Econometric Theory, **31**, 2015, 880–890.
doi:10.1017/S0266466614000632

ASYMPTOTIC INFERENCE FOR AR MODELS WITH HEAVY-TAILED G-GARCH NOISES

RONGMAO ZHANG
Zhejiang University

SHIQING LING
Hong Kong University of Science and Technology

It is well known that the least squares estimator (LSE) of an AR(p) model with i.i.d. (independent and identically distributed) noises is $n^{1/\alpha}L(n)$-consistent when the tail index α of the noise is within $(0,2)$ and is $n^{1/2}$-consistent when $\alpha \geq 2$, where $L(n)$ is a slowly varying function. When the noises are not i.i.d., however, the case is far from clear. This paper studies the LSE of AR(p) models with heavy-tailed G-GARCH(1,1) noises. When the tail index α of G-GARCH is within $(0,2)$, it is shown that the LSE is not a consistent estimator of the parameters, but converges to a ratio of stable vectors. When $\alpha \in [2,4]$, it is shown that the LSE is $n^{1-2/\alpha}$-consistent if $\alpha \in (2,4)$, $\log n$-consistent if $\alpha = 2$, and $n^{1/2}/\log n$-consistent if $\alpha = 4$, and its limiting distribution is a functional of stable processes. Our results are significantly different from those with i.i.d. noises and should warn practitioners in economics and finance of the implications, including inconsistency, of heavy-tailed errors in the presence of conditional heterogeneity.

1. INTRODUCTION

Since the seminal work by Engle (1982) and Bollerslev (1986), the G/ARCH-type models have been extensively applied in economics and finance. This paper considers the following strictly stationary autoregressive [AR(p)] process:

$$Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t, \tag{1.1}$$

where $\{\varepsilon_t : t = 1, 2, \ldots\}$ is generated by the general GARCH(1,1) process [G-GARCH(1,1)]:

$$\varepsilon_t = \eta_t h_t \text{ and } h_t^{\delta} = g(\eta_{t-1}) + c(\eta_{t-1}) h_{t-1}^{\delta}, \tag{1.2}$$

We thank Ms. Alice Cheng for her editing comments and three referees, the co-editor Giuseppe Cavaliere, and the editor Peter C.B. Phillips for their very helpful and professional comments. Zhang's research was supported by NSFC grants 11371318 and 11171074, the Fundamental Research Funds for the Central Universities, and Scientific Research Fund of Zhejiang Provincial Education Department (Y201009944). Ling's research was supported by the Hong Kong Research Grants Council (Grants HKUST641912, 603413 and FSGRF12SC12). Address correspondence to Rongmao Zhang, Zhejiang University, Hangzhou, 310027, China; e-mail: rmzhang@zju.edu.cn

 © Cambridge University Press 2014

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

where $\delta > 0$, $Pr\{h_t^\delta > 0\} = 1$, $c(0) < 1$, $c(\cdot)$ and $g(\cdot)$ are nonnegative functions, and $\{\eta_t\}$ is a sequence of i.i.d. (independent and identically distributed) symmetric white noises. The general model (1.2) was defined by He and Terasvirta (1999). It includes many models as special cases, for example, the GARCH(1,1) model of Bollerslev (1986), the absolute value GARCH(1,1) model of Taylor (1986) and Schwert (1989), the nonlinear GARCH(1,1) model of Engle (1990), the volatility switching GARCH(1,1) model of Fornari and Mele (1997), the threshold GARCH(1,1) model of Zakoian (1994), and the generalized quadratic ARCH(1,1) model of Sentana (1995).

Let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)'$ be the unknown parameter vector and its true value be $\boldsymbol{\phi}_0$. When ε_t is i.i.d. (i.e., h_t is a constant) with $\mathrm{E}\varepsilon_t^2 = \infty$, the estimated $\boldsymbol{\phi}_0$ has been well studied in the literature. Hannan and Kanter (1977) proved the least squares estimator (LSE) of $\boldsymbol{\phi}_0$ is $n^{1/\nu}$-consistent, where $\nu > \alpha$ and $\alpha \in (0, 2)$ is the tail index of ε_t (see also Knight, 1987) and n is the sample size. The same rate of convergence was obtained by An and Chen (1982) for the least absolute deviation (LAD) estimator of $\boldsymbol{\phi}_0$. The limiting distribution of the LSE was not available until Davis and Resnick (1986). Based on a point process technique and assuming that ε_t has a regular varying tail index α, they showed that the LSE converges weakly to a ratio of two stable random variables at the rate $n^{1/\alpha}L(n)$, where $L(n)$ is a slowly varying function. The asymptotic theory of the LAD and M-estimators of $\boldsymbol{\phi}_0$ was fully established by Davis, Knight, and Liu (1992). We refer to Mikosch, Gadrich, Klüppelberg, and Adler (1995) and Kokoszka and Taqqu (1996) for infinite variance ARMA and long-memory ARFIMA models. Up to date, it is well known that all the classical estimators have a faster rate of convergence when $\alpha \in (0, 2)$ than those when $\alpha \geq 2$ if the noises are i.i.d. However, when the noises are not i.i.d., the case is far from clear. A few exceptions can be found in Mikosch and Stărică (2000), Lange (2011), and the references therein.

In this paper, we show that due to the dependence, the cross-product terms related to $\varepsilon_t\varepsilon_{t-j}$, $j > 0$ do not vanish asymptotically, and the limiting distribution of the sample autocovariance $\{Y_tY_{t-k}\}$ and $\{Y_{t-k}\varepsilon_t\}$ depends on an infinite number of point processes, and thus it differs substantially from that in Davis and Resnick (1986) for i.i.d. noises. When the tail index α of G-GARCH(1,1) noise is within $(0, 2)$, $\sum_{t=1}^n Y_tY_{t-k}$ and $\sum_{t=1}^n Y_{t-k}\varepsilon_t$ have the same rate of convergence. As a result, the LSE is not a consistent estimator of the parameters, but tends to be a function of stable vectors. However, when $\alpha > 2$, the rate of convergence of $\sum_{t=1}^n Y_tY_{t-k}$ is controlled by its centralized constant whose rate is faster than that of $\sum_{t=1}^n Y_{t-k}\varepsilon_t$ (see Lemma 3.2). This leads to the consistency of the LSE. In particular, when $\alpha \in [2, 4)$, it is shown that the LSE is $n^{1-2/\alpha}$-consistent if $\alpha \in (2, 4)$, $\log n$-consistent if $\alpha = 2$, and its limiting distribution is a stable random vector. When $\alpha = 4$, the LSE is $n^{1/2}/\log n$-consistent and asymptotically normal. Our results are significantly different from those with i.i.d noises and should warn practitioners in economics and finance of the implications, including inconsistency, of heavy-tailed errors in the presence of conditional heterogeneity.

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

This paper is organized as follows. The main results are provided in Section 2 and the technical proofs are given in Section 3. Simulation results and additional proofs are reported in the online supplementary material.

2. MAIN RESULTS

Given observations $Y_1, \ldots, Y_n$, the LSE of ϕ_0 for model (1.1) is defined by

$$\widehat{\phi}_n = \left(\sum_{i=p+1}^{n} \mathbf{Y}_{i-1} \mathbf{Y}_{i-1}^T \right)^{-1} \left(\sum_{i=p+1}^{n} \mathbf{Y}_{i-1} Y_i \right), \tag{2.1}$$

where $\mathbf{Y}_i = (Y_i, Y_{i-1}, \ldots, Y_{i-p+1})$. Throughout the paper, we make the following assumptions:

H1. $\mathrm{E}\log(c(\eta_t)) < 0$.

H2. There exists a $k_0 > 0$ such that $\mathrm{E}(c(\eta_t))^{k_0} \geq 1$, $\mathrm{E}[(c(\eta_t))^{k_0} \log^+(c(\eta_t))] < \infty$, and $\mathrm{E}(g(\eta_t) + |\eta_t|^{\delta})^{k_0} < \infty$, where $\log^+(x) = \max\{0, \log(x)\}$.

H3. The density $f(x)$ of η_1 is positive in the neighborhood of zero.

Condition H1 is a necessary and sufficient condition for the existence of a stationary solution of h_t^2 (see Nelson, 1990). If condition H2 holds, then condition H1 is equivalent to $\mathrm{E}(c(\eta_1))^{\mu} < 1$ for some $\mu > 0$ (see Remark 2.9 of Basrak, Davis, and Mikosch, 2002). H2 also implies that h_t is not a constant and hence excludes the i.i.d. case. Suppose that there exists a $\iota_0 > 0$ such that $\mathrm{E}|\eta_t|^{\iota_0} = \infty$ and $\mathrm{E}|\eta_t|^{\iota} < \infty$ for all $\iota < \iota_0$, then conditions H1 and H2 are satisfied. Condition H3 can be relaxed to some certain (see Francq and Zakoïan, 2006). Furthermore, the symmetry assumption is used to simplify the proof for the case when $\alpha = 2$. Note that, when $\alpha = 2$, $\varepsilon_t \varepsilon_{t-j}$, $j \geq 1$ (see Lemma 3.1), is a regularly varying variable with tail index $\alpha/2 = 1$. If η_t is not symmetric, then a centralizing constant is required to derive the limiting distribution $\sum_{t=p}^{n} \varepsilon_t \varepsilon_{t-j}$. We first give a lemma for the tail index of $\{\varepsilon_t\}$. Its proof is similar to those of Lemmas A.1 and A.3 in Chan and Zhang (2010).

LEMMA 2.1. *Under conditions H1, H2, and H3, there exists a unique* $\alpha \in (0, \delta k_0]$ *such that* $\mathrm{E}(c(\eta_t))^{\alpha/\delta} = 1$ *and*

$$P(|\varepsilon_1| > x) \sim c_0^{(\alpha)} \mathrm{E}|\eta_1|^{\alpha} x^{-\alpha},$$

where

$$c_0^{(\alpha)} = \frac{E\left(\left[g(\eta_1) + c(\eta_1) \sigma_1^{\delta} \right]^{\alpha/\delta} - \left[c(\eta_1) \sigma_1^{\delta} \right]^{\alpha/\delta} \right)}{\alpha E\left(c(\eta_1)^{\alpha/\delta} \log^+(c(\eta_1)) \right)}.$$

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

We further make the following assumption:

H4. $\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i \neq 0$ for $|z| \leq 1$. Under condition H4, model (1.1) is stationary and has the following expansion:

$$Y_t = \sum_{l=0}^{\infty} \varphi_l \varepsilon_{t-l}, \tag{2.2}$$

and $\Sigma \equiv \mathrm{E}[\mathbf{Y}_1 \mathbf{Y}_1^T]$ exists and is positive definite when $\alpha > 2$.

Let $\xrightarrow{\mathcal{L}}$ denote convergence in distribution. We now state our main results as follows:

THEOREM 2.1. *Let α be given as in Lemma 2.1. Then under conditions H1–H4, it follows that*

(a) *when $\alpha \in (0, 2)$,*

$$\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0 \xrightarrow{\mathcal{L}} \Sigma_{\alpha/2}^{-1} \boldsymbol{Z}_{\alpha/2},$$

where $\boldsymbol{Z}_{\alpha/2}$ is a p-dimensional stable vector with index $\alpha/2$ and $\Sigma_{\alpha/2}$ is a $p \times p$ matrix whose elements are composed of stable variables with index $\alpha/2$;

(b) *when $\alpha = 2$,*

$$\log n(\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) \xrightarrow{\mathcal{L}} \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|} \right)_{p \times p}^{-1} \boldsymbol{Z}_{\alpha/2};$$

(c) *when $2 < \alpha < 4$,*

$$n^{1-2/\alpha}(\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) \xrightarrow{\mathcal{L}} \Sigma^{-1} \boldsymbol{Z}_{\alpha/2};$$

(d) *when $\alpha = 4$,*

$$(n/\log n)^{1/2}(\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) \xrightarrow{\mathcal{L}} \Sigma^{-1} \boldsymbol{N}(0, A),$$

where $A = (c_0^{(4)} \mathrm{E}\eta_1^2)(a_{ij})_{p \times p}$ is positive definite with $a_{ij} = \lim_{M \to \infty} E[u_{t,i,M} u_{t,jM}]$ and

$$u_{t,i,M} = \sum_{l=i}^{M} \varphi_{l-i} \eta_{t-l} \prod_{i=1}^{l} c^{\frac{1}{\delta}}(\eta_{t-i}) \prod_{k=l+1}^{M} c^{\frac{2}{\delta}}(\eta_{t-k}).$$

We should mention that when

$$h_t^2 = \omega + a\varepsilon_{t-1}^2 + \beta h_{t-1}^2, \tag{2.3}$$

where $\omega > 0$, $a > 0$, and $\beta > 0$, Chan and Zhang (2010) showed that for all $\alpha > 0$, the rate of convergence of the LSE in the unit-root case is of order n

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

and (c) was obtained by Lange (2011). Our theorem completely characterizes the feature of the LSE under a more general setup when $\mathrm{E}\varepsilon_t^4 = \infty$. Many empirical examples have shown the evidence that the fourth moment does not exist in economics and finance (see Mikosch and Stărică, 2000). Theorem 2.1 indicates that the statistical inference based on the LSE may be misleading and it is necessary to consider other approaches in this case. In fact, the tail trimming QMLE in Hill and Renault (2010) and modified QMLE in Lange, Rahbek, and Jensen (2011) are $\sqrt{n}$-consistent and asymptotically normal for models (1.1) and (1.2). Furthermore, if one can specify the form of G-GARCH model as (2.3), then the self-weighted QMLE in Ling (2007) and Zhu and Ling (2011) is also $\sqrt{n}$-consistent and asymptotically normal. When $\alpha < 2$, we conjecture that the M-estimators of Knight (1991), the LAD asymptotics of Phillips (1991), or the dummy-based estimators of Cavaliere and Georgiev (2013) with some kind of weights can achieve a fast rate of convergence.

The nonstandard results are mainly because of condition H2 which generates the heavy tails of volatility h_t. Let us consider the special T-CHARM model proposed by Chan, Li, Ling, and Tong (2014):

$$h_t = \sigma_1 I\{\varepsilon_{t-1} > r\} + \sigma_2 I\{\varepsilon_{t-1} \le r\}. \tag{2.4}$$

We can rewrite (2.4) as follows:

$$h_t = (\sigma_1 + \sigma_2 a_{t-1} - \sigma_1 b_{t-1}) + (b_{t-1} - a_{t-1}) h_{t-1},$$

where $a_t = I\{\eta_t \sigma_1 \le r\}$ and $b_t = I\{\eta_t \sigma_2 \le r\}$. We can see that $c(\eta_{t-1}) = b_{t-1} - a_{t-1}$ does not satisfy condition H2. Under (2.4), it is not difficult to show that

$$\sqrt{n}\left(\widehat{\phi}_n - \phi_0\right) \xrightarrow{\mathcal{L}} N\left(0, \Sigma^{-1}\Omega\Sigma^{-1}\right),$$

where $\Omega = \mathrm{E}[\mathbf{Y}_1 \mathbf{Y}_1^T h_t]$. This means that the classical statistical inference for the AR model is always valid under (2.4) specification. Otherwise, one needs to pay special attention to the tail index α of ε_t. The tail index α of ε_t is unknown in practice, but it is identical to that of Y_t. We can estimate it by using Hill's estimator.

3. TECHNICAL PROOFS

In this section, we prove Theorem 2.1. Note that

$$\widehat{\phi}_n - \phi = \left(\sum_{t=p+1}^{n} \mathbf{Y}_{t-1} \mathbf{Y}_{t-1}^T \right)^{-1} \left[\sum_{t=p+1}^{n} \mathbf{Y}_{t-1} \varepsilon_t \right].$$

The limiting distribution of $\widehat{\phi}_n - \phi$ will follow by the asymptotic behaviors of $\sum_{t=p+1}^{n} Y_{t-k}\varepsilon_t$ and $\sum_{t=p+1}^{n} Y_{t-k}Y_{t-j}$ for $1 \le k, j \le p$. Intuitively, by equation (2.2), we know that $\sum_{t=p+1}^{n} Y_{t-k}\varepsilon_t$ can be approximated by $\sum_{l=0}^{H} \varphi_l \sum_{t=p+1}^{n} \varepsilon_{t-k-l}\varepsilon_t$ and $\sum_{t=p+1}^{n} Y_{t-k}Y_{t-j}$ can be approximated by

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

$\sum_{l=0}^{H}\sum_{m=0}^{H}\varphi_l\varphi_m\sum_{t=p+1}^{n}\varepsilon_{t-k-l}\varepsilon_{t-j-m}$ as $H\to\infty$. So, by the continuous mapping theorem, it is enough to show:

(a) The union convergence of $\left(\sum_{t=p+1}^{n}\varepsilon_{t-k-l}\varepsilon_{t-j-m}, l=1,2,\ldots,H\right)$. This will be proved in Lemma 3.1 by using a point process convergence technique.

(b) The limiting distribution of $a_n^{-2}\lim_{H\to\infty}\sum_{l=0}^{H}\varphi_l\sum_{t=p+1}^{n}\varepsilon_{t-k-l}\varepsilon_t$ and $a_n^{-2}\lim_{H\to\infty}\sum_{l=0}^{H}\sum_{m=0}^{H}\varphi_l\varphi_m\sum_{t=p+1}^{n}\varepsilon_{t-k-l}\varepsilon_{t-j-m}$. This will be proved in Lemmas 3.2 and 3.3.

Denote $a_n^{(\alpha)}=\left(c_0^{(\alpha)}E|\eta_1|^{\alpha}n\right)^{1/\alpha}$. For simplicity, we write $c_0=c_0^{(\alpha)}$, $a_n=a_n^{(\alpha)}$ and assume $E\eta_1^2=1$ when $\alpha\geq 2$.

For any given integers l and H, we define two $(H+1)$-dimensional random vectors:

$$\begin{aligned}&\boldsymbol{X}_{t,l,H}=(\varepsilon_{t-l},\ldots,\varepsilon_{t-l-H})\text{ and}\\&\boldsymbol{Z}_{t,l,H}=(\varepsilon_{t-l}^2-c_n,\varepsilon_{t-l}\varepsilon_{t-l-1},\ldots,\varepsilon_{t-l}\varepsilon_{t-l-H}),\end{aligned}\tag{3.1}$$

where $c_n=0$ as $0<\alpha<2$, $c_n=E\varepsilon_1^2I_{(|\varepsilon_1|\leq\sqrt{c_0n})}=c_0\log n$ as $\alpha=2$, and $c_n=E\varepsilon_1^2$ as $\alpha>2$. Under the assumptions of Theorem 2.1, $\{\boldsymbol{X}_{t,l,H},t\geq 1\}$ is a varying regular random vector sequence with index α. By Theorem 2.8 of Davis and Mikosch (1998) (see also Theorem 3.1 of Mikosch and Stărică, 2000), there exists a Poisson process $\sum_{i=1}^{\infty}\delta_{P_i}$ defined on $\mathbb{R}_+$ with intensity measure $v(dy)=\Upsilon\alpha y^{-\alpha-1}\,dy$ and a sequence of i.i.d. point processes $\left\{\sum_{j=1}^{\infty}\delta_{\boldsymbol{Q}_{ij,H}}\right\}$ with distribution Q_H, which depends on α, such that

$$\sum_{t=1}^{n}\delta_{\boldsymbol{X}_{t,l,H}/a_n}\xrightarrow{\mathcal{L}}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\delta_{P_i\boldsymbol{Q}_{ij,H}},$$

where $\boldsymbol{Q}_{ij,H}=\left(Q_{ij}^0,Q_{ij}^1,\ldots,Q_{ij}^H\right)$, $\left\{\sum_{j=1}^{\infty}\delta_{\boldsymbol{Q}_{ij,H}}\right\}$ is independent of the process $\{P_i\}$, and Υ and Q_H are similarly defined as in Davis and Mikosch (1998). Thus, by Proposition 3.3 of Davis and Mikosch (1998) and the continuous mapping theorem, we have

$$\begin{aligned}\frac{1}{a_n^2}\sum_{t=1}^{n}\boldsymbol{Z}_{t,j,H}&\xrightarrow{\mathcal{L}}\left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_i^2(Q_{ij}^0)^2,\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_i^2Q_{ij}^0Q_{ij}^1,\ldots,\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_i^2Q_{ij}^0Q_{ij}^H\right)\\&=:\ (S^0(\alpha),S^1(\alpha),\ldots,S^H(\alpha))=\boldsymbol{S}_H(\alpha),\end{aligned}\tag{3.2}$$

if $0<\alpha<2$ and

$$\begin{aligned}\frac{1}{a_n^2}\sum_{t=1}^{n}\boldsymbol{Z}_{t,j,H}&\xrightarrow{\mathcal{L}}\left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_i^2Q_{ij}^0Q_{ij}^hI(|P_i^2Q_{ij}^0Q_{ij}^h|>0)\right.\\&\qquad\left.-\int_{(\boldsymbol{x}:|x_0x_h|>0)}x_0x_h\,d\mu(\boldsymbol{x})\right)_{h=0,\ldots,H}\\&=:\ (S^0(\alpha),S^1(\alpha),\ldots,S^H(\alpha))=\boldsymbol{S}_H(\alpha),\end{aligned}\tag{3.3}$$

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

if $2 \le \alpha < 4$, where $\mu(\cdot) = \lim_{n\to\infty} nP(\boldsymbol{Z}_{t,j,H}/n^{\frac{2}{\alpha}} \in \cdot)$ is a measure on $\mathbb{R}^{H+1}$ and $\boldsymbol{x} = (x_0, x_1, \ldots, x_H)$. $\boldsymbol{S}_H(\alpha)$ is a stable random vector with index $\alpha/2$.

LEMMA 3.1. *Under the conditions of Theorem 2.1, for any positive integer K and H,*

$$\frac{1}{a_n^2}\left(\sum_{t=1}^{n} \boldsymbol{Z}_{t,0,H}, \sum_{t=1}^{n} \boldsymbol{Z}_{t,1,H}, \ldots, \sum_{t=1}^{n} \boldsymbol{Z}_{t,K,H}\right) \xrightarrow{\mathcal{L}} (\boldsymbol{S}_H(\alpha), \boldsymbol{S}_H(\alpha), \ldots, \boldsymbol{S}_H(\alpha))_{1\times(K+1)},$$

as $n \to \infty$, where $0 < \alpha < 4$ and $\boldsymbol{S}_H(\alpha)$ is defined in (3.2) and (3.3).

Proof. We first have the expansion:

$$h_t^\delta = \sum_{i=1}^{h}\prod_{j=1}^{i-1} c(\eta_{t-j})g(\eta_{t-i}) + \prod_{i=1}^{h} c(\eta_{t-i})h_{t-h}^\delta, \tag{3.4}$$

where $\prod_{j=1}^{0} c(\eta_{t-j}) = 1$. Furthermore, we have

$$|\varepsilon_t\varepsilon_{t-h}|^\delta = |\eta_t|^\delta \sum_{i=1}^{h}\prod_{j=1}^{i-1} c(\eta_{t-j})g(\eta_{t-i})|\eta_{t-h}|^\delta h_{t-h}^\delta + |\eta_t|^\delta \prod_{i=1}^{h} c(\eta_{t-i})|\eta_{t-h}|^\delta h_{t-h}^{2\delta}.$$

By Lemma 2.1 and Proposition 3 of Breiman (1965), we have that when y is large enough,

$$P\left\{|\eta_t|^\delta \sum_{i=1}^{h}\prod_{j=1}^{i-1} c(\eta_{t-j})g(\eta_{t-i})|\eta_{t-h}|^\delta h_{t-h}^\delta > y\right\}$$
$$= c(\alpha)E\left\{|\eta_t|^\delta \sum_{i=1}^{h}\prod_{j=1}^{i-1} c(\eta_{t-j})g(\eta_{t-i})|\eta_{t-h}|^\delta\right\}^{\alpha/\delta} y^{-\alpha/\delta} =: c_0(\alpha)y^{-\alpha/\delta}$$

and

$$P\left\{|\eta_t|^\delta \prod_{i=1}^{h} c(\eta_{t-i})|\eta_{t-h}|^\delta h_{t-h}^{2\delta} > y\right\}$$
$$= c(\alpha)\mathrm{E}[\eta_0^2 c(\eta_0)]^{\frac{\alpha}{2\delta}}[\mathrm{E}c(\eta_0)^{\frac{\alpha}{2\delta}}]^{h-1}[\mathrm{E}|\eta_0|^{\frac{\alpha}{\delta}}]y^{-\frac{\alpha}{2\delta}},$$
$$:= c_1(\alpha)\rho^h y^{-\frac{\alpha}{2\delta}}, \text{ for some } 0 < \rho < 1.$$

Combining the previous two equations, we have, for any given h and a large enough x,

$$P(|\varepsilon_t\varepsilon_{t-h}| > x) = P(|\varepsilon_t\varepsilon_{t-h}|^\delta > x^\delta) \simeq c_1(\alpha)\rho^h x^{-\alpha/2}. \tag{3.5}$$

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

For any $0 \le i < j \le K$, (3.5) yields that

$$\frac{1}{a_n^2}\sum_{t=1}^{n}(\boldsymbol{Z}_{t,i,H} - \boldsymbol{Z}_{t,j,H}) \xrightarrow{p} 0.$$

Thus, by (3.2) and (3.3), we can show that the conclusion holds. ■

Let $S^i(\alpha)$ be given as in (3.2) for $\alpha < 2$ and (3.3) for $2 \le \alpha < 4$ and define

$$Z_{\alpha/2}^{(k)}(Y) = \left(\sum_{l=0}^{\infty}\varphi_l\varphi_{l+k}\right)S^0(\alpha) + \sum_{h=1}^{\infty}\left(\sum_{l=0}^{\infty}\varphi_l\varphi_{l+h+k} + \sum_{l=0\vee(k-h)}^{\infty}\varphi_l\varphi_{l+h-k}\right)S^h(\alpha)$$

and $Z_{\alpha/2}^{(k)}(\varepsilon) = \sum_{h=k+1}^{\infty}\varphi_{h-k-1}S^h(\alpha)$. We have the following lemma.

LEMMA 3.2. *Under the conditions of Theorem 2.1, we have for* $0 < \alpha < 4$,

$$\frac{1}{a_n^2}\left\{\sum_{t=1}^{n}\left[Y_tY_{t-k} - \left(\sum_{l=0}^{\infty}\varphi_l\varphi_{l+k}\right)c_n\right], \sum_{t=1}^{n}Y_{t-1-k}\varepsilon_t, 0 \le k \le p\right\}$$
$$\xrightarrow{\mathcal{L}} \left\{Z_{\alpha/2}^{(k)}(Y), Z_{\alpha/2}^{(k)}(\varepsilon), 0 \le k \le p\right\},$$

where c_n is defined as in (3.1).

Proof. By Cramér–Wold's lemma, it is sufficient to show that, for any real number $f_k, g_k, 0 \le k \le p$,

$$\frac{1}{a_n^2}\left\{\sum_{k=0}^{p}\sum_{t=1}^{n}f_k\left[Y_tY_{t-k} - \left(\sum_{l=0}^{\infty}\varphi_l\varphi_{l+k}\right)c_n\right] + \sum_{k=0}^{p}\sum_{t=1}^{n}g_kY_{t-1-k}\varepsilon_t\right\}$$
$$\xrightarrow{\mathcal{L}} \sum_{k=0}^{p}f_kZ_{\alpha/2}^{(k)}(Y) + \sum_{k=0}^{p}g_kZ_{\alpha/2}^{(k)}(\varepsilon). \tag{3.6}$$

Let $d_j = \sum_{k=0}^{j\wedge p}f_k\varphi_{j-k}$ and $e_j = \sum_{k=0}^{j\wedge p}g_k\varphi_{j-k}$. By (2.2), we have

$$\sum_{k=0}^{p}f_k\left[Y_tY_{t-k} - \left(\sum_{l=0}^{\infty}\varphi_l\varphi_{l+k}\right)c_n\right] + \sum_{k=0}^{p}g_kY_{t-1-k}\varepsilon_t$$
$$= \sum_{j=-\infty}^{t}\sum_{h=-\infty,h\ne j}^{t}d_{t-h}\varphi_{t-j}\varepsilon_h\varepsilon_j + \sum_{h=-\infty,h\ne j}^{t-1}e_{t-1-h}\varepsilon_h\varepsilon_t + \sum_{j=0}^{\infty}d_j\varphi_j(\varepsilon_{t-j}^2 - c_n)$$
$$= \left\{\sum_{j=0}^{K}d_j\varphi_j(\varepsilon_{t-j}^2 - c_n) + \sum_{h=1}^{H}\sum_{j=0}^{K}(d_{j+h}\varphi_j + d_j\varphi_{j+h})\varepsilon_{t-j-h}\varepsilon_{t-j} + \sum_{h=1}^{H}e_{h-1}\varepsilon_{t-h}\varepsilon_t\right\}$$

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

$$+\left\{\left(\sum_{h=H+1}^{\infty}\sum_{j=0}^{\infty}+\sum_{h=1}^{H}\sum_{j=K+1}^{\infty}\right)\left(d_{j+h}\varphi_j+d_j\varphi_{j+h}\right)\varepsilon_{t-j-h}\varepsilon_{t-j}\right.$$

$$\left.+\sum_{j=K+1}^{\infty}d_j\varphi_j(\varepsilon_{t-j}^2-c_n)+\sum_{h=H+1}^{\infty}e_{h-1}\varepsilon_{t-h}\varepsilon_t\right\}$$

$$=: I_t^1(H,K)+I_t^2(H,K).$$

Thus, by Lemma 3.1, along the lines of the proof for Theorem 3.1 of Zhang, Sin, and Ling (2014), we can show that

$$\frac{1}{a_n^2}\sum_{t=1}^{n}I_t^1(H,K)\xrightarrow{\mathcal{L}}\sum_{l=0}^{K}d_l\varphi_lS^0(\alpha)+\sum_{h=1}^{H}\left[\sum_{l=0}^{K}(d_{h+l}\varphi_l+d_l\varphi_{h+l})+e_{h-1}\right]S^h(\alpha)$$

$$\xrightarrow{p}\sum_{k=0}^{p}f_kZ_{\alpha/2}^{(k)}(Y)+\sum_{k=0}^{p}g_kZ_{\alpha/2}^{(k)}(\varepsilon)$$

by letting $H\to\infty$ and $K\to\infty$ and for any $\delta>0$

$$\lim_{H,K\to\infty}\lim_{n\to\infty}P\left\{\left|\frac{1}{a_n^2}\sum_{t=1}^{n}I_t^2(H,K)\right|>\delta\right\}=0.$$

Combining the previous equations, we can see that (3.6) holds. ■

LEMMA 3.3. *Let A be the matrix given in Theorem 2.1. If $\alpha=4$, then*

$$\frac{1}{\sqrt{n\log n}}\sum_{t=1}^{n}\{Y_{t-k}\varepsilon_t,\,1\le k\le p\}\xrightarrow{\mathcal{L}}\boldsymbol{N}(0,A). \tag{3.7}$$

Proof. Since the proof is extremely technical, we put it in the supplementary material. ■

The proof of Theorem 2.1 (i) When $0<\alpha<2$,

$$\widehat{\boldsymbol{\phi}}_n-\boldsymbol{\phi}=\left(\frac{1}{a_n^2}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}^T\right)^{-1}\left[\frac{1}{a_n^2}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\varepsilon_t\right].$$

By Lemma 3.2 and continuous mapping theorem, (a) holds. (ii) When $\alpha=2$,

$$(\log n)(\widehat{\boldsymbol{\phi}}_n-\boldsymbol{\phi})=\left(\frac{1}{c_0n\log n}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}^T\right)^{-1}\left[\frac{1}{c_0n}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\varepsilon_t\right].$$

By Lemma 3.2, (b) holds. (iii) When $2<\alpha<4$,

$$\frac{n}{a_n^2}(\widehat{\boldsymbol{\phi}}_n-\boldsymbol{\phi})=\left(\frac{1}{n}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}^T\right)^{-1}\left[\frac{1}{a_n^2}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\varepsilon_t\right].$$

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

By Lemma 3.2, (c) holds. (iv) When $\alpha = 4$,

$$\frac{n}{\sqrt{n\log n}}(\widehat{\phi}_n - \phi) = \left(\frac{1}{n}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\mathbf{Y}_{t-1}^T\right)^{-1}\left[\frac{1}{\sqrt{n\log n}}\sum_{t=p+1}^{n}\mathbf{Y}_{t-1}\varepsilon_t\right].$$

By the ergodic theorem and Lemma 3.3, (d) holds. This completes the proof. ■

REFERENCES

An, H.Z. & Z.G. Chen (1982) On convergence of LAD estimates in autoregression with infinite variance. *Journal of Multivariate Analysis* 12, 335–345.

Basrak, B., R.A. Davis, & T. Mikosch (2002) Regular variation of GARCH processes. *Stochastic Processes and Their Applications* 99, 95–115.

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

Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory of Probability and Its Applications* 10, 323–331.

Cavaliere, G. & I. Georgiev (2013) Exploiting infinite variance through dummy variables in nonstationary autoregressions. *Econometric Theory* 29, 1162–1195.

Chan, K.S., D. Li, S. Ling, & H. Tong (2014) On Conditionally Heteroscedastic AR Models with Thresholds. Statistica Sinica 24, 625–652.

Chan, N.H. & R.M. Zhang (2010) Inference for unit-root models with infinite variance GARCH errors. *Statistica Sinica* 20, 1363–1393.

Davis, R.A., K. Knight, & J. Liu (1992) M-estimation for autoregressions with infinite variance. *Stochastic Processes and Their Applications* 40, 145–180.

Davis, R.A. & T. Mikosch (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics* 26, 2049–2080.

Davis, R.A. & S. Resnick (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* 14, 533–558.

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica* 50, 987–1008.

Engle, R.F. (1990) Discussion: Stock market volatility and the crash of 1987. *Review of Financial Studies* 3, 103–106.

Fornari, F. & A. Mele (1997) Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics* 12, 1779–1801.

Francq, C. & J.M. Zakoïan (2006) Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* 22, 815–834.

Hannan, E.J. & M. Kanter (1977) Autoregressive processes with infinite variance. *Journal of Applied Probability* 14, 411–415.

He, C. & T. Terasvirta (1999) Properties of moments of a family of GARCH processes. *Journal of Econometrics* 92, 173–192.

Hill, J.B. & E. Renault (2010) Generalized Method of Moments with Tail Trimming. Working paper, University of North Carolina.

Knight, K. (1987) Rate of convergence of centred estimates of autoregressive parameters for infinite variance autoregressions. *Journal of Time Series Analysis* 8, 51–60.

Knight, K. (1991) Limit theory for M-estimates in an integrated infinite variance. *Econometric Theory* 7, 200–212.

Kokoszka, P.S. & M.S. Taqqu (1996) Parameter estimation for infinite variance fractional ARIMA. *Annals of Statistics* 24, 1880–1913.

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press

Lange, T. (2011) Tail behavior and OLS estimation in AR-GARCH models. *Statistica Sinica* 21, 1191–1200.

Lange, T., A. Rahbek, & S.T. Jensen (2011) Estimation and asymptotic inference in the AR-ARCH model. *Econometric Reviews* 30, 129–153.

Ling, S. (2007) Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. *Journal of Econometrics* 140, 849–873.

Mikosch, T., T. Gadrich, C. Klüppelberg, & R.J. Adler (1995) Parameter estimation for ARMA models with infinite variance innovations. *Annals of Statistics* 23, 305–326.

Mikosch, T. & C. Stărică (2000) Limit theory for the sample autocorrelations and extremes of a GARCH (1,1) process. *Annals of Statistics* 28, 1427–1451.

Nelson, D.B. (1990) Stationary and persistence in GARCH (1,1) models. *Econometric Theory* 6, 318–334.

Phillips, P.C.B. (1991) A shortcut to LAD asymptotics. *Econometric Theory* 7, 450–463.

Schwert, G.W. (1989) Why does stock market volatility change over time? *Journal of Finance* 45, 1129–1155.

Sentana, E. (1995) Quadratic ARCH models. *Review of Economic Studies* 62, 639–661.

Taylor, S. (1986) *Modelling Financial Time Series*. Wiley.

Zakoian, J.M. (1994) Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* 18, 931–955.

Zhang, R.M., C.Y. Sin, & S. Ling (2014) On functional limits of short- and long-memory linear processes with GARCH(1,1) noises. *Stochastic Processes and Their Applications*, to appear.

Zhu, K. & S. Ling (2011) Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models. *Annals of Statistics* 39, 2131–2163.

https://doi.org/10.1017/S0266466614000632 Published online by Cambridge University Press