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 \(\alpha\) 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 \(\alpha\) 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 GARCH-type models have been extensively applied in economics and finance. This paper considers the following strictly stationary autoregressive [AR\((p)\)] process:

\[
Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + \epsilon_t, \tag{1.1}
\]

where \(\{\epsilon_t : t = 1, 2, \ldots\}\) is generated by the general GARCH\((1,1)\) process

\[
\begin{align*}
\epsilon_t & = \eta_t h_t, \\
h_{t}^1 & = g(\eta_{t-1}) + c(\eta_{t-1}) h_{t-1}^1,
\end{align*} \tag{1.2}
\]

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).
where $\delta > 0$, $Pr[h_{ij}^\delta > 0] = 1$, $c(0) < 1$, $c(\cdot)$ 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 $\phi = (\phi_1, \ldots, \phi_p)'$ be the unknown parameter vector and its true value be $\phi_0$. When $\epsilon_t$ is i.i.d. (i.e., $h_t$ is a constant) with $E\epsilon_t^2 = \infty$, the estimated $\hat{\phi}_n$ has been well studied in the literature. Hannan and Kanter (1977) proved the least squares estimator (LSE) of $\phi_0$ is $n^{1/\nu}$-consistent, where $\nu > \alpha$ and $\alpha \in (0, 2)$ is the tail index of $\epsilon_t$ (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 $\phi_0$. The limiting distribution of the LSE was not available until Davis and Resnick (1986). Based on a point process technique and assuming that $\epsilon_t$ has a regular varying tail index $\alpha$, they showed that under condition $\alpha \in (0, 2)$, the LSE converges weakly to a ratio of two stable random variables at the rate $n^{1/\nu} L(n)$, where $L(n)$ is a slowly varying function. The asymptotic theory of the LAD and M-estimators of $\phi_0$ 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 Stáricá (2000), Lange (2011), and the references therein.

In this paper, we show that due to the dependence, the cross-product terms related to $\epsilon_t \epsilon_{t-j}$, $j > 0$ do not vanish asymptotically, and the limiting distribution of the sample autocovariance $\{Y_t Y_{t-k}\}$ and $\{Y_{t-k} \epsilon_t\}$ 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 $\alpha$ of G-GARCH(1,1) noise is within $(0, 2)$, $\sum_{t=1}^n Y_t Y_{t-k}$ and $\sum_{t=1}^n Y_{t-k} \epsilon_t$ 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 $\sum_{t=1}^n Y_t Y_{t-k}$ is controlled by its centralized constant whose rate is faster than that of $\sum_{t=1}^n Y_{t-k} \epsilon_t$ (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.
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 $\phi_0$ for model (1.1) is defined by

$$
\hat{\phi}_n = \left( \sum_{i=p+1}^{n} Y_{i-1}Y_{i-1}' \right)^{-1} \left( \sum_{i=p+1}^{n} Y_{i-1}Y_{i} \right), \tag{2.1}
$$

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

H1. $\log(c(\eta_t)) < 0$.

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

H3. The density $f(x)$ of $\eta_1$ 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 $E(c(\eta_1))^{\mu} < 1$ for some $\mu > 0$ (see Remark 2.9 of Basrak, Davis, and Mikosch, 2002). H3 also implies that $h_t$ is not a constant and hence excludes the i.i.d. case. Suppose that there exists a $t_0 > 0$ such that $E|\eta_t|^{k_0} = \infty$ and $E|\eta_t|^{\mu} < \infty$ for all $t < t_0$, then conditions H1 and H2 are satisfied. Condition H3 is a mixing condition for model (1.2) and 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 $\alpha = 2$, $e_{i-1-j}$, $j \geq 1$ (see Lemma 3.1), is a regularly varying variable with tail index $\alpha/2 = 1$. If $\eta_t$ is not symmetric, then a centralizing constant is required to derive the limiting distribution $\sum_{i=p}^{n} e_{i-1-j}$. We first give a lemma for the tail index of $\{e_{i}\}$. 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 $E(c(\eta_1))^{\alpha/\delta} = 1$ and

$$
P(|e_1| > x) \sim c_0^{(\alpha)} E|\eta_1|^{\alpha} x^{-\alpha},$$

where

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

\[
Y_t = \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}, \tag{2.2}
\]

and \( \Sigma \equiv E[Y_i Y_i'] \) exists and is positive definite when \( \alpha > 2 \).

Let \( \overset{L}{\longrightarrow} \) denote convergence in distribution. We now state our main results as follows:

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

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

\[
\hat{\phi}_n - \phi_0 \overset{L}{\longrightarrow} \Sigma_{-1/2} Z_{\alpha/2},
\]

where \( Z_{\alpha/2} \) is a \( p \)-dimensional stable vector with index \( \alpha/2 \) and \( \Sigma_{-1/2} \) is a \( p \times p \) matrix whose elements are composed of stable variables with index \( \alpha/2 \);

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

\[
\log(n) (\hat{\phi}_n - \phi_0) \overset{L}{\longrightarrow} \left( \sum_{i=0}^{\infty} \phi_i (\varepsilon_{t+i} - \varepsilon_t) \right)^{-1} \Sigma_{-1/2} Z_{\alpha/2};
\]

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

\[
n^{1-2/\alpha} (\hat{\phi}_n - \phi_0) \overset{L}{\longrightarrow} \Sigma_{-1} Z_{\alpha/2};
\]

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

\[
(n/\log n)^{1/2} (\hat{\phi}_n - \phi_0) \overset{L}{\longrightarrow} \Sigma_{-1} N(0, A),
\]

where \( A = (c_n^0 \eta_1^{-2}) (a_{ij})_{p \times p} \) is positive definite with \( a_{ij} = \lim_{M \to \infty} E[n_{t,i,M} n_{t,j,M}] \) and

\[
u_{t,i,M} = \sum_{l=i}^{M} \phi_l \varepsilon_{t-l} \prod_{i=1}^{l} \varepsilon_{t-i} \prod_{k=l+1}^{M} \varepsilon_{t-k}.
\]

We should mention that when

\[
h^2_t = \omega + a \varepsilon^2_{t-1} + \beta h^2_{t-1}, \tag{2.3}
\]

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 \( n \)
and (c) was obtained by Lange (2011). Our theorem completely characterizes the feature of the LSE under a more general setup when $E\varepsilon_t^4 = \infty$. 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 (2012):

$$h_t = \sigma_1 I[\varepsilon_{t-1} > r] + \sigma_2 I[\varepsilon_{t-1} \leq r].$$

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

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

where $\Omega = E[Y_t Y_t^T h_t]$. 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 $\alpha$ of $\varepsilon_t$. The tail index $\alpha$ of $\varepsilon_t$ 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

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

The limiting distribution of $\hat{\phi}_n - \phi$ 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 $1 \leq k, j \leq p$. Intuitively, by equation (2.2), we know that $\sum_{t=p+1}^{n} Y_{t-k} \varepsilon_t$ can be approximated by
If \( \alpha > 0 \), then \( \sum_{i=0}^{H} \sum_{j=m}^{n} \epsilon_{i-k} \epsilon_{j-m} \) can be approximated by \( \sum_{i=0}^{H} \sum_{m=0}^{n} \epsilon_{i-k} \epsilon_{j-m} \) as \( H \to \infty \). So, by the continuous mapping theorem, it is enough to show:

(a) The union convergence of \( \left\{ \sum_{i=0}^{H} \epsilon_{i-k} \epsilon_{j-m}, l = 1, 2, \ldots, H \right\} \). This will be proved in Lemma 3.1 by using a point process convergence technique.

(b) The limiting distribution of \( a_n^{-1} \lim_{H \to \infty} \sum_{i=0}^{H} \epsilon_{i-k} \epsilon_{j-m} \) and \( a_n^{-2} \lim_{H \to \infty} \sum_{i=0}^{H} \sum_{m=0}^{n} \epsilon_{i-k} \epsilon_{j-m} \). This will be proved in Lemmas 3.2 and 3.3.

Denote \( a_n^{(\alpha)} = \left( c_0^{(\alpha)} |\eta|^{1/n} \right)^{1/\alpha} \). 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:

\[
X_{l,H} = (\epsilon_{l-1}, \ldots, \epsilon_{l-H}) \quad \text{and} \quad Z_{l,H} = (\epsilon_{l-1} \epsilon_{l-2}, \ldots, \epsilon_{l-H}),
\]

where \( c_n = 0 \) as \( 0 < \alpha < 2 \), \( c_n = E \epsilon_1^2 I_{[|\eta| \leq \sqrt{n}]} = c_0 \log n \) as \( \alpha = 2 \), and \( c_n = E \epsilon_1^2 \) as \( \alpha > 2 \). Under the assumptions of Theorem 2.1, \( \{X_{l,H}, l \geq 1\} \) is a varying regular random vector sequence with index \( \alpha \). By Theorem 2.8 of Davis and Mikosch (1998) (see also Theorem 3.1 of Mikosch and Stărică, 2000), there exists a Poisson process \( \sum_{i=0}^{\infty} \delta_{P_i} \) defined on \( \mathbb{R}_+ \) with intensity measure \( \nu(dy) = \Upsilon \alpha y^{\alpha-1} dy \) and a sequence of i.i.d. point processes \( \left\{ \sum_{j=1}^{\infty} \delta_{Q_{ij}}, \sum_{i=1}^{\infty} \delta_{Q_{ij}} \right\} \) with distribution \( Q_H \),

which depends on \( \alpha \), such that

\[
\frac{1}{a_n} \sum_{i=1}^{n} \sum_{j=1}^{\infty} \delta_{Q_{ij}},
\]

where \( Q_{ij} = (Q_{ij}^0, Q_{ij}^1, \ldots, Q_{ij}^H) \), \( \left\{ \sum_{i=1}^{\infty} \delta_{Q_{ij}} \right\} \) is independent of the process \( \{P_i\} \), and \( \Upsilon \) 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

\[
\frac{1}{a_n} \sum_{i=1}^{n} Z_{l,H} \xrightarrow{\mathcal{L}} \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 (Q_{ij}^0)^2, \sum_{i=1}^{\infty} P_i^2 Q_{ij}^0 Q_{ij}^1, \ldots, \sum_{i=1}^{\infty} P_i^2 Q_{ij}^0, Q_{ij}^H \right)
\]

(3.2)

if \( 0 < \alpha < 2 \) and

\[
\frac{1}{a_n} \sum_{i=1}^{n} Z_{l,H} \xrightarrow{\mathcal{L}} \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P_i^2 Q_{ij}^0 Q_{ij}^1 I(|P_i^2 Q_{ij}^0 Q_{ij}^1| > 0) \right.
\]

\[
\left. - \int_{\{x : |x_0 x_0| > 0\}} x_0 x_0 d \mu(x) \right)_{h=0,\ldots,H}
\]

(3.3)
if $2 \leq \alpha < 4$, where $\mu() = \lim_{\alpha \to \infty} nP(Z_{t,j,H}/n^{\frac{\alpha}{2}} \in \cdot)$ is a measure on $\mathbb{R}^{H+1}$ and $x = (x_0, x_1, \ldots, x_H)$. $S_H(\alpha)$ 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 $H$, 

$$
\frac{1}{n^2} \left( \sum_{i=1}^{n} Z_{i,0;H;1} \sum_{i=1}^{n} Z_{i,1;H;1}, \ldots, \sum_{i=1}^{n} Z_{i,K;H} \right)^2 \to (S_H(\alpha), S_H(\alpha), \ldots, S_H(\alpha))_{1 \times (K+1)},
$$

as $n \to \infty$, where $0 < \alpha < 4$ and $S_H(\alpha)$ is defined in (3.2) and (3.3).

**Proof.** We first have the expansion:

$$
h^\delta = \sum_{i=1}^{h} \prod_{j=1}^{i-1} c(\eta_{i-j}) g(\eta_{i-i}) + \sum_{i=1}^{h} c(\eta_{i-i}) h^\delta_{i-i-h},
$$

(3.4)

where $\prod_{i=1}^{h} c(\eta_{i-i}) = 1$. Furthermore, we have

$$
|e_{t-t-h}^\delta| = |\eta_1^\delta \prod_{j=1}^{h} c(\eta_{i-j}) g(\eta_{i-i}) | |\eta_{i-i-h}^\delta h^\delta_{i-i-h} + |\eta_1^\delta \prod_{i=1}^{h} c(\eta_{i-i}) | |\eta_{i-i-h}^\delta h^\delta_{i-i-h}.
$$

By Lemma 2.1 and Proposition 3 of Breiman (1965), we have that

$$
\lim_{y \to \infty} P \left\{ |\eta_1^\delta \prod_{i=1}^{h} c(\eta_{i-i}) g(\eta_{i-i}) | |\eta_{i-i-h}^\delta h^\delta_{i-i-h} > y \right\}
= c(\alpha) E \left\{ \prod_{i=1}^{h} c(\eta_{i-i}) g(\eta_{i-i}) | |\eta_{i-i-h}^\delta \right\}^{\alpha/\delta} y^{-\alpha/\delta} = c_0(\alpha) y^{-\alpha/\delta}
$$

and

$$
\lim_{y \to \infty} P \left\{ |\eta_1^\delta \prod_{i=1}^{h} c(\eta_{i-i}) g(\eta_{i-i}) | |\eta_{i-i-h}^\delta h^\delta_{i-i-h} > y \right\}
= c(\alpha) E[\eta_0^2 c(\eta_0) | |\eta_0| |E| \eta_0| |E| \eta_0| |E|]^{h-1} E| \eta_0| |E| \eta_0| |E| y^{-\rho}= c_1(\alpha) \rho y^{-\rho},
$$

for some $0 < \rho < 1$.

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

$$
P(|e_{t-t-h}| > x) = P(|e_{t-t-h}| > x^\delta) \to c_1(\alpha) \rho x^{-\alpha/2}.
$$

(3.5)

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

$$
\frac{1}{a_n^i} \sum_{i=1}^{n} (Z_{t,i,H} - Z_{i,j,H}) \to 0.
$$
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 $2 \leq \alpha < 4$ and define

$$Z_{\alpha/2}^{(k)}(Y) = \left( \sum_{l=0}^{\infty} \phi_l \psi_{l+k} \right) S_{\alpha}^{(k)} + \sum_{h=1}^{\infty} \left( \sum_{l=0}^{h-1} \phi_l \psi_{l+h+k} + \sum_{l=h}^{\infty} \phi_l \psi_{l+h-k} \right) S_{\alpha}^{(h)}(\alpha).$$

and $Z_{\alpha/2}^{(k)}(c) = \sum_{h=k+1}^{\infty} \phi_h \psi_{h-1} S_{\alpha}^{(h)}(\alpha)$. We have the following lemma.

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

$$\frac{1}{\alpha} \left\{ \sum_{i=1}^{n} \left[ Y_i Y_{i-k} - \left( \sum_{l=0}^{\infty} \phi_l \psi_{l+k} \right) c_n \right] + \sum_{i=1}^{n} g_k Y_{i-1-k} e_i \right\} \overset{L}{\longrightarrow} \left\{ Z_{\alpha/2}^{(k)}(Y), Z_{\alpha/2}^{(k)}(c), 0 \leq k \leq p \right\},$$

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 $f_k$, $g_k$, $0 \leq k \leq p$,

$$\frac{1}{\alpha^2} \left\{ \sum_{i=1}^{n} f_k Y_{i-k} - \left( \sum_{l=0}^{\infty} \phi_l \psi_{l+k} \right) c_n \right\} + \sum_{i=1}^{n} g_k Y_{i-1-k} e_i \overset{L}{\longrightarrow} \sum_{k=0}^{p} f_k Z_{\alpha/2}^{(k)}(Y) + \sum_{k=0}^{p} g_k Z_{\alpha/2}^{(k)}(c). \quad (3.6)$$

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

$$\sum_{k=0}^{p} f_k \left( Y_i Y_{i-k} - \left( \sum_{l=0}^{\infty} \phi_l \psi_{l+k} \right) c_n \right) + \sum_{k=0}^{p} g_k Y_{i-1-k} e_i$$

$$= \sum_{j=-\infty}^{1} \sum_{h=-\infty, h \neq j} \sum_{j=-\infty}^{1} d_j \psi_{j-h} e_{j-h} e_j + \sum_{j=-\infty}^{H} \sum_{h=-\infty, h \neq j} e_{i-1-h} e_{h-l} e_j + \sum_{j=0}^{K} d_j \psi_{j} (e_{i-j}^2 - c_n)$$

$$= \sum_{j=0}^{K} d_j \psi_{j} (e_{i-j}^2 - c_n) + \sum_{h=1}^{H} \sum_{j=0}^{K} (d_j + d_j \psi_{j+h}) e_{i-j} e_{h-l} e_j + \sum_{h=1}^{H} e_{h-1} e_{h-l} e_j$$

$$+ \left( \sum_{h=H+1}^{\infty} \sum_{j=0}^{K} \sum_{h=1}^{H} \sum_{j=K+1}^{\infty} (d_j + d_j \psi_{j+h} + d_j \psi_{j+h}) e_{i-j} e_{h-l} e_j + \sum_{h=H+1}^{\infty} e_{h-1} e_{h-l} e_j \right)$$

$$= l_1^1 (H, K) + l_1^2 (H, K).$$
Thus, by Lemma 3.1, along the lines of the proof for Theorem 3.1 of Zhang, Sin, and Ling (2013), we can show that

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

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

\[
\frac{1}{\sqrt{n \log n}} \sum_{t=1}^{n} (Y_{t-k} - \{1 \leq k \leq p\}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, A).
\] (3.7)

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

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

\[
\hat{\phi}_n - \phi = \left( \frac{1}{a_n^2} \sum_{t=p+1}^{n} Y_{t-1}Y_{t-1}^T \right)^{-1} \left[ \frac{1}{a_n^2} \sum_{t=p+1}^{n} Y_{t-1} \varepsilon_t \right].
\]

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

\[
(\log n)(\hat{\phi}_n - \phi) = \left( \frac{1}{\log n} \sum_{t=p+1}^{n} Y_{t-1}Y_{t-1}^T \right)^{-1} \left[ \frac{1}{\log n} \sum_{t=p+1}^{n} Y_{t-1} \varepsilon_t \right].
\]

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

\[
\frac{n}{a_n^2} (\hat{\phi}_n - \phi) = \left( \frac{1}{n} \sum_{t=p+1}^{n} Y_{t-1}Y_{t-1}^T \right)^{-1} \left[ \frac{1}{n} \sum_{t=p+1}^{n} Y_{t-1} \varepsilon_t \right].
\]
By Lemma 3.2, (c) holds. (iv) When $\alpha = 4$,
\[
\frac{1}{n \log n} (\hat{\phi}_n - \phi) = \left( \frac{1}{n} \sum_{t=p+1}^{n} Y_{t-1} Y_{t-1}^T \right)^{-1} \left[ \frac{1}{\sqrt{n \log n}} \sum_{t=p+1}^{n} Y_{t-1} \varepsilon_t \right].
\]

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

REFERENCES


ASYMPTOTIC INFERENCE FOR AR MODELS

Author Query Form

Journal: Econometric Theory

Article: S0266466614000632

Dear Author, During the preparation of your manuscript for publication, the questions listed below have arisen. The numbers pertain to the numbers in the margin of the proof. Please attend to these matters and return the form with this proof.

Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query no.</th>
<th>Query</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Please check that the suggested running head is correct.</td>
<td></td>
</tr>
<tr>
<td>Q2</td>
<td>Please provide corresponding author details for this article.</td>
<td></td>
</tr>
<tr>
<td>Q3</td>
<td>Please check that the inserted definition for “i.i.d.” is correct.</td>
<td></td>
</tr>
<tr>
<td>Q4</td>
<td>The meaning of the sentence “Condition H3 is...” is not clear. Please clarify.</td>
<td></td>
</tr>
<tr>
<td>Q5</td>
<td>Please update the references “Cavaliere &amp; Georgiev (2013),” “Chan et al. (2012),” and “Zhang et al. (2013).”</td>
<td></td>
</tr>
</tbody>
</table>