| Statistica Sinica Preprint No: SS-2022-0013 | |
|---|---|
| **Title** | Test for Zero Mean of Errors In An ARMA-GGARCH Model After Using A Median Inference |
| **Manuscript ID** | SS-2022-0013 |
| **URL** | http://www.stat.sinica.edu.tw/statistica/ |
| **DOI** | 10.5705/ss.202022.0013 |


| | |
|---|---|
| **Complete List of Authors** | Yaolan Ma,<br>Mo Zhou,<br>Liang Peng and<br>Rongmao Zhang |
| **Corresponding Authors** | Rongmao Zhang |
| **E-mails** | rmzhang@zju.edu.cn |

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



# Test for Zero Mean of Errors In An ARMA-GGARCH Model After Using A Median Inference

Yaolan Ma[1], Mo Zhou[2], Liang Peng[3] and Rongmao Zhang[2*]

[1]*North Minzu University,* [2]*Zhejiang University and* [3]*Georgia State University*

*Abstract:*

Median inferences are appealing for fitting an ARMA model with heteroscedastic errors to financial returns, because such returns are known to have heavy tails. To ensure that the model is still related to the conditional mean, we test for a zero mean of the errors by using a random weighted bootstrap method to quantify the estimation uncertainty. The proposed test is robust against heteroscedasticity and heavy tails, because we do not infer heteroscedasticity and need fewer finite moments. Simulations confirm that the proposed test exhibits good finite-sample performance in terms of size and power. Empirical applications show that we need to exercise caution when interpreting the model after using a median inference.

*Key words and phrases:* ARMA model, heteroscedasticity, weighted estimation, zero mean.

## 1. Introduction

Corresponding author: Rongmao Zhang, rmzhang@zju.edu.cn

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



Consider the following ARMA$(r, s)$ model with general GARCH (G-GARCH) errors:

$$\begin{cases} X_t = \mu + \sum_{i=1}^{r} \phi_i X_{t-i} + \sum_{j=1}^{s} \psi_j \varepsilon_{t-j} + \varepsilon_t, \\ \varepsilon_t = \sigma_t \eta_t, \ \sigma_t^2 = h(\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots), \end{cases} \tag{1.1}$$

where $\mu \in \mathbb{R}$, $\phi_i \in \mathbb{R}$ for $i = 1, \cdots, r$, $\psi_j \in \mathbb{R}$ for $j = 1, \cdots, s$, $\{\eta_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and variance one, and $h$ is a positively measurable function. Because $\mathbb{E}(\eta_t) = 0$ and $\mathbb{E}(\eta_t^2) = 1$, equation (1.1) models the conditional mean and conditional standard deviation of $\{X_t\}$. Examples of GGARCH models include the ARCH models of Engle (1982), GARCH models of Bollerslev (1986), absolute value GARCH models of Taylor (1986) and Schwert (1989), nonlinear GARCH models of Engle (1990), volatility switching GARCH models of Fornari and Mele (1997), threshold GARCH models of Zakoian (1994), and generalized quadratic ARCH models of Sentana (1995).

For ARMA-GARCH models, the quasi-maximum likelihood estimator (QMLE) is often used for statistical inference, requiring $\mathbb{E}(\varepsilon_t^4) < \infty$ and $\mathbb{E}(\eta_t^4) < \infty$ to ensure a normal limit; see Francq and Zakoïan (2004). Recently, by assuming that the median of $\eta_t$ is zero instead of $\mathbb{E}(\eta_t) = 0$, Zhu and Ling (2011) proposed the self-weighted quasi-maximum exponential

likelihood estimator (SWQMELE), deriving its asymptotic normality without $\mathbb{E}(\varepsilon_t^4) < \infty$ and $\mathbb{E}(\eta_t^4) < \infty$. This relaxation of the moment constraints is vital when analyzing financial returns, which are known to have heavy tails. However, changing $\mathbb{E}(\eta_t) = 0$ in (1.1) to a zero median implies that one is interested in modeling the conditional median rather than the conditional mean, which is contrary to the purpose of the classical ARMA-GARCH models. See Fan, Qi, and Xiu (2014) for a discussion on the transformation effect of skewed data.

Recently, Zhou, Peng, and Zhang (2021) developed an empirical likelihood test for a zero mean of $\eta_t$ using a median inference for an ARMA-GARCH model under the assumption that the GARCH errors have a zero median. When we do not reject the null hypothesis, the ARMA-GARCH model is still related to the conditional mean, and can be inferred from fewer finite moments. Let $\mathcal{F}_t$ denote the $\sigma$-field generated by $\{\eta_s : s \leq t\}$. When $\{w_t > 0\}$ is stationary, $w_t$ is $\mathcal{F}_t$-measurable, and $\mathbb{E}(\sigma_t/w_{t-1}) \in (0, \infty)$, then the null hypothesis of a zero mean of $\eta_t$ in (1.1) is equivalent to that of a zero mean of $\varepsilon_t/w_{t-1}$. This motivates us to test the zero mean of $\varepsilon_t/w_{t-1}$ without estimating $\sigma_t$ and requiring $\mathbb{E}(\varepsilon_t^4) < \infty$. Hence, unlike the test of Zhou, Peng, and Zhang (2021), the proposed test is robust against the specification of heteroscedasticity. Specifically, we use the median inference

of Zhu and Ling (2015) to estimate the ARMA model, which has a normal limit when $\mathbb{E}(|\varepsilon_t|^\delta) < \infty$, for some $\delta > 0$. Using the sample mean of the estimated $\varepsilon_t$, we test for a zero mean of $\varepsilon_t/w_{t-1}$. To effectively combine the estimation uncertainties, we use a profile empirical likelihood test based on the estimating equation method of Qin and Lawless (1994); see Owen (2001) for an overview of the empirical likelihood method. Applications of the empirical likelihood method to ARMA-GARCH models include the works of Chan and Ling (2006) for a GARCH model, Li, Liang, and He (2012) for an AR-ARCH model, and Chan, Peng, and Zhang (2012) and Zhang, Li, and Peng (2019) for the tail index of a GARCH(1,1) sequence. However, the results of a simulation study show that the profile empirical likelihood test exhibits poor finite-sample performance, possibly because using the median inference complicates the computation of the profile empirical likelihood method. Therefore, we propose using the random weighted bootstrap method of Jin, Ying, and Wei (2001) and Zhu (2016) to conduct the test for a zero mean. Note that we cannot use the residual-based bootstrap method because we do not infer heteroscedasticity.

A related study is that of Ma, Zhou, Peng, and Zhang (2021), who develop an empirical likelihood test for a zero median of $\varepsilon_t$ after estimating the ARMA model under the assumption of $\mathbb{E}(\eta_t) = 0$. Their proposed

profile empirical likelihood method performs well, because it uses a weighted least squares estimation to fit the ARMA model.

The remainder of this paper is organized as follows. Section 2 presents the proposed methodologies and asymptotic results. Sections 3 and 4 discuss a simulation study and a data analysis, respectively. Section 5 concludes the paper. All proofs are provided in the Appendix.

## 2. Methodologies and Theoretical Results

Consider the ARMA$(r, s)$-GGARCH model (1.1) with a zero median of $\eta_t$ instead of a zero mean. Recall that $\mathcal{F}_t$ denotes the $\sigma$-field generated by $\{\eta_s : s \leq t\}$. Put $\boldsymbol{\theta} = (\mu, \phi_1, \cdots, \phi_r, \psi_1, \cdots, \psi_s)'$, with $\boldsymbol{\theta}_0$ denoting the true value. For ease of notation, we write $\eta_t(\boldsymbol{\theta}_0) = \eta_t$, $\varepsilon_t(\boldsymbol{\theta}_0) = \varepsilon_t$, and $\sigma_t(\boldsymbol{\theta}_0) = \sigma_t$, and define $\phi(z) = 1 - \sum_{i=1}^{r} \phi_i z^i$ and $\psi(z) = 1 + \sum_{j=1}^{s} \psi_j z^j$.

Given the observations $\{X_1, \cdots, X_n\}$ and the initial values $\{X_0, X_{-1}, \cdots\}$, taken as zero in our simulation study and data analysis, we express the parametric form of (1.1) as

$$\begin{cases} \varepsilon_t(\boldsymbol{\theta}) = X_t - \mu - \sum_{i=1}^{r} \phi_i X_{t-i} - \sum_{j=1}^{s} \psi_j \varepsilon_{t-j}(\boldsymbol{\theta}), \\ \sigma_t^2(\boldsymbol{\theta}) = h(\varepsilon_{t-1}(\boldsymbol{\theta}), \varepsilon_{t-2}(\boldsymbol{\theta}), \cdots),\ \eta_t(\boldsymbol{\theta}) = \varepsilon_t(\boldsymbol{\theta})/\sigma_t(\boldsymbol{\theta}). \end{cases}$$

In this study, we test

$$H_0 : \mathbb{E}(\varepsilon_t/w_{t-1}) = 0 \quad \text{vs.} \quad H_a : \mathbb{E}(\varepsilon_t/w_{t-1}) \neq 0, \tag{2.1}$$

which is equivalent to

$$H_0 : \mathbb{E}(\eta_t) = 0 \quad \text{vs.} \quad H_a : \mathbb{E}(\eta_t) \neq 0,$$

where $\{w_t = w(X_t, X_{t-1}, \ldots) > 0\}$ is defined later. To estimate $\varepsilon_t$ under the zero median assumption, we employ the weighted least absolute deviation estimator (LADE) of Zhu and Ling (2015), defined as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{n} w_{t-1}^{-1} |\varepsilon_t(\boldsymbol{\theta})|.$$

The resulting estimator $\hat{\boldsymbol{\theta}}$ has a normal limit when $\mathbb{E}(|\varepsilon_t|^{\delta}) < \infty$, for some $\delta > 0$. Using this estimator, we estimate $\varepsilon_t$ by $\varepsilon_t(\hat{\boldsymbol{\theta}})$, and estimate $\nu = \mathbb{E}(\varepsilon_t / w_{t-1})$ by

$$\hat{\nu} = \frac{1}{n} \sum_{t=1}^{n} w_{t-1}^{-1} \varepsilon_t(\hat{\boldsymbol{\theta}}).$$

To avoid estimating the complicated asymptotic variance of $\hat{\nu}$ when testing $H_0$, we adopt the random weighted bootstrap method of Jin, Ying, and Wei (2001) and Zhu (2016):

- Step 1) Draw a random sample of size $n$ from a distribution with mean one and variance one, for example, the standard exponential distribution. Denote the observations by $\{\delta_t^b\}_{t=1}^n$.

- Step 2) Solve

$$\hat{\boldsymbol{\theta}}^b = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{n} \delta_t^b w_{t-1}^{-1} |\varepsilon_t(\boldsymbol{\theta})|,$$

and compute

$$\hat{\nu}^b = \frac{\sum_{t=1}^n \delta_t^b w_{t-1}^{-1} \varepsilon_t(\hat{\boldsymbol{\theta}}^b)}{\sum_{t=1}^n \delta_t^b}.$$

- Step 3) Repeat the first two steps $B$ times to obtain $\{\hat{\nu}^b\}_{b=1}^B$.

Therefore, we reject the null hypothesis $H_0 : \nu = 0$ at the level $a$ whenever

$$\hat{\nu}^2 / \{\frac{1}{B} \sum_{b=1}^B (\hat{\nu}^b - \hat{\nu})^2\} \geq \chi^2_{1,1-a},$$

where $\chi^2_{1,1-a}$ denotes the $(1-a)$th quantile of a chi-squared distribution with one degree of freedom.

To validate the above test theoretically, we introduce some regularity conditions.

**Assumption 1.** $\boldsymbol{\theta}_0$ is an interior point in $\Theta$, and for each $\boldsymbol{\theta} \in \Theta$, $\phi(z) \neq 0$ and $\psi(z) \neq 0$ when $|z| \leq 1$, and $\phi(z)$ and $\psi(z)$ have no common root with $\phi_r \neq 0$ or $\psi_s \neq 0$.

**Assumption 2.** $\varepsilon_t$ is strictly stationary and ergodic.

**Assumption 3.** $\mathbb{E}(w_{t-1}^{-4} \xi_{\rho,t-1}^4) < \infty$, for any $\rho \in (0,1)$, where $\xi_{\rho,t} = 1 + \sum_{i=0}^{\infty} \rho^i |X_{t-i}|$, $\{w_t = w(X_t, X_{t-1}, \ldots)\}$ is a stationary sequence satisfying $\inf_{t \geq 1} w_t > c_0 > 0$, and $w_t$ is $\mathcal{F}_t$-measurable.

**Assumption 4.** $\{\eta_t\}$ is a sequence of i.i.d. random variables with median zero and $\mathbb{E}(\eta_t^2) = 1$.

**Assumption 5.** $\{\eta_t\}$ has a continuous density function $g(x)$ satisfying $g(0) > 0$ and $\sup_{x\in\mathbb{R}} g(x) < \infty$.

Assumptions 1 and 2 ensure that there exists a unique, strictly stationary causal solution to the first and second equations, respectively, of (1.1) (see Zhu and Ling (2015)). When $\varepsilon_t$ follows a GARCH$(p, q)$ model, Theorem 3.1 of Basrak, Davis, and Mikosch (2002) ensures Assumption 2 if the Lyapunov exponent of the random coefficient matrices $A_t$ is negative, where $\boldsymbol{\varepsilon}_t = A_t\boldsymbol{\varepsilon}_{t-1} + B_t$ and $\boldsymbol{\varepsilon}_t = (\sigma^2_{t+1}, \sigma^2_t, \cdots, \sigma^2_{t-q+2}, \varepsilon^2_t, \cdots, \varepsilon^2_{t-p+2})'$. The weight $w_t$ in Assumption 3 reduces the moment effect of $\sigma_t$, and is defined later. Assumptions 4 and 5 allow us to use a median inference for the ARMA model.

**Theorem 1.** *Under Assumptions 1–5 and the null hypothesis of (2.1),*

$$\sqrt{n}\hat{\nu} \xrightarrow{d} N(0, \sigma^2) \text{ and } \frac{n}{B}\sum_{b=1}^{B}(\hat{\nu}^b - \hat{\nu})^2/\sigma^2 \xrightarrow{p} 1 \tag{2.2}$$

*as $B \to \infty$ and $n \to \infty$, where*

$$\sigma^2 = (-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1)\mathbb{E}[\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}_1'](-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1)',$$

$$\boldsymbol{\Gamma} = \mathbb{E}\{w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\},\ \boldsymbol{\Sigma} = \mathrm{E}[(w_{t-1}\sigma_t)^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}],$$

*and*

$$\tilde{\boldsymbol{D}}_t = (w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} sgn(\varepsilon_t),\ w_{t-1}^{-1}\varepsilon_t)'.$$

Theorem 1 shows that our proposed test for $H_0 : \nu = 0$ has the asymptotically correct size. To investigate the local power of the proposed test, we consider the following local alternative hypothesis:

$$H_a : \nu = \frac{M}{\sqrt{n}}, \text{ for some constant } M. \tag{2.3}$$

Theorem 2 shows that the power of the proposed test goes to one as $|M| \to \infty$.

**Theorem 2.** *Suppose that Assumptions 1–5 hold for model (1.1). Under the alternative hypothesis of (2.3),*

$$\frac{\sqrt{n}\hat{\nu}}{\sqrt{nB^{-1}\sum_{t=1}^{B}(\hat{\nu}^b - \hat{\nu})^2}} \xrightarrow{d} N(M/\sigma, 1)$$

*as $n \to \infty$ and $B \to \infty$, where $\sigma$ is defined in Theorem 1.*

As in Ling (2007) and Zhu and Ling (2011, 2015), the key idea when choosing $w_t$ is to bound $\xi_{\rho,t}$, defined in Assumption 3. There are many different choices, including the one in Ling (2007). Because $\sum_{i=0}^{\infty} e^{\log(h)\log^2(i+1)} < \infty$ and $e^{\log(h)\log^2(t+1)} \geq \rho^t$, for sufficiently large $t$ and any given $\rho \in (0,1)$ and $h \in (0,1)$, we can use $\sum_{i=0}^{t-1} e^{\log(h)\log^2(i+1)}|X_{t-i}|$ to bound $\xi_{\rho,t}$, and thus control the moment effect of $\sigma_{t+1}$. To avoid overweighting, we use the weight function

$$w_t(h) = \max(C, \sum_{i=0}^{t-1} e^{\log(h)\log^2(i+1)}|X_{t-i}|), \text{ for some } h \in (0,1) \text{ and } t = 1, \cdots, n, \tag{2.4}$$

where $C$ is chosen as the 90% quantile of $|X_t|$ and $w_0(h) = 1$. As in He, Hou, Peng, and Shen (2020), we can show that the aforementioned weight function, with $C$ replaced by the corresponding sample quantile, does not change the asymptotic distribution. As in the case of a kernel density estimation, choosing an optimal $h$ in terms of the coverage probability is challenging, requiring the Edgeworth expansion for the proposed test statistic. In our simulation study and data analysis, we use $h = 0.2$ and 0.4, which exhibit good finite-sample performance.

## 3. Simulation study

In this section, we examine the finite-sample performance of the proposed test in terms of size and power.

We generate 5000 random samples with sample size $n = 1000$ and 2500 from ARMA(1,0)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1) models, with $\mu = 0.1, \phi_1 = 0.5, \psi_1 = 0.2$, $\omega = 0.1$, $a_1 = 0.1$, $b_1 = 0.8$, and $\eta_t = V/\sqrt{\mathbb{E}(V^2)}$, where

$$V = I(U < 1/2)V_1 - I(U \geq 1/2)V_2,$$

with $U \sim \text{Uniform}(0, 1)$, $V_1 \sim \text{Pareto}(1, \alpha_1)$ (i.e., $P(V_1 \leq x) = 1-(1+x)^{-\alpha_1}$ for $x \geq 0$), and $V_2 \sim \text{Pareto}(1, \alpha_2)$ being independent. It is easy to check

that

$$\mathbb{E}(V) = \frac{1}{2}\{\frac{1}{\alpha_1 - 1} - \frac{1}{\alpha_2 - 1}\} \text{ and } \mathbb{E}(V^2) = \frac{1}{(\alpha_1 - 1)(\alpha_1 - 2)} + \frac{1}{(\alpha_2 - 1)(\alpha_2 - 2)},$$

and that $\eta_t$ has a zero median, right tail index $\alpha_1$, left tail index $\alpha_2$, a zero mean if $\alpha_1 = \alpha_2$, and a nonzero mean if $\alpha_1 \neq \alpha_2$.

We take $\alpha_1 = \alpha_2 = 2.2, 2.5, 3$ to compute the size, and $\alpha_1 = 3.2$ or $3.5$ with $\alpha_2 = 3$ to calculate the power, giving $\mathbb{E}(\eta_t) = -0.024$. To implement the proposed test for a zero mean after using the median inference, we use $B = 1000$ in the random weighted bootstrap method, and the weight function $w_t(h)$ in (2.4), with $h = 0.2$ and $0.4$. We report the empirical size of the test at levels 10% and 5% in Table 1, and the empirical power in Table 2, showing that the proposed test has an accurate size and nontrivial power. Furthermore, using $h = 0.2$ and $h = 0.4$ gives robust results, and the power increases when the sample size becomes large or $\alpha_1$ is away from the null hypothesis.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



Table 1: Test size for the ARMA(1,0)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1) models.

| | | | ARMA(1,0)-GARCH(1,1) | | ARMA(1,1)-GARCH(1,1) | |
|---|---|---|---|---|---|---|
| $\alpha_1$ | $n$ | Level | $h = 0.2$ | $h = 0.4$ | $h = 0.2$ | $h = 0.4$ |
| 2.2 | 1000 | 10% | 0.0980 | 0.0944 | 0.0962 | 0.0944 |
| | | 5% | 0.0466 | 0.0464 | 0.0468 | 0.0470 |
| | 2500 | 10% | 0.0988 | 0.1040 | 0.0972 | 0.1018 |
| | | 5% | 0.0486 | 0.0468 | 0.0494 | 0.0450 |
| 2.5 | 1000 | 10% | 0.0966 | 0.0974 | 0.0966 | 0.0986 |
| | | 5% | 0.0472 | 0.0468 | 0.0474 | 0.0464 |
| | 2500 | 10% | 0.1024 | 0.1032 | 0.1002 | 0.1016 |
| | | 5% | 0.0528 | 0.0502 | 0.0510 | 0.0494 |
| 3 | 1000 | 10% | 0.0980 | 0.0990 | 0.1002 | 0.1008 |
| | | 5% | 0.0442 | 0.0434 | 0.0468 | 0.0446 |
| | 2500 | 10% | 0.0975 | 0.0990 | 0.0996 | 0.1010 |
| | | 5% | 0.0472 | 0.0476 | 0.0472 | 0.0498 |

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



Table 2: Test power for the ARMA(1,0)-GARCH(1,1) and ARMA(1,1)-GARCH(1,1) models.

| | | | ARMA(1,0)-GARCH(1,1) | | ARMA(1,1)-GARCH(1,1) | |
|---|---|---|---|---|---|---|
| $\alpha_1$ | $n$ | Level | $h = 0.2$ | $h = 0.4$ | $h = 0.2$ | $h = 0.4$ |
| 3.2 | 1000 | 10% | 0.2244 | 0.2168 | 0.2182 | 0.2140 |
| | | 5% | 0.1386 | 0.1338 | 0.1292 | 0.1292 |
| | 2500 | 10% | 0.3672 | 0.3648 | 0.3676 | 0.3668 |
| | | 5% | 0.2500 | 0.2526 | 0.2546 | 0.2570 |
| 3.5 | 1000 | 10% | 0.6568 | 0.6508 | 0.6490 | 0.6484 |
| | | 5% | 0.5188 | 0.5242 | 0.5200 | 0.5180 |
| | 2500 | 10% | 0.9334 | 0.9322 | 0.9330 | 0.9324 |
| | | 5% | 0.8842 | 0.8862 | 0.8842 | 0.8842 |

## 4. Application: Exchange Rates

In this section, we examine the daily log-returns ($\times 100$) of the following six exchange rates for the period May 3, 2011, to May 2, 2021: HKD/USD, EUR/USD, CNY/USD, CAD/USD, MXN/USD, and INR/USD. We plot these exchange rates in Figure 1.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



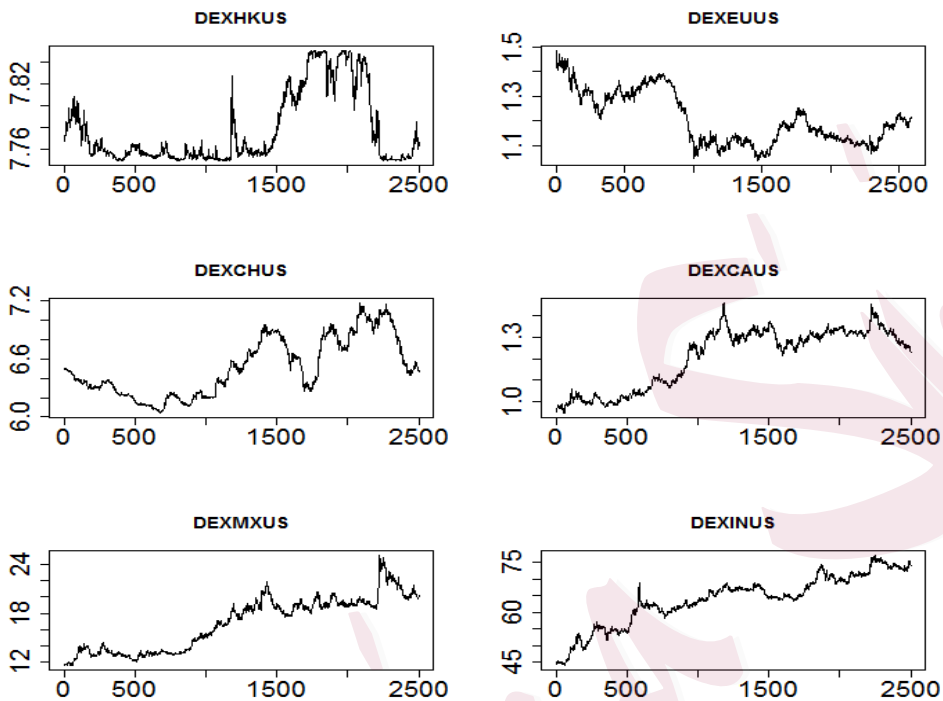


Figure 1: Plots of the HKD/USD, EUR/USD, CNY/USD, CAD/USD, MXN/USD, and INR/USD exchange rates for the period May 3, 2011, to May 2, 2021.

We first estimate the tail index of $\{|X_t|\}$ using the Hill estimator of Hill (1975), defined as

$$\hat{\alpha}(k) = \left[\frac{1}{k}\sum_{i=1}^{k}\log\left(\frac{X_{(n-i+1)}}{X_{(n-k)}}\right)\right]^{-1},$$

with $\{X_{(t)}\}$ being the ascending order statistics of $\{|X_t|\}$. We plot the Hill estimates against various $k$ in Figure 2, which shows that the tail indices of

all the log-returns of the exchange rates considered here, except CAD/USD, may be less than four, that is, $EX_t^4 = \infty$. Therefore, the inference may not have a normal limit when we fit an ARMA-GARCH model using the QMLE. To explore the possibility of using the SWQMELE to fit an ARMA-GARCH model, we test whether the GARCH model has a zero mean after using a median inference. If we do not reject the null hypothesis, then the fitted ARMA-GARCH model using a median inference is still related to the conditional mean. Otherwise, the fitted ARMA-GARCH model focuses on a conditional median rather than a conditional mean, which is contrary to the conventional purpose of using a ARMA-GARCH model.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



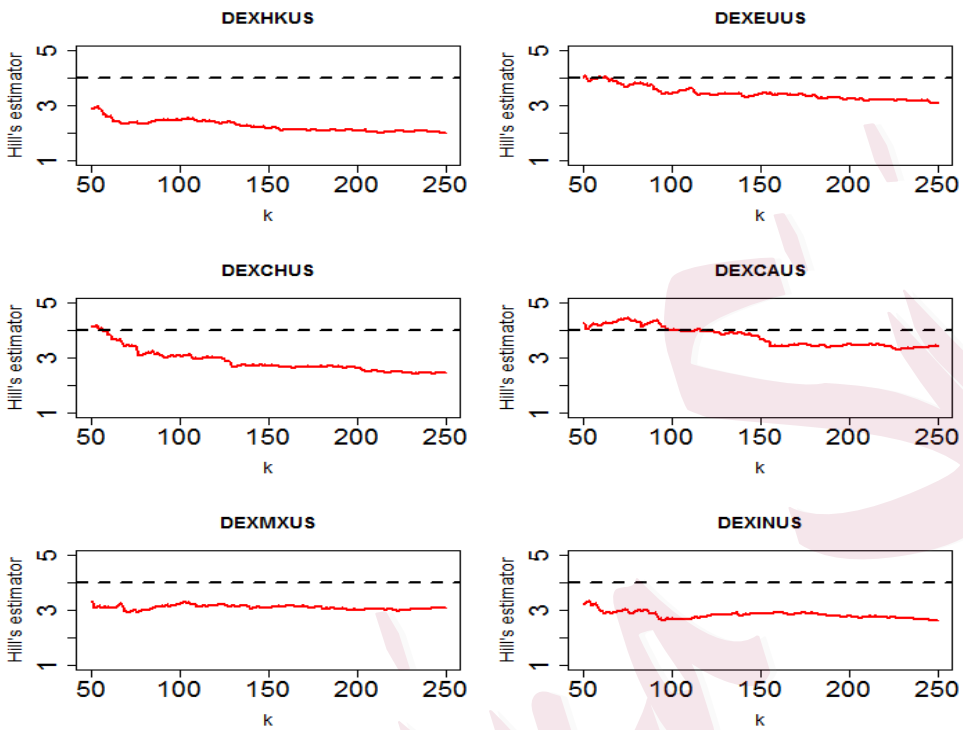


Figure 2: The Hill estimates for the HKD/USD, EUR/USD, CNY/USD, CAD/USD, MXN/USD, and INR/USD exchange rates for the period May 3, 2011, to May 2, 2021.

To apply our proposed test, we use the function "auto.arima" in the R package "forecast" with the Akaike information criterion (AIC) to obtain the appropriate orders of the employed ARMA model. We report the fitted models in Table 3, and plot the ACFs of the residuals in Figure 3. Using the selected orders of the ARMA model, we compute the p-value of the

proposed test using the weight function $w_t(h)$ in (2.4), with $h = 0.2$ and 0.4, and $B = 5000$ in the random weighted bootstrap method. Table 3 shows that we strongly and weakly reject the null hypothesis of a zero mean for the INR/USD and MXN/USD exchange rates, respectively, but do not reject the null hypothesis for the other exchange rates. Therefore, one should be cautious when interpreting data analyses for INR/USD and MXN/USD exchange rates when using the SWQMELE to fit an ARMA-GARCH model to the log-returns, because the ARMA part no longer models the conditional mean. Note that the aforementioned procedure for selecting the ARMA model uses the zero mean of the errors, rather than the null hypothesis of a zero median of the errors, and ignores the possibility that the least squares estimate has a nonnormal limit because there are too few finite moments. Thus, it would be useful to develop an order-selection procedure that uses a median inference and allows heavy-tailed residuals. This is left as a topic for future research.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



Table 3: The fitted ARMA models and computed p-values of the proposed test for a zero mean of the log-returns of daily exchange rates for the period May 3, 2011, to May 2, 2021.

| Exchange rate | ARMA model | $h = 0.2$ | $h = 0.4$ |
|---|---|---|---|
| HKD/USD | ARMA(1,3) | 0.3931 | 0.2159 |
| EUR/USD | ARMA(0,1) | 0.5625 | 0.4433 |
| CNY/USD | ARMA(1,2) | 0.9138 | 0.9024 |
| CAD/USD | ARMA(2,0) | 0.2306 | 0.1622 |
| MXN/USD | ARMA(0,1) | 0.0932 | 0.1044 |
| INR/USD | ARMA(3,1) | 0.0025 | 0.0009 |

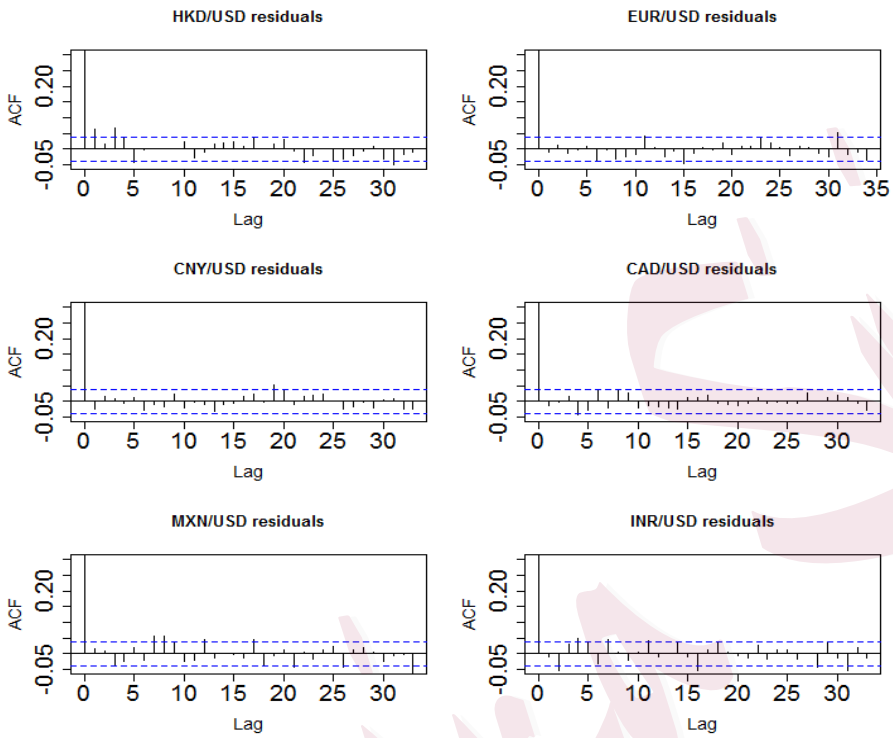


Figure 3: The autocorrelation functions of the residuals.

## 5. Conclusion

The heavy tails of financial returns make median statistical inferences popular for fitting an ARMA model with heteroscedastic errors to such returns. To ensure that the ARMA model is still related to the conditional mean after using a median inference, we test for a zero mean of the errors by using a random weighted bootstrap method to quantify the uncertainty. The proposed test is robust against heteroscedasticity and heavy tails, because

it does not infer the heteroscedasticity and requires fewer finite moments. A simulation study confirms the good finite-sample performance of the proposed test in terms of size and power. The results of our empirical analysis show that we need to exercise caution when using a median inference to fit an ARMA-GARCH model to the log-returns of INR/USD and MXN/USD exchange rates, because we reject the null hypothesis of a zero mean of the errors in this case.

## Acknowledgments

We thank the co-editor, associate editor, and two reviewers for their helpful comments, which led to a much improved version of this paper. Ma's research was partly supported by the Fundamental Research Funds for the Central Universities, North Minzu University (No. 2021JCYJ06), the Natural Science Foundation of Ningxia (No. 2022AAC03232), the Scientific Research Project of Ningxia Higher Education Institutions (No. NGY2020064), and NSFC (No. 62066001/62163001). Peng's research was partly supported by the Simons Foundation and the NSF grant of DMS-2012448. Zhang's research was supported by grants from the NSFC (No. 12171427/U21A20426), Zhejiang Provincial Natural Science Foundation (No. LZ21A010002), and Fundamental Research Funds for the Central Univer-

sities (No. 2021XZZX002).

## References

B. Basrak, R.A. Davis and T. Mikosch (2002). Regular variation of GARCH processes. *Stochastic Processes and their Applications 99*, pp. 95–116.

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

N.H. Chan, L. Peng and R. Zhang (2012). Interval estimation of the tail index of a Garch(1,1) model. *Test 21*, pp. 546–565.

N.H. Chan and S. Ling (2006). Empirical likelihood for GARCH models. *Econometric Theory 22*, pp. 403–428.

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

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

J. Fan, L. Qi and D. Xiu (2014). Quasi Maximum Likelihood Estimation of GARCH Models with Heavy-Tailed Likelihoods. *Journal of Business and Economic Statistics 32*, pp. 178–205.

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

C. Francq and J. Zakoïan (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli 10*, pp. 847–860.

P. Hall and C. C. Heyde (1980). *Martingale Limit Theory and its Application.*, Academic Press, New York.

B. Hill (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics 3*, pp. 1163–1174.

Y. He, Y. Hou, L. Peng and H. Shen (2020). Inference for conditional Value-at-Risk of a predictive regression. *Annals of Statistics 48*, pp. 3442–3464.

Z. Jin, Z. Ying and L.J. Wei (2001). A simple resampling method by perturbing the minimand. *Biometrika 88*, pp. 381–390.

J. Li, W. Liang and S. He (2012). Empirical likelihood for AR-ARCH models based on LAD estimation. *Acta Mathematicae Applicatae Sinica 28*, pp. 371-382.

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

Y. Ma, M. Zhou, L. Peng and R. Zhang (2021). Test for Zero Median of Errors in an ARMA-GARCH Model. *Econometric Theory*, to appear.

A.B. Owen (2001). *Empirical Likelihood.* Chapman Hall, New York.

J. Qin and J. Lawless (1994). Empirical likelihood and general estimating equations. *Annals of Statistics 22*, pp. 300–325.

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

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

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

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

R. Zhang, C. Li and L. Peng (2019). Inference for the Tail Index of a GARCH(1,1) Model and an AR(1) Model with ARCH(1) Errors. *Econometric Reviews 38*, pp. 151–169.

M. Zhou, L. Peng and R. Zhang (2021). Empirical likelihood test for the application of SWQMELE in fitting an ARMA-GARCH model. *Journal of Time Series Analysis 42*, pp. 222-239.

K. Zhu (2016). Bootstrapping the portmanteau tests in weak auto-regressive moving average models. *Journal of the Royal Statistical Society: Series A 78*, pp. 463–485.

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

K. Zhu and S. Ling (2015). LADE-based inference for ARMA models with unspecified and heavy-tailed heteroscedastic noises. *Journal of the American Statistical Association 110*, pp. 784–794.

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



## Appendix: Proofs of Theorems 1 and 2

Throughout, define

$$\tilde{\boldsymbol{D}}_{t,1} = w_{t-1}^{-1}\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t),\ \tilde{D}_{t,2} = w_{t-1}^{-1}\varepsilon_t,\ \tilde{\boldsymbol{D}}_t = (\tilde{\boldsymbol{D}}_{t,1}', \tilde{D}_{t,2})',$$

and recall $\varepsilon_t = \varepsilon_t(\boldsymbol{\theta}_0)$. First, we need two lemmas below.

**Lemma 1.** *Under conditions of Theorem 1, there exist a constant $\rho \in (0,1)$, a constant $C > 0$, and a neighborhood $\Theta_0$ of $\boldsymbol{\theta}_0$ such that*

$$\sup_{\Theta_0}|\varepsilon_t(\boldsymbol{\theta})| \le C\xi_{\rho,t-1},\ \sup_{\Theta_0}\left\|\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right\| \le C\xi_{\rho,t-1},\ \textit{and}\ \sup_{\Theta_0}\left\|\frac{\partial^2 \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right\| \le C\xi_{\rho,t-1},$$

*where $\xi_{\rho,t}$ is defined in Assumption 3.*

*Proof.* See Lemma A.1 of Ling (2007). □

**Lemma 2.** *Under the conditions of Theorem 1, we have as $n \to \infty$,*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}_t \xrightarrow{d} N\Big(\mathbf{0}, \mathbb{E}\{\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}_1'\}\Big).$$

*Proof.* Recall that $\mathcal{F}_t$ is the $\sigma$-filed generated by the sequence $\{\eta_t, \eta_{t-1}, \cdots\}$. It is straightforward to verify that

$$\mathbb{E}\big(\tilde{\boldsymbol{D}}_{t,1}\big|\mathcal{F}_{t-1}\big) = \mathbb{E}\left(w_{t-1}^{-1}\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t)\middle|\mathcal{F}_{t-1}\right) = w_{t-1}^{-1}\frac{\partial \varepsilon_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathbb{E}\big(\mathrm{sgn}(\eta_t)\big) = \mathbf{0}$$

and

$$\mathbb{E}\big(\tilde{D}_{t,2}\big|\mathcal{F}_{t-1}\big) = \mathbb{E}\big(w_{t-1}^{-1}\varepsilon_t\big|\mathcal{F}_{t-1}\big) = w_{t-1}^{-1}\sigma_t\mathbb{E}(\eta_t) = 0,$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



i.e., $\{\tilde{\boldsymbol{D}}_t\}$ is a sequence of Martingale differences. It follows from Assumption 3, Lemma 1, and the dominated convergence theorem that

$$\max_{1\le t\le n}\|\frac{1}{\sqrt{n}}\tilde{\boldsymbol{D}}_t\| = o_p(1),\ \frac{1}{n}\sum_{t=1}^{n}\{\tilde{\boldsymbol{D}}_t\tilde{\boldsymbol{D}}_t^{'}\} = \mathbb{E}\{\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}_1^{'}\} + o_p(1),\ \mathbb{E}\left\|\max_{1\le t\le n}\frac{1}{n}\tilde{\boldsymbol{D}}_t\tilde{\boldsymbol{D}}_t^{'}\right\| = o(1).$$

Hence, the conditions of the central limit theorem for Martingale differences are satisfied (see Theorem 3.2 of Hall and Heyde (1980)), i.e., the theorem follows. □

***Proof of Theorem 1.*** It follows from Theorem 2 of Zhu and Ling (2015) that

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t) + o_p(1), \tag{A.1}$$

where $\boldsymbol{\Sigma}$ is given in Theorem 1. Using Taylor expansion, Lemma 1, and (A.1), we have

$$\begin{aligned}
\sqrt{n}\hat{\nu} &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\varepsilon_t(\hat{\boldsymbol{\theta}}) \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\left[\varepsilon_t(\hat{\boldsymbol{\theta}}) - \varepsilon_t\right] + \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\varepsilon_t \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) + \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\varepsilon_t + o_p(1), \\
&= -\boldsymbol{\Gamma}\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t) + \frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_{t-1}^{-1}\varepsilon_t + o_p(1) \\
&\xrightarrow{d} N\Big(0, (-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1)\mathbb{E}[\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}_1^{'}](-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1)'\Big),
\end{aligned} \tag{A.2}$$

where $\boldsymbol{\Gamma} = \mathbb{E}\Big\{w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\Big\}$ is given in Theorem 1. Hence, the first equation of (2.2) follows.

Similar to the proof of Theorem 2 of Zhu and Ling (2015), we have

$$\sqrt{n}(\hat{\boldsymbol{\theta}}^b - \boldsymbol{\theta}_0) = -\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t) + o_p(1). \tag{A.3}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



Following the proof of (A.2) and using (A.3), we have

$$\begin{aligned}
\sqrt{n}\hat{\nu}^b &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\varepsilon_t(\hat{\boldsymbol{\theta}}^b)+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}(\hat{\boldsymbol{\theta}}^b-\boldsymbol{\theta}_0)+\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\varepsilon_t+o_p(1),\\
&= -\boldsymbol{\Gamma}\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t)+\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\delta_t^b w_{t-1}^{-1}\varepsilon_t+o_p(1) \quad \text{(A.4)}
\end{aligned}$$

By (A.2) and (A.4), we have

$$\begin{aligned}
&\sqrt{n}(\hat{\nu}^b-\hat{\nu})\\
=\ & -\boldsymbol{\Gamma}\frac{(2g(0)\boldsymbol{\Sigma})^{-1}}{\sqrt{n}}\sum_{t=1}^{n}(\delta_t^b-1)w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t)+\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\delta_t^b-1)w_{t-1}^{-1}\varepsilon_t+o_p(1).
\end{aligned} \quad \text{(A.5)}$$

Put

$$\boldsymbol{Z}_{t,1}^b=(\delta_t^b-1)w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t),\ Z_{t,2}^b=(\delta_t^b-1)w_{t-1}^{-1}\varepsilon_t,$$

$$\boldsymbol{Z}_{t,1}=w_{t-1}^{-1}\frac{\partial\varepsilon_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\mathrm{sgn}(\varepsilon_t),\ \text{and}\ Z_{t,2}=w_{t-1}^{-1}\varepsilon_t.$$

Using (A.5) and letting $B\to\infty$ and $n\to\infty$, we can show that

$$\begin{aligned}
&\frac{n}{B}\sum_{b=1}^{B}(\hat{\nu}^b-\hat{\nu})^2\\
=\ & \frac{1}{B}\sum_{b=1}^{B}\big\{\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{Z}_{t,1}^b(\boldsymbol{Z}_{t,1}^b)'(\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1})'\\
&+\frac{1}{n}\sum_{t=1}^{n}(Z_{t,2}^b)^2-2\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{Z}_{t,1}^b Z_{t,2}^b\big\}+o_p(1)\\
=\ & \boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{Z}_{t,1}(\boldsymbol{Z}_{t,1})'(\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1})'\\
&+\frac{1}{n}\sum_{t=1}^{n}(Z_{t,2})^2-2\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{Z}_{t,1}Z_{t,2}+o_p(1)\\
=\ & (-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1},1)\mathbb{E}[\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}_1'](-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1},1)'+o_p(1),
\end{aligned}$$

i.e., the second equation of (2.2) holds. □

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0013



***Proof of Theorem 2.*** Put $\tilde{D}^*_{t,2} = w^{-1}_{t-1}[\varepsilon_t - \mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})]$ and $\tilde{\boldsymbol{D}}^*_t = (\tilde{\boldsymbol{D}}'_{t,1}, \tilde{D}^*_{t,2})'$. Then

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}_t &= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\tilde{\boldsymbol{D}}'_{t,1}, w^{-1}_{t-1}[\varepsilon_t - \mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})]\right)' + \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\mathbf{0}', w^{-1}_{t-1}\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})\right)' \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}^*_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\mathbf{0}', w^{-1}_{t-1}\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})\right)' \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}^*_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\mathbf{0}', w^{-1}_{t-1}\sigma_t\mathbb{E}(\eta_t)\right)' \\
&= \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}^*_t + \frac{M}{n}\sum_{t=1}^{n}\left(\mathbf{0}', w^{-1}_{t-1}\sigma_t/\mathbb{E}(w^{-1}_{t-1}\sigma_t)\right)', \qquad \text{(A.6)}
\end{aligned}$$

where $\mathbf{0}$ is a $(r+s+1)$-vector, and the last equation follows by $\mathbb{E}(w^{-1}_{t-1}\varepsilon_t) = \mathbb{E}(\eta_t)\mathbb{E}(w^{-1}_{t-1}\sigma_t) = M/\sqrt{n}$.

Because $\mathbb{E}\{w^{-1}_{t-1}[\varepsilon_t - \mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})]|\mathcal{F}_{t-1}\} = 0$, $\{\tilde{\boldsymbol{D}}^*_t\}$ is a sequence of Martingale differences. Like the proof of Lemma 2, we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\tilde{\boldsymbol{D}}^*_t \xrightarrow{d} N\Big(\mathbf{0}, \mathbb{E}\{\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}'_1\}\Big) \text{ as } n\to\infty.$$

On the other hand, by the weak law of large numbers for stationary series, we have

$$\frac{M}{n}\sum_{t=1}^{n}w^{-1}_{t-1}\sigma_t/\mathbb{E}(w^{-1}_{t-1}\sigma_t) \xrightarrow{p} M \text{ as } n\to\infty.$$

Thus, similar to the proof of Theorem 1, as $B\to\infty$ and $n\to\infty$,

$$\sqrt{n}\hat{\nu} \xrightarrow{d} N\Big(M, \{-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1\}\mathbb{E}[\tilde{\boldsymbol{D}}_1\tilde{\boldsymbol{D}}'_1]\{-\boldsymbol{\Gamma}(2g(0)\boldsymbol{\Sigma})^{-1}, 1\}'\Big),$$

and

$$\frac{n}{B}\sum_{b=1}^{B}(\hat{\nu}^b - \hat{\nu})^2 = \sigma^2 + o_p(1),$$

i.e., Theorem 2 holds. □