

# On the threshold hyperbolic GARCH models\*

WILSON KWAN, WAI KEUNG LI AND GUODONG LI†

In the financial market, the volatility of financial assets plays a key role in the problem of measuring market risk in many investment decisions. Insights into economic forces that may contribute to or amplify volatility are thus important. The financial market is characterized by regime switching between phases of low volatility and phases of high volatility. Nonlinearity and long memory are two salient features of volatility. To jointly capture the features of long memory and nonlinearity, a new threshold time series model with hyperbolic generalized autoregressive conditional heteroscedasticity is considered in this article. A goodness of fit test is derived to check the adequacy of the fitted model. Simulation and empirical results provide further support to the proposed model.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 91B84; secondary 62M10.

KEYWORDS AND PHRASES: hyperbolic GARCH model, long memory, threshold model, volatility.

## 1. INTRODUCTION

The volatility of financial returns has been shown to exhibit long memory, and its correlations remain positive for long lags and decay slowly to zero, see Greene and Fielitz (1977), Ding, Granger and Engle (1993), Kokoszka and Taqqu (1996), Cont (2001) and among others. Although price changes appear to be unpredictable, the magnitude of those changes, measured either by the absolute values or the squares of the return series, appears to be predictable in the sense that large changes are likely to be followed by large changes – of either sign – whereas small changes are likely to be followed by small changes. Price fluctuations are then characterized by periods of low volatility irregularly interspersed by periods of high volatility. This phenomenon was first investigated by Mandelbrot (1963) for commodity prices. The modeling of volatility has been presented in a wide variety of financial assets such as stocks, market indices and exchange rates, and the autoregressive conditional heteroscedasticity (ARCH) model (Engle, 1982) or the generalized ARCH (GARCH) model (Bollerslev, 1986) is one of the most popular tools.

\*Dedicated to Professor Howell Tong in celebration of his 65th birthday.

†Corresponding author.

The GARCH model can be rewritten as the following ARCH( $\infty$ ) form,

$$(1) \quad e_t = \varepsilon_t h_t^{1/2}, \quad h_t = \gamma + \left(1 - \frac{\delta^*(L)}{\beta(L)}\right) e_t^2,$$

where  $L$  is the back shift operator, and  $\delta^*(\cdot)$  and  $\beta(\cdot)$  are polynomials, see Davidson (2004). The standard GARCH model was introduced originally to describe the dependence structure of volatility, and failed to capture the long memory feature of volatility in the returns. Note that  $\delta^*(L) = (1 - L)\delta(L)$  for the case of the integrated GARCH (Engle and Bollerslev, 1986), i.e.  $x = 1$  is the root of  $\delta^*(x)$ . Following the idea of the fractional integrated autoregressive moving average (ARFIMA) model, Baillie, Bollerslev and Mikkeslen (1996) extended the common GARCH model to the fractional integrated GARCH (FIGARCH) model,

$$e_t = \varepsilon_t h_t^{1/2}, \quad h_t = \gamma + \left(1 - \frac{\delta(L)}{\beta(L)}(1 - L)^{d_{FG}}\right) e_t^2,$$

where  $0 < d_{FG} < 1$ , and the item  $(1 - L)$  in (1) is replaced by  $(1 - L)^{d_{FG}}$ . However, the FIGARCH process has always infinite variance, and there also exists a paradox for the definition of the long memory for ARCH( $\infty$ ) processes. To overcome these problems, Davidson (2004) derived a new definition of the memory for ARCH( $\infty$ ) processes based on the concept of the near-epoch dependence, and proposed the hyperbolic GARCH (HYGARCH) model,

$$e_t = \varepsilon_t h_t^{1/2}, \quad h_t = \gamma + \left(1 - \frac{\delta(L)}{\beta(L)}[1 + \alpha((1 - L)^{d_{FG}} - 1)]\right) e_t^2.$$

The main motivation of the above model is that it nests the FIGARCH and integrated GARCH models. For more discussions about the HYGARCH models, we can refer to Kwan, Li and Li (2010a).

In financial practice, the main feature for measuring risk in many investment decisions is the volatility of financial assets. Theoretical and observational insights into economic forces that may contribute to or amplify volatility are thus important. Lo and MacKinlay (1990) suggested a possible relation with technical trading rules and stock market over-reaction. Shefrin (2000) and Hirshleifer (2001) emphasized the role of market psychology and investor sentiment in financial markets. In the trading process, it is observed that price changes are driven by a combination of news about fundamentals and evolutionary forces. Under different trading

strategies, expectations about future prices and dividends of a risky asset, a nonlinear phenomenon structure may occur in the price volatility caused by the interaction between traders, fundamentalists and technical analysts. The financial market is thus characterized by an irregular regime switching between phases of low volatility and phases of high volatility, and some recent works can be referred to Schwert (1989), Rabemanjara and Zakoian (1993) and Li and Li (1996). This motivates us to propose a threshold HYGARCH model in section 2, and it allows for a different HYGARCH structure for each regime.

In the process of modeling volatility with ARCH-type models, more and more evidences have shown that many financial time series may be so heavy-tailed that the distribution of the innovation  $\varepsilon_t$  is far from normality, see Mikosch and Starica (2000) and Hall and Yao (2003). The commonly used Gaussian quasi-maximum likelihood estimation (MLE) may be lack of efficiency, and the Student's  $t$  MLE seems preferable in analyzing finance and economic time series although it is harder to be extended to the quasi-MLE in theory, see Bollerslev (1987). Davidson (2004) also employed the Student's  $t$  MLE to fit the ARFIMA-HYGARCH models. Section 3 derives the asymptotic properties of its student's  $t$  MLE, and a diagnostic tool based on the squared residual autocorrelations is considered to check whether the fitted threshold HYGARCH model is adequate in section 4. Two simulation experiments are performed in section 5, and section 6 analyzes the daily exchange rates of Korean Won against US dollar.

## 2. THE THRESHOLD ARFIMA-HYGARCH MODELS

Let  $\{y_t\}$  be a stationary and ergodic time series generated by the threshold ARFIMA-HYGARCH model with  $P$  regimes,

$$(2) \quad (1 - L)^{d_{ARF}} \phi^{(i)}(L) y_t = \psi^{(i)}(L) e_t, \quad e_t = \varepsilon_t h_t^{1/2}$$

and

$$(3) \quad h_t = \gamma^{(i)} + \{1 - [1 - \alpha^{(i)} + \alpha^{(i)}(1 - L)^{d_{FG}}] \frac{\delta^{(i)}(L)}{\beta^{(i)}(L)}\} e_t^2,$$

as  $r_{i-1} \leq z_{t-d} < r_i$ , where  $d$  is the delay parameter,  $-\infty = r_0 < \dots < r_P = +\infty$ ,  $L$  is the back-shift operator,  $\phi^{(i)}(x) = 1 - \sum_{k=1}^{p^{(i)}} \phi_k^{(i)} x^k$ ,  $\psi^{(i)}(x) = 1 - \sum_{k=1}^{q^{(i)}} \psi_k^{(i)} x^k$ ,  $\delta^{(i)}(x) = 1 - \sum_{k=1}^{m^{(i)}} \delta_k^{(i)} x^k$ ,  $\beta^{(i)}(x) = 1 - \sum_{k=1}^{s^{(i)}} \beta_k^{(i)} x^k$ ,  $p^{(i)}$ ,  $q^{(i)}$ ,  $m^{(i)}$  and  $s^{(i)}$  with  $i = 1, \dots, P$  are known positive integers, the innovation sequence  $\{\varepsilon_t\}$  is identically and independently distributed (*i.i.d.*) with mean zero and variance one, and

$$(1 - L)^d = 1 - \sum_{j=1}^{\infty} \frac{d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} L^j \quad \text{as } 0 < d < 1.$$

For the  $i$ th regime, denoting  $\theta_V^{(i)} = (\gamma^{(i)}, \beta_1^{(i)}, \dots, \beta_{s^{(i)}}^{(i)}, \delta_1^{(i)}, \dots, \delta_{m^{(i)}}^{(i)}, d_{FG}^{(i)}, \alpha^{(i)})'$ , we can rewrite model (3) into the following ARCH( $\infty$ ) form,

$$(4) \quad h_t = \gamma^{(i)} + \pi^{(i)}(L) e_t^2 = \gamma^{(i)} + \sum_{j=1}^{\infty} \pi_j^{(i)} e_{t-j}^2$$

as  $r_{i-1} \leq z_{t-d} < r_i$ , where the  $\pi_j^{(i)}$ 's are functions of  $\theta_V^{(i)}$ . The threshold variable  $z_t$  may be the observed time series  $y_t$  or the innovation sequence  $e_t$ , or even some exogenous variable, see Tong (1990) and Liu, Li and Li (1997).

**Assumption 1.** For all  $1 \leq i \leq P$ ,  $\alpha^{(i)} \geq 0$ ,  $0 < d_{FG}^{(i)} \leq 1$  and  $\pi_j^{(i)} \geq 0$  with  $j = 1, 2, \dots$ ; the polynomials  $\delta^{(i)}(x)$  and  $\beta^{(i)}(x)$  have no common root.

Davidson (2004) pointed out that one of the main motivations for using (3) is that it nests the FIGARCH and integrated GARCH models. It is observed that when  $d_{FG}^{(i)} = 0$ , the conditional variance model becomes an ordinary GARCH model. Thus, the focus of this article will be on the range  $0 < d_{FG}^{(i)} \leq 1$ . As argued by Davidson (2004), there are two kinds of memory to be recognized: hyperbolic decaying memory and geometric decaying memory, with the former being defined as long memory. For model (3), when  $0 < d_{FG}^{(i)} < 1$ ,  $\pi_j^{(i)} = O(j^{-1-d})$ , i.e. the coefficients decay hyperbolically, and the conditional variance  $h_t$  in (4) or (3) will exhibit the long-memory effect. Conrad (2010) discussed the non-negativity conditions for the conditional variance of a HYGARCH process, and these constraints are required for each regime to make sure that the conditional variance of the threshold models (2) and (3) is always non-negative.

**Assumption 2.** For all  $1 \leq i \leq P$ ,  $d_{ARF}^{(i)} > -0.5$ ;  $\sum_{k=1}^{q^{(i)}} |\psi_k^{(i)}| < 1$ ; and the polynomials  $\phi^{(i)}(x)$  and  $\psi^{(i)}(x)$  have no common root.

The parameter  $d_{ARF}^{(i)}$  is usually used to describe the extent of long memory in the ARFIMA processes. The ARFIMA process is short memory or long memory respectively when  $d_{ARF}^{(i)} \in (-0.5, 0)$  or  $(0, \infty)$ , and is stationary or nonstationary respectively when  $d_{ARF}^{(i)} \in (-0.5, 0.5)$  or  $(0.5, \infty)$ , see Ling and Li (1997). However, it may be difficult to discuss the memory properties and the stationarity of the models (2) and (3) since the effects are mixed for the threshold models. We leave it for possible future research. Note that the definition of long memory here is different from that in the conditional variance in (3) or (4). The condition  $\max_i \sum_{k=1}^{q^{(i)}} |\psi_k^{(i)}| < 1$  is sufficient to make sure that the threshold model given by (2) and (3) is invertible, see Ling and Tong (2005) and Ling, Tong and Li (2007).

In the following sections, we will concentrate our discussion on two-regime threshold models, i.e.  $P = 2$  in models (2) and (3), and the notation  $r_1$  is replaced by  $r$  for simplicity.

### 3. ESTIMATION OF THRESHOLD ARFIMA-HYGARCH MODELS

As in Bollerslev (1987) and Davidson (2004), this section will consider a Student's  $t$  MLE for the threshold ARFIMA-HYGARCH models proposed in the previous section.

Let  $\theta_M^{(i)} = (\phi_1^{(i)}, \dots, \phi_{p^{(i)}}^{(i)}, \psi_1^{(i)}, \dots, \psi_{q^{(i)}}^{(i)}, d_{ARF}^{(i)})'$  with  $i = 1$  and  $2$ ,  $\theta_M = (\theta_M^{(1)'}, \theta_M^{(2)'})'$ ,  $\theta_V = (\theta_V^{(1)'}, \theta_V^{(2)'})'$  and  $\theta = (\theta_M', \theta_V')'$ . Denote by  $\Theta$  a suitable compact set in  $R^l$  with  $l = \sum_{i=1}^2 (p^{(i)} + q^{(i)} + r^{(i)} + s^{(i)}) + 8$ . Suppose that the delay parameter  $d$  and the threshold value  $r$  are known, and  $y_1, \dots, y_n$  are generated by the threshold models (2) and (3) with the true parameter vector  $\theta_0$ , an interior point of  $\Theta$ . Define functions

$$e_t(\theta) = \left[ \sum_{k=1}^{q^{(1)}} \psi_k^{(1)} e_{t-k}(\theta) + (1-L)^{d_{ARF}^{(1)}} \phi^{(1)}(L) y_t \right] I(z_{t-d} < r) \\ + \left[ \sum_{k=1}^{q^{(2)}} \psi_k^{(2)} e_{t-k}(\theta) + (1-L)^{d_{ARF}^{(2)}} \phi^{(2)}(L) y_t \right] \\ \cdot I(z_{t-d} \geq r)$$

and

$$h_t(\theta) = [\gamma^{(1)} + \sum_{j=1}^{\infty} \pi_j^{(1)} e_{t-j}^2(\theta)] I(z_{t-d} < r) \\ + [\gamma^{(2)} + \sum_{j=1}^{\infty} \pi_j^{(2)} e_{t-j}^2(\theta)] I(z_{t-d} \geq r).$$

Note that the above two functions are all dependent on past observations infinitely far away, however, there are only  $n$  values available in real applications. Hence, some initial values are needed, and we may simply assume  $y_s = e_s(\theta) = 0$  for  $s \leq 0$ . The effect of the initial values can be shown to be negligible asymptotically, see Weiss (1986) and Ling and Li (1997). Without loss of generality, we assume that the values  $y_s$  for  $s \leq 0$  are all observable in deriving the asymptotic results in this and the next sections.

Suppose the innovation  $\varepsilon_t$  follows  $t_{\nu_0}$ , where  $t_{\nu}$  is the normalized Student's  $t$  distribution with mean zero, variance one, and  $\nu$  degrees of freedom, and its density is given by

$$f_{\nu}(x) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi(\nu-2)}\Gamma(\nu/2)} \left(1 + \frac{x^2}{(\nu-2)}\right)^{-(\nu+1)/2},$$

where  $\nu > 2$ . Then the log likelihood function of  $y_1, \dots, y_n$ ,

conditional on  $y_0, y_{-1}, \dots$ , is

$$L_n(\theta, \nu) \\ = \sum_{t=1}^n \left[ \log f_{\nu} \left( \frac{e_t(\theta)}{\sqrt{h_t(\theta)}} \right) - \frac{1}{2} \log h_t(\theta) \right] \\ = n \log \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi(\nu-2)}\Gamma(\nu/2)} \\ - \frac{1}{2} \sum_{t=1}^n \left[ \log h_t(\theta) + (\nu+1) \log \left( 1 + \frac{e_t^2(\theta)}{(\nu-2)h_t(\theta)} \right) \right],$$

and the MLE can be defined as

$$\hat{\tau}_n = (\hat{\theta}_n, \hat{\nu}_n)' = \underset{\theta \in \Theta, \nu \in V}{\operatorname{argmax}} L_n(\theta, \nu),$$

where  $V \subset (2, \infty)$  is a compact set, and the true degrees of freedom  $\nu_0$  is an interior point of  $V$ .

Define

$$c_1(x) = \frac{\partial}{\partial \nu} \log f_{\nu_0}(x), \quad c_2(x) = \frac{\partial}{\partial x} \log f_{\nu_0}(x)$$

and  $c_3(x) = -0.5[1 + xc_2(x)]$ . It holds that  $E[c_1(\varepsilon_t)] = 0$ ,  $E[c_2(\varepsilon_t)] = 0$ ,  $E[c_3(\varepsilon_t)] = 0$ ,  $E[c_1(\varepsilon_t)c_2(\varepsilon_t)] = 0$ ,  $E[c_2(\varepsilon_t)c_3(\varepsilon_t)] = 0$ ,

$$E[c_2^2(\varepsilon_t)] = \frac{\nu_0(\nu_0+1)}{(\nu_0-2)(\nu_0+3)} \quad \text{and} \quad E[c_3^2(\varepsilon_t)] = \frac{\nu_0}{2(\nu_0+3)},$$

see Kwan, Li and Li (2010a). For quantities  $E[c_1^2(\varepsilon_t)]$  and  $E[c_1(\varepsilon_t)c_3(\varepsilon_t)]$ , it is difficult to get their explicit forms, and we may calculate them using sample averages.

Consider the derivative functions of the likelihood function  $L_n(\theta, \nu)$ ,

$$\frac{\partial L_n(\theta_0, \nu_0)}{\partial \nu} = \sum_{t=1}^n c_1(\varepsilon_t)$$

and

$$\frac{\partial L_n(\theta_0, \nu_0)}{\partial \theta} = \sum_{t=1}^n \left[ c_2(\varepsilon_t) \frac{1}{\sqrt{h_t}} \frac{\partial e_t}{\partial \theta} + c_3(\varepsilon_t) \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \right].$$

Then, the Fisher information matrix is denoted by

$$I_{\tau} = \begin{pmatrix} I_{\theta} & J \\ J' & E[c_1^2(\varepsilon_t)] \end{pmatrix},$$

where

$$I_{\theta} = E[c_2^2(\varepsilon_t)] E \left[ \frac{1}{h_t} \frac{\partial e_t}{\partial \theta} \frac{\partial e_t}{\partial \theta'} \right] + E[c_3^2(\varepsilon_t)] E \left[ \frac{1}{h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \right],$$

and

$$J = E[c_1(\varepsilon_t)c_3(\varepsilon_t)] E \left[ \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \right].$$

Suppose that the process  $\{y_t\}$  generated by models (2) and (3) is strictly stationary and ergodic with finite fourth moment, and Assumptions 1 and 2 are satisfied. Then it holds that  $I_\tau < \infty$ . By a method similar to Ling and Li (1997) and Straumann (2005), we can show that the information matrix  $I_\tau$  is also positive definite. By the Taylor expansion and the routine arguments for MLEs, we have the following asymptotic results.

**Theorem 3.1.** *Suppose  $\{y_t\}$  is strictly stationary and ergodic with finite fourth moment. If Assumptions 1 and 2 hold, then*

$$\sqrt{n}(\hat{\tau}_n - \tau_0) \longrightarrow N(0, I_\tau^{-1})$$

in distribution as  $n \rightarrow \infty$ , where  $\tau_0 = (\theta'_0, \nu_0)'$ .

When the delay parameter  $d$  and the threshold value  $r$  are unknown, we may estimate them by

$$(\hat{d}, \hat{r}) = \underset{d \in D, r \in [a, b]}{\operatorname{argmax}} L_n(\hat{\theta}_n, \hat{\nu}_n),$$

where  $D$  is a set including all possible values of  $d$ , and  $[a, b]$  is a predetermined interval, say 0.1 and 0.9-quantiles of the observed sequence, see Tong (1990).

In the Fisher information matrix  $I_\tau$ , the quantities  $E[c_1^2(\varepsilon_t)]$ ,  $E[c_2^2(\varepsilon_t)]$ ,  $E[c_3^2(\varepsilon_t)]$  and  $E[c_1(\varepsilon_t)c_3(\varepsilon_t)]$  are all functions of  $\nu_0$ , and can be estimated in practice by replacing  $\nu_0$  by  $\hat{\nu}_n$ . Let  $\hat{\varepsilon}_t = e_t(\hat{\theta}_n)/h_t^{1/2}(\hat{\theta}_n)$ . We may alternatively estimate these four quantities by their respective sample averages, e.g. using  $n^{-1} \sum_{t=1}^n c_1^2(\hat{\varepsilon}_t)$  and  $n^{-1} \sum_{t=1}^n [c_1(\hat{\varepsilon}_t)c_3(\hat{\varepsilon}_t)]$  respectively to estimate  $E[c_1^2(\varepsilon_t)]$  and  $E[c_1(\varepsilon_t)c_3(\varepsilon_t)]$ , where  $\nu_0$  is replaced by  $\hat{\nu}_n$ . Furthermore,

$$\frac{1}{n} \sum_{t=1}^n \frac{1}{h_t(\hat{\theta}_n)} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta} \frac{\partial e_t(\hat{\theta}_n)}{\partial \theta'} = E \left[ \frac{1}{h_t} \frac{\partial e_t}{\partial \theta} \frac{\partial e_t}{\partial \theta'} \right] + o_p(1),$$

$$\frac{1}{n} \sum_{t=1}^n \frac{1}{h_t^2(\hat{\theta}_n)} \frac{\partial h_t(\hat{\theta}_n)}{\partial \theta} \frac{\partial h_t(\hat{\theta}_n)}{\partial \theta'} = E \left[ \frac{1}{h_t^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \right] + o_p(1),$$

and

$$\frac{1}{n} \sum_{t=1}^n \frac{1}{h_t(\hat{\theta}_n)} \frac{\partial h_t(\hat{\theta}_n)}{\partial \theta} = E \left[ \frac{1}{h_t} \frac{\partial h_t}{\partial \theta} \right] + o_p(1).$$

Correspondingly, we can consistently estimate  $I_\tau$  by  $\hat{I}_\tau$ .

## 4. DIAGNOSTIC CHECKING

Residual autocorrelations from traditional autoregressive moving average models have been found useful in model diagnostic checking, see Box and Jenkins (1976). Li and Mak (1994) derived the asymptotic distribution of the square residual autocorrelations, and constructed a diagnostic tool for the adequacy of the conditional heteroscedastic models, see also Li (2004). Following their ideas, we derive

a portmanteau test to check whether the fitted threshold ARFIMA-HYGARCH model is adequate.

Without confusion, we use the notations  $e_t(\tau)$  and  $h_t(\tau)$  in this section, and denote  $e_t(\hat{\tau}_n)$  and  $h_t(\hat{\tau}_n)$  respectively by  $\hat{e}_t$  and  $\hat{h}_t$  for simplicity, where  $\tau = (\theta', \nu)'$  and  $\hat{\tau}_n$  is the Student's  $t$  MLE in section 3. It is observed that  $e_t(\tau)$  and  $h_t(\tau)$  do not depend on  $\nu$  and further  $e_t(\tau)$  does not depend on  $\theta_V$ . Hence,  $\partial e_t(\tau)/\partial \nu = 0$ ,  $\partial e_t(\tau)/\partial \theta_V = 0$  and  $\partial h_t(\tau)/\partial \nu = 0$ . Note that  $\{\hat{e}_t/\hat{h}_t^{1/2}\}$  is the residual sequence for models (2) and (3),  $n^{-1} \sum_{t=1}^n (\hat{e}_t/\hat{h}_t^{1/2}) = o_p(1)$  and  $n^{-1} \sum_{t=1}^n (\hat{e}_t^2/\hat{h}_t) = 1 + o_p(1)$ . Then, for a positive integer  $k$ , the lag- $k$  squared residual autocorrelation is

$$\hat{r}_k = \frac{\sum_{t=k+1}^n (\hat{e}_t^2/\hat{h}_t - 1)(\hat{e}_{t-k}^2/\hat{h}_{t-k} - 1)}{\sum_{t=1}^n (\hat{e}_t^2/\hat{h}_t - 1)^2}.$$

Let  $\hat{R} = (\hat{r}_1, \dots, \hat{r}_K)'$ , and we next consider the asymptotic distributions of  $\hat{R}$ .

Let  $\hat{C} = (\hat{C}_1, \dots, \hat{C}_K)'$  and  $C = (C_1, \dots, C_K)'$ , where

$$\hat{C}_k = \frac{1}{n} \sum_{t=k+1}^n \left( \frac{\hat{e}_t^2}{\hat{h}_t} - 1 \right) \left( \frac{\hat{e}_{t-k}^2}{\hat{h}_{t-k}} - 1 \right),$$

and  $C_k = n^{-1} \sum_{t=k+1}^n (\varepsilon_t^2 - 1)(\varepsilon_{t-k}^2 - 1)$  is the corresponding values when  $\hat{\tau}_n$  in  $\hat{C}_k$  is replaced by the true parameter vector  $\tau_0$ . Note that  $n^{-1} \sum_{t=1}^n (\hat{e}_t^2/\hat{h}_t - 1)^2 = \kappa_1 + o_p(1)$ , where  $\kappa_1 = E(\varepsilon_t^4) - 1 = 6(\nu_0 - 4)^{-1} + 2$ . Then it is sufficient to derive the asymptotic distribution of  $\hat{C}$ . Taking the Taylor expansion, we can obtain that

$$(5) \quad \hat{C} = C + X'(\hat{\tau}_n - \tau_0) + o_p(n^{-1/2}),$$

where  $X = (X_1, \dots, X_K)$ , and  $X_k = -E[h_t^{-1}(e_{t-k}^2/h_{t-k} - 1)(\partial h_t/\partial \tau)]$ . It holds that, by Theorem 3.1,

$$(6) \quad \begin{aligned} & \sqrt{n}(\hat{\tau}_n - \tau_0) \\ &= I_\tau^{-1} \cdot \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[ c_1(\varepsilon_t) + \frac{c_2(\varepsilon_t)}{\sqrt{h_t}} \frac{\partial e_t}{\partial \tau} + \frac{c_3(\varepsilon_t)}{h_t} \frac{\partial h_t}{\partial \tau} \right] \\ & \quad + o_p(1), \end{aligned}$$

where  $c_i(\varepsilon_t)$  with  $i = 1, 2$  and  $3$  are defined as in section 3. Note that  $E[(\varepsilon_t^2 - 1)(\varepsilon_{t-k}^2 - 1)c_1(\varepsilon_t)] = 0$  for  $k > 0$ ,  $E[(\varepsilon_t^2 - 1)c_2(\varepsilon_t)] = 0$  and  $E[(\varepsilon_t^2 - 1)c_3(\varepsilon_t)] = 1$ . By (5), (6), the central limit theorem and the Cramér-Wold device, under the conditions of Theorem 3.1, we can show that

$$\sqrt{n}\hat{R} \rightarrow N(0, \Sigma)$$

in distribution as  $n \rightarrow \infty$ , where

$$\Sigma = I - \frac{1}{\kappa_1^2} X' I_\tau^{-1} X.$$

Table 1. Estimation result from 500 simulated series of model

1							
regime	$\gamma$	$\beta$	$\alpha$	$d_{FG}$	$\delta$	$\nu$	
$n = 1000$							
Bias	1	0.0239	0.0877	0.1083	0.0862	0.0232	0.3464
$\sqrt{\text{MSE}}$		0.0439	0.1975	0.1498	0.2076	0.1623	1.8189
Bias	2	0.0215	0.0757	0.1095	0.1464	-0.0668	
$\sqrt{\text{MSE}}$		0.0465	0.2319	0.1548	0.2538	0.2176	
$n = 2000$							
Bias	1	0.0188	0.0771	0.1048	0.0704	0.0178	0.1401
$\sqrt{\text{MSE}}$		0.0338	0.1697	0.1135	0.1638	0.1290	1.5796
Bias	2	0.0156	0.0748	0.1033	0.1374	-0.0595	
$\sqrt{\text{MSE}}$		0.0404	0.2023	0.1225	0.2164	0.1940	
$n = 4000$							
Bias	1	0.0163	0.0697	0.1046	0.0694	0.0172	0.1376
$\sqrt{\text{MSE}}$		0.0272	0.1726	0.0912	0.1560	0.1223	1.3321
Bias	2	0.0056	0.0725	0.1017	0.1356	-0.0349	
$\sqrt{\text{MSE}}$		0.0319	0.1792	0.1037	0.1926	0.1428	

Table 2. Estimation result from 500 simulated series of model

2							
regime	$\gamma$	$\beta$	$\alpha$	$d_{FG}$	$\delta$	$\nu$	
$n = 1000$							
Bias	1	0.0242	0.0725	0.1171	0.0794	0.0138	0.3705
$\sqrt{\text{MSE}}$		0.0455	0.1656	0.3588	0.2519	0.1835	1.7752
Bias	2	0.0220	0.0473	0.1113	0.0763	-0.0068	
$\sqrt{\text{MSE}}$		0.0458	0.2244	0.3922	0.2726	0.2419	
$n = 2000$							
Bias	1	0.0208	0.0579	0.1123	0.0735	0.0103	0.2099
$\sqrt{\text{MSE}}$		0.0369	0.1590	0.3110	0.2144	0.1576	1.5142
Bias	2	0.0186	0.0540	0.1090	0.0601	-0.0041	
$\sqrt{\text{MSE}}$		0.0409	0.1966	0.3315	0.2481	0.2088	
$n = 4000$							
Bias	1	0.0201	0.0697	0.1121	0.0732	0.0085	0.1850
$\sqrt{\text{MSE}}$		0.0258	0.1292	0.2835	0.1891	0.1301	1.2535
Bias	2	0.0146	0.0594	0.1067	0.0535	-0.0017	
$\sqrt{\text{MSE}}$		0.0294	0.1507	0.2757	0.2184	0.1698	

Denote  $\widehat{X} = (\widehat{X}_1, \dots, \widehat{X}_K)$ , where

$$\widehat{X}_k = -\frac{1}{n} \sum_{t=k+1}^n \frac{1}{\widehat{h}_t} \left( \frac{\widehat{e}_{t-k}^2}{\widehat{h}_{t-k}} - 1 \right) \frac{\partial h_t(\widehat{\tau}_n)}{\partial \tau}.$$

It can be shown that  $\widehat{X} = X + o_p(1)$ . Let  $\widehat{\kappa}_1 = 6(\widehat{\nu}_n - 4)^{-1} + 2$ , and then we can construct consistent estimators of  $\Sigma$ , denoted by  $\widehat{\Sigma}$ . Based on the asymptotic normality of  $\widehat{R}$ , we consider the portmanteau test,

$$Q_R(K) = n\widehat{R}'\widehat{\Sigma}^{-1}\widehat{R}.$$

Under the conditions of Theorem 3.1, if the threshold ARFIMA-HYGARCH model proposed in this paper is correctly specified, the quantity  $Q_R(K)$  will be asymptotically distributed as  $\chi_K^2$ , the chi-square distribution with  $K$  degrees of freedom. Similarly, we can consider the portmanteau test based on the residual autocorrelations, however, it is usually less powerful in detecting departures of the conditional variance specifications, see Li and Li (2008). Hence, we concentrate on  $Q_R(K)$  only.

## 5. SIMULATION RESULTS

In this section, we conduct two simulation experiments to demonstrate the usefulness of the asymptotic results. The following threshold HYGARCH(1,  $d_{FG}$ , 1) model is considered in both experiments,  $y_t = \varepsilon_t h_t^{1/2}$ ,

$$h_t = \gamma^{(1)} + \left( 1 - \frac{1 - \delta^{(1)}L}{1 - \beta^{(1)}L} \left( 1 - \alpha^{(1)} + \alpha^{(1)}(1 - L)^{d_{FG}^{(1)}} \right) \right) e_t^2,$$

if  $y_{t-1} \leq 0$ , and

$$h_t = \gamma^{(2)} + \left( 1 - \frac{1 - \delta^{(2)}L}{1 - \beta^{(2)}L} \left( 1 - \alpha^{(2)} + \alpha^{(2)}(1 - L)^{d_{FG}^{(2)}} \right) \right) e_t^2,$$

if  $y_{t-1} > 0$ , where  $\varepsilon_t$  follows Student's  $t$  distribution with mean zero, variance one and  $\nu$  degrees of freedom. Note that the conditional variance can be rewritten as  $h_t = \gamma^{(1)} + \sum_{k=1}^{\infty} \pi_k^{(1)} e_{t-k}^2$  if  $y_{t-d} \leq 0$ , or  $\gamma^{(2)} + \sum_{k=1}^{\infty} \pi_k^{(2)} e_{t-k}^2$  if  $y_{t-d} > 0$ . In practice, some truncations are often applied to  $\pi^{(i)}(L) = \sum_{k=1}^{\infty} \pi_k^{(i)} L^k$ , see Chung (1999) and Lombardi and Gallo (2002). It is remarked that 50 terms of  $\pi_k^{(i)}$  are used.

Table 3. Estimation result from 500 simulated series of model

3							
regime	$\gamma$	$\beta$	$\alpha$	$d_{FG}$	$\delta$	$\nu$	
$n = 1000$							
Bias	1	0.0162	0.1108	0.1342	0.0463	0.0750	0.3604
$\sqrt{\text{MSE}}$		0.0449	0.2109	0.3045	0.2478	0.1982	1.8102
Bias	2	0.0179	0.0488	0.1248	0.0474	-0.0184	
$\sqrt{\text{MSE}}$		0.0459	0.2184	0.3591	0.2984	0.2473	
$n = 2000$							
Bias	1	0.0155	0.0818	0.1268	0.0458	0.0480	0.2051
$\sqrt{\text{MSE}}$		0.0376	0.1678	0.2531	0.1993	0.1562	1.5358
Bias	2	0.0163	0.0457	0.1215	0.0442	-0.0142	
$\sqrt{\text{MSE}}$		0.0399	0.1904	0.2931	0.2749	0.2253	
$n = 4000$							
Bias	1	0.0118	0.0509	0.1228	0.0347	0.0291	0.1717
$\sqrt{\text{MSE}}$		0.0245	0.1128	0.2361	0.1587	0.1337	1.2453
Bias	2	0.0117	0.0404	0.1200	0.0419	-0.0102	
$\sqrt{\text{MSE}}$		0.0310	0.1621	0.2539	0.2613	0.1873	

In the first experiment, three sub-models of the threshold HYGARCH(1,  $d_{FG}$ , 1) are considered, and the values of the parameter vector  $\theta^{(i)} = (\gamma^{(i)}, \beta^{(i)}, \alpha^{(i)}, d_{FG}^{(i)}, \delta^{(i)})$  and  $\theta =$

$(\theta^{(1)}; \theta^{(2)}; \nu)$  are listed as follows,

Model 1: (0.1, 0.1, 0.80, 0.45, 0.4; 0.1, 0.3, 0.80, 0.45, 0.6; 10),

Model 2: (0.1, 0.1, 0.80, 0.20, 0.4; 0.1, 0.3, 0.80, 0.45, 0.6; 10),

Model 3: (0.1, 0.1, 0.80, 0.45, 0.4; 0.1, 0.3, 0.65, 0.45, 0.6; 10).

We consider three different sample sizes, 1000, 2000 and 4000, and there are 500 replications for each sample size. The Student's  $t$  likelihood-based MLE in section 3 were calculated, and the biases (Bias) and the empirical root mean squared errors ( $\sqrt{\text{MSE}}$ ) are summarized in Tables 1-3. It is observed that the biases are generally small. As sample size increases, all biases and empirical root mean squared errors decrease.

The second experiment is conducted to check the empirical sizes and powers of the test statistic  $Q_R(K)$  in section 5. The generating processes are

Pair 1:

$$\theta = (0.1, -, 0.80, 0.25, -, 0.1, -, 0.80, 0.45, -, 10),$$

or (0.1, 0.1, 0.80, 0.45, 0.4; 0.1, 0.3, 0.80, 0.45, 0.6; 10),

Pair 2:

$$\theta = (0.1, 0.1, 0.80, 0.20, -, 0.1, 0.3, 0.80, 0.45, -, 10),$$

or (0.1, 0.1, 0.80, 0.20, 0.4; 0.1, 0.3, 0.80, 0.45, 0.6; 10),

Pair 3:

$$\theta = (0.1, 0.1, -, -, 0.4; 0.1, 0.3, -, -, 0.6; 10),$$

or (0.1, 0.1, 0.80, 0.45, 0.4; 0.1, 0.3, 0.65, 0.45, 0.6; 10),

where, for each pair, the first model is called the true model, and the second one is the misspecified model. We generated the samples by both the true model and the misspecified model, however the generated samples were always estimated by the true model. For samples generated by the misspecified models, the estimated results are expected to be worse since a sub-model was considered to fit the samples, and hence they are used to assess the powers. Note that the first two pairs correspond to the misspecification of  $\delta$ , while the last pair correspond to the misspecification of the  $\alpha$  and  $d_{FG}$ . As in the first experiment, the sample sizes are set to be 500, 1000, or 2000, and there are 500 replications for each sample size. We consider five different values for  $K$ : 3, 6, 10, 15 and 25. The empirical sizes and powers are presented in Table 4 for  $Q_R(K)$ , and they are based on the upper fifth percentile of the chi-squared distribution with the corresponding degrees of freedom. It can be seen that all empirical sizes are close to the nominal value 0.05, and the test is more powerful as the sample size  $n$  increases.

As is well known, the proposed test statistic can be regarded as a pure significance test and hence its power would not be very high with general departures from the null

Table 4. Empirical size and power of  $Q_R(K)$

$n$	Size			Power		
	500	1000	2000	500	1000	2000
Pair 1						
$Q_R(3)$	0.040	0.044	0.044	0.624	0.634	0.642
$Q_R(6)$	0.040	0.046	0.046	0.622	0.632	0.646
$Q_R(10)$	0.046	0.050	0.050	0.618	0.624	0.638
$Q_R(15)$	0.042	0.054	0.050	0.622	0.628	0.634
$Q_R(25)$	0.042	0.050	0.050	0.622	0.628	0.634
Pair 2						
$Q_R(3)$	0.042	0.044	0.044	0.622	0.636	0.644
$Q_R(6)$	0.044	0.044	0.044	0.624	0.634	0.642
$Q_R(10)$	0.042	0.044	0.046	0.618	0.628	0.640
$Q_R(15)$	0.046	0.046	0.046	0.620	0.628	0.638
$Q_R(25)$	0.040	0.040	0.046	0.624	0.626	0.634
Pair 3						
$Q_R(3)$	0.040	0.042	0.042	0.628	0.646	0.656
$Q_R(6)$	0.042	0.042	0.042	0.636	0.648	0.654
$Q_R(10)$	0.052	0.048	0.048	0.644	0.650	0.656
$Q_R(15)$	0.046	0.054	0.054	0.632	0.646	0.652
$Q_R(25)$	0.044	0.046	0.052	0.648	0.654	0.660

model. Like many of the tests proposed for testing linearity (Luukkonen, Saikkonen and Terasvirta, 1988; Saikkonen and Luukkonen, 1988), its power might be substantially improved if it can be shown to be equivalent to some classical tests under certain alternative hypotheses.

## 6. EMPIRICAL RESULTS

In the asset pricing model, temporary bubbles with prices deviating from the fundamental may arise when the fractions of traders believing in those bubbles is sufficiently large. Driven by evolutionary competition between different trading strategies, the model generates excess volatility, see Timmermann (1993, 1996), Routledge (1999) and Farmer and Joshi (2002). Traders may believe that, in a heterogeneous world, prices will deviate from their fundamental value. Traders can be recognized into two main types: fundamentalists and trend followers. Fundamentalists are Efficient Market Hypothesis (EMH) believers. The EMH states that it is impossible, except through luck, to consistently outperform the market by using any information that the market already knows. Trend followers are chartists. They extrapolate the latest observed price change. Fundamentalists believe that prices will move towards its fundamental rational expectations (RE) value. In contrast, the trend followers are not entirely unaware of the fundamental price. When prices move far away from the fundamental value, they start believing that a price correction towards the fundamental price is about to occur. The fractions of the two different trader types change over time upon price deviations from the RE fundamental value. The financial market is thus characterized by an irregular regime switching.

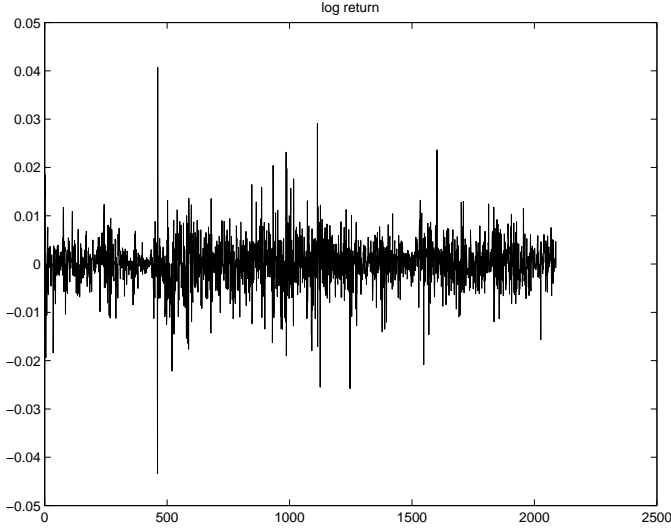


Figure 1. Log-returns of the daily exchange rates of the Korean Won again the US dollar from 2 January, 1999 to 3 January, 2007.

In this section, we apply the proposed threshold HYGARCH model to the daily exchange rates of the Korean Won against the US dollar. There are 2089 observed values for the Won/USD exchange rates from 2 January, 1999 to 3 January, 2007.

We consider a self-excited threshold model, and the threshold value is assumed to be zero. In Kwan, Li and Ng (2010b), it is found that the value of the delay parameter,  $d$ , equals to one is a natural choice. The purpose here is to measure the asymmetric behavior of positive returns and negative returns. In practice, the threshold value is not too far from zero. As in Kwan, Li and Li (2010a), the long memory characteristic is found in the volatility. The fitted threshold model is  $y_t = \varepsilon_t h_t^{1/2}$ , where

$$h_t = 2.826 \times 10^{-6} \times 10^{0.0880 \times 10^{-6}} + \left\{ 1 - \frac{1 - 0.0316_{0.0181} L}{1 - 0.0794_{0.0388} L} \cdot (1 + 2.8203_{0.9882} [(1 - L)^{0.1038_{0.0436}} - 1]) \right\} y_t^2$$

if  $y_{t-1} \leq 0$ , or

$$h_t = 1.129 \times 10^{-6} \times 10^{0.1070 \times 10^{-6}} + \left\{ 1 - \frac{1 - 0.8404_{0.3024} L}{1 - 0.8074_{0.3859} L} \cdot (1 + 1.2922_{0.3462} [(1 - L)^{0.3115_{0.1028}} - 1]) \right\} y_t^2$$

if  $y_{t-1} > 0$ , the fitted degrees of freedom is  $3.8574_{0.2191}$  and the standard errors are given in the corresponding subscripts. The fitted conditional variances  $\hat{h}_t$  are plotted in Figure 2.

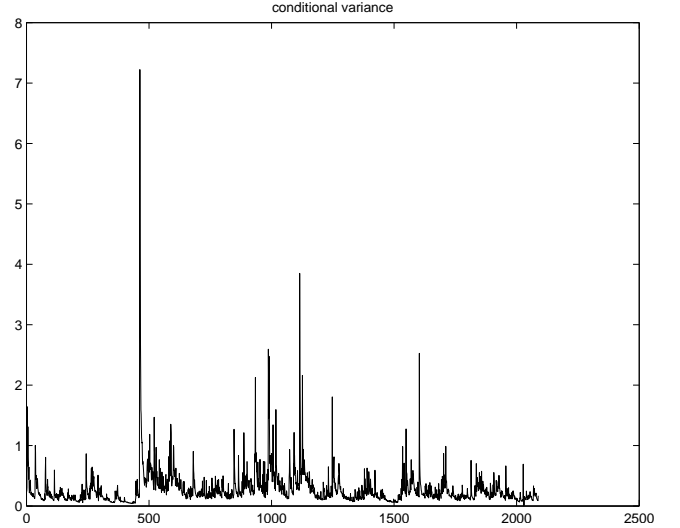


Figure 2. Fitted conditional variance, multiplied by  $10^4$ , of the daily exchange rates of the Korean Won again the US dollar from 2 January, 1999 to 3 January, 2007.

The diagnostic tool  $Q_R(K)$  was also performed here, and the calculated  $p$ -values for  $K = 3, 6, 10, 15$  and  $25$  are respectively  $0.89, 0.97, 0.99, 0.97$  and  $0.99$ . This result shows that the exchange rates possess an hyperbolic memory. The log likelihood has the value of  $8204.28$  for the fitted threshold model, while it is  $8185.57$  for the HYGARCH model as stated in Kwan, Li and Li (2010a). A likelihood ratio test was performed with the null hypothesis of no threshold, and the calculated  $p$ -value is  $0.0022$ . This demonstrates that the threshold HYGARCH model is better than the model without threshold in this application.

## 7. CONCLUSION

In this article, a threshold HYGARCH model is proposed, and some properties, estimation and a diagnostic procedure were presented. The application of this model to the exchange rates of Korean Won against US dollar seems to suggest that the long-memory phenomenon with a threshold structure may exist in the volatility of some financial time series. In other words, the volatility exhibits a nonlinear long memory feature. This phenomenon may also exist in many other time series. The proposed threshold HYGARCH model should be useful in modeling time series that exhibit nonlinear long memory feature.

## ACKNOWLEDGMENTS

The authors are most grateful to the editor and two anonymous referees for suggestions and valuable comments that led to the improvement of this paper. W.K. Li thanks

the Hong Kong Research Grants Councils GRF grant HKU702908P for partial support.

Received August 2010

## REFERENCES

- BAILLIE, R. T., BOLLERSLEV, T. and MIKKESLEN, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **74** 3-30.
- BOLLERSLEV, T. (1986). Generalized autoregression conditional heteroscedasticity. *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** 542-547.
- BOX, G. E. P. and JENKINS, G. M. (1976). *Time Series Analysis Forecasting and Control*, 2nd edition ed. Holden-Day, San Francisco.
- CHUNG, C. F. (1999). Estimating the fractionally integrated GARCH model. *Technical Report, Institute of Economics, Academia Sinica, Taipei*.
- CONRAD, C. (2010). Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics* **157** 441-457.
- CONT, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1** 223 - 236.
- DAVIDSON, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. *Journal of Business and Economic Statistics* **22** 16-29.
- DING, Z. X., GRANGER, C. W. J. and ENGLE, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* **1** 83-106.
- ENGLE, R. F. (1982). Autoregression conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50** 987-1008.
- ENGLE, R. F. and BOLLERSLEV, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews* **5** 1-50.
- FARMER, J. D. and JOSHI, S. (2002). The price dynamics of common trading strategies. *Journal of Economic Behavior and organization* **49** 149 - 171.
- GREENE, M. T. and FIELITZ, B. D. (1977). Long-term dependence in common stock returns. *Journal of Financial Economics* **4** 339 - 349.
- HALL, P. and YAO, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71** 285-317.
- HIRSHLEIFER, D. (2001). Investor psychology and asset pricing. *Journal of Finance* **56** 1533 - 1597.
- KOKOSZKA, P. S. and TAQQU, M. S. (1996). Infinite variance stable moving averages with long memory. *Journal of Econometrics* **73** 79 - 99.
- KWAN, W., LI, W. K. and LI, G. (2010a). On the estimation and diagnostic checking of the ARFIMA-GARCH model. *Computational Statistics and Data Analysis* doi:10.1016/j.csda.2010.07.010.
- KWAN, W., LI, W. K. and NG, K. W. (2010b). A multivariate threshold varying conditional correlation model. *Econometric Reviews* **29** 20-38.
- LI, W. K. (2004). *Diagnostic Checks in Time Series*. Chapman & Hall, Boca Raton.
- LI, C. W. and LI, W. K. (1996). On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics* **11** 253-274.
- LI, G. and LI, W. K. (2008). Least absolute deviation estimation for fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Biometrika* **95** 399-414.
- LI, W. K. and MAK, T. K. (1994). On the squared residual autocorrelations in nonlinear time series with conditional heteroskedasticity. *Journal of Time Series Analysis* **15** 627-636.
- LING, S. and LI, W. K. (1997). On fractionally integrated autoregressive moving average time series models with conditional heteroscedasticity. *Journal of the American Statistical Association* **92** 1184-1193.
- LING, S. and TONG, H. (2005). Testing for a linear MA model against threshold MA models. *The Annals of Statistics* **33** 2529-2552.
- LING, S., TONG, H. and LI, D. (2007). Ergodicity and invertibility of threshold moving-average models. *Bernoulli* **13** 161-168.
- LIU, J., LI, W. K. and LI, C. W. (1997). On a threshold autoregression with conditional heteroscedastic variances. *Journal of Statistical Planning and Inference* **62** 279-300.
- LO, A. W. and MACKINLAY, A. C. (1990). When are contrarian profits due to stock market overreaction? *The Review of Financial Studies* **3** 175 - 205.
- LOMBARDI, M. J. and GALLO, G. M. (2002). Analytic Hessian matrices and the computation of FIGARCH estimates. *Statistical Methods & Applications* **11** 247-264.
- LUUKKONEN, R., SAIKKONEN, P. and TERASVIRTA, T. (1988). Testing Linearity in Univariate Time Series Models. *Scandinavian Journal of Statistics* **15** 161-175.
- MANDELBROT, B. (1963). The variation of certain speculative prices. *Journal of Business* **36** 394 - 419.
- MIKOSCH, T. and STARICA, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *The Annals of Statistics* **28** 1427-1451.
- RABEMANJARA, R. and ZAKOIAN, J. M. (1993). Threshold ARCH model and asymmetries in volatility. *Journal of Applied Econometrics* **8** 31 - 49.
- ROUTLEDGE, B. R. (1999). Adaptive learning in financial markets. *The Review of Financial Studies* **12** 1165 - 1202.
- SAIKKONEN, P. and LUUKKONEN, R. (1988). Lagrange Multiplier Tests for Testing Non-linearities in Time Series Models. *Scandinavian Journal of Statistics* **15** 55-68.
- SCHWERT, G. W. (1989). Why do stock market volatility change over time? *Journal of Finance* **44** 1115 - 1153.
- SHEFRIN, H. (2000). *Beyond Greed and Fear. Understanding Behavioral Finance and the Psychology of Investing*. Harvard Business School Press, Boston.
- STRAUMANN, D. (2005). *Estimation in conditionally heteroscedastic time series models*. Springer, Berlin.
- TIMMERMANN, A. (1993). How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics* **108** 1135 - 1145.
- TIMMERMANN, A. (1996). Excess volatility and predictability of stock prices in autoregressive dividend models with learning. *Review of Economic Studies* **63** 523 - 557.
- TONG, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford.
- WEISS, A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory* **2** 107-131.

Wilson Kwan  
Hong Kong Community College  
The Hong Kong Polytechnic University, Hong Kong  
E-mail address: ccwilson@hkcc-polyu.edu.hk

Wai Keung Li  
Department of Statistics and Actuarial Science  
University of Hong Kong, Hong Kong  
E-mail address: hrntlwk@hku.hk

Guodong Li  
Department of Statistics and Actuarial Science  
University of Hong Kong, Hong Kong  
E-mail address: gdli@hku.hk