 1979 due to the shift in Federal Reserve monetary regimes. See, for example, Huizinga and Mishkin (1984) and Clarida and Friedman (1984). However, Chan *et al.* (1992) report that they could not find evidence for such a structural change based on their GMM estimation of model (2). They interpret their results as follows: ' . . . [their] interest rate models may be rich enough to capture the change in interest rate behaviour evident in the post-1979 period. These results also raise the possibility that previous tests for structural breaks may be misspecified because of their failure to model the conditional heteroscedasticity in interest rate changes correctly' (p. 1222). We also examine this issue by estimating model Equations 11 and 12 separately with data before and after October 1979. The estimation results are presented in Tables 5 and 6. It is surprising to see that the estimates of both κ and λ for each of the two sub-periods are much larger than those from the whole period in Table 2. We also note the t-statistics for the κ estimates are smaller, obviously due to smaller sample sizes in each sub-period.

In Table 7 we present the log-likelihood values. In the column under the title 'without structural break' are the ones corresponding to the estimates for the entire sample that were reported in Table 2. Those in the next column under the title 'with structural break' are the sums of two log-likelihood values obtained from the two sub-period estimations. The entries in the last column are two times the differences between the previous two columns and are the likelihood ratio test statistics for testing the hypothesis of no structural break. The degree of freedom of the χ^2 test is 3 for the normal distribution cases. The critical values at the 95% level and the 99% level are 7.82 and 9.35, respectively. The degree of freedom is 4 for the t-distribution cases and the critical values at the 95% level and the 99% level are 9.49 and 11.14, respectively. From Table 7 we conclude that the hypothesis of no structural break is strongly rejected for the two-month and three-month rates, while it is marginally rejected for the one-month rate. Conseqently, our estimation results appear to be in conflict with the conclusion of Chan *et al.* on one-month rate but consistent with many earlier results obtained by other researchers. Since our model Equations 11 and 12 differ from the model of Chan *et al.* mainly in the additional GARCH effect, we may argue that their finding of no structural break is due to their insufficient consideration of conditional heteroscedasticity, which, ironically, was precisely the same criticism they made on previous works in the literature.

Table 5. *The MLE for the first subsample from February 1964 to October 1979*

	γ	κ	λ	d.f.
One-month interest rate				
Normal distribution	0.3586	3.3630	0.7845	
	(3.6063)	(9.0404)	(11.5879)	
t-distribution	0.4267	3.5470	0.7876	4.9627
	(1.9547)	(5.8980)	(9.2487)	(1.9221)
Two-month interest rate				
Normal distribution	0.2724	3.7624	0.8334	
	(4.3126)	(8.2375)	(18.8842)	
t-distribution	1.6443	3.0277	0.9203	2.0450
	(0.3634)	(3.4780)	(20.2536)	(16.2446)
Three-month interest rate				
Normal distribution	0.2542	4.3650	0.8697	
	(4.7966)	(7.4079)	(32.8638)	
t-distribution	6.3153	3.9679	0.9355	2.0136
	(0.3720)	(4.9215)	(29.0881)	(55.2955)

Notes: This table presents the approximate MLE of the model Equations 11 and 12:

$$\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$$

$$\sigma^2_{t+1} = \gamma r_t^{\kappa} + \lambda\sigma_t^2$$

under the normal and t-distributional assumptions. The estimation is based on 187 monthly data on three types of short-term interest rates from February 1964 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^{κ} term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^{κ}.

Table 6. *The MLE for the second subsample from November 1979 to December 1986*

	γ	κ	λ	d.f.
One-month interest rate				
Normal distribution	0.5108 (2.7863)	4.4869 (8.4381)	0.6091 (4.9833)	
t-distribution	0.4618 (2.1782)	4.4393 (6.9563)	0.6411 (4.1517)	12.1273 (0.4841)
Two-month interest rate				
Normal distribution	0.3997 (2.1521)	4.9023 (9.6723)	0.4075 (1.6779)	
t-distribution	0.3487 (1.9832)	4.9771 (7.9035)	0.4730 (2.1024)	14.3891 (0.4201)
Three-month interest rate				
Normal distribution	0.2075 (2.0683)	4.9738 (9.4855)	0.6176 (3.8909)	
t-distribution	0.1506 (1.6297)	5.2170 (6.6641)	0.7043 (4.5802)	8.3116 (0.7436)

Notes: This table presents the approximate MLE of the model Equations 11 and 12:

$$\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$$

$$\sigma_{t+1}^2 = \gamma r_t^\kappa + \lambda\sigma_t^2$$

under the normal and t-distributional assumptions. The estimation is based on 88 monthly data on three types of short-term interest rates from November 1979 to December 1986. Numbers in parentheses are t-statistics. The value of r_t in the γr_t^κ term is divided by 10 to stabilize the iterative estimation. As a result, the γ estimates are magnified by the factor of 10^κ.

Table 7. *Log-likelihood values*

	Without structural break	With structural break	Likelihood-ratio test statistics
One-month interest rate			
Normal distribution	− 231.5513	− 225.2441	12.6144*
t-distribution	− 226.7645	− 221.2635	11.0020†
Two-month interest rate			
Normal distribution	− 202.2414	− 189.4505	25.5818*
t-distribution	− 195.2152	− 183.3483	23.7338*
Three-month interest rate			
Normal distribution	− 200.8301	− 186.7173	28.2256*
t-distribution	− 188.8084	− 173.8307	29.9554*

Notes: This table present the log-likelihood values under the normal and t-distributional assumptions. The second column gives the log-likelihood values based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. (The corresponding parameter estimates are in Table 2.) The entries in the third column are the sums of two log-likelihood values of the two subsample estimations: one from February 1964 to October 1979 and the other from November 1979 to December 1986. (The corresponding parameter estimates are in Tables 5 and 6, respectively.) So the log-likelihood values in the third column are from estimations that allow for structural break. The models under consideration are Equations 11 and 12:

$$\Delta r_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$$

$$\sigma_{t+1}^2 = \gamma r_t^\kappa + \lambda\sigma_t^2$$

The fourth column gives the likelihood-ratio test statistics. The degree of freedom of the χ^2 test is 3 for the normal distribution cases. The critical values at the 95% level and the 99% level are 7.82 and 9.35, respectively. The degree of freedom is 4 for the t-distribution cases and the critical values at the 95% level and the 99% level are 9.49 and 11.14, respectively. The test statistics with * are significant at 99% level. The one with † is insignificant at 99% level but significant at 95% level.

Table 8. *Log-likelihood values*

	Without structural break	With structural break	Likelihood-ratio test statistics
One-month interest rate			
Normal distribution	−233.6993	−231.0451	5.3084†
t-distribution	−230.6871	−228.2474	4.8794†
Two-month interest rate			
Normal distribution	−202.6113	−192.6287	19.9652*
t-distribution	−194.1166	−185.9173	16.3986*
Three-month interest rate			
Normal distribution	−201.7283	−192.8603	17.7360*
t-distribution	−186.6077	−180.8673	11.4808*

Notes: This table presents the log-likelihood values under the normal and t-distributional assumptions. The second column gives the log-likelihood values based on 275 monthly data on three types of short-term interest rates from February 1964 to December 1986. The entries in the third column are the sums of two log-likelihood values of the two subsample estimations: one from February 1964 to October 1979 and the other from November 1979 to December 1986. So the log-likelihood values in the third column are from estimations that allow for structural break. The model under consideration is Equation 13:

$$[1-(\beta+1)L]\Delta r_{t+1}=\sqrt{\gamma}r_t^{\kappa/2}\varepsilon_{t+1}$$

which is essentially the same as Cehn *et al.*'s model based on the first-differenced data. The fourth column gives the likelihood-ratio test statistics. The degree of freedom of the χ^2 test is 3 for the normal distribution cases. The critical values at the 95% level and the 99% level are 7.82 and 9.35, respectively. The degree of freedom is 4 for the t-distribution cases and the critical values at the 95% level and the 99% level are 9.49 and 11.14, respectively. The test statistics with * are significant at 99% level. The one with † is insignificant at 99% level but significant at 95% level.

To reaffirm that the different testing results Chan *et al.* get is caused by the omission of the statistically significant GARCH term in their model of volatility, we conduct further testing based on the model of Chan *et al.* for the first-differenced data:

$$[1-(\beta+1)L]\Delta r_{t+1}=\sqrt{\gamma}r_t^{\kappa/2}\varepsilon_{t+1} \tag{13}$$

The testing procedure is the same as that given in Table 7 and the results are shown in Table 8. Unsurprisingly, for the one-month rate the hypothesis of no structural break cannot be rejected at 95% level, just like what Chan *et al.* have reported. This result clearly suggests the inference of no structural break does have something to do with the omission of the GARCH term in their specification of volatility, even though they have correctly included the r_t^{κ} term. Incidentally, from Table 8 we also note, contrary to the case of the one-month rate, that the structural changes are again present in the two-month and three-month rates, just like the earlier result in Table 7 based on model Equations 11 and 12.

VI. DISCUSSION

The main conclusion drawn from the estimates of our preferred model Equations 11 and 12 in Tables 2, 5 and 6 is as follows. Over the entire sampled period from February 1964 to December 1986, the power function of the interest rate is highly significant in the determination of the conditional variances of all three types of short-term interest rates. A significant GARCH effect is also present, but it is less apparent in the two-month rate than the others. In the absence of the GARCH effect as in the model of Chan *et al.*, the unambiguously significant estimate of κ such as 2.9998 indicates striking sensitivity of the interest rate volatility with respect to the level changes in the interest rate. However, the presence of the positive GARCH effect in our model lessens the influence of the κ estimate on the effect of the level changes in the interest rate. The implication of an additional GARCH effect is that a change in the interest rate volatility in the next period results not only from the change in the current level of the interest rate but also from the persistence in volatility itself. Consequently, the effects of the level changes in the interest rate on its volatility is not fixed at the value of the κ estimate, but may in fact fluctuate around some values below the κ estimates ranging from 2.8 to 3.2. Obviously, it is this extra flexibility that makes the detection of the structural breaks possible. From Tables 5 and 6, we notice for all three types of interest rates the κ estimates from the first subsample and the λ estimates from the second subsample are consistently lower than their respective counterparts from the other subsamples. It appears that after 1979 the interest rate volatility becomes more sensitive to the level of the interest rate and less persistent.

Also note that the estimates in Tables 5 and 6 differ from those in Table 2 in a systematic way: the subsample estimates of κ are higher while the subsample estimates of λ are lower in comparison with those of the whole sample. In other words, the estimates of the κ and λ parameters from the whole sample are not averages of those from the two subsamples. The reason for this somewhat counter-intuitive result may lie in the possible effect cancellation between the κ and λ estimates when the whole sample period is considered. It should also be pointed out that such effect cancellation may result in less fluctuating estimates of the effect of the level changes in the interest rate. That is, estimations based on the longer sampled period produce smoother effects of the level of interest rate on volatility than those in the subsamples.

As a concluding remark, we believe our findings in this paper are substantive contributions to the literature of the interest rate modelling and our results can be useful in the valuation of interest rate contingent claim and optimal hedging strategies. We note, in particular, valuation of itnerest rate contingent claim is very sensitive to the volatility of the interest rates, while flexible modelling of the volatility is the centrepiece of our analysis. An interesting future research subject concerns how the valuation of interest rate contingent claim can take into account both the power function of interest rate and the GARCH effect.

ACKNOWLEDGEMENTS

We are very grateful to a referee for his valuable suggestions.

REFERENCES

Avram, F. and Taqqu, M. S. (1987) Noncentral limit theorems and Appell polynomials, *Annals of Probability*, **15**, 767–75.

Backus, D. K. and Zin, S. E. (1993) Long-memory inflation uncertainty: evidence from the term structure of interest rates, *Journal of Money, Credit, and Banking*, **25**, 681–700.

Baillie, R. T. (1996) Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, **73**, 6–59.

Baillie, R. T., Chung, C.-F. and Tieslau, M. (1996) Analyzing inflation by the fractionally integrated ARFIMA-GARCH model, *Journal of Applied Econometrics*, **11**, 23–40.

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

Bollerslev, T. (1987) A conditionally heteroskedastic time series model for speculative prices and rates of return, *The Review of Economics and Statistics*, **69**, 542–7.

Bollerslev, T., Engle, R. F. and Nelson D. B. (1994) ARCH models, in *Handbook of Econometrics*, Vol. 4, Chapter 49, (Eds) R. F. Engle and D. L. McFadden, North-Holland, Amsterdam.

Bollerslev, T. and Wooldridge, J. M. (1992) Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews*, **11**, 143–72.

Brennan, M. and Schwartz, E. (1977) Savings bonds, retractable bonds, and callable bonds, *Journal of Financial Economics*, **3**, 231–62.

Brennan, M. and Schwartz, E. (1979) A continuous time approach to the pricing of bonds, *Journal of Banking and Finance*, **3**, 133–55.

Brennan, M. and Schwartz, E. (1980) Analyzing convertible bonds, *Journal of Financial and Quantitative Analysis*, **15**, 907–29.

Chan, K., Karolyi, A., Longstaff, F. and Sanders, A. (1992) An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance*, **47**, 1209–27.

Chung, C.-F. (1995) Calculating and analyzing impulse responses and their asymptotic distributions for the ARFIMA and VARMA models, Econometrics and Economic Theory Paper No. 9402, Michigan State University.

Chung, C.-F. and Baillie, R. T. (1993) Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models, *Empirical Economics*, **18**, 791–806.

Clarida, R. H. and Friedman, B. M. (1984) The behavior of U.S. short term interest rates since October 1979, *Journal of Finance*, **39**, 671–84.

Cox, J. C., Ingersoll, J. E., Jr. and Ross, S. A. (1981) A re-examination of traditional hypotheses about the term structure of interest rates, *Journal of Finance*, **36**, 769–99.

Cox, J. C., Ingersoll, J. E., Jr. and Ross, S. A. (1985) A theory of the term structure of interest rates, *Econometrica*, **53**, 385–407.

Davydov, Y. A. (1970) The invariance principle for stationary processes, *Theory of Probability and Its Applications*, **15**, 487–9.

Granger, C. W. J. (1980) Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics*, **14**, 227–38.

Granger, C. W. J. and Joyeux, R. (1980) An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis*, **1**, 15–39.

Hosking, J. R. M. (1981) Fractional differencing, *Biometrika*, **68**, 165–76.

Huizinga, J. and Mishkin, F. (1984) Inflation and real interest rates on assets with different risk characteristics, *Journal of Finance*, **39**, 699–712.

Jonas, A. B. (1983) Persistent memory random processes, PhD Dissertation, Harvard University.

Li, W. K. and McLeod, A. I. (1986) Fractional time series modeling, *Biometrika*, **73**, 217–21.

Mandelbrot, B. and Van Ness, J. W. (1968) Fractional Brownian motions, fractional noises and applications, *SIAM Review*, **10**, 422–37.

Mankiw, N. G. and Miron, J. A. (1986) The changing behavior of the term structure of interest rates, *Quarterly Journal of Economics*, **101**, 211–28.

Mishkin, F. S. (1990) What does the term structure tell us about future inflation? *Journal of Monetary Economics*, **25**, 77–95.

Mishkin, F. S. (1992) Is the Fisher effect for real? A reexamination of the relationship between inflation and interest rates, *Journal of Monetary Economics*, **30**, 195–215.

Øksendal, B. (1992) *Stochastic Differential Equations*, third edition (New York: Springer-Verlag).

Vasicek, O. (1977) An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5**, 177–88.