Robust inference in conditionally heteroskedastic autoregressions

Pedersen, Rasmus Søndergaard

Published in:
Econometric Reviews

Publication date:
2019

Document version
Peer reviewed version

Citation for published version (APA):
Pedersen, R. S. (2019). Robust inference in conditionally heteroskedastic autoregressions. *Econometric Reviews*.

Robust inference in conditionally heteroskedastic autoregressions

Rasmus Søndergaard Pedersen*

University of Copenhagen and Danish Finance Institute

December 17, 2018

Abstract

We consider robust inference for an autoregressive parameter in a stationary linear autoregressive model with GARCH innovations. As the innovations exhibit GARCH, they are by construction heavy-tailed with some tail index κ. This implies that the rate of convergence 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 based on least squares estimates for a fixed number of groups of the original sample. The merit of this approach is that no knowledge about the value of κ nor about the rate of convergence and the limiting distribution of the least squares estimator is required. We verify that the two-sided t-test is asymptotically a level α test whenever $\alpha \leq 5\%$ for any $\kappa \geq 2$, 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.

*I thank Esfandiar Maasoumi (the editor), an anonymous associate editor, and three referees for very constructive comments and suggestions that led to a much improved paper. I thank Philipp Ketz and Olivier Wintenberger for valuable discussions. I gratefully acknowledge financial support from the Carlsberg Foundation.
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`

1 Introduction

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

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

where ε_t follows a general GARCH(1,1) (GGARCH) process, proposed by He and Teräsvirta (1999),

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

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 consider a robust method for testing a hypothesis about an element of the vector $\phi := (\phi_1, ..., \phi_p)' \in \mathbb{R}^p$. Our test exploits the asymptotic properties of the OLS estimator for ϕ given by

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

where $Y_t = (y_t, ..., y_{t-p+1})'$ and n is the length of the sample. 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 is regularly varying with some tail index $\kappa > 0$. As recently demonstrated by ZL, the value of κ determines the rate of convergence as well as the limiting distribution of the (suitably scaled and normalized) OLS estimator. Specifically, 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), Davis and Mikosch (1998), and Mikosch and Stărică (2000). As pointed out by Lange (2011, Remark 3), we do not have an expression for the dispersion parameters or for the dependence structure of the stable vector. Importantly, our test is robust in the sense that we are able to make inference about ϕ_i without requiring any knowledge about (or estimation of) the index κ, the limiting distribution, or the rate of convergence of the OLS estimator.

We show that, under suitable conditions, each element of the OLS estimator has a mixed Gaussian distribution. This property allows us to apply a two-sided t-

statistic based on a fixed number of groups of the sample, as considered by Ibragimov and Müller (2010, 2016) and Ibragimov et al. (2015, Chapter 3.3). Specifically, we split our original sample of size n into a *fixed* number, $q \geq 2$, of equi-sized groups $(y_t : t = 1 + (i-1)\lfloor n/q \rfloor, ..., i\lfloor n/q \rfloor)$, $i = 1, ..., q$, where $\lfloor x \rfloor$ denotes the integer part of $x \in \mathbb{R}$. For each group we obtain the OLS estimator for ϕ,

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

and let

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

in order to obtain the t-statistic based on q "observations",

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

where $\bar{X} := q^{-1}\sum_{j=1}^{q} X_j$ and $s_X^2 := (q-1)^{-1}\sum_{j=1}^{q}(X_j - \bar{X})^2$. Let T_{q-1} be a random variable with a Student's t-distribution with $q-1$ degrees of freedom, and let $\text{cv}_q(\alpha)$ satisfy $P(|T_{q-1}| > \text{cv}_q(\alpha)) = \alpha$ for some $\alpha \leq 2\Phi(-\sqrt{3}) = 0.08326...$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution. We show that whenever $\kappa \geq 2$ (which is the region of the tail index for which the OLS estimator is consistent for ϕ), $\limsup_{n\to\infty} P(|\tau_{\phi_i=\phi_{i,0}}| > \text{cv}_q(\alpha)) \leq \alpha$ under H_0. Hence the two-sided group-based t-test - that does not require any knowledge about κ, the rate of convergence, or the limiting distribution of $\hat{\phi}^{(j)}$ - is asymptotically a level α test for any choice of $\alpha \leq 0.08326...$, which includes the most commonly used $\alpha = 5\%$. A simulation study shows that the robust group-based test has attractive finite-sample size and power properties, superior to those of alternative HAC-based and (infeasible) asymptotic tests. We are not aware of any other robust methods with appealing finite-sample properties in the context of AR-GARCH models that do not require any knowledge or estimation of κ, the rate of convergence, or the asymptotic distribution.

It is by now well-established that many economic and financial time series exhibit heavy-tail behavior, large downfalls, nonlinear dependence, and volatility clustering, see e.g. Loretan and Phillips (1994), Cont (2001), and Ibragimov et al. (2015) and the references therein. Such features of the data may invalidate standard statistical methods. Specifically, as mentioned, standard Gaussian-based $\sqrt{n}$-asymptotics of the OLS estimator break down in the case of non-linearities and heavy-tailedness, such as in the AR-GARCH model in (1.1)-(1.2), even in the case of finite variances where $\kappa \in (2,4)$. The property that $\kappa \in (2,4)$, where variances are finite but fourth moments are infinite, are typically found in returns on stocks and exchange rates in developed markets, whereas issues with infinite variances ($\kappa \leq 2$) are more prominent in such series in the case of emerging markets (Ibragimov et al., 2015, Sections 1.2 and 3.2). 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. (2016). In terms of dependent heavy-tailed innovations, Mikosch and Stărică (2000), Lange (2011), and ZL have investigated the properties of the least squares estimator. More recently, Cavaliere et al. (2018) have considered bootstrap inference in non-stationary linear time series with innovations driven by a heavy-tailed linear process.

In order to show that the two-sided t-test based on (1.6) is asymptotically valid, we establish that the suitably scaled X_j, defined in (1.5), is asymptotically mixed Gaussian and that X_j and X_k are asymptotically independent for $j \neq k$. Our theoretical results are derived under conditions that are in line with the assumptions in ZL under an additional restriction on h_t. In particular, we assume that the distribution of z_t is symmetric and we impose a technical Assumption 2.4 (given below) that holds in the case where h_t does not depend on the sign of past values of y_t, i.e. in the absence of leverage effects.[1]

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 group-based 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 group estimators in (1.5) are asymptotically independent. Section 6 provides some concluding remarks.

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, $\int_0^\infty \phi\left((x-\mu)/\sigma\right) dF\left(\sigma\right)$, where $\phi(\cdot)$ is the standard normal pdf, and $F(\cdot)$ is an arbitrary cdf on $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$. A random vector Y is said to have a symmetric distribution if Y and $-Y$ have the same distribution. 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 $\lim_{x\to\infty} f(x)/g(x) = 1$. With $\|\cdot\|$ the Euclidean norm,

[1]As kindly pointed out by a referee, the symmetry condition may not hold in the context of financial stock returns, whereas it is typically found that foreign exchange rate returns are much more symmetric, see e.g. Cont (2001).

let $\mathbb{S}^{p-1} = \{x \in \mathbb{R}^p : \|x\| = 1\}$. Lastly, for $x \in \mathbb{R}$, $\operatorname{sign}(x) = 1$ if $x > 0$, $\operatorname{sign}(x) = -1$ if $x < 0$ and $\operatorname{sign}(x) = 0$ if $x = 0$.

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.[2]

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 (see e.g. Lemma 4.2.(3) of Jessen and Mikosch, 2006) 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

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

Likewise (under Assumption 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

[2] 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.

implies that the process in (1.1)-(1.2) has a strictly stationary and ergodic solution satisfying

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

where the sum converges absolutely with probability one. We will assume throughout that Assumption 2.1 is satisfied such that the process (y_t) is stationary and ergodic. Moreover, as in ZL, we 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 which is due to ZL.

Theorem 2.2 (Theorem 2.1 of ZL). *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 $\kappa = 2$,*

$$a_n^{(\kappa)}(\hat{\phi} - \phi_0) \xrightarrow{w} \left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|} \right)_{i,j=1,\ldots,p}^{-1} Z_1,$$

where $\left(\sum_{l=0}^{\infty} \varphi_l \varphi_{l+|i-j|}\right)_{i,j=1,\ldots,p}$ is a $p \times p$ matrix and Z_1 is a p-dimensional 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 $\kappa \geq 4$,

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

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

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 convergence of the OLS estimator as well as its limiting distribution. Note 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 ($\kappa = 2$).

In order to make inference based on the group-based two-sided t-test, it is essential that each element of the limiting random vector $Z_{\kappa/2}$ is mixed Gaussian for $\kappa \in [2, 4)$. In order to ensure this, we make the following assumption.

Assumption 2.4. *For $\kappa \in [2, 4)$, for any $u \in \mathbb{S}^{p-1}$, $\mathrm{sign}(u'Y_{t-1})$ and h_t are independent.*

Remark 2.5. The assumption is sufficient for $Y_{t-1}\varepsilon_t$ being symmetric, which is important for the proof of Lemma 2.6 below. The assumption imposes additional restrictions on h_t. In particular, the assumption holds in the case where h_t does not depend on the sign of lagged values of y_t, and hence when there are no leverage effects. As an example, consider the case where $p = 1$ and $\phi_1 = 0$ and where h_t has a GJR-type specification, $h_t^2 = b + c_- 1_{(\eta_{t-1}<0)}\eta_{t-1}^2 h_{t-1}^2$ where $b, c_- > 0$. Here h_t depends on $\mathrm{sign}(\eta_{t-1}) = \mathrm{sign}(y_{t-1})$, which violates Assumption 2.4. In a simulation experiment in Section 4, we consider the size properties the group-based test when Assumption 2.4 is violated. The simulations indicate that the assumption is important in order for the test to control size.

We obtain the following lemma.

Lemma 2.6. *Suppose that the assumptions of Theorem 2.2 and Assumption 2.4 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 $\kappa \in [2, 4)$ $Z_{\kappa/2}$ is the weak limit of the suitably scaled $\sum_{t=p+1}^n Y_{t-1}\varepsilon_t$. Note that $Y_{t-1}\varepsilon_t = Y_{t-1}h_t\eta_t$. For any $u \in \mathbb{S}^{p-1}$ and any $x \in \mathbb{R}$, $P(u'Y_{t-1}h_t \leq x) = P(|u'Y_{t-1}|h_t \leq x, \mathrm{sign}(u'Y_{t-1}) = 1) + P(-|u'Y_{t-1}|h_t \leq x, \mathrm{sign}(u'Y_{t-1}) = -1)$, as the events $\{\mathrm{sign}(u'Y_{t-1}) = 1\}$ and $\{\mathrm{sign}(u'Y_{t-1}) = -1\}$ are disjoint, $P(u'Y_{t-1} = 0) = 0$, and h_t is strictly positive almost surely. Under Assumption 2.1 Y_{t-1} is symmetric, which by Zuo and Serfling (2000, Lemma 2.1) implies that $u'Y_{t-1}$ is symmetric. Hence, $|u'Y_{t-1}|$ and $\operatorname{sign}(u'Y_{t-1})$ are independent. The symmetry of $u'Y_{t-1}$ and Assumption 2.4 imply that $P(u'Y_{t-1}h_t \leq x) = P(|u'Y_{t-1}|h_t \leq x)P(\operatorname{sign}(u'Y_{t-1}) = 1) + P(-|u'Y_{t-1}|h_t \leq x)P(\operatorname{sign}(u'Y_{t-1}) = -1) = P(|u'Y_{t-1}|h_t \leq x)/2 + P(-|u'Y_{t-1}|h_t \leq x)/2$. By similar arguments, we obtain that $P(-u'Y_{t-1}h_t \leq x) = P(|u'Y_{t-1}|h_t \leq x)/2 + P(-|u'Y_{t-1}|h_t \leq x)/2$, and we conclude that $u'Y_{t-1}h_t$ is symmetric, which implies that $Y_{t-1}h_t$ is symmetric by Zuo and Serfling (2000, Lemma 2.1). Since $Y_{t-1}h_t$ and η_t are independent and since η_t is symmetric, we have 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. $\square$

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 $q \geq 2$ groups 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 2.6, 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).

Lemma 3.1 (Ibragimov and Müller 2010, Theorem 1 and the comments thereafter)**.** *Let $(Z_j : j = 1, ..., q)$ be a collection of $q \geq 2$ independent mixed Gaussian variables with zero median. Let*

$$\tau = \sqrt{q}\frac{\bar{Z}}{s_Z}, \tag{3.1}$$

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 $P(|T_{q-1}| > \text{cv}_q(\alpha)) = \alpha$. *Then if* $\alpha \leq 5\%$,

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

As pointed out by Ibragimov and Müller (2010), the result holds for $\alpha \leq 2\Phi(-\sqrt{3}) = 0.08326...$ where $\Phi(\cdot)$ is the cdf of the standard normal distribution. Moreover, the result holds for $q \in \{2, ..., 14\}$ if $\alpha \leq 10\%$ and $q \in \{2, 3\}$ if $\alpha \leq 20\%$. Throughout we focus on the most commonly used case $\alpha \leq 5\%$. Note that the idea is to use the above lemma in an asymptotic sense, as the weak limit of $\tau_{\phi_i=\phi_{i,0}}$ in (1.6) is of the form (3.1); see Ibragimov and Müller (2010, Section 2.2) for a discussion of the asymptotic applicability of the lemma.

The following lemma contains sufficient conditions for asymptotic independence between the normalized group estimators, $a^{(\kappa)}_{\lfloor n/q \rfloor} X_j$ and $a^{(\kappa)}_{\lfloor n/q \rfloor} X_k$ for $j \neq k$.[3]

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. The 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 processes. We refer to Chapter 5 of Rio (2017) for more details on mixing processes and coupling. In Assumption 5.1 in Section 5 we give sufficient conditions for β-mixing. 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 may in some cases be relaxed. As an example, consider the case $\delta = 1$, $p = 1$ and $\phi = 0$ where $h_t = b + \gamma|\eta_{t-1}|h_{t-1} = b + \gamma|y_{t-1}|$, $b, \gamma > 0$. Then Assumption 5.1.2 in Section 5 is violated, as $c(x) = \gamma|x|$ is not differentiable. However, it can be shown that y_t has a β-mixing stationary solution for suitable values of γ, see e.g. Kristensen and Rahbek (2005, Section 2).

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 $\tilde{a}^{-1}_{\lfloor n/2 \rfloor} \sum_{t=2}^{\lfloor n/2 \rfloor} y_{t-1}\varepsilon_t$ and $\tilde{a}^{-1}_{\lfloor n/2 \rfloor} \sum_{t=2+\lfloor n/2 \rfloor}^{2\lfloor n/2 \rfloor} y_{t-1}\varepsilon_t$ 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 $(k_1, k_2) \in \mathbb{R}^2$, $k_1 \tilde{a}^{-1}_{\lfloor n/2 \rfloor} \sum_{t=2}^{\lfloor n/2 \rfloor} y_{t-1}\varepsilon_t + k_2 \tilde{a}^{-1}_{\lfloor n/2 \rfloor} \sum_{t=2+\lfloor n/2 \rfloor}^{2\lfloor n/2 \rfloor} y_{t-1}\varepsilon_t \xrightarrow{w} k_1 Z^{(1)}_{2/\kappa} + k_2 Z^{(2)}_{2/\kappa}$ where $Z^{(1)}_{2/\kappa}$ and $Z^{(2)}_{2/\kappa}$ are independent and identically distributed stable random

[3] Let $(X_n : n \in \mathbb{N})$ and $(Y_n : n \in \mathbb{N})$ be two sequences of random vectors. Then X_n and Y_n are said to be asymptotically independent, if $(X_n, Y_n) \xrightarrow{w} (X, Y)$ as $n \to \infty$ and X and Y are independent.

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 ZL (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. $\square$

Note that by Lemmas 2.6 and 3.2 and an application of the continuous mapping theorem, with $\tau_{\phi_i=\phi_{i,0}}$ defined in (1.6), $\tau_{\phi_i=\phi_{i,0}} \xrightarrow{w} \tau := \sqrt{q}\bar{Z}/s_Z$ where $\bar{Z} := q^{-1}\sum_{j=1}^q Z_j$, $s_Z^2 := (q-1)^{-1}\sum_{j=1}^q (Z_j - \bar{Z})^2$ and $(Z_j : j = 1, ..., q)$ is a collection of independent, mixed Gaussian random variables with zero median. Hence, $\limsup_{n\to\infty} P(|\tau_{\phi_i=\phi_{i,0}}| > \mathrm{cv}_q(\alpha) = P(|\tau| > \mathrm{cv}_q(\alpha) \le P(|T_{q-1}| > \mathrm{cv}_q(\alpha)) = \alpha$, where the inequality holds by Lemma 3.1. We obtain the following theorem.

Theorem 3.4. *Under the assumptions of Theorem 2.2 and Assumption 2.4, 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* $\alpha \leq 5\%$,

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

The theorem states that the usual two-sided t-test, based on a fixed number of $q \geq 2$ groups, is asymptotically a level α test for $\alpha \leq 5\%$. The property holds for any value of 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 groups, and it does not require any knowledge about κ.

The group-based t-test is particularly useful for the cases where one expects that the tail index $\kappa \in [2, 4]$, such as in modelling of financial return data, as mentioned in the introduction. An alternative to the group-based approach is subsampling, which, however, is typically found to have rather poor finite-sample size and power properties, see e.g. Ibragimov et al. (2015, Section 3.3) for simulation results for inference for the mean (or location) of heavy-tailed (linearly) dependent time series models. For the case $\kappa > 4$, alternative methods exist. These methods include t-tests based on the least squares estimator and HAC-type standard errors, as well as bootstrap methods, see e.g. Gonçalves and Kilian (2004, 2007). In the next section, we investigate the finite-sample properties of the group-based t-test. The properties of the test are compared to standard t-tests based on HAC standard errors as well as to a test based on the asymptotic distribution of the OLS estimator.

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 use the following AR(1)-ARCH(1),

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

$$\varepsilon_t = \eta_t h_t, \tag{4.2}$$

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

$$\eta_t \sim i.i.d.N(0, 1). \tag{4.4}$$

The tail properties of ε_t have been studied in Embrechts et al. (2012, Chapters 8.4.2-8.4.3). Specifically, whenever $\gamma > 0$ and $E[\log(\gamma\eta_t^2)] < 0$, ε_t is regularly varying with index $\kappa > 0$ satisfying $E[(\gamma\eta_t^2)^{\kappa/2}] = 1$. 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 group-based t-test for testing the hypothesis $H_0 : \phi = 0$ for various cases of tail heaviness at the 5% nominal level. Motivated by empirical findings discussed in the introduction, we consider the tail indices $\kappa = 2, 3, 4$, corresponding to $\gamma = 1, \pi^{1/3}/2, 3^{-1/2}$, 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 $q = 2, 4, 8, 16$.

For comparison, we consider the performance of a two-sided t-test where ϕ is estimated by OLS and the standard error is estimated using a HAC estimator, and where the critical value is from the standard normal distribution. Specifically, and in line with Ibragimov and Müller (2010, Section 3.1), we use a quadratic spectral estimator with an automatic bandwidth selection (HAC-QA) as well as a prewhitened variance estimator with a second stage automatic bandwidth quadratic spectral kernel estimator (HAC-PW); see Ibragimov and Müller (2010, p.459) for additional details and references. We emphasize, that the standard normal distribution is potentially a poor approximation of the test statistics for the case $\kappa \in [2, 4]$, in light of Theorem 2.2 and since the consistency of HAC variance estimators is typically derived for $\kappa > 4$. Lastly, we consider the performance of the asymptotic distribution for a known κ, i.e., with $\hat{\phi}$ the OLS estimator and for a given n, we compare $|a_n^{(\kappa)}\hat{\phi}|$ with its (approximate) limiting distribution (Asymp.); see Theorem 2.2. Note that even for a known $\kappa \in [2, 4)$ the limiting stable distributions stated in Theorem 2.2 are infeasible in the sense that the spectral measure (or dispersion parameter for the case $p = 1$) of the distributions is hard to obtain from the DGP, as the distributions are stated in terms of limiting point processes. Instead we determine the critical values (at the 5 % nominal level) as the 95 percentile of 10,000 independent draws of $|a_n^{(\kappa)}\hat{\phi}|$ for $n = 50{,}000$. All simulations are based on 10,000 Monte Carlo replications and burn-in periods of 1,000 observations.

[Table 1 about here]

Table 1 contains the empirical rejection frequencies under $H_0 : \phi = 0$. Overall, the rejection frequencies for the group-based t-test seem very reasonable for any κ. The only situations with remarkable over-rejection are for the cases with 100 observations and 16 groups. One may note that in this case, the OLS estimator within each group is based on six observations, and hence the underlying assumptions of asymptotic mixed Gaussian distributions and asymptotic independence of the group estimators are potentially very poor approximations. Even for this case, the robust test has better size properties than the HAC-QA and HAC-PW t-tests that over-reject for $\kappa = 2$ as well as for all small sample sizes. This finding is in line with the simulation results of Ibragimov and Müller (2010, Table 1) for AR models with heavy-tailed linear dependent errors. The (infeasible) test based on the asymptotic distribution is conservative for all values of κ and sample sizes. Figures 1-3 contain the 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 groups performs quite poorly, even for a large sample length. On the other hand, the tests based on 8 and 16 groups seem to have quite good finite-sample power properties, and in most cases better than those of the alternative tests. Lastly, the empirical power seems to be slightly increasing in κ. The overall appealing finite-sample properties of the group-based t-test are in line with the general theory on efficiency of such tests, as discussed in detail in Ibragimov and Müller (2010, Section 4).

[Figure 1 about here]

[Figure 2 about here]

[Figure 3 about here]

As a robustness check, we consider the rejection frequencies of the tests when the innovation η_t is asymmetric, such that the assumptions of Theorem 3.4 are violated. We consider the DGP in (4.1)-(4.3) with η_t standardized skewed Student's t-distributed,

$$\eta_t \sim i.i.d.\text{SKST}(0, 1, \xi, \nu), \tag{4.5}$$

as for instance considered in Giot and Laurent (2004).[4]

For the simulations, we choose $\xi = 0.5$ and $\nu = 50$, such that the distribution of z_t is rather thin-tailed but highly asymmetric. We consider the cases $\kappa = 2, 3, 4$ for the tail index of ε_t. As $\gamma = (E[|\eta|^\kappa])^{-2/\kappa}$, we have that $\gamma = 1$ for $\kappa = 2$, $\gamma \approx 0.70254$ for $\kappa = 3$, and $\gamma \approx 0.5183$ for $\kappa = 4$. We investigate the properties of the group-based t-test as well as the HAC-based t-test. We leave out considering

[4]Specifically, η_t has a standardized skewed Student's t-distribution with skewness parameter $\xi > 0$ and degrees of freedom $\nu > 2$, if it has density

$$f(x) = \begin{cases} \frac{2s}{\xi+\xi^{-1}} g[(sx+m)\xi] & \text{if } x < -m/s, \\ \frac{2s}{\xi+\xi^{-1}} g[(sx+m)\xi^{-1}] & \text{if } x \geq -m/s, \end{cases}$$

the performance of the the t-test based on the asymptotic distribution of the OLS estimator, as the limiting distributions stated in Theorem 2.2 are derived under the assumption of symmetric η_t.

[Table 2 about here]

Table 2 contains the empirical rejection frequencies under $H_0 : \phi = 0$ with asymmetric η_t. The size-properties are comparable to those reported in Table 1 for the case of symmetric η_t. One may note that that the group-based t-tests seem to slightly over-reject for $q = 8, 16$. The performance of the HAC-based tests is qualitatively the same as in the symmetric case.

As a last robustness check, we consider the case where h_t has the following GJR-type specification,

$$h_t^2 = 1 + \gamma 1_{(\eta_{t-1}<0)} \eta_{t-1}^2 h_{t-1}^2. \tag{4.6}$$

As discussed in Remark 2.5, this specification of h_t violates Assumption 2.4 as h_t depends on the sign of y_{t-1}.[5] Similar to the previous simulations, we focus on the tail indices $\kappa = 2, 3, 4$, corresponding to $\gamma = 2, (4\pi)^{1/3}/2, (2/3)^{1/2}$, respectively.

[Table 3 about here]

Table 3 contains the empirical rejection frequencies under $H_0 : \phi = 0$ and with asymmetric h_t given in (4.6). It is apparent that the group-based t-test severely over-rejects for $\kappa = 2, 3$ and $q = 4, 8, 16$. The HAC-based tests do also seem to perform quite poorly for these values of κ. For $\kappa = 4$ the rejection frequencies are overall more reasonable, which is likely to be explained by the fact that the limiting distribution of the suitably normalized OLS estimator is Gaussian, in light

where

$$m = \frac{\Gamma\left(\frac{\nu+1}{2}\right)\sqrt{\nu-2}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi}}(\xi - \xi^{-1}),$$
$$s = \sqrt{(\xi^2 + \xi^{-2} - 1) - m^2},$$
$$g(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}}\left(1 + \frac{x^2}{\nu-2}\right)^{-\left(\frac{\nu+1}{2}\right)},$$

with $\Gamma(\cdot)$ the Gamma function. By construction $E[\eta_t] = 0$ and $E[\eta_t^2] = 1$, and it holds that $\xi^2 = P(\eta_t \geq 0)/P(\eta_t < 0)$. The distribution is symmetric if $\xi = 1$, left-skewed if $\xi < 1$, and right-skewed if $\xi > 1$.

[5]The specification in (4.6) violates Assumption 5.1.2 in Section 5, as $c(x) = \gamma 1_{(x<0)} x^2$ is not differentiable. However, it can be shown that the Markov chain $y_t = (1 + \gamma 1_{(\eta_{t-1}<0)} \eta_{t-1}^2 h_{t-1}^2)^{1/2} \eta_t = (1 + \gamma 1_{(y_{t-1}<0)} y_{t-1}^2)^{1/2} \eta_t$ has a β-mixing stationary solution for the chosen values of γ.

of Theorem 2.2. To conclude, the absence of leverage effects (i.e. that h_t does not depend on the lagged values of y_t) appears to be important for the size-control of the group-based t-test for $\kappa \in [2,4)$.

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 MS 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). $\square$

6 Concluding remarks

We have considered a robust method for testing a hypothesis about an autoregressive parameter in a general class of heavy-tailed AR-GARCH models. Importantly, the method does not require any knowledge or estimation of the tail-heaviness κ or any knowledge of the rate of convergence or the asymptotic distribution of the (OLS) estimators entering the test statistic. The method is found to have appealing finite-sample size and power properties. We are not aware of any other robust methods with such appealing finite-sample properties in the context of AR-GARCH models.

The theoretical results are derived under two symmetry conditions that impose that the stationary distribution of the data-generating process is symmetric and that essentially rule out leverage effects. An important direction for future research is to combine the present approach with some symmetrization approach such that the group-based t-test is valid under asymmetry. Specifically, this could be done by obtaining an alternative estimator that (when suitably scaled and centered) is asymptotically mixed Gaussian under asymmetry.

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 (2016): "Sieve-based inference for infinite-variance linear processes," *The Annals of Statistics*, 44, 1467–1494.

——— (2018): "Unit root inference for non-stationary linear processes driven by infinite variance innovations," *Econometric Theory*, 34, 302–348.

Cont, R. (2001): "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative Finance*, 1, 223–236.

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.

Giot, P. and S. Laurent (2004): "Modelling daily Value-at-Risk using realized volatility and ARCH type models," *Journal of Empirical Finance*, 11, 379–398.

Gonçalves, S. and L. Kilian (2004): "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 123, 89–120.

——— (2007): "Asymptotic and bootstrap inference for AR(∞) processes with conditional heteroskedasticity." *Econometric Reviews*, 26, 609–641.

He, C. and T. Teräsvirta (1999): "Properties of moments of a family of GARCH processes," *Journal of Econometrics*, 92, 173–192.

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.

Jessen, A. H. and T. Mikosch (2006): "Regularly varying functions," *Publications de l'Institute Mathématique, Nouvelle Série*, 80(94), 171–192.

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

Kristensen, D. and A. Rahbek (2005): "Asymptotics of the QMLE for a class of ARCH(q) models," *Econometric Theory*, 21, 946–961.

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.

Zuo, Y. and R. Serfling (2000): "On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry," *Journal of Statistical Planning and Inference*, 84, 55–79.

$\kappa = 2$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW	Asymp.
100	0.0490	0.0575	0.0631	0.0839	0.1468	0.1436	0.0093
500	0.0461	0.0443	0.0473	0.0543	0.1352	0.1189	0.0243
1,000	0.0446	0.0436	0.0467	0.0493	0.1346	0.1113	0.0291
10,000	0.0403	0.0379	0.0416	0.0450	0.1267	0.1042	0.0396

$\kappa = 3$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW	Asymp.
100	0.0474	0.0565	0.0636	0.0832	0.1148	0.1115	0.0092
500	0.0475	0.0461	0.0501	0.0554	0.0843	0.0742	0.0251
1,000	0.0475	0.0450	0.0466	0.0479	0.0725	0.0623	0.0309
10,000	0.0480	0.0442	0.0444	0.0507	0.0607	0.0547	0.0394

$\kappa = 4$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW	Asymp.
100	0.0505	0.0585	0.0656	0.0848	0.1021	0.0996	0.0229
500	0.0510	0.0463	0.0508	0.0553	0.0678	0.0618	0.0375
1,000	0.0481	0.0462	0.0461	0.0506	0.0581	0.0528	0.0396
10,000	0.0498	0.0466	0.0488	0.0520	0.0522	0.0514	0.0477

Table 1: Empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, the HAC-PW t-test, and the asymptotic test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.4). The tests are done at the 5% nominal level.

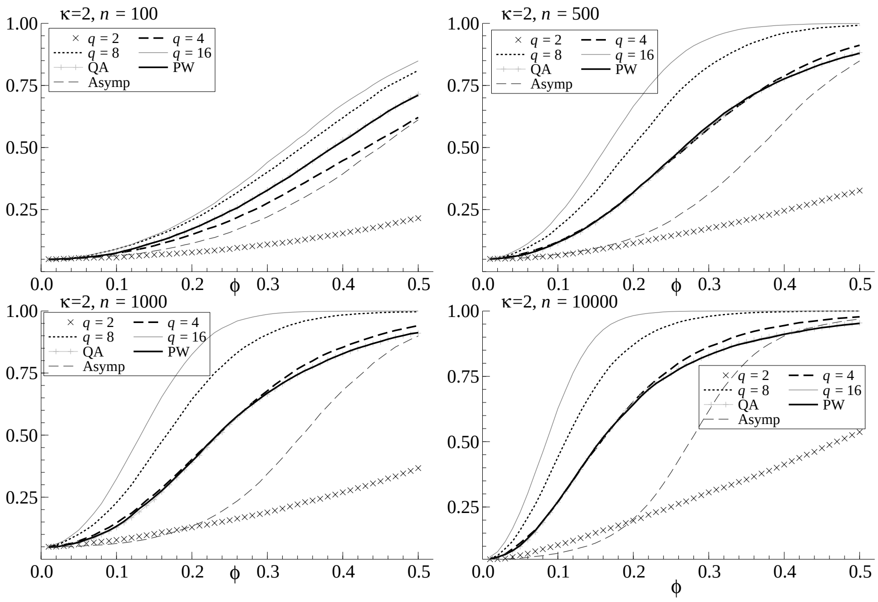

Figure 1: Size-corrected empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, the HAC-PW t-test, and the asymptotic test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.4), and the alternatives are $\phi \in \{0.01, 0.02, ..., 0.5\}$. Tail index, $\kappa = 2$.

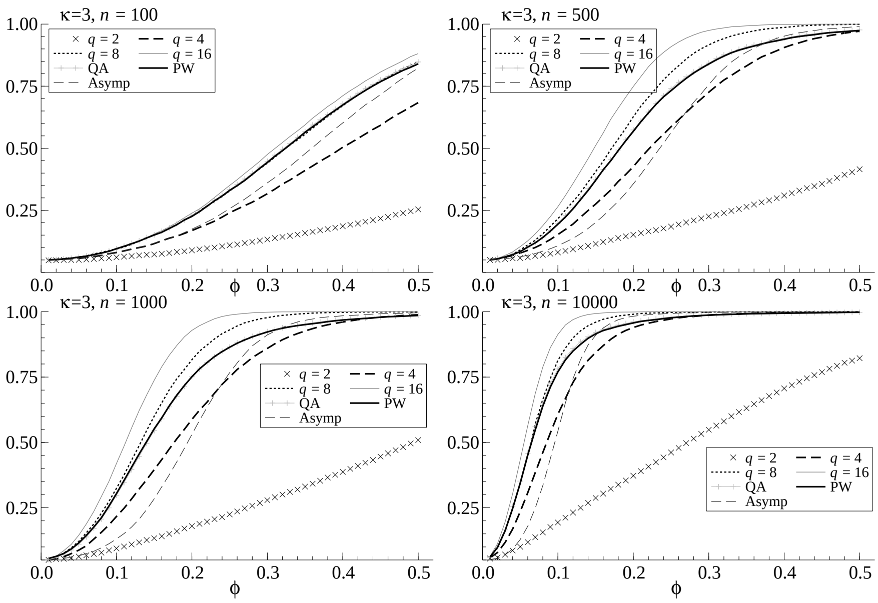

Figure 2: Size-corrected empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, the HAC-PW t-test, and the asymptotic test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.4), and the alternatives are $\phi \in \{0.01, 0.02, ..., 0.5\}$. Tail index, $\kappa = 3$.

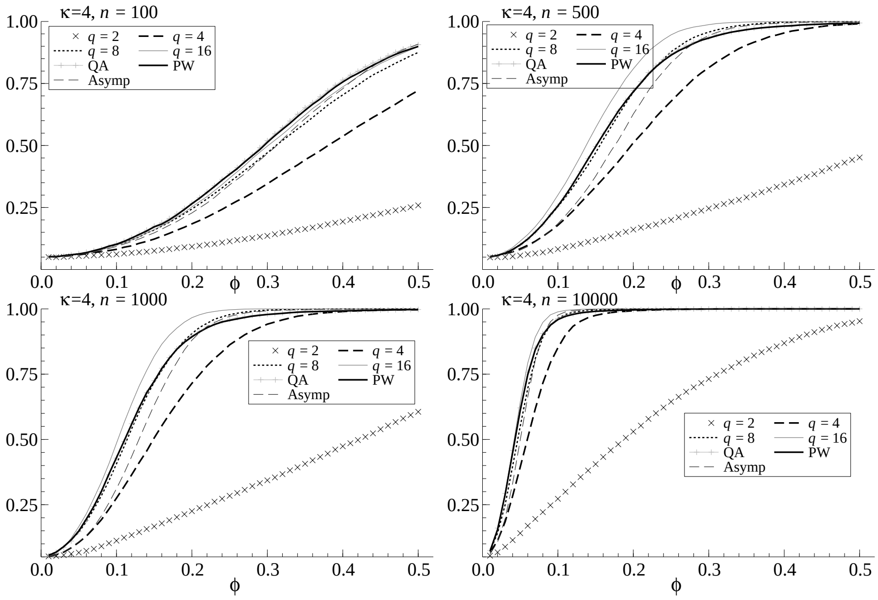

Figure 3: Size-corrected empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, the HAC-PW t-test, and the asymptotic test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.4), and the alternatives are $\phi \in \{0.01, 0.02, ..., 0.5\}$. Tail index, $\kappa = 4$.

$\kappa = 2$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0508	0.0560	0.0690	0.0857	0.1451	0.1421
500	0.0481	0.0460	0.0564	0.0614	0.1334	0.1177
1,000	0.0471	0.0448	0.0545	0.0618	0.1343	0.1154
10,000	0.0425	0.0370	0.0455	0.0544	0.1339	0.1089

$\kappa = 3$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0497	0.0560	0.0696	0.0866	0.1144	0.1128
500	0.0484	0.0466	0.0559	0.0634	0.0857	0.0764
1,000	0.0494	0.0485	0.0565	0.0617	0.0783	0.0708
10,000	0.0470	0.0427	0.0476	0.0583	0.0604	0.0568

$\kappa = 4$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0521	0.0557	0.0720	0.0888	0.1026	0.1017
500	0.0476	0.0485	0.0563	0.0627	0.0704	0.0663
1,000	0.0508	0.0494	0.0552	0.0626	0.0655	0.0637
10,000	0.0464	0.0479	0.0498	0.0555	0.0533	0.0511

Table 2: Empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, and the HAC-PW t-test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) model in (4.1)-(4.3),(4.5) with $\xi = 0.5$ and $\nu = 50$. The tests are done at the 5% nominal level.

$\kappa = 2$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0707	0.1016	0.0963	0.0941	0.1984	0.1958
500	0.0727	0.1478	0.1785	0.1875	0.2733	0.2434
1,000	0.0703	0.1518	0.2032	0.2372	0.2768	0.2455
10,000	0.0675	0.1445	0.2294	0.2902	0.2862	0.2546

$\kappa = 3$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0540	0.0732	0.0746	0.0854	0.1198	0.1209
500	0.0580	0.0779	0.0871	0.0950	0.1208	0.1093
1,000	0.0560	0.0741	0.0895	0.0988	0.1074	0.0954
10,000	0.0540	0.0731	0.0949	0.1118	0.0998	0.0830

$\kappa = 4$

n	t-statistic (q) 2	4	8	16	HAC-QA	HAC-PW
100	0.0509	0.0660	0.0688	0.0859	0.0972	0.1012
500	0.0534	0.0606	0.0652	0.0691	0.0783	0.0754
1,000	0.0499	0.0572	0.0648	0.0650	0.0699	0.0646
10,000	0.0508	0.0570	0.0616	0.0646	0.0622	0.0556

Table 3: Empirical rejection frequencies for the t-test ($q = 2, 4, 8, 16$), the HAC-QA t-test, and the HAC-PW t-test. The null hypothesis is $\phi = 0$ in the AR(1)-ARCH(1) process in (4.1),(4.2),(4.4),(4.6). The tests are done at the 5% nominal level.