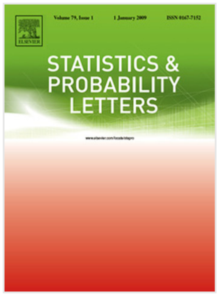



Dynamic behavior of volatility in a nonstationary generalized regime-switching GARCH model

Won-Tak Hong, Eunju Hwang

PII: S0167-7152(15)30350-3
DOI: http://dx.doi.org/10.1016/j.spl.2016.04.002
Reference: STAPRO 7583

To appear in: *Statistics and Probability Letters*

Received date: 18 November 2015
Revised date: 1 April 2016
Accepted date: 1 April 2016

# Dynamic behavior of volatility in a nonstationary generalized regime-switching GARCH model[1]

by Won-Tak Hong and Eunju Hwang[2]

*Department of Mathematical Finance and Department of Applied Statistics, Gachon University*

**Abstract**

In this paper we consider a generalized regime-switching GARCH model with a wide class of dependent innovations, and establish asymptotic normality of the logarithm of volatility in the nonstationary generalized regime-switching GARCH model. This extends existing results for volatilities in nonstationary GARCH model with mixing sequences of innovations. A Monte-Carlo experiment is conducted to validate the main theory for the dynamics of the nonstationary volatilities.

**AMS Classification numbers** Primary: 62E20

Secondary: 60E05, 37M10

**Keywords:** regime-switching GARCH model; volatility; asymptotic normality

[1] This work was supported by the National Research Foundation of Korea (NRF-2015-1006133).

[2] Corresponding author. Mailing Address: Dept. of Applied Statistics, Gachon University, Seongnam, South Korea. Tel: +82-31-750-5373; *E-mail address*: ehwang@gachon.ac.kr

## 1 Introduction

Volatility forecasting is one of important issues in financial markets because volatility is commonly used as a measure of risk associated with financial returns as well as volatility estimates are used as inputs for portfolio optimization, derivative pricing and risk management. To analyze and forecast the volatility, many time series models are proposed and among them the most popular and successful models are ARCH (autoregressive conditional heteroscedastic) models by Engle (1982) and GARCH (generalized autoregressive conditional heteroscedastic) models by Bollerslev (1986).

Since the seminal works of Engle (1982) and Bollerslev (1986), much attention and numerous efforts have been focused on modeling the dynamic volatility. One of the dynamic volatility models is a Markov-regime switching GARCH model, in which different states affect the evolution of time series and dynamic properties depend on the present regime. Markov-regime switching models were introduced by Hamilton (1988, 1989), and continued to be applied to financial time series analysis like foreign exchange rates by Engle and Hamilton (1990), Engle (1994), Cai (1994), Hamilton and Susmel (1994), Bollen et al. (2000), etc. Econometricians such as Gray (1996), Klaassen (2002), Haas et al. (2004), and Bauwens et al. (2010) studied regime-switching GARCH models to describe the dynamics of time series and to improve volatility forecasting. In particular, Gray (1996) proposed a GRS (generalized regime-switching)-GARCH models for short-term interest rates, with time-varying switching probabilities depending on the level of the short rate.

This paper considers the GRS-GARCH of Gray (1996) and discusss asymptotic behavior of the volatility in the nonstationary GRS-GARCH model. Nelson (1990) first studied dynamic behavior of volatility of nonstationary GARCH model and showed that the conditional variance tends to infinity with exponential rate. Linton, et al. (2010) proved the divergence of volatility in a semi-strong GARCH model which has a mixing sequence of innovations. Li et al. (2014) pointed out that it is an interesting question whether or not volatilities in nonstationary GARCH models can be renormalized to converge in distribution to a non-degenerate limit, and characterized precisely the dynamic behavior of volatilities in the nonstationary GARCH(1,1) models with $\varphi$-mixing innovations.

In this work we extend the result of Li et al. (2014) to the GRS-GARCH(1,1) model of Gray (1996) with time-varying switching probabilities. Also, a much wider class of innovations for the GRS-GARCH model is considered to include more general time series models. We adopt $L_2$-near epoch dependent (NED) sequences for the innovations of the GRS-GARCH model. Time series models satisfying the $L_2$-NED are, for examples, ARMA models, bilinear models, GARCH models, threshold models, unit-root SETAR models, HAR models, etc, and include much wider classes than mixing sequences. In the GRS-GARCH(1,1) model with $L_2$-NED innovations, we prove the convergence in distribution of the conditional variance renormalized in a logarithm form to the standard normal distribution. This work extends existing results for volatilities of nonstationary GARCH models such as Nelson (1990), Linton et al. (2010) and Li et al. (2014) by including much wider classes of time series models as well as by considering the regime-switching model.

The remaining of the paper is organized as follows. Section 2 describes the GRS-GARCH model and presents assumptions and a main result. In Section 3, a Monte Carlo study is conducted to verify the dynamics of volatility. Proof is given in Section 4.

## 2 Model and main result

We consider the GRS-GARCH(1,1) model, proposed by Gray (1996). To describe the GRS-GARCH(1,1) model, we start with the following GARCH(1,1) given in (1) and (2) below:

$$y_t = \mu_t + \eta_t\sqrt{h_t}, \quad t = 1, 2, \dots \tag{1}$$

where $\mu_t = \mu(\theta_\mu, \Psi_{t-1})$, $h_t = h(\theta_h, \Psi_{t-1})$ with $\theta_\mu, \theta_h$ being vectors of parameters and $\Psi_{t-1}$ being the entire past history of the data up to time $t-1$, and $\{\eta_t\}$ is a stationary sequence of random variables. The conditional variance $h_t$ is assumed to follow a nonstationary GARCH(1,1) as follows:

$$h_t = h(\theta_h, \Psi_{t-1}) = \omega + \alpha(y_{t-1} - \mu_{t-1})^2 + \beta h_{t-1} \tag{2}$$

with $\theta_h = (\omega, \alpha, \beta)$, $\omega > 0$ and $\alpha, \beta \geq 0$, satisfying

$$E \log(\beta + \alpha\eta_1^2) \geq 0. \tag{3}$$

Condition (3) implies that the GARCH(1,1) model is nonstationary and $h_t \to \infty$ a.s. as $t \to \infty$, (see Nelson (1990)). Gray (1996) proposed the GRS-GARCH model by adopting Markov switching-regime with time-varying switching probabilities. The model has the following general form:

$$y_t = \mu[\theta_\mu(S_t), \Psi_{t-1}] + \eta_t\sqrt{h[\theta_h(S_t), \Psi_{t-1}]}$$

where $S_t$ is the unobserved regime at time $t$. As seen in Gray (1996), $\Psi_{t-1}$ does not contain $S_t$ or lagged values of $S_t$. In this paper, we consider two-state Markov process for regimes, i.e., $S_t \in \{1, 2\}$, (multiple-regimes case can be given in the same way), and we are interested in dynamic behavior of the volatility by observing the asymptotic of renormalized $h_t = h[\theta_h(S_t), \Psi_{t-1}]$.

Let $p_{1t} = Pr(S_t = 1|\Psi_{t-1})$ and $p_{2t} = 1 - p_{1t} = Pr(S_t = 2|\Psi_{t-1})$. These probabilities are determined by transition probabilities of first-order Markov process, following Hamilton (1988, 1989, 1990), which are assumed to be time-dependent as in Gray (1996). Let

$$\begin{aligned} &Pr(S_t = 1|S_{t-1} = 1) = P_t, \quad Pr(S_t = 2|S_{t-1} = 1) = 1 - P_t, \\ &Pr(S_t = 2|S_{t-1} = 2) = Q_t, \quad Pr(S_t = 1|S_{t-1} = 2) = 1 - Q_t. \end{aligned} \tag{4}$$

We adopt the recursive formula of probability $p_{1t}$ with $P_t$, $Q_t$ and the likelihood function of the GRS-GARCH model in Eq.(11) on p. 37 of Gray (1996), which is rewritten on (10) below in Appendix. In this work we assume that $0 < p_{1t} < 1$ for taking account of two distinct regimes.

For $i = 1, 2$, let

$$h_{it} = h[\theta_h(S_t), \Psi_{t-1}] \quad \text{if } S_t = i.$$

Then we have the conditional variance;

$$h_t = Var[y_t|\Psi_{t-1}] = Var\left[y_t - E(y_t|\Psi_{t-1})|\Psi_{t-1}\right] = E[(y_t - \mu_t)^2|\Psi_{t-1}]$$

$$= P(S_t = 1|\Psi_{t-1})E[(y_t - \mu_t)^2|\Psi_{t-1}, S_t = 1] + P(S_t = 2|\Psi_{t-1})E[(y_t - \mu_t)^2|\Psi_{t=1}, S_t = 2]$$

$$= p_{1t}h_{1t} + p_{2t}h_{2t} =: \mathbf{p}_t'\mathbf{h}_t$$

where $\mathbf{p}_t = (p_{1t}, p_{2t})'$ and $\mathbf{h}_t = (h_{1t}, h_{2t})'$. Note that $h_{it}$ is given by $\theta_h(S_t) := (\omega(S_t), \alpha(S_t), \beta(S_t))$ with $S_t = i$. Denote $(\omega_i, \alpha_i, \beta_i) = (\omega(S_t), \alpha(S_t), \beta(S_t))$ with $S_t = i$ for $i = 1, 2$. Then we have

$$h_{it} = \omega_i + \alpha_i(y_{t-1} - \mu_{t-1})^2 + \beta_i h_{t-1}$$

where $h_{t-1} = p_{1,t-1}h_{1,t-1} + p_{2,t-1}h_{2,t-1} = \mathbf{p}_{t-1}'\mathbf{h}_{t-1}$.

If conditional normality is assumed for each regime, then, (w.p. below stands for 'with probability')

$$y_t|\Psi_{t-1} \sim \begin{cases} N(\mu_{1t}, h_{1t}) & \text{w.p. } p_{1t} \\ N(\mu_{2t}, h_{2t}) & \text{w.p. } 1 - p_{1t} \end{cases}$$

where $\mu_{it}$ are defined similarly for $i = 1, 2$: $\mu_{it} = \mu[\theta_\mu(S_t), \Psi_{t-1}]$ if $S_t = i$.

We see

$$\mathbf{h}_t = \begin{bmatrix} h_{1t} \\ h_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 + \alpha_1(y_{t-1} - \mu_{t-1})^2 + \beta_1 h_{t-1} \\ \omega_2 + \alpha_2(y_{t-1} - \mu_{t-1})^2 + \beta_2 h_{t-1} \end{bmatrix}$$

$$= \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + (y_{t-1} - \mu_{t-1})^2 \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + h_{t-1} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$=: \mathbf{w} + (y_{t-1} - \mu_{t-1})^2\mathbf{a} + h_{t-1}\mathbf{b}$$

where $\mathbf{w} = (\omega_1, \omega_2)'$, $\mathbf{a} = (\alpha_1, \alpha_2)'$, $\mathbf{b} = (\beta_1, \beta_2)'$.

We multiply by $\mathbf{p}_t'$ to obtain

$$\mathbf{p}_t'\mathbf{h}_t = \mathbf{p}_t'\mathbf{w} + (y_{t-1} - \mu_{t-1})^2\mathbf{p}_t'\mathbf{a} + h_{t-1}\mathbf{p}_t'\mathbf{b}$$

$$h_t = W_t + (y_{t-1} - \mu_{t-1})^2 A_t + h_{t-1} B_t$$

where $W_t = \mathbf{p}_t'\mathbf{w}$, $A_t = \mathbf{p}_t'\mathbf{a}$ and $B_t = \mathbf{p}_t'\mathbf{b} \in \mathbb{R}^1$.

Rewrite

$$h_t = W_t + A_t(y_{t-1} - \mu_{t-1})^2 + B_t h_{t-1} \tag{5}$$

where

$$W_t = p_{1t}\omega_1 + (1 - p_{1t})\omega_2, \quad A_t = p_{1t}\alpha_1 + (1 - p_{1t})\alpha_2, \quad B_t = p_{1t}\beta_1 + (1 - p_{1t})\beta_2. \tag{6}$$

We have the following expressions for $h_t$, which are similar forms to those of the conditional variance of a GARCH(1,1) model except for parameters. Note that the parameters in (5) are dependent on $t$ due to time-varying switching probabilities of the GRS-GARCH model.

**Lemma 2.1** *The conditional variance $h_t$ in (5) of the GRS-GARCH model satisfies the following two expressions.*

$$(a) \quad h_t = W_t + \sum_{k=1}^{t-1} W_k \prod_{j=1}^{t-k} (B_{t-j+1} + A_{t-j+1}\eta_{t-j}^2) + \prod_{i=0}^{t-1} (B_{i+1} + A_{i+1}\eta_i^2) h_0$$

$$(b) \quad h_t = \left[\prod_{i=0}^{t-1} (B_{i+1} + A_{i+1}\eta_i^2)\right] \left[h_0 + \sum_{k=1}^{t} \prod_{j=0}^{k-1} \frac{W_k}{B_{j+1} + A_{j+1}\eta_j^2}\right]$$

*where $W_j$, $A_j$ and $B_j$ are given in (6) above.*

Now in order to derive a main result, we make the following assumptions.

**(A1)** $\{\eta_t\}$ is strictly stationary and ergodic with $E[\log(\beta_i + \alpha_i\eta_1^2)]^r < \infty$ for $i = 1, 2$, for some $r \geq 2$.

**(A2)** The GARCH parameters $(\omega_i, \alpha_i, \beta_i)$ with $\omega_i > 0$, $\alpha_i, \beta_i \geq 0$, for $i = 1, 2$, satisfy

$$E \log(\beta_i + \alpha_i\eta_1^2) \geq 0.$$

**(A3)** $\{\eta_t\}$ is $L_2$-NED (near epoch dependent) on $\{e_t\}$, where $\{e_t\}$ is a sequence of random variables with zero mean, that is, $\{\eta_t\}$ satisfies

$$\left\| \eta_t - E\left[\eta_t | \mathcal{F}_{t-\ell}^{t+\ell}\right] \right\|_2 \leq d_t \nu(\ell)$$

where $\mathcal{F}_{t-\ell}^{t+\ell} = \sigma\{e_{t-\ell}, \ldots, e_t, e_{t+1}, \ldots, e_{t+\ell}\}$, $d_t$ is a sequence of positive constants and $\nu(\ell) \to 0$ as $\ell \to \infty$.

See Davidson (2002) for the functional central limit theorem (FCLT) and examples of $L_2$-NED sequences. According to Davidson (2002), ARMA models, bilinear models, GARCH models, threshold models, unit-root SETAR models, etc are $L_2$-NED. Lee (2014) showed that heterogeneous autoregressive model of order $\infty$, HAR($\infty$), proposed by Hwang and Shin (2014) as an extension of Corsi (2009)'s HAR(3) model, is $L_2$-NED. Hence this work includes much wider classes of time series models than the existing literature for volatilities of nonstationary GARCH models such as Nelson (1990), Linton et al. (2010) and Li et al. (2014).

Let $X_t = \log(B_{t+1} + A_{t+1}\eta_t^2)$, $m_t = E\log(B_{t+1} + A_{t+1}\eta_t^2)$ and

$$\sigma^2 = \lim_{T\to\infty} Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(X_t - m_t)\right].$$

Note that $\{X_t : t = 1, 2, \dots\}$ is $L_2$-NED sequence.

**Theorem 2.2** *Suppose that assumptions (A1)–(A3) hold. Then $\sigma^2 < \infty$ and as $T \to \infty$,*

$$\frac{1}{\sigma\sqrt{T}}\left[\log h_T - \sum_{t=0}^{T-1} m_t\right] \xrightarrow{\text{d}} N(0,1).$$

**Remark 2.3** The result in Theorem 2.2 extends Theorems of Li et al. (2014) in two-fold: first, Li et al. (2014) considered $\varphi$-mixing sequence $\eta_t$ while this present work deals with a wider class of weak dependence, $L_2$-NED sequence $\eta_t$. Secondly, the model of Li et al. (2014) is a GARCH model, while we adopt the GRS-GARCH model of Gray (1996) with time-varying regime-switching probabilities.

**Remark 2.4** In this work it is assumed that probabilities $p_{1t} = Pr(S_t = 1|\Psi_{t-1})$ and $p_{2t} = Pr(S_t = 2|\Psi_{t-1})$ are neither zero nor one, that is, two distinct regimes are adopted. If one of $p_{1t}$ and $p_{2t}$ for all $t$ is zero, say, $p_{2t} = 0$, the situation is the same as that in Li et al. (2014) along with $m_t$ replaced by the top Lyapunov exponent $\gamma := E\log(\beta_1 + \alpha_1\eta_1^2)$. If $\gamma > 0$, then we follow Theorem 2.1 of Li et al. (2014), and if $\gamma = 0$, then we follow their Theorem 2.2. Hence in this paper we need the condition $0 < p_{1t} < 1$.

Under two distinct regimes with $0 < p_{1t} < 1$, note that $\log(B_{t+1} + A_{t+1}\eta_t^2) =$

$$\log[p_{1,t+1}(\beta_1 + \alpha_1\eta_t) + p_{2,t+1}(\beta_2 + \alpha_2\eta_t)] > p_{1,t+1}\log(\beta_1 + \alpha_1\eta_t) + p_{2,t+1}\log(\beta_2 + \alpha_2\eta_t) \quad \text{a.s.},$$

where the inequality does not include the equality by the concavity of log function unless $p_{1,t+1}$ is zero or one, and thus

$$m_t = E\log(B_{t+1} + A_{t+1}\eta_t^2) > p_{1,t+1}E\log(\beta_1 + \alpha_1\eta_t) + p_{2,t+1}E\log(\beta_2 + \alpha_2\eta_t).$$

Even if both $E\log(\beta_i + \alpha_i\eta_1^2) = 0$ for $i = 1, 2$, it implies that $m_t > 0$, and thus $\sum_{t=0}^{T-1} m_t > 0$. Therefore, in any cases of either $E\log(\beta_i + \alpha_i\eta_1^2) > 0$ or $E\log(\beta_i + \alpha_i\eta_1^2) = 0$ under condition (A2) we have the asymptotic normality result in Theorem 2.2.

# 3 Monte Carlo study

This section provides numerical validations to our main result in Theorem 2.2 by generating a nonstationary GRS-GARCH(1,1) process. For the numerical simulation, an AR process $\eta_t = \rho\eta_{t-1} + \epsilon_t$, $t = 1, \ldots, T$, with $|\rho| < 1$ and with i.i.d. standard normal $\epsilon_t$, is adopted as a stationary sequence, and constant transition probabilities $P_t \equiv P$ and $Q_t \equiv Q$ are chosen. We choose the initial values of $p_{1t}, h_{1t}$ and $\mu_{1t}$ randomly from uniform distribution $U(0, 1)$ and generate values of $p_{1t}, h_{1t}$ and $\mu_{1t}$ recursively by using (10) in Appendix below. Also to enforce conditions for the parameters in (A2) so that RS-GARCH(1,1) becomes nonstationary we use $\alpha_1 = 0.8$, $\alpha_2 = 0.9$, $\beta_1 = 0.9$, $\beta_2 = 0.8$, $\omega_1 = 1$ and $\omega_2 = 1.2$. From these values and $p_{1t}$, we obtain $A_t$, $B_t$ and $W_t$ in (6) and thus we get $X_t$ as well as $h_t$, $t = 1, \ldots, T$. To estimate $m_t$ and $\sigma^2$, we use sample size $N$ to obtain their consistent estimates.

For numerical evaluations, we use $T = 100$, $1500$ and $N = 200$, $2000$. Figure 1 indicates histograms of $\log h_T - \sum_{t=0}^{T-1} m_t$. In Figure 1, we see that the mean tends to be centered as $T$ gets larger. Figure 1(a)(b) depict deviation of the mean from zero for smaller $T$ and Figure 1(c)(d) indicate mean closer to zero for larger $T$.

To further validate the convergence in normal distribution we compare the empirical cumulative distribution function (CDF) to the standard normal CDF, as seen in Figure 2. As time $T$ gets larger from $T = 100$ to $T = 1500$, we see the CDF gets closer to standard normal CDF. The convergence for smaller sample size is pictured in Figure 2(a)(c). Figure 2(b)(d) show the CDF convergence of larger sample size. Other qualitative way to see the convergence is to see the normal probability plot given in Figure 3. Curvature or large deviation from the guided linear line means non-normality. We see that Figure 3(d) with $N = 2000$ and $T = 1500$ shows clearer linearity than other cases.

We further investigate the convergence in normal distribution given in Theorem 2.2 using three well-known normality tests; Kolmogorov-Smirnov, Lilliefors, and Jarque-Bera tests. $N = 2000$ and $T = 100, 1500$ are considered for the three tests. The significant level for the tests is set to be $\alpha = 0.1$, and $p$-values are obtained in Table 1, that confirms the normality of our main theorem for different sets of parameters under the nonstationary RS-GARCH(1,1) process. As seen in Table 1, most of cases with larger $T = 1500$ provide $p$-values larger than the significant level 0.1.

# 4 Proof

**Proof of Theorem 2.2:** Note that

$$\frac{1}{\sigma\sqrt{t}}\left[\log h_t - \sum_{j=0}^{t-1} m_j\right] = \frac{1}{\sigma\sqrt{t}}\log\left(\frac{h_t}{\exp\sum_{j=0}^{t-1} m_j}\right).$$

We recall the result of Lemma 2.1(b) and write

$$h_t = \left[\prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\right]\left[h_0 + \sum_{k=1}^{t}\prod_{j=0}^{k-1}\frac{W_k}{B_{j+1} + A_{j+1}\eta_j^2}\right] = \left(\prod_{i=0}^{t-1} e^{X_i}\right)\left[h_0 + \sum_{k=1}^{t} W_k \prod_{j=0}^{k-1} e^{-X_j}\right]$$

$$= \exp\left(\sum_{j=0}^{t-1} X_j\right)\left[h_0 + \sum_{k=1}^{t} W_k \cdot \exp\left(-\sum_{j=0}^{k-1} X_j\right)\right].$$

Thus we write

$$\frac{h_t}{\exp\sum_{j=0}^{t-1} m_j} = \exp\left(\sum_{j=0}^{t-1}(X_j - m_j)\right)\left[h_0 + \sum_{k=1}^{t} W_k \cdot \exp\left(-\sum_{j=0}^{k-1} X_j\right)\right]$$

and

$$\frac{1}{\sigma\sqrt{t}}\log\left(\frac{h_t}{\exp\sum_{j=0}^{t-1} m_j}\right) = \frac{1}{\sigma\sqrt{t}}\sum_{j=0}^{t-1}(X_j - m_j) + \frac{1}{\sigma\sqrt{t}}\log\left[h_0 + \sum_{k=1}^{t} W_k \cdot \exp\left(-\sum_{j=0}^{k-1} X_j\right)\right]. \quad (7)$$

First we consider the first term of the right-hand side of (7). We apply Theorem 1.2 of Davidson (2002), of which conditions will be checked. Note that for $r \geq 2$,

$$E|X_j|^r = E\left|\log[p_{1,j+1}(\beta_1 + \alpha_1\eta_j^2) + p_{2,j+1}(\beta_2 + \alpha_2\eta_j^2)]\right|^r$$

$$\leq E\left|\log[(\beta_1 + \alpha_1\eta_j^2) + (\beta_2 + \alpha_2\eta_j^2)]\right|^r \leq E\left|\log(\beta_1 + \alpha_1\eta_j^2) + \log(\beta_2 + \alpha_2\eta_j^2)]\right|^r$$

$$\leq \left(\{E|\log(\beta_1 + \alpha_1\eta_j^2)|^r\}^{1/r} + \{E|\log(\beta_2 + \alpha_2\eta_j^2)|^r\}^{1/r}\right)^r < \infty$$

by (A1). Also, according to Theorem 21.1 of Billingsley (1968), we can see $\sigma^2 < \infty$. Therefore, by Theorem 1.2 of Davidson (2002), the CLT for $\{X_j : j = 1, 2, \dots\}$ holds:

$$\frac{1}{\sigma\sqrt{t}}\sum_{i=0}^{t-1}(X_i - m_i) \xrightarrow{\text{d}} N(0,1). \quad (8)$$

Now we show that the second term of the right-hand side of (7) tends to 0 in probability:

$$\frac{1}{\sigma\sqrt{t}}\log\left[h_0 + \sum_{k=1}^{t} W_k \cdot \exp\left(-\sum_{j=0}^{k-1} X_j\right)\right] \xrightarrow{\text{p}} 0. \quad (9)$$

Let

$$U_t := \sum_{k=1}^{t} W_k \cdot \exp\left(-\sum_{j=0}^{k-1} X_j\right) = W_1 + \frac{W_2}{e^{X_1}} + \frac{W_3}{e^{X_1+X_2}} + \cdots + \frac{W_t}{e^{X_1+\cdots+X_{t-1}}}.$$

Note that $U_t$ increases monotonically and

$$|U_t| \leq \bar{w}\left(1 + \frac{1}{e^{X_1}} + \frac{1}{e^{X_1+X_2}} + \cdots + \frac{1}{e^{X_1+\cdots+X_{t-1}}}\right) =: \bar{w}V_t$$

where $\bar{w} = 2\max\{\omega_1, \omega_2\}$, and note that

$$V_t = V_{t-1} + \frac{1}{e^{X_1+\cdots+X_{t-1}}}.$$

By (8) and by the fact that $m_i > 0$ in the discussion of Remark 2.4, we have that $X_1+\cdots+X_t \to \infty$ a.s. as $t \to \infty$. Hence $\lim_{t\to\infty} V_t$ exists a.s. and so $U_t$ converges a.s. Hence, the desired result in (9) holds: $\frac{1}{\sigma\sqrt{t}} \log\left[h_0 + U_t\right] \xrightarrow{\text{p}} 0$ as $t \to \infty$. Therefore, we complete the proof of Theorem 2.2 $\square$

## Appendix

In this appendix, (i): we present the recursive formula of probability $p_{1t}$, with $P_t$, $Q_t$ in (4) and the likelihood function of the GRS-GARCH model, given in Grey (1996), and (ii): we give proofs of Lemma 2.1.

(i) *Derivation of the recursive formula of $p_{it}$:* we let $\Psi_{t-1} = \{\tilde{y}_{t-1}\}$ where $\tilde{y}_{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$ and observe

$$p_{1t} = Pr(S_t = 1|\tilde{y}_{t-1}) = P_t Pr(S_{t-1} = 1|\tilde{y}_{t-1}) + (1 - Q_t)Pr(S_{t-1} = 2|\tilde{y}_{t-1})].$$

For $j = 1, 2,$ $Pr(S_{t-1} = j|\tilde{y}_{t-1}) = Pr(S_{t-1} = j|y_{t-1}, \tilde{y}_{t-2})$

$$= \frac{f(y_{t-1}|S_{t-1} = j, \tilde{y}_{t-2})Pr(S_{t-1} = j|\tilde{y}_{t-2})}{\sum_{i=1}^{2} f(y_{t-1}|S_{t-1} = i, \tilde{y}_{t-2})Pr(S_{t-1} = i|\tilde{y}_{t-2})} = \frac{f(y_{t-1}|S_{t-1} = j)p_{j,t-1}}{\sum_{i=1}^{2} f(y_{t-1}|S_{t-1} = i)p_{i,t-1}}$$

by Bayes Rule, where $f(y_{t-1}|S_{t-1} = i)$ is the likelihood function of GRS-GARCH model at time $t-1$ given $S_{t-1} = i$. The conditional distribution of $y_t$ given $\Psi_{t-1}$ can be written by

$$y_t|\Psi_{t-1} \sim \begin{cases} f(y_t|S_t = 1,\ \Psi_{t-1}) & \text{w.p.} \quad p_{1t}, \\ f(y_t|S_t = 2,\ \Psi_{t-1}) & \text{w.p.} \quad p_{2t}. \end{cases}$$

If the conditional normality is assumed, then

$$f(y_t|S_t = i) = \frac{1}{\sqrt{2\pi h_{it}}} \exp\left\{\frac{-(y_t - \mu_{it})^2}{2h_{it}}\right\}, \qquad i = 1, 2.$$

Therefore we have

$$p_{1t} = P_t\left[\frac{g_{1,t-1}p_{1,t-1}}{g_{1,t-1}p_{1,t-1} + g_{2,t-1}(1 - p_{1,t-1})}\right] + (1 - Q_t)\left[\frac{g_{2,t-1}(1 - p_{1,t-1})}{g_{1,t-1}p_{1,t-1} + g_{2,t-1}(1 - p_{1,t-1})}\right] \tag{10}$$

where $g_{1t} = f(y_t|S_t = 1)$ and $g_{2t} = f(y_t|S_t = 2)$.

(ii) *Proof of Lemma 2.1:*

(a) $h_t = W_t + A_t(y_{t-1} - m_{t-1})^2 + B_t h_{t-1} = W_t + A_t(\eta_{t-1}^2 h_{t-1}) + B_t h_{t-1} = W_t + (B_t + A_t\eta_{t-1}^2)h_{t-1}$

$$= W_t + (B_t + A_t\eta_{t-1}^2)[W_{t-1} + (B_{t-1} + A_{t-1}\eta_{t-2}^2)h_{t-2}]$$

$$= W_t + W_{t-1}(B_t + A_t\eta_{t-1}^2) + (B_t + A_t\eta_{t-1}^2)(B_{t-1} + A_{t-1}\eta_{t-2}^2)h_{t-2}$$

$$\vdots$$

$$= W_t + W_{t-1}(B_t + A_t\eta_{t-1}^2) + \cdots + W_1(B_t + A_t\eta_{t-1}^2)(B_{t-1} + A_{t-1}\eta_{t-2}^2)\cdots(B_2 + A_2\eta_1^2)$$

$$+(B_t + A_t\eta_{t-1}^2)(B_{t-1} + A_{t-1}\eta_{t-2}^2)\cdots(B_1 + A_1\eta_0^2)h_0.$$

Thus

$$h_t = W_t + \sum_{k=1}^{t-1} W_k \prod_{j=1}^{t-k}(B_{t-j+1} + A_{t-j+1}\eta_{t-j}^2) + \prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)h_0.$$

(b) Note that the right-hand side of equality in (b) of Lemma 2.1 is equal to

$$h_0\prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2) + \sum_{k=1}^{t} W_k \prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\prod_{j=0}^{k-1}\frac{1}{B_{j+1} + A_{j+1}\eta_j^2}.$$

Since $h_0\prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)h_0$ is common in (a) and (b), we may show that

$$W_t + \sum_{k=1}^{t-1} W_k \prod_{j=1}^{t-k}(B_{t-j+1} + A_{t-j+1}\eta_{t-j}^2) = \sum_{k=1}^{t} W_k \prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\prod_{j=0}^{k-1}\frac{1}{B_{j+1} + A_{j+1}\eta_j^2}. \tag{11}$$

Start with the right term of (11):

$$\sum_{k=1}^{t} W_k \prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\prod_{j=0}^{k-1}\frac{1}{B_{j+1} + A_{j+1}\eta_j^2}$$

$$= \sum_{k=1}^{t-1} W_k \prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\prod_{j=0}^{k-1}\frac{1}{B_{j+1} + A_{j+1}\eta_j^2} + W_t\prod_{i=0}^{t-1}(B_{i+1} + A_{i+1}\eta_i^2)\prod_{j=0}^{t-1}\frac{1}{B_{j+1} + A_{j+1}\eta_j^2}$$

$$= W_t + \sum_{k=1}^{t-1} W_k \frac{(B_1 + A_1\eta_0^2)\cdots(B_{k+1} + A_{k+1}\eta_k^2)\cdots(B_t + A_t\eta_{t-1}^2)}{(B_1 + A_1\eta_0^2)\cdots(B_k + A_k\eta_{k-1}^2)}$$

$$= W_t + \sum_{k=1}^{t-1} W_k(B_{k+1} + A_{k+1}\eta_k^2)\cdots(B_t + A_t\eta_{t-1}^2) = W_t + \sum_{k=1}^{t-1} W_k \prod_{l=k+1}^{t}(B_l + A_l\eta_{l-1}^2).$$

Note that in the left term of (11), $\prod_{j=1}^{t-k}(B_{t-j+1} + A_{t-j+1}\eta_{t-j}^2) = \prod_{l=k+1}^{t}(B_l + A_l\eta_{l-1}^2)$. Hence we have the desired equality in (11) and complete proof of Lemma 2.1. $\square$

## References


Bauwens, L., Preminger, A., Rombouts, J.V.K., 2010. Theory and inference for a Markov switching GARCH model. Econometrics Journal 13, 218–244.

Billingsley, P., 1968. Convergence of Probability Measures. John Wiley & Sons, New York.

Bollen, N.P.B., Gray, S.F., Whaley, R.E., 2000. Regime switching in foreign exchange rates: Evidence from currency option prices. Journal of Econometrics 94, 239–276.

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

Cai, J., 1994. A Markov model of switching-regime ARCH. Journal of Business and Economic Statistics 12, 309–316.

Corsi, F., 2009. A simple approximate long memory model of realized volatility. Journal of Financial Economics 7, 174–196.

Davidson, J., 2002. Establishing condition for the functional central limit theorem in nonlinear and semi-parametric time series processes. Journal of Econometrics 106, 243–269.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. Econometrica 50, 987–1008.

Engle, C., 1994. Can the Markov switching model forecast exchange rates? Journal of International Economics 36, 151–165.

Engle, C., Hamilton, J.D., 1990. Long swings in the dollar: Are they in the data and do markets know it? American Economic Review 80, 689–713.

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

Haas, M., Mittnik, S., Palolella, M.S., 2004. Mixed normal conditional heteroskedasticity. Journal of Financial Econometrics 2, 211–250.

Hamilton, J.D., 1988. Rational-expectations econometric analysis of changes in regime. An investigation of the term structure of interest rates. Journal of Economic Dynamics and Control 12, 385–423.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57, 357–384.

Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39–70.

Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64, 307–333.

Hwang, E., Shin, D.W., 2014. Infinite order, long memory heterogeneous autoregressive models. Computational Statistics and Data Analysis 76, 339–358.

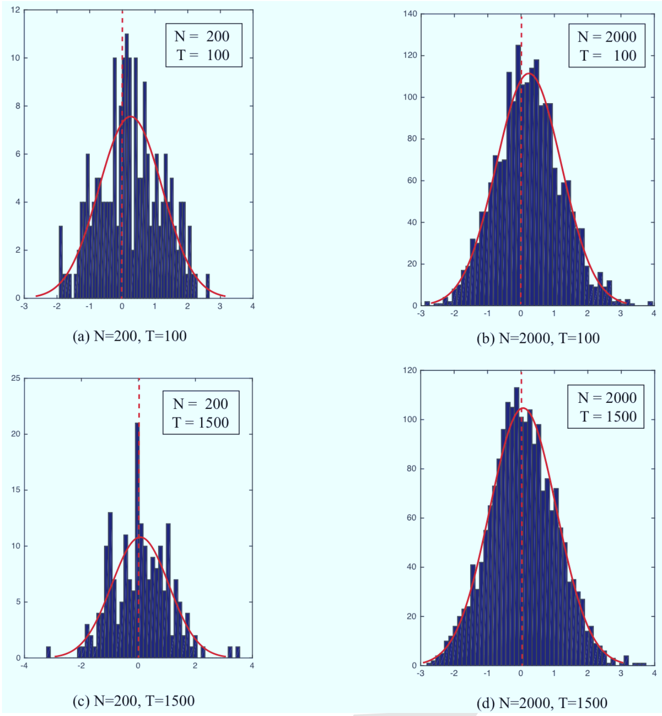
Klaassen, F., 2002. Improving GARCH volatility forecasts with regime-switching GARCH. Empirical Economics 27, 363–394.

Lee, O., 2014. The functional central limit theorem and structural change test for the HAR($\infty$) model. Economics Letters 124, 370–373.

Li, D., Li, M., Wu, W., 2014. On dynamics of volatilities in nonstationary GARCH models. Statistics and Probability Letters 94, 86–90.

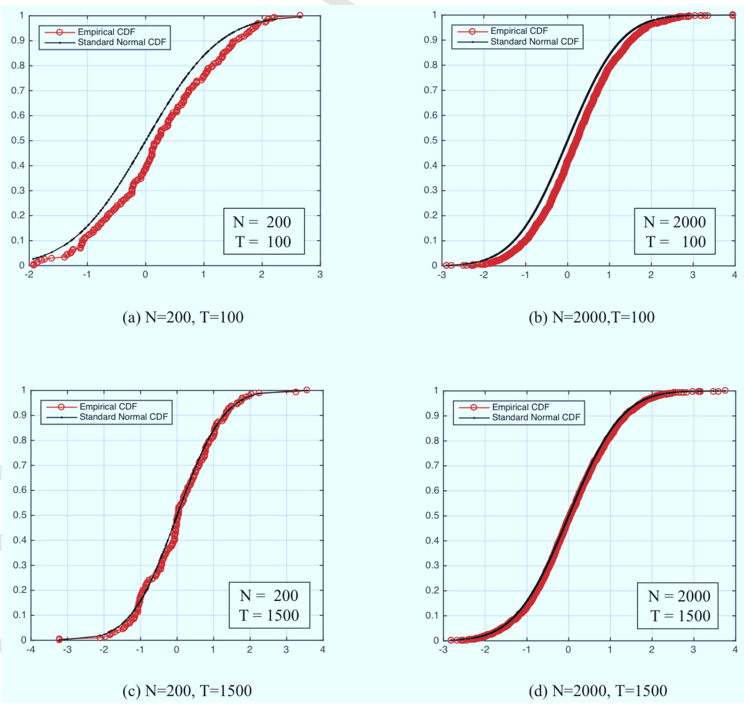
Linton, O., Pan, J., Wang, H., 2010. Estimation for a nonstationary semi-strong GARCH(1,1) with heavy-tailed errors. Econometric Theory 26, 1–28.

Nelson, D.B., 1990. Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318–334.



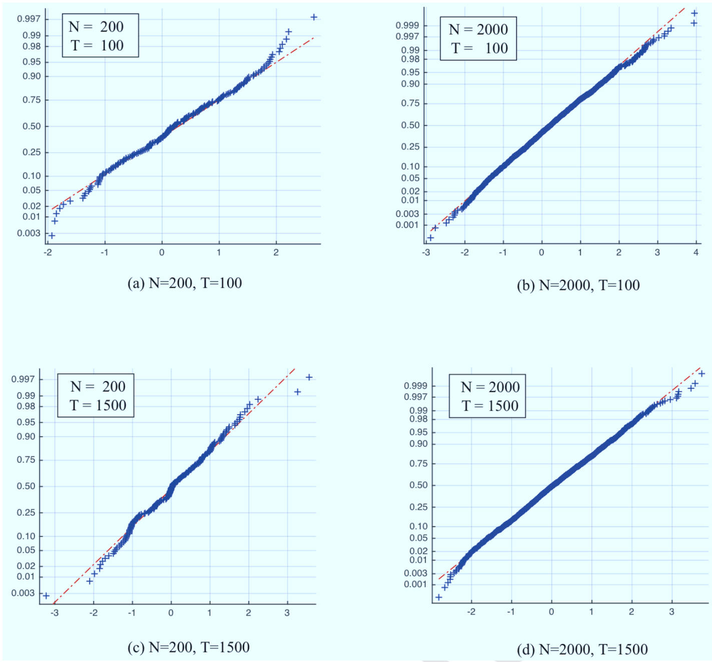

(a) N=200, T=100 (b) N=2000, T=100 (c) N=200, T=1500 (d) N=2000, T=1500

Figure 1: Histogram of the centered logarithmic volatilities for $N = 200, 2000$, and $T = 100, 1500$. Fitted distribution is shown in solid red line and the location of zero is pictured in dotted line to show the shift of the mean.




(a) N=200, T=100 (b) N=2000,T=100 (c) N=200, T=1500 (d) N=2000, T=1500

Figure 2: Cumulative distribution function (CDF) for $N = 200, 2000$, and $T = 100, 1500$. Empirical CDF is generated by simulated data and compared to the standard normal CDF.




(a) N=200, T=100 (b) N=2000, T=100

(c) N=200, T=1500 (d) N=2000, T=1500

Figure 3: Normal probability plot of simulated data for $N = 200, 2000$, and $T = 100, 1500$. Linearity of the data is guided by red dotted line.

Table 1: Normality test results for $T = 100, 1500$ and $N = 2000$. Kolmogorov-Smirnov, Lilliefors, and Jarque-Bera test statistics are used with significant level $\alpha = 0.1$.

| $T$ | $P_t$ | $Q_t$ | $\rho$ | $p-$value | | |
|---|---|---|---|---|---|---|
| | | | | Kolmogorov-Smirnov Test | Lilliefors Test | Jarque-Bera Test |
| 100 | 0.2 | 0.7 | 0.1 | 0.000 | 0.166 | 0.056 |
| | | | 0.3 | 0.000 | 0.029 | 0.006 |
| | | 0.5 | 0.1 | 0.000 | 0.163 | 0.055 |
| | | | 0.3 | 0.000 | 0.025 | 0.006 |
| | 0.8 | 0.7 | 0.1 | 0.000 | 0.148 | 0.056 |
| | | | 0.3 | 0.000 | 0.024 | 0.005 |
| | | 0.5 | 0.1 | 0.000 | 0.187 | 0.053 |
| | | | 0.3 | 0.000 | 0.015 | 0.005 |
| 1500 | 0.2 | 0.7 | 0.1 | 0.130 | 0.227 | 0.124 |
| | | | 0.3 | 0.126 | 0.500 | 0.097 |
| | | 0.5 | 0.1 | 0.126 | 0.247 | 0.123 |
| | | | 0.3 | 0.130 | 0.500 | 0.097 |
| | 0.8 | 0.7 | 0.1 | 0.141 | 0.276 | 0.124 |
| | | | 0.3 | 0.177 | 0.499 | 0.099 |
| | | 0.5 | 0.1 | 0.121 | 0.254 | 0.123 |
| | | | 0.3 | 0.188 | 0.500 | 0.099 |