Robust inference in conditionally heteroskedastic autoregressions

Pedersen, Rasmus Søndergaard

Department of Economics, University of Copenhagen

4 October 2017

Robust inference in conditionally heteroskedastic autoregressions

Rasmus Søndergaard Pedersen[*]

October 4, 2017

Abstract

We consider robust inference for an autoregressive parameter in a stationary autoregressive model with GARCH innovations when estimation is based on least squares estimation. As the innovations exhibit GARCH, they are by construction heavy-tailed with some tail index κ. The rate of consistency as well as the limiting distribution of the least squares estimator depend on κ. In the spirit of Ibragimov and Müller ("t-statistic based correlation and heterogeneity robust inference", *Journal of Business & Economic Statistics,* 2010, vol. 28, pp. 453-468), we consider testing a hypothesis about a parameter based on a Student's t-statistic for a fixed number of subsamples of the original sample. The merit of this approach is that no knowledge about the value of κ nor about the rate of consistency and the limiting distribution of the least squares estimator is required. We verify that the one-sided t-test is asymptotically a level α test whenever $\alpha \leq 5\%$ uniformly over $\kappa \geq 2$, which includes cases where the innovations have infinite variance. A simulation experiment suggests that the finite-sample properties of the test are quite good.

Keywords: t-test, AR-GARCH, regular variation, least squares estimation.

JEL Classification: C12, C22, C46, C51.

1 Introduction

We consider, as in Zhang and Ling (2015) (ZL hereafter), the AR(p) model,

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

[*]Address: University of Copenhagen, Department of Economics, Oester Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark. Phone: (+45)35323074. Fax: (+45)35323000. E-mail: rsp@econ.ku.dk

where ε_t follows a general GARCH (GGARCH) process,

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

with $(\eta_t : t \in \mathbb{Z})$ an I.I.D. process, $\delta > 0$, and $b, c : \mathbb{R} \to \mathbb{R}_+$ such that $P(h_t^\delta > 0) = 1$ and $c(0) < 1$. The objective of this paper is to propose a robust method for testing a hypothesis about an element of the vector $\phi := (\phi_1, ..., \phi_p)' \in \mathbb{R}^p$ when ϕ is estimated with the OLS estimator,

$$\hat{\phi} = \left(\sum_{t=p+1}^{n} Y_{t-1} Y_{t-1}' \right)^{-1} \left(\sum_{t=p+1}^{n} Y_{t-1} y_t \right), \tag{1.3}$$

where $Y_t = (y_t, ..., y_{t-p+1})'$ and n is the length of the sample. Specifically, suppose that we want to test $H_0 : \phi_i = \phi_{i,0}$ for some $i = 1, ..., p$ against the alternative $\phi_i \neq \phi_{i,0}$. What complicates inference in the model is that (under suitable conditions) the distribution of ε_t will be regularly varying with some tail index $\kappa > 0$. As recently demonstrated by ZL, the value of κ will determine the rate of consistency as well as the limiting distribution of the (suitably scaled) OLS estimator. The limiting distribution is given by the distribution of some function of a stable random vector with index $\kappa/2 \wedge 2$. We note that the tail index may be estimated, by e.g. a Hill estimator, but even for a known $\kappa \in (0, 4)$ the limiting distribution of the OLS estimator is only partly known, in the sense that the parameters of the limiting stable distributions are stated in terms of limiting point processes, see e.g. Davis and Hsing (1995) and Davis and Mikosch (1998). As pointed out by Lange (2011, Remark 3), we do not have an expression for the dispersion parameter or for the dependence structure of the stable vector.

Under suitable conditions in line with the assumptions by ZL, we show that each element of the OLS estimator has a mixed Gaussian distribution. This property will show up to be very useful, as it allows us to apply a two-sided t-statistic based on a fixed number of subsamples, as recently considered by Ibragimov and Müller (2010, 2016) and Ibragimov et al. (2015, Chapter 3.3). Specifically, we split our original sample into $q \geq 2$ (without loss of generality) equi-sized subsamples $(y_t : t = 1 + (i-1)\lfloor n/q \rfloor, ..., i\lfloor n/q \rfloor)$, $i = 1, ..., q$, where $\lfloor x \rfloor$ denotes the integer part of $x \in \mathbb{R}$. For each subsample we obtain the OLS estimator for ϕ,

$$\hat{\phi}^{(j)} = \left(\sum_{t=p+1+(j-1)\lfloor n/q \rfloor}^{j\lfloor n/q \rfloor} Y_{t-1} Y_{t-1}' \right)^{-1} \left(\sum_{t=p+1+(j-1)\lfloor n/q \rfloor}^{j\lfloor n/q \rfloor} Y_{t-1} y_t \right), \quad j = 1, ..., q. \tag{1.4}$$

Let

$$X_j := (\hat{\phi}_i^{(j)} - \phi_{i,0}), \quad j = 1, ..., q, \tag{1.5}$$

and obtain the t-statistic based on q "observations",

$$\tau_{\phi_i=\phi_{i,0}} = \sqrt{q}\frac{\bar{X}}{s_X}, \tag{1.6}$$

where $\bar{X} := q^{-1}\sum_{j=1}^{q} X_j$ and $s_X^2 := (q-1)^{-1}\sum_{j=1}^{q}(X_j - \bar{X})^2$. With T_{q-1} a random variable with a Student's t-distribution with $q-1$ degrees of freedom, and with $\mathrm{cv}_q(\alpha)$ satisfying $P(|T_{q-1}| > \mathrm{cv}_q(\alpha)) = \alpha$ for $\alpha \leq 5\%$, we show that whenever $\kappa \geq 2$ (which is the region of the tail index for which the OLS estimator is consistent for ϕ), $P(|\tau_{\phi_i=\phi_{i,0}}| > \mathrm{cv}_q(\alpha)) \leq \alpha$ as $n \to \infty$ under H_0. Hence the two-sided t-test is asymptotically a level α test, and is robust in the sense that we are able to make inference about ϕ_i without requiring any knowledge about the rate of consistency of the OLS estimator as well as any knowledge about its limiting distribution. This property relies on showing that the suitably scaled X_j is asymptotically mixed Gaussian and that X_j and X_k are asymptotically independent for $j \neq k$.

It is by now well-known that many time series exhibit heavy-tail behavior, as investigated in e.g. Loretan and Phillips (1994) in terms of financial time series. Least squares estimation of the autoregressive parameters in stationary AR models driven by heavy-tailed independent innovations has been studied by Davis and Resnick (1986) and bootstrap-based inference has been considered by Davis and Wu (1997) and Cavaliere et al. (2016a). In terms of dependent heavy-tailed innovations, Mikosch and Stărică (2000), Lange (2011), and Zhang and Ling (2015) have investigated the properties of the least squares estimator. More recently, Cavaliere et al. (2016b) have considered bootstrap inference in non-stationary linear time series with innovations driven by a heavy-tailed linear process. We are not aware of any other papers on robust inference in stationary autoregressions with heavy-tailed GARCH-type innovations.

The remainder of the paper is organized as follows. In Section 2 we present the asymptotic properties of the OLS estimator. In Section 3 we show that the two-sided t-test is asymptotically a level α test. Section 4 contains a short simulation experiment where we investigate the finite-sample properties of the t-test when testing for a zero-valued autoregressive coefficient in an AR(1)-ARCH(1) model with potential infinite variance. Section 5 states sufficient conditions for β-mixing for the process (1.1)-(1.2). This property is used for showing that the subsample estimators are asymptotically independent.

Notation: We say that a random variable has a mixed Gaussian distribution with median $\mu \in \mathbb{R}$, if it has pdf of the form, $\int_0^\infty \phi\left((x-\mu)/\sigma\right) dF(\sigma)$, where $\phi(\cdot)$ is the standard normal pdf, and $F(\cdot)$ is an arbitrary cdf on $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$. Unless stated otherwise, all limits are taken as the sample size n tends to infinity, and "$\xrightarrow{w}$" denotes convergence in distribution. For two functions $f, g : \mathbb{R} \to \mathbb{R}_+$, $f(x) \sim g(x)$ if $\lim_{x\to\infty} f(x)/g(x) = 1$.

2 Properties of the OLS estimator

In this section we present the properties of the OLS estimator for ϕ in (1.3). We make the following assumptions about the model in (1.1)-(1.2), in line with ZL.[1]

Assumption 2.1.

1. $E[\log(c(\eta_t))] < 0$.
2. There exists a $k_0 > 0$ such that $E[(c(\eta_t))^{k_0}] \geq 1$, $E[(c(\eta_t))^{k_0} \log^+(c(\eta_t))] < \infty$, where $\log^+(x) = \max\{0, \log(x)\}$. Moreover, $P(b(\eta_t) = 0) < 1$, $E[b(\eta_t)^{k_0}] < \infty$, and $E[|\eta_t|^{\delta k_0}] < \infty$.
3. The distribution of η_t is symmetric and has a Lebesgue density that is strictly positive on a neighborhood of zero, such that the conditional distribution of $\log c(\eta_t)$ given $\{c(\eta_t) > 0\}$ is non-arithmetic.
4. $1 - \sum_{i=1}^p \phi_i z^i \neq 0$ for $|z| \leq 1$.

Note that Assumptions 2.1.1-2 imply that there exists an almost surely unique, strictly stationary, and ergodic solution to $h_t^\delta = b(\eta_{t-1}) + c(\eta_{t-1})h_{t-1}^\delta$, see e.g. Buraczewski et al. (2016, Theorem 2.1.3). Due to the Kesten-Goldie theorem, see e.g. Kesten (1973, Theorem 4), Assumptions 2.1.1-3 imply that there exists a unique $\kappa \in (0, \delta k_0]$ such that $E[(c(\eta_t))^{\kappa/\delta}] = 1$ and $P(|h_t| > x) \sim c_0 x^{-\kappa}$ for some constant $c_0 > 0$ as $x \to \infty$. Breiman's lemma then ensures that $P(|\varepsilon_t| > x) \sim c_0 E[|\eta_t|^\kappa] x^{-\kappa}$, see also Lemma 2.1 in ZL. By the symmetry of η_t, the distribution of ε_t is symmetric and satisfies

$$P(\varepsilon_t > x) \sim (c_0/2)E[|\eta_t|^\kappa]x^{-\kappa} \quad \text{and} \quad P(-\varepsilon_t > x) \sim (c_0/2)E[|\eta_t|^\kappa]x^{-\kappa}.$$

[1] Compared to assumptions H1-H3 in ZL, we have included some slightly stronger conditions. We have added that $P(b(\eta_t) = 0) < 1$ and that the conditional distribution of $\log c(\eta_t)$ given $\{c(\eta_t) > 0\}$ is non-arithmetic, which appears to be required in order to apply Theorem 4 of Kesten (1973) in the proof of Lemma 2.1 in the supplementary material to ZL.

Likewise (under Assumptions 2.1), y_t has a symmetric distribution, and by arguments given in Lange (2011), y_t has the same tail index as ε_t. In particular, Assumption 2.1 implies that the process in (1.1)-(1.2) has a strictly stationary and ergodic solution satisfying

$$y_t = \sum_{i=0}^{\infty} \varphi_i \varepsilon_{t-i}.$$

We will assume throughout that Assumptions 2.1 is satisfied such that the process (y_t) is stationary and ergodic. Moreover, as in ZL, we will assume that $E[\eta_t^2] = 1$ if $\kappa \geq 2$. Lastly, note that if $\kappa > 2$,

$$\Sigma := E[Y_t Y_t'] \text{ exists and is positive definite,} \tag{2.1}$$

such that $(n-p)^{-1} \sum_{t=p+1}^{n} Y_{t-1} Y_{t-1}' = \Sigma + o(1)$ almost surely. Assumption 2.1 implies the following result, due to Theorem 2.1 of ZL.

Theorem 2.2. *Under Assumption 2.1, let* $\kappa > 0$ *satisfy* $E[(c(\eta_t))^{\kappa/\delta}] = 1$. *Moreover, define*

$$a_n^{(\kappa)} := \begin{cases} \log(n) & \text{if} \quad \kappa = 2, \\ n^{1-2/\kappa} & \text{if} \quad \kappa \in (2,4), \\ (n/\log(n))^{1/2} & \text{if} \quad \kappa = 4, \\ n^{1/2} & \text{if} \quad \kappa > 4. \end{cases}$$

With $\hat{\phi}$ *defined in (1.3) and* ϕ_0 *the true value of* ϕ,

1. *if* $\kappa \in (0,2)$,
$$(\hat{\phi} - \phi_0) \xrightarrow{w} \Sigma_{\kappa/2}^{-1} \tilde{Z}_{\kappa/2},$$
where $\tilde{Z}_{\kappa/2}$ *is a p-dimensional stable vector with index* $\kappa/2$ *and* $\Sigma_{\kappa/2}$ *is a* $p \times p$ *matrix with elements containing stable variables with index* $\kappa/2$,

2. *if* $\kappa = 2$,
$$a_n^{(\kappa)}(\hat{\phi} - \phi_0) \xrightarrow{w} \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|} \right)_{i,j=1,\dots,p}^{-1} Z_1,$$
where $\left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|} \right)_{i,j=1,\dots,p}$ *is a* $p \times p$ *matrix and* Z_1 *is a stable vector with index one;*

3. *if* $\kappa \in (2,4)$,
$$a_n^{(\kappa)}(\hat{\phi} - \phi_0) \xrightarrow{w} \Sigma^{-1} Z_{\kappa/2},$$
where $Z_{\kappa/2}$ *is a p-dimensional stable vector with index* $\kappa/2$ *and* Σ *is given by (2.1);*

4. if $\kappa = 4$,

$$a_n^{(\kappa)}(\hat{\phi} - \phi_0) \xrightarrow{w} \Sigma^{-1} N(0, A),$$

where A is some positive definite constant $p \times p$ matrix;

5. if $\kappa > 4$,

$$a_n^{(\kappa)}(\hat{\phi} - \phi_0) \xrightarrow{w} \Sigma^{-1} N(0, \tilde{A}),$$

where $\tilde{A}$ is some positive definite constant $p \times p$ matrix.

Remark 2.3. The limiting distribution for the case $\kappa > 4$ in the above theorem is not stated in ZL, but is immediate by noting that $(\hat{\phi}-\phi_0) = (\sum_{t=p+1}^n Y_{t-1}Y_{t-1}')^{-1}(\sum_{t=p+1}^n Y_{t-1}\varepsilon_t)$ and by an application of a CLT for martingales to the quantity $n^{-1/2}\sum_{t=p+1}^n Y_{t-1}\varepsilon_t$.

The above theorem states the rate of consistency of the OLS estimator as well as its limiting distribution. Notice that the estimator is inconsistent for $\kappa \in (0, 2)$, and we will throughout focus on the case $\kappa \geq 2$, which includes the possibility that ε_t has infinite variance ($\kappa = 2$). Due to the symmetry of ε_t and y_t, we have that the skewness and location parameters of the elements of $Z_{\kappa/2}$ are equal to zero, and hence that each element of $Z_{\kappa/2}$ has a symmetric stable distribution with index $\kappa/2$. This is stated in the following lemma that will be essential for making inference based on the two-sided t-test.

Lemma 2.4. *Suppose that the assumptions of Theorem 2.2 hold. For $\kappa \geq 2$, each marginal of the limiting distribution of $a_n^{(\kappa)}(\hat{\phi} - \phi_0)$, stated in Theorem 2.2, is mixed Gaussian with zero median.*

Proof. Note that $(\hat{\phi} - \phi_0) = (\sum_{t=p+1}^n Y_{t-1}Y_{t-1}')^{-1}(\sum_{t=p+1}^n Y_{t-1}\varepsilon_t)$. For $\kappa \in [2, 4)$ is $Z_{\kappa/2}$ is the weak limit of the suitably scaled $\sum_{t=p+1}^n Y_{t-1}\varepsilon_t$. The symmetry of ε_t implies that $Y_{t-1}\varepsilon_t$ is symmetric, and hence that $Z_{\kappa/2}$ has a symmetric stable distribution. By Samorodnitsky and Taqqu (1994, Theorem 2.1.2), $\left(\sum_{l=0}^\infty \varphi_l \varphi_{l+|i-j|}\right)^{-1} Z_1$ and $\Sigma^{-1}Z_{\kappa/2}$ have symmetric marginals. The result then follows by noting that any univariate symmetric stable distribution is mixed Gaussian (Samorodnitsky and Taqqu, 1994, Proposition 1.3.1) with zero median. For $\kappa \geq 4$ the result is immediate. □

3 Inference based on the t-statistic

We seek to test the hypothesis

$$H_0 : \phi_i = \phi_{i,0},$$

against $\phi_i \neq \phi_{i,0}$ for some $i = 1, ..., p$. This will be done by relying on a t-statistic based on $q \geq 2$ subsamples of the original sample. Specifically, let X_j and $\tau_{\phi_i=\phi_{i,0}}$ be defined as in (1.5) and (1.6), respectively. By Lemma 2.4, we have that $a^{(\kappa)}_{\lfloor n/q \rfloor} X_j$ is asymptotically mixed Gaussian, and, as will be shown below, $a^{(\kappa)}_{\lfloor n/q \rfloor} X_j$ and $a^{(\kappa)}_{\lfloor n/q \rfloor} X_k$ are asymptotically independent for $j \neq k$. This motivates an application of the following lemma due to Ibragimov and Müller (2010, Theorem 1 and comments in Section 2.2).

Lemma 3.1. *Let $(Z_j : j = 1, ..., q)$ be a sequence of $q \geq 2$ independent mixed Gaussian variables with zero median. Let*

$$\tau = \sqrt{q}\frac{\bar{Z}}{s_Z},$$

where $\bar{Z} := q^{-1}\sum_{j=1}^{q} Z_j$ and $s_Z^2 := (q-1)^{-1}\sum_{j=1}^{q}(Z_j - \bar{Z})^2 > 0$. With T_{q-1} a Student's t-distributed random variable with degrees of freedom $q-1$, let $\mathrm{cv}_q(\alpha)$ satisfy $P(|T_{q-1}| > \mathrm{cv}_q(\alpha)) = \alpha$. Then if $\alpha \leq 5\%$,

$$P(|\tau| > \mathrm{cv}_q(\alpha)) \leq P(|T_{q-1}| > \mathrm{cv}_q(\alpha)) = \alpha.$$

As pointed out by Ibragimov and Müller (2010), the result holds for $\alpha \leq 2\Phi(-\sqrt{3}) = 0.08326...$ where Φ is the cdf of the standard normal distribution. Moreover, the result does also hold for $q \in \{2, ..., 14\}$ if $\alpha \leq 10\%$ and $q \in \{2, 3\}$ if $\alpha \leq 20\%$. We will throughout focus on the case $\alpha \leq 5\%$.

The following lemma contains sufficient conditions for asymptotic independence between the normalized subsample estimators, $a^{(\kappa)}_{\lfloor n/q \rfloor} X_j$ and $a^{(\kappa)}_{\lfloor n/q \rfloor} X_k$ for $j \neq k$.

Lemma 3.2. *Suppose that Assumption 2.1 holds and that the process (y_t) is β-mixing. With X_j defined in (1.5) and $a^{(\kappa)}_n$ defined in Theorem 2.2, for $\kappa \geq 2$, $a^{(\kappa)}_{\lfloor n/q \rfloor} X_j$ and $a^{(\kappa)}_{\lfloor n/q \rfloor} X_k$* are asymptotically independent for $j, k = 1, .., q$, with $j \neq k$.

Remark 3.3. The lemma relies on assuming that (y_t) is β-mixing. In Section 5 we give sufficient conditions for this property to hold. These conditions impose additional smoothness restrictions on the functions b and c driving h_t in (1.2). We emphasize that the conditions are sufficient, and we conjecture that they can be relaxed. Moreover, it might be possible to relax the assumption about β-mixing. This assumption is used for making a coupling argument in the proof of Lemma 3.2 below, which might be adapted to e.g. strongly mixing process. We refer to Chapter 5 of Rio (2017) for more details on mixing processes and coupling.

Proof. Without loss of generality we may assume that $p = 1$ and $q = 2$ such that $\phi = \phi_1$ and $Y_{t-1} = y_{t-1}$. In light of the proof of Theorem 2.1 in ZL, it suffices to show that $\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2}^{\lfloor n/2\rfloor} y_{t-1}\varepsilon_t$ and $\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2\lfloor n/2\rfloor} y_{t-1}\varepsilon_t$ are asymptotically independent, where $\tilde a_n = n^{2/\kappa}$ if $\kappa\in[2,4)$, $\tilde a_n=\sqrt{n\log(n)}$ if $\kappa=4$, and $\tilde a_n=\sqrt{n}$ for $\kappa>4$. Due to the Cramér-Wold device, the asymptotic independence holds, if we show that for any $(k_1,k_2)\in\mathbb{R}^2$, $k_1\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2}^{\lfloor n/2\rfloor} y_{t-1}\varepsilon_t + k_2\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2\lfloor n/2\rfloor} y_{t-1}\varepsilon_t \xrightarrow{w} k_1 Z_{2/\kappa}^{(1)}+k_2 Z_{2/\kappa}^{(2)}$ where $Z_{2/\kappa}^{(1)}$ and $Z_{2/\kappa}^{(2)}$ are independent and identically distributed stable random variables with index $\kappa/2\wedge 2$. Let $\tilde n := \tilde n(n)$ be an increasing sequence of positive integers satisfying $\tilde n = o(n)$ as $n\to\infty$. It holds that

$$\begin{aligned}\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2\lfloor n/2\rfloor} y_{t-1}\varepsilon_t &= \tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2+\lfloor n/2\rfloor+\tilde n} y_{t-1}\varepsilon_t + \tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=3+\lfloor n/2\rfloor+\tilde n}^{2\lfloor n/2\rfloor} y_{t-1}\varepsilon_t\\ &=: S_n^{(1)}+S_n^{(2)}.\end{aligned}$$

Note that

$$S_n^{(1)} = \frac{\tilde a_{\lfloor \tilde n\rfloor}}{\tilde a_{\lfloor n/2\rfloor}}\tilde a_{\lfloor \tilde n\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2+\lfloor n/2\rfloor+\tilde n} y_{t-1}\varepsilon_t.$$

By Lemmas 3.2 and 3.3 of ZL, $\tilde a_{\lfloor \tilde n\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2+\lfloor n/2\rfloor+\tilde n} y_{t-1}\varepsilon_t = O_p(1)$ for $\kappa\in[2,4]$. By a CLT for martingales the same property holds for $\kappa>4$. Since $\tilde a_{\lfloor \tilde n\rfloor}/\tilde a_{\lfloor n/2\rfloor} = o(1)$, we conclude that $S_n^{(1)} = o_p(1)$. Since (y_t) is β-mixing it follows by a result for exact coupling, see e.g. Theorem 5.1 of Rio (2017), that as $\tilde n\to\infty$, $S_n^{(2)} = \tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=3+\lfloor n/2\rfloor+\tilde n}^{2\lfloor n/2\rfloor} y_{t-1}^\star\varepsilon_t^\star + o_p(1)$, where $(y_t^\star : t\in\mathbb{Z})$ is a copy of $(y_t : t\in\mathbb{Z})$ and independent of $\mathcal{F}_{\lfloor n/2\rfloor} := \sigma(y_t : t\le\lfloor n/2\rfloor)$. By Lemmas 3.2 and 3.3 of Zhang and Ling (2015) (for the case $\kappa\in[2,4]$) and a CLT for martingales (for the case $\kappa>4$),

$$\begin{aligned}&k_1\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2}^{\lfloor n/2\rfloor} y_{t-1}\varepsilon_t + k_2\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2+\lfloor n/2\rfloor}^{2\lfloor n/2\rfloor} y_{t-1}\varepsilon_t\\ &= k_1\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=2}^{\lfloor n/2\rfloor} y_{t-1}\varepsilon_t + k_2\tilde a_{\lfloor n/2\rfloor}^{-1}\sum_{t=3+\lfloor n/2\rfloor+\tilde n}^{2\lfloor n/2\rfloor} y_{t-1}^\star\varepsilon_t^\star + o_p(1)\\ &\xrightarrow{w} k_1 Z_{2/\kappa}^{(1)} + k_2 Z_{2/\kappa}^{(2)}.\end{aligned}$$

Since $(\varepsilon_t^\star : t\in\mathbb{Z})$ and $\sum_{t=2}^{\lfloor n/2\rfloor} y_{t-1}\varepsilon_t$ are independent, we conclude that $Z_{2/\kappa}^{(1)}$ and $Z_{2/\kappa}^{(2)}$ are independent. □

The following theorem is immediate from Lemmas 2.4, 3.1, 3.2, and an application of the continuous mapping theorem.

Theorem 3.4. *Under the assumptions of Theorem 2.2, suppose that $\kappa\ge 2$, that (y_t) is β-mixing, and that H_0 is true. With T_{q-1} a Student's t-distributed random variable with degrees of freedom $q-1$, let $\mathrm{cv}_q(\alpha)$ satisfy $P(|T_{q-1}|>\mathrm{cv}_q(\alpha))=\alpha$.*

With $\tau_{\phi_i=\phi_{i,0}}$ defined in (1.6), if $\alpha \leq 5\%$,

$$P(|\tau_{\phi_i=\phi_{i,0}}| > \mathrm{cv}_q(\alpha)) \leq \alpha,$$

as $n \to \infty$.

The theorem states that the usual two-sided t-test, based on a fixed number of $q \geq 2$ subsamples, is asymptotically a level α test for $\alpha \leq 5\%$. The property holds uniformly over the tail index $\kappa \geq 2$. We emphasize that the test is straightforward to carry out in practice, it does not rely on any data-driven choices of number of subsamples, and it does not require any knowledge about κ (in addition to the assumption that it is not less than two). Notice that the theorem does not contain any information about the finite-sample properties of the test. These are investigated in a simulation experiment in the next section.

4 Simulation experiment

In this section we consider the finite-sample properties of the t-test in a simulation experiment. As a data-generating process (DGP), we rely on the following AR(1)-ARCH(1),

$$y_t = \phi y_{t-1} + \varepsilon_t, \tag{4.1}$$

$$\varepsilon_t = \eta_t h_t, \quad \eta_t \sim I.I.D.N(0,1), \tag{4.2}$$

$$h_t^2 = 1 + \alpha\eta_{t-1}^2 h_{t-1}^2, \quad \alpha \geq 0. \tag{4.3}$$

The tail properties of ε_t have been studied in Embrechts et al. (2012, Chapters 8.4.2-8.4.3). Specifically, whenever $\alpha > 0$ and $E[\log(\alpha\eta_t^2)] < 0$, ε_t is regularly varying with index $\kappa > 0$ satisfying $E[(\alpha\eta_t^2)^{\kappa/2}] = 1$. If in addition $|\phi| < 1$, the DGP in (4.1)-(4.3) satisfies Assumption 2.1. Moreover, it can be shown that the DGP satisfies Assumption 5.1 in the next section, which ensures that the stationary version of the DGP is β-mixing, and hence that Theorem 3.4 applies. We investigate the properties of the t-test for testing the hypothesis $H_0 : \phi = 0$ for various cases of tail heaviness for the 5% significance level. Specifically, we consider the tail indices $\kappa = 2, 3, 4$, corresponding to $\alpha = 1, (\pi^{1/3}/2), 3^{-1/2}$, respectively. We compute the empirical rejection frequencies under the null hypothesis as well as under the alternative for $\phi = 0.01, 0.02, ..., 0.5$. Similar to the simulation experiments in Ibragimov et al. (2015, Chapter 3.3) we choose $q = 2, 4, 8, 16$.

Table 1 contains the empirical rejection frequencies under H_0. Overall the rejection frequencies seem very reasonable uniformly over κ. The only situations with remarkable overrejection are for the cases with 100 observations and 16 subsamples. Figures 4.1-4.6 contain the non-size-corrected as well as size-corrected empirical power curves under the alternatives $\phi = 0.01, 0.02, ..., 0.5$. Unsurprisingly, the rejection frequency is increasing in ϕ and n. Moreover, the empirical power is increasing in q, and we see that the test based on two subsamples performs quite poorly, even for a large sample length. On the other hand, the tests based on 8 and 16 subsamples seem to have quite good finite-sample power properties. Lastly, the empirical power seems to be slightly increasing in the tail heaviness, κ.

	$\kappa = 2$			
n/q	2	4	8	16
100	0.0499	0.0575	0.0631	0.0839
500	0.0461	0.0443	0.0473	0.0543
1000	0.0446	0.0436	0.0467	0.0493
10,000	0.0403	0.0379	0.0416	0.0450
	$\kappa = 3$			
n/q	2	4	8	16
100	0.0474	0.0565	0.0636	0.0832
500	0.0475	0.0461	0.0501	0.0554
1000	0.0475	0.0450	0.0466	0.0479
10,000	0.0480	0.0442	0.0444	0.0507
	$\kappa = 4$			
n/q	2	4	8	16
100	0.0505	0.0585	0.0656	0.0848
500	0.0510	0.0463	0.0508	0.0553
1000	0.0481	0.0462	0.0461	0.0506
10,000	0.0498	0.0466	0.0488	0.0520

Table 1: Empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3). The size properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail indices $\kappa \in \{2, 3, 4\}$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

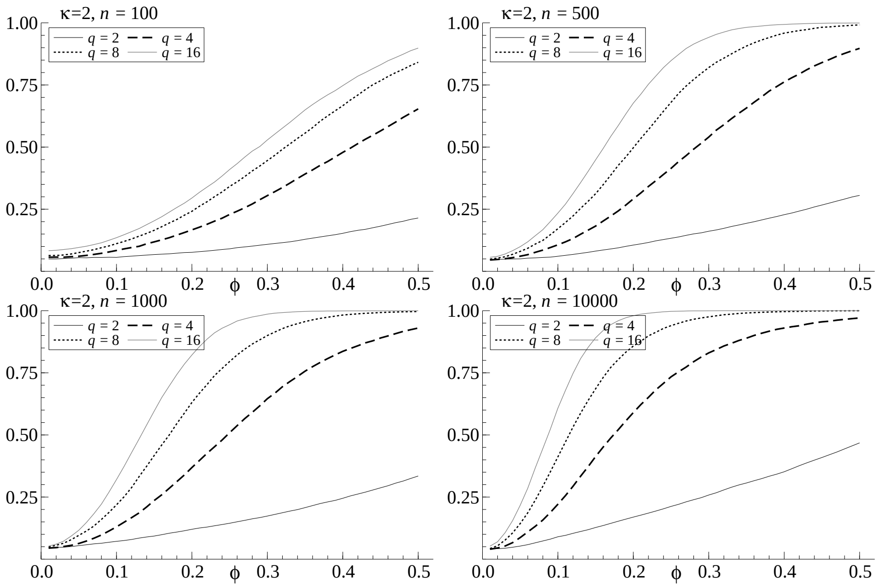

Figure 4.1: Empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 2$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

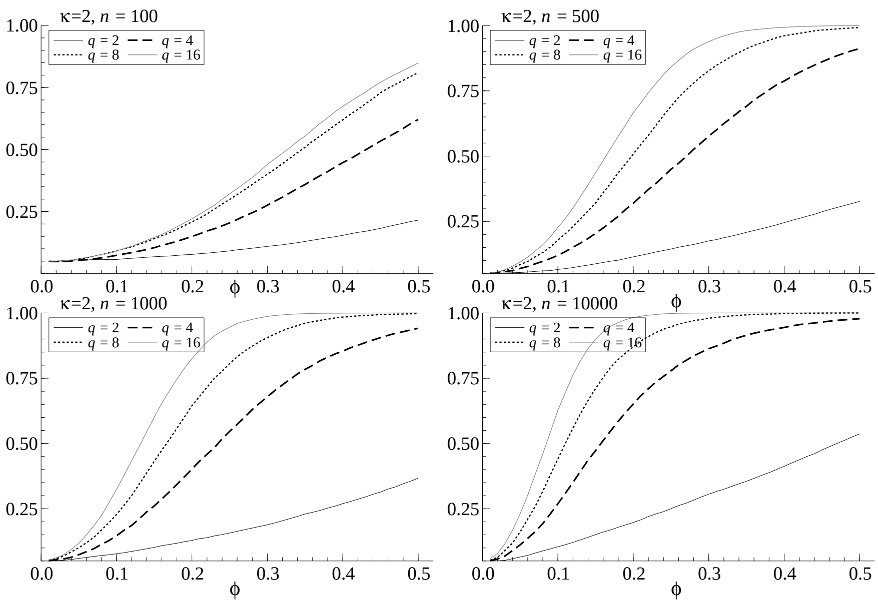

Figure 4.2: Size-corrected empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 2$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

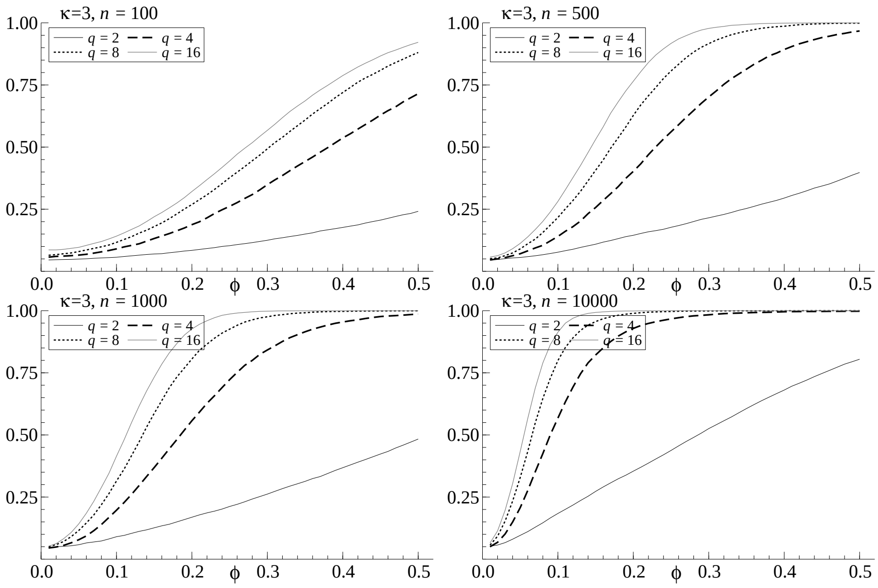

Figure 4.3: Empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 3$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

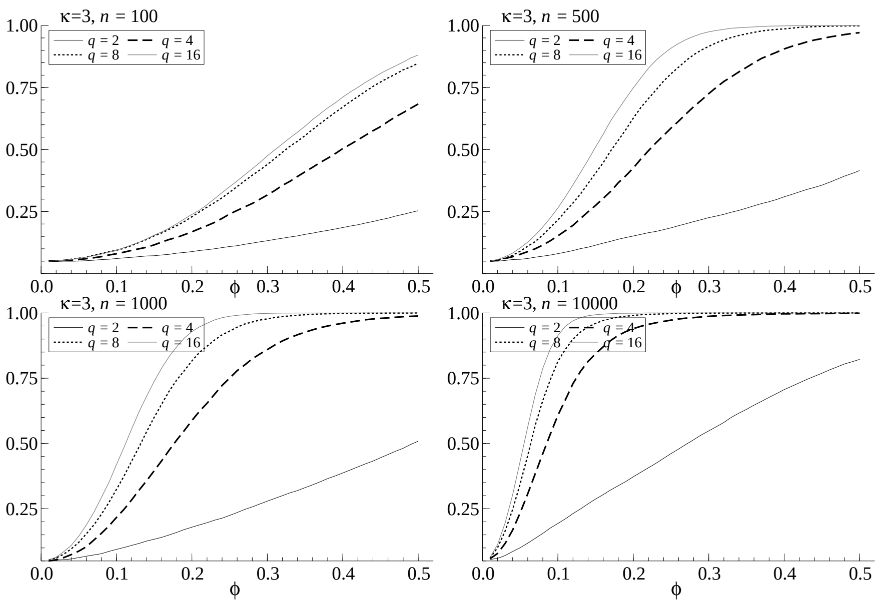

Figure 4.4: Size-corrected empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 3$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

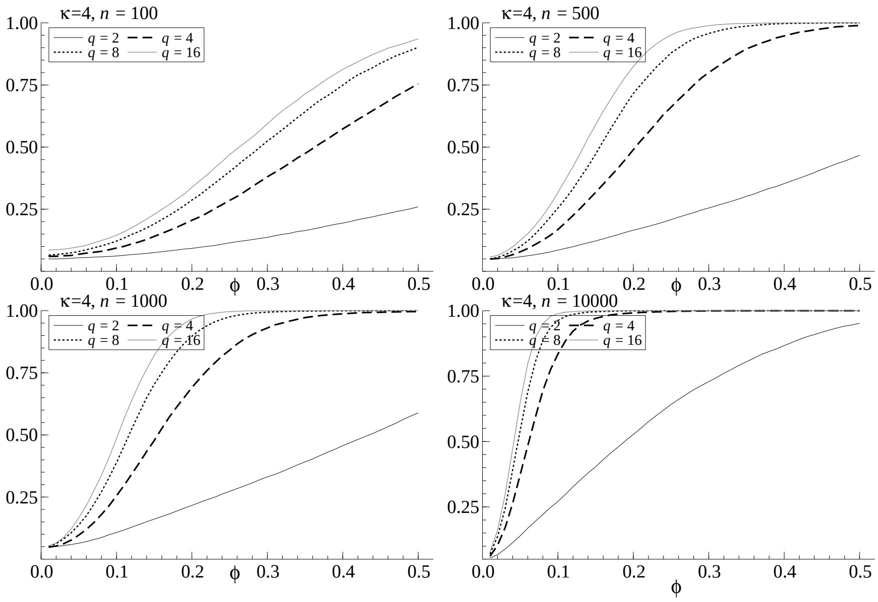

Figure 4.5: Empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 4$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

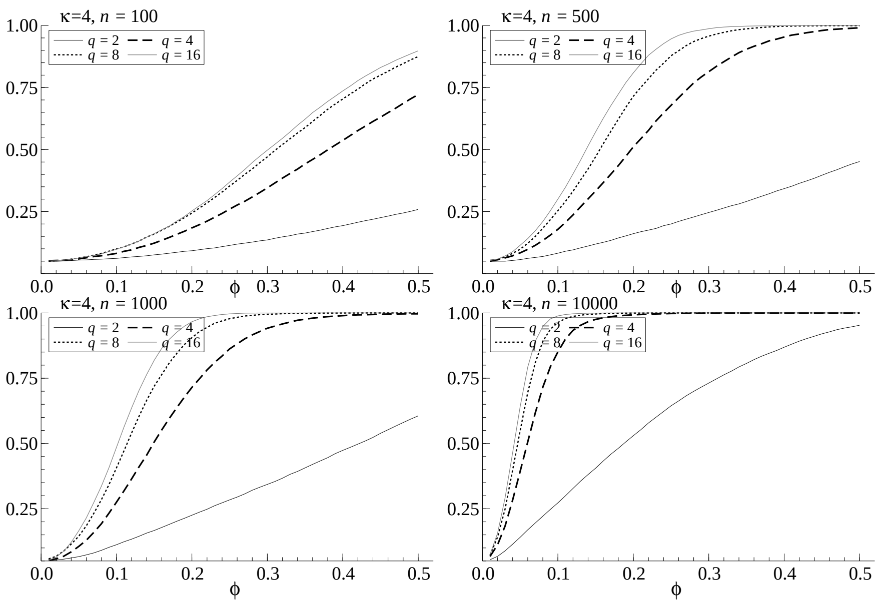

Figure 4.6: Size-corrected empirical rejection frequencies for the t-test for $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3) under the alternatives $\phi \in \{0.01, 0.02, ..., 0.5\}$. The power properties is for the 5% nominal level with sample length $n \in \{100, 500, 1000, 10000\}$, number of subsamples $q = \{2, 4, 8, 16\}$, and for tail index $\kappa = 4$. Based on 10,000 Monte Carlo replications and burn-in periods of 1000 observations.

5 Sufficient conditions for β-mixing

We now state sufficient conditions for the process (y_t) being β-mixing. This relies on applying results for Markov chains, due to Meitz and Saikkonen (2008) (MS hereafter). Define

$$Z_t := (y_t, .., y_{t-p}, h_t)' \in \mathcal{Z} := \mathbb{R}^{p+1} \times \mathbb{R}_{++},$$

where $\mathbb{R}_{++} := \{x \in \mathbb{R} : x > 0\}$, and let $g : \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}_{++}$ satisfy

$$g(\varepsilon, h) = [b(\varepsilon/h^{1/2}) + c(\varepsilon/h^{1/2})h^{\delta/2}]^{2/\delta}. \tag{5.1}$$

Noting that $\varepsilon_t = y_t - \sum_{i=1}^{p} \phi_i y_{t-i}$, we have that $h_t = g(\varepsilon_{t-1}, h_{t-1})$. We define the function $h : \mathcal{Z} \to \mathbb{R}_{++}$ such that $h_t = h(Z_{t-1}) = g(y_{t-1} - \sum_{i=1}^{p} \phi_i y_{t-1-i}, h_{t-1})$. Then define the function $F : \mathcal{Z} \times \mathbb{R} \to \mathcal{Z}$ such that

$$Z_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p} \\ h_t \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{p} \phi_i y_{t-i} \\ y_{t-1} \\ \vdots \\ y_{t-p} \\ h(Z_{t-1}) \end{bmatrix} + \begin{bmatrix} h^{1/2}(Z_{t-1})\eta_t \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = F(Z_{t-1}, \eta_t). \tag{5.2}$$

Clearly, (Z_t) is a Markov chain on $\mathcal{Z}$. In the following we show that the chain is geometrically ergodic in the sense of Liebscher (2005, Definition 1). This ensures that the stationary version of the chain is β-mixing. In addition to Assumption 2.1, we make the following assumptions.

Assumption 5.1.

1. *The distribution of η_t has a Lebesgue density which is positive and lower semi-continuous on $\mathbb{R}$.*

2. *The functions $b, c \in \mathcal{C}^\infty$, i.e. all their derivatives are continuous on $\mathbb{R}$. The function b satisfies $\inf_{x\in\mathbb{R}} b(x) > 0$ and $\sup_{x\in\mathbb{R}} b(x) < \infty$. The function c satisfies $\lim_{x\to\infty} c(x) = \infty$.*

3. *If $\delta < 2$, with $\tilde{\varphi} : \mathbb{R} \to \mathbb{R}_+$ given by $\varphi(x) = 2^{2/\delta-1}c(x)^{2/\delta}$, it holds that $\varphi(0) < 1$. Moreover, there exists $r > 0$ such that $E[\tilde{\varphi}(\eta_t)^r] < 1$.*

4. *There exists $x_1 \in \mathbb{R}$ such that $b'(x_1) = c'(x_1) = 0$. For any $x_2 \in \mathbb{R}_+$ there exists $x_3 \in \mathbb{R}$ such that $b'(x_3) + c'(x_3)x_2 \neq 0$.*

The above assumptions together with Assumption 2.1 yield the following result.

Theorem 5.2. *Let Z_t satisfy (5.2) for $t = 1, 2, ...$ with some initial $Z_0 \in \mathcal{Z}$. Under Assumptions 2.1 and 5.1, the Markov chain $\{Z_t : t \in \mathbb{N}_0\}$ is Q-geometrically ergodic. If the chain is initiated from the invariant distribution, then it is β-mixing with geometric decay.*

Proof. The Q-geometric ergodicity follows by Theorem 1 of MS, provided that Assumptions 1-6 of MS hold. Assumption 1 of MS holds by Assumptions 2.1.2 and 5.1.1. Noting that the function $f : \mathbb{R}^p \to \mathbb{R}$, introduced on p.455 in MS, corresponds to $f(x) = \phi' x$, we have that Assumption 2 of MS is satisfied. Moreover, Assumption 3 of MS holds by Assumption 2.1.4 and Lemma 1 of MS. With g defined in (5.1), we have by Assumption 5.1.2 that g is smooth (i.e. it belongs to $\mathcal{C}^\infty$) and that $\inf_{(\varepsilon,h)\in\mathbb{R}\times\mathbb{R}_{++}} g(\varepsilon, h) > 0$. Hence, Assumption 4(a) of MZ is satisfied. Moreover, by Assumption 5.1.2, for any $h \in \mathbb{R}_{++}$, $\lim_{\varepsilon\to\infty} g(\varepsilon, h) = \infty$, which ensures that Assumption 4(b) of MS is satisfied. With $b := b(0) > 0$ and $c := c(0) < 1$, we have, in light of Assumption 5.1.2, that the sequence $(h_k : k = 1, 2, ...)$ defined by $h_k = g(0, h_{k-1})$ converges to $[a/(1-b)]^{2/\delta}$ for any $h_0 \in \mathbb{R}_{++}$. This gives that Assumption 4(c) of MS is satisfied. For $\delta \geq 2$, $g(h^{1/2}\eta_t, h) \leq \bar{b} + \varphi(\eta_t)h$ where $\bar{b} := \sup_{x\in\mathbb{R}} b(x)^{2/\delta} < \infty$ and $\varphi(x) = c(x)^{2/\delta}$ with $\varphi(0) < 1$, since $c(0) < 1$. Likewise, for $\delta < 2$, $g(h^{1/2}\eta_t, h) \leq \tilde{b} + \tilde{\varphi}(\eta_t)h$ where $\tilde{b} := 2^{2/\delta-1}\sup_{x\in\mathbb{R}} b(x)^{2/\delta} < \infty$ and $\tilde{\varphi}(x) = 2^{2/\delta-1}c(x)^{2/\delta}$ with $\tilde{\varphi}(0) < 1$, by Assumption 5.1.3. Hence Assumption 4(d) of MS is satisfied. Turning to Assumption 5 of MS, for the case $\delta \geq 2$, we have that Assumption 2.2.1 and the fact that there exists $\kappa > 0$ such that $E[(c(\eta_t))^{\kappa/\delta}] = 1$ imply that there exists $r > 0$ such that $E[(c(\eta_t))^{r2/\delta}] < 1$. Hence Assumption 5 of MS is satisfied if $\delta \geq 2$. If $\delta < 2$, Assumption 5 of MS holds by Assumption 5.1.3. Assumption 6 of MS holds by Assumption 5.1.4 and the comments on p. 460 of MS. The β-mixing holds by Proposition 2 of Liebscher (2005). □

References

Buraczewski, D., E. Damek, and T. Mikosch (2016): *Stochastic Models with Power-Law Tails: The Equation $X = AX + B$*, Springer Series in Operations Research and Financial Engineering, Springer International Publishing.

Cavaliere, G., I. Georgiev, and A. M. R. Taylor (2016a): "Sieve-based inference for infinite-variance linear processes," *The Annals of Statistics*, 44, 1467–1494.

——— (2016b): "Unit root inference for non-stationary linear processes driven by infinite variance innovations," *Econometric Theory*, forthcoming.

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

Davis, R. A. and T. Hsing (1995): "Point process and partial sum convergence for weakly dependent random variables with infinite variance," *The Annals of Probability*, 23, 879–917.

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

Davis, R. A. and W. Wu (1997): "Bootstrapping M-estimates in regression and autoregression with infinite variance," *Statistica Sinica*, 7, 1135–1154.

Embrechts, P., C. Klüppelberg, and T. Mikosch (2012): *Modelling Extremal Events: For Insurance and Finance*, Applications of Mathematics, Heidelberg: Springer, 4 ed.

Ibragimov, M., R. Ibragimov, and J. Walden (2015): *Heavy-Tailed Distributions and Robustness in Economics and Finance*, Lecture Notes in Statistics, Heidelberg: Springer.

Ibragimov, R. and U. K. Müller (2010): "t-statistic based correlation and heterogeneity robust inference," *Journal of Business & Economic Statistics*, 28, 453–468.

——— (2016): "Inference with few heterogeneous clusters," *The Review of Economics and Statistics*, 98, 83–96.

Kesten, H. (1973): "Random difference equations and renewal theory for products of random matrices," *Acta Mathematica*, 131, 207–248.

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

Liebscher, E. (2005): "Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes," *Journal of Time Series Analysis*, 26, 669–689.

Loretan, M. and P. C. B. Phillips (1994): "Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets," *Journal of Empirical Finance*, 1, 211–248.

Meitz, M. and P. Saikkonen (2008): "Stability of nonlinear AR-GARCH models," *Journal of Time Series Analysis*, 29, 453–475.

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

Rio, E. (2017): *Asymptotic Theory of Weakly Dependent Random Processes*, Berlin: Springer-Verlag.

Samorodnitsky, G. and M. Taqqu (1994): *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Stochastic modeling, New York, NY: Chapman & Hall.

Zhang, R. and S. Ling (2015): "Asymptotic inference for AR models with heavy-tailed G-GARCH noises," *Econometric Theory*, 31, 880–890.