Journal of Financial Econometrics, 2019, Vol. 17, No. 3, 495–515
doi: 10.1093/jjfinec/nbx037
Advance Access Publication Date: 19 December 2017
Article

OXFORD

Subsampling Inference for the Autocorrelations of GARCH Processes*

Tucker McElroy[1] and Agnieszka Jach[2]

[1]U.S. Census Bureau and [2]Hanken School of Economics

Address correspondence to A. Jach, Hanken School of Economics, 00101 Helsinki, Finland, or e-mail: agnieszka.jach@hanken.fi.

Received January 28, 2017; revised November 16, 2017; editorial decision November 16, 2017; accepted November 20, 2017

Abstract

We provide self-normalization for the sample autocorrelations of power GARCH(*p*, *q*) processes whose higher moments might be infinite. To validate the studentization, whose goal is to match the growth rate dependent on the index of regular variation of the process, we substantially extend existing weak-convergence results. Since asymptotic distributions are non-pivotal, we construct subsampling-based confidence intervals for the autocorrelations and cross-correlations, which are shown to have satisfactory empirical coverage rates in a simulation study. The methodology is further applied to daily returns of CAC40 and FTSA100 indices and their squares.

Key words: conditional heteroskedasticity, heavy tails, parameter-dependent convergence rates, self-normalization, studentization

JEL classification: C13, C14

In the exploratory analysis of time series, it is common practice to examine the sample autocorrelations (acs) of the observed data (suitably transformed to remove nonstationarity) $X_1, X_2, \ldots, X_n$, to see whether the process differs significantly from white noise. For financial time series, such as log-returns, there is also interest in studying the acs of the squared data. In order to ascertain whether a sample autocorrelation at a particular lag differs significantly from zero, it is necessary to obtain an accurate construction of the parameter's confidence interval, and this in turn requires some knowledge of the asymptotic behavior of the sample

* The authors are grateful for helpful comments from the Associate Editor and anonymous referees. This work was partially supported by the Spanish Ministry of Economy and Competitiveness [grant number ECO2012-38442 to AJ]. This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

acs of $\{X_t\}$ (or of $\{X_t^2\}$). For linear processes with finite variance, the sample acs (under mild conditions) are asymptotically normal with the asymptotic variance-covariance matrix given by the standard Bartlett's formula (Brockwell and Davis, 1991), whereas for nonlinear processes with potentially infinite variance the asymptotic behavior is much more complex.

In this article we focus on nonlinear processes, namely, the popular class of GARCH(p,q) processes $\{X_t\}$ (Bollerslev, 1986), which are widely used for modeling financial log-returns. For GARCH processes, the asymptotic normality of the sample acs with the standard $\sqrt{n}$-rate holds if $\mathbb{E}[X_t^4] < \infty$. Since the marginal distributions of GARCH processes are regularly varying with index $\kappa > 0$ (Davis and Mikosch, 1998; Mikosch and Stărică, 2000; Basrak, Davis, and Mikosch, 2002), this moment requirement is met if $\kappa > 4$. However, the asymptotic variance-covariance matrix can no longer be expressed via the standard Bartlett's formula, and instead is given via the general Bartlett's formula (Chapter 5 of Francq and Zakoian, 2010; Kokoszka and Politis, 2011). Nevertheless, this matrix can be consistently estimated (Francq and Zakoian, 2010; Kokoszka and Politis, 2011) so long as $\mathbb{E}[X_t^4] < \infty$. This is in tension with the empirical evidence suggesting that the tails of financial log-returns are heavier, having an infinite fourth moment (Mittnik, Paolella, and Rachev, 2002; Carrasco and Chen, 2002; Tully and Lusey, 2007).

Statistical inference for the acs of GARCH processes in the absence of a finite fourth moment ($2 < \kappa < 4$) is rather challenging. Under this scenario, the convergence rate of the sample acs is slower than the classical $\sqrt{n}$ and is determined by the abovementioned index κ, which depends on the model parameters and the distribution of the innovations $\{Z_t\}$ and is difficult to estimate (Wagner and Marsh, 2004; Baek et al., 2009). Closed-form expressions for κ exist only for ARCH(1) and GARCH(1,1). Additionally, if $\mathbb{E}[Z_t^4] = \infty$, estimation of model coefficients via the quasi maximum likelihood poses difficulties (nonstandard rates of convergence and non-normal asymptotics) and other methods have been proposed in the literature (Hall and Yao, 2003; Peng and Yao, 2003; Huang, Wang, and Yao, 2008). The limit distributions of sample acs when $2 < \kappa < 4$ involve the infinite-variance stable laws (Basrak, Davis, and Mikosch, 2002) and thus the asymptotic quantiles cannot be determined analytically. When $0 < \kappa < 2$, the population acs of GARCH processes are not well-defined and the sample acs are inconsistent. For similar reasons, the asymptotics and the convergence rates for the sample acs of squares of GARCH processes also show trichotomous behavior, which is subject to $\kappa > 8$, $\kappa \in (4, 8)$ and $\kappa \in (0, 4)$, respectively.

Our primary goal in the current article is to construct confidence intervals for the acs of GARCH(p,q) processes, of their squares and of the cross-correlations between the process and its squares. Recall that the acs of a GARCH process are zero, whereas the acs of squares decay with geometric rate. Moreover, as shown herein, the cross-correlations between values and squares for the GARCH process are zero whenever the marginal distributions are symmetric. Therefore, any empirical evidence against these three features will render dubious the hypothesis that the data's dynamics can be adequately captured through a GARCH model. Thus, our procedures for the construction of confidence intervals can be viewed as several misspecification tests for the GARCH hypothesis. Also in the case that the acs of the absolute values of the process are of interest, we establish convergence results for power GARCH (PGARCH) processes.

In prior literature, Kokoszka, Teyssière, and Zhang (2004) compared several resampling methods of constructing confidence intervals for lag-1-autocorrelation of squares in GARCH-type models, recommending residual bootstrap as the best approach. In contrast, we examine all the lags of the acs, and employ a nonparametric approach that combines the concepts of self-normalization and subsampling. This approach is valid irrespective of whether the asymptotic distributions for the acs are Gaussian or not (i.e., without assuming finiteness of the fourth moment), and does not require knowledge of model orders p and q. Our procedure does not involve parameter estimation, which under some scenarios can be particularly troublesome, for example, when the error distribution is heavy-tailed with an infinite fourth moment.

Self-normalization (McElroy and Politis, 2007; Jach, McElroy, and Politis, 2012) addresses the issue of parameter-dependent convergence rates and is accomplished by dividing the sample ac of $\{X_t\}$ (resp. $\{X_t^2\}$) by a quantity that correctly matches its asymptotic growth rate, without knowing a priori whether the fourth (resp. eighth) population moment is finite or not. We show that the fourth (resp. eighth) sample moment is suitable for such studentization; we also provide a studentization for the cross-correlations. The identification of suitable studentizations is nontrivial, and is a novel facet of this work. In order to validate this technique, it is necessary to substantially extend some of the weak-convergence results of Mikosch and Stărică (2000), which is a stand-alone contribution of this article.

Clearly, self-normalization can only resolve half of the problem—namely, eliminating the need to know the convergence rate to compute the studentized statistic—since the limit distributions will be unconventional and non-pivotal. Subsampling (Politis, Romano, and Wolf, 1999) can be used to empirically estimate the quantiles of the sampling distribution. This scheme operates by computing the same statistics on a small subsample—typically a contiguous stretch—drawn from the original time series data, with the unknown parameter being replaced by its best large-sample estimate. Consistency of the resulting empirical distribution for the sampling (or asymptotic) distribution is typically established through a strong mixing assumption, together with strict stationarity, which in the context of GARCH processes is immediate.

The article is organized as follows: Section 1 develops the statistical methodology of self-normalization for sample acs, as well as other unobserved quantities. While these results are not of direct statistical applicability, they are necessary for establishing the subsequent subsampling methodology. Section 2 provides a detailed asymptotic theory, upon which the statistical methodology relies. An application to stock returns is given in Section 3, while Section 4 concludes. Simulations that explore finite-sample performance of the subsampling estimators, as well as proofs of technical results, are provided in the appendices (see Supplementary Data).

1 Self-normalization for Autocovariances and Autocorrelations

1.1 Process

The GARCH process satisfies $X_t = \sigma_t Z_t$ for an iid sequence $\{Z_t\}$ that are only assumed to be symmetric about zero, and

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{t-j}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{1}$$

We refer to $\{\sigma_t\}$ as the volatility process. A discussion of the conditions for stationarity are summarized in Lindner (2009); necessary and sufficient conditions in terms of the process' Lyapunov exponent are given in Bougerol and Picard (1992), and a sufficient condition—in the case that Z_0 has finite variance—in terms of the coefficients α_j and β_j is given in Bollerslev (1986). In this article, we consider $\{Z_t\}$ that are heavy-tailed, but are interested in stationary GARCH processes—see the discussion in Remark 3.2 of Basrak, Davis, and Mikosch (2002) (henceforth BDM) for stationarity conditions.

Theorem 3.1 of BDM discusses the properties of the GARCH process, and in their Corollary 3.5 they show that $U_t = (X_t, \sigma_t)$ has heavy-tailed marginal distributions of index κ, for some $\kappa > 0$ that depends on distributional properties of Z_0 and the GARCH coefficients in a complicated fashion. The bivariate process $\{U_t\}$ is also strong mixing with geometric rate, and is regularly varying with some rate a_n. The rate a_n is related to the tail index, being given by $a_n = c\,n^{1/\kappa}$ for some constant $c > 0$—see Remark 2.1 of BDM.

The PGARCH process $\{X_t\}$ is defined via $X_t = \sigma_t Z_t$ together with volatility satisfying

$$|\sigma_t|^\nu = \alpha_0 + \sum_{j=1}^{p} \alpha_j |X_{t-j}|^\nu + \sum_{j=1}^{q} \beta_j |\sigma_{t-j}|^\nu, \tag{2}$$

where $\nu > 0$ is the exponent of the process ($\nu = 2$ corresponds to the GARCH). For references on PGARCH, see Mittnik, Paolella, and Rachev, 2002; Carrasco and Chen, 2002; and Tully and Lusey, 2007. Interest focuses on the transformed variables

$$Y_t = |X_t|^\nu \qquad \text{and} \qquad S_t = \sigma_t^\nu. \tag{3}$$

Define the following sample quantities for k integer:

$$\hat{\gamma}_{Y,Y}(k) = n^{-1} \sum_{t=1}^{n} Y_t Y_{t+k}$$

$$\hat{\gamma}_{Y,S}(k) = n^{-1} \sum_{t=1}^{n} Y_t S_{t+k}$$

$$\hat{\gamma}_{S,S}(k) = n^{-1} \sum_{t=1}^{n} S_t S_{t+k}.$$

(The latter two quantities are not statistics, because $\{S_t\}$ is not observed.) Also $\hat{\gamma}_{S,Y}(k)$ is defined by swapping the order of Y and S. Up to negligible errors (i.e., terms that converge to zero in probability—see below for the precise discussion), $\hat{\gamma}_{Y,Y}(-k) \approx \hat{\gamma}_{Y,Y}(k)$, $\hat{\gamma}_{Y,S}(-k) \approx \hat{\gamma}_{S,Y}(k)$, and $\hat{\gamma}_{S,S}(-k) \approx \hat{\gamma}_{S,S}(k)$. If we remove the ˆ symbol, we refer to expectations (whenever these exist), and the above relations in the lags are exact. Because X_0 is regularly varying of index κ, the autocovariances and cross-covariances for $\{Y_t\}$ and $\{S_t\}$ exist whenever $\nu < \kappa/2$.

We are interested in weak convergence of the so-called *roots* defined by $\tilde{\gamma}_{Y,Y}(k) = \hat{\gamma}_{Y,Y}(k) - \gamma_{Y,Y}(k)$, and secondarily in weak convergence of the analogous quantities for the volatilities S and cross-covariances. Although the latter roots involving the volatilities are theoretical quantities, the distribution of $\tilde{\gamma}_{Y,Y}(k)$ depends upon $\tilde{\gamma}_{Y,S}(k)$ and $\tilde{\gamma}_{S,S}(k)$, so it behooves us to analyze these objects together. When $\kappa < 2\nu$ the theoretical quantity $\gamma_{Y,Y}(k)$ does not exist, and the root is just defined via $\tilde{\gamma}_{Y,Y}(k) = \hat{\gamma}_{Y,Y}(k)$. We extend this definition to $\tilde{\gamma}_{Y,S}(k)$, and $\tilde{\gamma}_{S,S}(k)$ in the obvious fashion.

The appropriate rate of convergence actually depends on the regular variation rate a_n. Trivially, regular variation for the νth power implies a rate of a_n^ν for Y_t, and the product of two variables (in the autocovariances or cross-covariances) indicates a rate of $a_n^{2\nu}$. In the case of a GARCH(1,1), previous work indicates that $na_n^{-4}\tilde{\gamma}_{Y,Y}(k)$ converges weakly to a nondegenerate random variable, jointly in k, when $\kappa < 8$; here we utilize the rate $a_n^{2\nu}$ in lieu of a_n^4 (i.e., replacing the exponent 2 by ν). When $\kappa > 2\nu$ the autocovariances and cross-covariances exist, and centering of the sample quantities becomes possible, but when $\kappa < 2\nu$ no centering is necessary.

1.2 Self-normalization

For our applications we investigate sample autocovariances, acs, and cross-covariances and cross-correlations for the process $\{X_t\}$ with its powers $\{Y_t\}$, where $Y_t = |X_t|^\nu$. Recall that $\gamma_{X,X}(h) = 0$ for $h \neq 0$, whenever $\kappa > \nu$. With the notation for $\hat{\gamma}(k)$ of the previous subsection, and with $\hat{\rho}(k) = \hat{\gamma}(k)/\hat{\gamma}(0)$ by definition, we here study

$$\hat{\gamma}_{X,X},\ \hat{\rho}_{X,X},\ \hat{\gamma}_{X,Y},\ \hat{\rho}_{X,Y},\ \hat{\gamma}_{Y,Y},\ \hat{\rho}_{Y,Y}.$$

In the case of the cross-correlation, the normalization is $\sqrt{\hat{\gamma}_{X,X}(0)\hat{\gamma}_{Y,Y}(0)}$. In each case, there are rates of convergence (sometimes with a mean centering) for each statistic with the results dependent on κ and ν. Since we are interested in developing simple studentizations for the statistics, we prove joint results involving absolute sample moments, abbreviated by $\hat{\mu}_{X^j} = n^{-1}\sum_{t=1}^{n} |X_t|^j$ for j real.

Here we discuss self-normalized statistics, such that the studentized quantity's rate of convergence does not depend on unknown quantities. First consider self-normalization for the acs of a GARCH(p,q) process. It follows from Theorem 2 below that $c_n^{-1} n \hat{\rho}_{X,X}(k) = O_P(1)$, where

$$c_n = \begin{cases} n^1 & \text{if } \kappa \in (0,2) \\ a_n^2 & \text{if } \kappa \in (2,4) \\ n^{1/2} & \text{if } \kappa \in (4,\infty). \end{cases}$$

Excluding $\kappa \in \{2,4\}$, the rate c_n equals a constant times n raised to the power $1 \wedge (2/\kappa \vee 1/2)$. A suitable choice for the self-normalization that matches this growth rate c_n is

$$\hat{\sigma}_{X,X} = \left(n^{-1} + (n\hat{\mu}_{X^4})^{-1/2}\right)^{-1},$$

as is demonstrated in Theorem 5 below. Similarly, the growth rate for the acs of powers takes the form of n raised to the power $1 \wedge (2\nu/\kappa \vee 1/2)$, and hence we can normalize with

$$\hat{\sigma}_{Y,Y} = \left(n^{-1} + (n\hat{\mu}_{Y^4})^{-1/2}\right)^{-1}.$$

The case of cross-correlations is more complex. In the case that $\nu \in (0,1)$, the growth rate c_n takes the form of n raised to the power $[1 \wedge (1/2 + \nu/\kappa)] \wedge [(1+\nu)/\kappa \vee 1/2]$. Therefore, we can use the studentization

$$\hat{\sigma}_{X,Y} = \left(n^{-1} + n^{-1/2}\,(n\hat{\mu}_{X^2})^{-\nu/2} + (n\hat{\mu}_{X^2Y^2})^{-1/2}\right)^{-1}.$$

On the other hand, if $\nu \geq 1$ then the growth rate c_n takes the form of n raised to the power $[1 \wedge (1/2 + 1/\kappa)] \wedge [(1+\nu)/\kappa \vee 1/2]$. Then we can use the studentization

$$\hat{\sigma}_{X,Y} = \left(n^{-1} + n^{-1/2}\,(n\hat{\mu}_{Y^2})^{-1/2\nu} + (n\hat{\mu}_{X^2Y^2})^{-1/2}\right)^{-1}.$$

Each self-normalization converges jointly with the respective correlation statistics, and is bounded in probability under the assumptions of Theorem 5 given below.

1.3 Subsampling

We now proceed to the statistical portion of the article, namely conducting inference for the process' (and its powers') acs. For a PGARCH(p,q) process, the acs and cross-correlations with the powers are zero, while the acs of the powers decay exponentially.

Our testing paradigm is as follows: we assume as null hypothesis that the process is PGARCH(p, q) and check whether the subsampling-based confidence intervals capture zero (for the process acs and for the cross-correlations with the powers) or decay exponentially fast (in case of the powers). In both cases, we replace the value of the population parameter in the root by its full-sample estimate, although in the former case we have the choice of substituting it with zero. So if zero is not contained in a $1-\alpha$ level confidence interval for the acs or cross-cs at some lag, then we can reject the GARCH hypothesis with Type I error rate α. We begin with the definitions, and then consider consistency of the estimators.

To construct confidence intervals for $\rho_{X,X}(k)$ (although for a PGARCH process this parameter is always zero) we need to approximate the sampling distribution of

$$T_{X,X}(k) = n\left(\frac{\hat{\rho}_{X,X}(k) - \rho_{X,X}(k)}{\hat{\sigma}_{X,X}}\right),$$

that is, $L_{X,X,k}(x) = \mathbb{P}[T_{X,X}(k) \leq x]$. Let $L_{\infty,X,X,k}$ denote the cumulative distribution function (cdf) of the corresponding limiting random variable—which by Equation (17) below depends on κ. The use of this asymptotic distribution is impractical, as $L_{\infty,X,X,k}$ depends on the unknown parameter and there is no known analytic formula for it. Hence, we propose to approximate $L_{X,X,k}$ (and $L_{\infty,X,X,k}$) nonparametrically via subsampling (Politis, Romano, and Wolf, 1999). According to this procedure we divide the sample into overlapping blocks of size b ($b \to \infty$, $b/n \to 0$), containing $X_t, X_{t+1}, \ldots, X_{t+b-1}$ for $t = 1, 2, \ldots, n-b+1$, and calculate the self-normalized statistic upon each block, treating each block as if it were a full sample. Moreover, the parameter $\rho_{X,X}(k)$ is replaced by its large-sample estimate $\hat{\rho}_{X,X}(k)$. This produces $n-b+1$ subsampling statistics

$$T_{X,X,t}(k) = b\left(\frac{\hat{\rho}_{X,X,t}(k) - \hat{\rho}_{X,X}(k)}{\hat{\sigma}_{X,X,t}}\right),$$

where $\hat{\rho}_{X,X,t}(k) = \hat{\gamma}_{X,X,t}(k)/\hat{\gamma}_{X,X,t}(0)$ with $\hat{\gamma}_{X,X,t}(k) = \frac{1}{b}\sum_{\ell=t}^{t+b-1-k}(X_\ell - \overline{X}_t)(X_{\ell+k} - \overline{X}_t)$ and $\overline{X}_t = \sum_{\ell=t}^{t+b-1-k} X_\ell/b$. Note that we could also replace $\hat{\rho}_{X,X}(k)$ by zero, if we wish to utilize the "null hypothesis" that the process is PGARCH, but we elect instead to utilize a large-sample estimate, which makes the confidence interval construction more intellectually consistent with the case of the powers' acs. The normalization $\hat{\sigma}_{X,X,t}$ is given by

$$\hat{\sigma}_{X,X,t} = \left(b^{-1} + \left[\sum_{\ell=t}^{t+b-1-k} X_\ell^4\right]^{-1/2}\right)^{-1}$$

and the sampling distribution $L_{X,X,k}(x)$ is approximated by

$$\hat{L}_{X,X,k}(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} \mathbf{1}_{\{T_{X,X,t}(k) \le x\}}.$$

If by $c_{X,X,k}(1-p) = \inf \{x : \hat{L}_{X,X,k}(x) \ge 1-p\}$ we denote its lower $1 - p$ quantile, we obtain for $\rho_{X,X}(k)$ a $(1-p)$ equal-tailed subsampling confidence interval $CI_{et;1-p}(\rho_{X,X}(k))$, defined as

$$\left[\hat{\rho}_{X,X}(k) - \frac{\hat{\sigma}_{X,X}}{n} c_{X,X,k}(1-p/2), \hat{\rho}_{X,X}(k) - \frac{\hat{\sigma}_{X,X}}{n} c_{X,X,k}(p/2)\right]. \quad (4)$$

Introducing $\hat{L}_{X,X,k,|\cdot|}(x) = \sum_{t=1}^{n-b+1} \mathbf{1}_{\{|T_{X,X,t}(k)| \le x\}}/(n-b+1)$, another related distribution, and its quantile $c_{X,X,k,|\cdot|}(1-p) = \inf \{x : \hat{L}_{X,X,k,|\cdot|}(x) \ge 1-p\}$ offers an $(1 - p)$ symmetric subsampling confidence interval for $\rho_{X,X}(k)$,

$$CI_{s;1-p}(\rho_{X,X}(k)) = \left[\hat{\rho}_{X,X}(k) \mp \frac{\hat{\sigma}_{X,X}}{n} c_{X,X,k,|\cdot|}(1-p)\right]. \quad (5)$$

For the powers $\{Y_t\}$, we have the following, analogous definitions, starting with the statistic and its subsampling version

$$T_{Y,Y}(k) = n\left(\frac{\hat{\rho}_{Y,Y}(k) - \rho_{Y,Y}(k)}{\hat{\sigma}_{Y,Y}}\right), \quad T_{Y,Y,t}(k) = b\left(\frac{\hat{\rho}_{Y,Y,t}(k) - \hat{\rho}_{Y,Y}(k)}{\hat{\sigma}_{Y,Y,t}}\right),$$

where $\hat{\rho}_{Y,Y,t}(k) = \hat{\gamma}_{Y,Y,t}(k)/\hat{\gamma}_{Y,Y,t}(0)$ with $\hat{\gamma}_{Y,Y,t}(k) = \frac{1}{b} \sum_{\ell=t}^{t+b-1-k} (Y_\ell - \overline{Y}_t)(Y_{\ell+b} - \overline{Y}_t)$ and $\overline{Y}_t = \sum_{\ell=t}^{t+b-1-k} Y_\ell/b$. Unlike the previous case of the regular acs, we do not presume that $\rho_{Y,Y}(k)$ equals zero, so we must estimate it instead. The normalization $\hat{\sigma}_{Y,Y,t}$ is given by

$$\hat{\sigma}_{Y,Y,t} = \left(b^{-1} + \left[\sum_{\ell=t}^{t+b-1-k} Y_\ell^4\right]^{-1/2}\right)^{-1}.$$

The sampling distribution $L_{Y,Y,k}(x) = \mathbb{P}[T_{Y,Y}(k) \le x]$ is approximated by

$$\hat{L}_{Y,Y,k}(x) = \frac{1}{n-b+1} \sum_{t=1}^{n-b+1} \mathbf{1}_{\{T_{Y,Y,t}(k) \le x\}},$$

and is an estimator of $L_{\infty,Y,Y,k}(x)$, the cdf of the limit variable given in (18). For $\rho_{Y,Y}(k)$ a $(1-p)$ equal-tailed subsampling confidence interval $CI_{et;1-p}(\rho_{Y,Y}(k))$ is given by

$$\left[\hat{\rho}_{Y,Y}(k) - \frac{\hat{\sigma}_{Y,Y}}{n} c_{Y,Y,k}(1-p/2), \hat{\rho}_{Y,Y}(k) - \frac{\hat{\sigma}_{Y,Y}}{n} c_{Y,Y,k}(p/2)\right], \quad (6)$$

where $c_{Y,Y,k}(1-p) = \inf \{x : \hat{L}_{Y,Y,k}(x) \ge 1-p\}$. An $(1-p)$ symmetric subsampling confidence interval for $\rho_{Y,Y}(k)$ is then

$$CI_{s;1-p}(\rho_{Y,Y}(k)) = \left[\hat{\rho}_{Y,Y}(k) \mp \frac{\hat{\sigma}_{Y,Y}}{n} c_{Y,Y,k,|\cdot|}(1-p)\right], \quad (7)$$

with $c_{Y,Y,k,|\cdot|}(1-p) = \inf\{x : \hat{L}_{Y,Y,k,|\cdot|}(x) \geq 1-p\}$ and $\hat{L}_{Y,Y,k,|\cdot|}(x) = \sum_{t=1}^{n-b+1} \mathbf{1}_{\{|T_{Y,Y,t}(k)| \leq x\}}/(n-b+1)$. For the cross-correlations between the data process and its powers, the normalized difference and its subsampling counterpart are

$$T_{X,Y}(k) = n\left(\frac{\hat{\rho}_{X,Y}(k) - \rho_{X,Y}(k)}{\hat{\sigma}_{X,Y}}\right), \quad T_{X,Y,t}(k) = b\left(\frac{\hat{\rho}_{X,Y,t}(k) - \hat{\rho}_{X,Y}(k)}{\hat{\sigma}_{X,Y,t}}\right),$$

where $\hat{\rho}_{X,Y,t}(k) = \hat{\gamma}_{X,Y,t}(k)/\sqrt{\hat{\gamma}_{X,X,t}(0)\,\hat{\gamma}_{Y,Y,t}(0)}$. Again, we could replace $\rho_{X,Y}(k)$ by zero under the PGARCH hypothesis, but in keeping with the above treatment we utilize the parameter's large-sample estimate. The normalization $\hat{\sigma}_{X,Y,t}$ is given (in the case of $\nu \geq 1$) by

$$\hat{\sigma}_{X,Y,t} = \left(b^{-1} + b^{-1/2}\left[\sum_{\ell=t}^{t+b-1-k} Y_\ell^2\right]^{-1/2\nu} + \left[\sum_{\ell=t}^{t+b-1-k} |X_\ell|^{2+2\nu}\right]^{-1/2}\right)^{-1}.$$

The sampling distribution $L_{X,Y,k}(x) = \mathbb{P}[T_{X,Y}(k) \leq x]$ is approximated by

$$\hat{L}_{X,Y,k}(x) = \frac{1}{n-b+1}\sum_{t=1}^{n-b+1} \mathbf{1}_{\{T_{X,Y,t}(k) \leq x\}},$$

and is an estimator of $L_{\infty,X,Y,k}(x)$, the cdf of the limit variable given in (19). For $\rho_{X,Y}(k)$ a $(1-p)$ equal-tailed subsampling confidence interval $CI_{et;1-p}(\rho_{X,Y}(k))$ is defined as

$$\left[\hat{\rho}_{X,Y}(k) - \frac{\hat{\sigma}_{X,Y}}{n} c_{X,Y,k}(1-p/2), \hat{\rho}_{X,Y}(k) - \frac{\hat{\sigma}_{X,Y}}{n} c_{X,Y,k}(p/2)\right], \tag{8}$$

where $c_{X,Y,k}(1-p) = \inf\{x : \hat{L}_{X,Y,k}(x) \geq 1-p\}$. An $(1-p)$ symmetric subsampling confidence interval for $\rho_{X,Y}(k)$ is then

$$CI_{s;1-p}(\rho_{X,Y}(k)) = \left[\hat{\rho}_{X,Y}(k) \mp \frac{\hat{\sigma}_{X,Y}}{n} c_{X,Y,k,|\cdot|}(1-p)\right], \tag{9}$$

with $c_{X,Y,k,|\cdot|}(1-p) = \inf\{x : \hat{L}_{X,Y,k,|\cdot|}(x) \geq 1-p\}$ and $\hat{L}_{X,Y,k,|\cdot|}(x) = \sum_{t=1}^{n-b+1} \mathbf{1}_{\{|T_{X,Y,t}(k)| \leq x\}}/(n-b+1)$.

2 Asymptotic Theory

We expand the theoretical results of Davis and Mikosch (1998) and Mikosch and Stărică (2000) in two substantial directions: we derive asymptotic relations for process and volatility autocovariances and cross-covariances (where the volatility is defined to be $\{\sigma_t\}$) for the PGARCH process, and we derive cross-covariance results for a PGARCH process and its power. Our main objective is to obtain nondegenerate weak limits of all studentized roots, so that the subsampling method is applicable. The first subsection reviews some background concepts and notation (also see Appendix A in Supplementary Data), while the second subsection applies these results to certain statistics, providing our main theorems. The third subsection applies these theorems to produce studentized statistics, and the fourth subsection establishes the consistency of subsampling.

2.1 GARCH and PGARCH

The GARCH(p,q) process $\{X_t\}$ has been studied in Mikosch and Stărică (2000) and BDM. In addition to establishing results on autocovariances, BDM provides results on Stochastic Recurrence Equations (SREs), from which they derive regular variation properties of GARCH processes; by adapting their methods of proof, many of the results can be trivially extended to PGARCH processes. We first review BDM's GARCH results.

The vector process of dimension $d = q + p - 1$ defined by

$$\vec{W}_t = [\sigma_{t+1}, \sigma_t, \cdots, \sigma_{t-q+2}, X_t, X_{t-1}, \cdots, X_{t-p+2}]' \tag{10}$$

satisfies a SRE in the squares of its components, as described in Equation (3.1) of BDM. Let $\varepsilon_{\mathbf{x}}$ be the point measure concentrated at $\mathbf{x}$, and let $\stackrel{\mathcal{L}}{\Rightarrow}$ denote convergence in distribution of point measures on $\overline{\mathbb{R}}^{d(m+1)} \backslash \{\mathbf{0}\}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and $d = q + p - 1$. BDM shows that $\vec{W}_t(m) = \text{vec}(\vec{W}_t, \cdots, \vec{W}_{t+m})$ yields a point process convergence

$$N_n = \sum_{t=1}^{n} \varepsilon_{\vec{W}_t(m) a_n^{-1}} \stackrel{\mathcal{L}}{\Rightarrow} N_\infty = \sum_{i,j \geq 1} \varepsilon_{P_i \vec{Q}_{ij}} \tag{11}$$

as $n \to \infty$, where a_n is the rate of regular variation, being implicitly defined by

$$n\,\mathbb{P}[|\vec{W}_0(m)| > a_n] \to 1. \tag{12}$$

The $\{P_i\}_{i \geq 1}$ are the points of a Poisson process defined on $(0, \infty)$ with intensity measure given in terms of κ and the extremal index. The Poisson process $\sum_{i \geq 1} \varepsilon_{P_i}$ is independent of each iid point process $\sum_{i \geq 1} \varepsilon_{Q_{ij}}$ (for each $j \geq 1$). The points $\{P_i\}$ and $\{Q_{ij}\}$ correspond to the radial and spherical portions of the limiting points $\vec{W}_t(m) a_n^{-1}$; see Corollary 2.4 of Davis and Mikosch (1998) for more detail. It is apparent that $\vec{Q}_{ij}$ has $m+1$ blocks of d components, which we denote via

$$\vec{Q}_{ij} = \text{vec}\{Q_{ij}^{(0)}, Q_{ij}^{(1)}, \cdots, Q_{ij}^{(m)}\}, \tag{13}$$

and each $Q_{ij}^{(\ell)}$ itself has $d = q + p - 1$ components, where $0 \leq \ell \leq m$. To refer to the kth component, with $1 \leq k \leq d$, of $Q_{ij}^{(\ell)}$, we write $Q_{ij}^{(\ell,k)}$.

BDM then makes an application to establishing weak convergence of sample autocovariances of $\{X_t\}$, extending earlier results by Davis and Mikosch (1998) and Mikosch and Stărică (2000) on the GARCH(1,1). These latter articles also discuss results for the autocovariances of $\{|X_t|\}$ and $\{X_t^2\}$, which are objects of interest for financial applications; BDM claim that these results can be extended to the GARCH(p,q) case without proof. For the sample autocovariances of $\{X_t^2\}$, (asymptotic) recursive relations can be derived that extend ideas in Davis and Mikosch (1998) and Mikosch and Stărică (2000) for the GARCH(1,1), but these relations seem hard to extend to the autocovariances of $\{|X_t|\}$—something rather special happens when $p, q \leq 1$ that prevents the method of proof to be easily generalized for $p > 1$ or $q > 1$.

Essentially, these same results also hold for the PGARCH, which we prove below. Set $\vec{X}_t = [S_{t+1}, \cdots, S_{t-q+2}, Y_t, \cdots, Y_{t-p+2}]'$. Then $\{Y_t\}$ and $\{S_t\}$, defined via (3), satisfy the following SRE:

$$\vec{X}_t = A_t \vec{X}_{t-1} + B_t, \tag{14}$$

with $B_t = [\alpha_0, 0, \cdots, 0]'$ and

$$A_t = \begin{bmatrix} \alpha_1 |Z_t|^\nu + \beta_1 & \beta_2 & \cdots & \beta_{q-1} & \beta_q & \alpha_2 & \alpha_3 & \cdots & \alpha_p \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ |Z_t|^\nu & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The key assumptions needed to establish the existence of a stationary solution, and its regular variation properties, are given below. The Lyapunov exponent of the SRE is given in terms of the L_1 matrix norm by

$$\gamma = \inf \{n^{-1}\, \mathbb{E}\log ||A_1 \cdots A_n||, n \in \mathbb{N}\},$$

and sufficient conditions (in the GARCH case) for the exponent's negativity are discussed in Remark 3.2 of BDM.

Assumption A: $\alpha_0 > 0$ and the Lyapunov exponent $\gamma < 0$.

Assumption B: Z_0 has a positive density on $\mathbb{R}$ such that $\mathbb{E}|Z_0|^h < \infty$ for all $h < h_0$ and $\mathbb{E}|Z_0|^{h_0} = \infty$ for some $h_0 \in (0, \infty]$; not all the parameters α_j and β_k for $1 \le j \le p$ and $1 \le k \le q$ are zero.

Assumption C: The density of Z_0 is positive in an interval containing zero.

Assumption D: $\{U_t = (X_t, \sigma_t)\}$ is strong mixing with geometric rate.

Theorem 1: *For the SRE (14), assume Assumptions A and B. Then there exists a unique strictly stationary solution to the SRE, and there exists ρ such that for all $x \in \mathbb{R}^d \backslash \{0\}$ the inner product $\langle x, \vec{X}_1 \rangle$ is regularly varying with index ρ. Furthermore, if ρ is not an even integer, then $\vec{X}_1$ is regularly varying with index ρ.*

Remark 1: This theorem is proved by exactly following the proof of Theorem 3.1 of BDM, only substituting $|Z_t|^\nu$ for Z_t^2 and consolidating parts A and B; therefore, its proof is omitted. Because it is unknown whether the PGARCH process is strong mixing—though in part C of Theorem 3.1 of BDM it is stated that for a GARCH process Assumption C implies Assumption D—we omit discussion of this point in Theorem 1. Also see Fryzlewicz and Subba Rao (2011).

Corollary 1: *For the SRE (14), assume Assumptions A, B, and D. Then a stationary version of the process $\{U_t = (X_t, \sigma_t)\}$ exists, and if the index ρ of Theorem 1 is not an even* *integer, then the finite-dimensional distributions of* $\{U_t\}$ *are regularly varying with index* $\kappa = \nu\rho$.

Remark 2: The proof of Corollary 1 follows almost verbatim from the proof of Corollary 3.5 of BDM, the only change being that $\kappa = \nu\rho$ and ν need not be 2, as in the GARCH case. Note we assume the strong mixing property, although in the GARCH case it suffices to assume the weaker Assumption C instead.

Hence using Theorem 2.8 of Davis and Mikosch (1998), from Corollary 1 we can conclude that—under our working assumptions A, B, and D—with $\vec{W}_t$ given by (10) the convergence (11) holds, along with (12). Note that the points P_i and $\vec{Q}_{ij}$ pertain to the process $\{U_t\}$, not $\{Y_t\}$ and $\{S_t\}$. The regular variation is of index κ, such that κ/ν is not an even integer.

2.2 Sampling Behavior for PGARCH Processes

Under Assumptions A, B, and D we can state our convergence results for the roots of interest. Each theorem below splits into three main cases: (i) κ is sufficiently low that the mean centering does not exist, and a stable limit is obtained; (ii) κ is a bit larger so that centerings are utilized, and a stable limit for the root is obtained; (iii) κ is large enough that a central limit theorem holds, so that the limit of the root is Gaussian. (The third theorem splits case (ii) into two sub-cases.) In cases (i) and (ii) the limit distributions are stated in terms of the points of (11), which are explained in more detail in Theorem 2.10 of BDM; also see Corollary 2.4 of Davis and Mikosch (1998). In case (iii) we only need to know the joint covariance of the limiting Gaussian variables.

In the first Theorem below, case (ii) involves a limit variable that is described as follows (also see Proposition 3.3 of Davis and Mikosch, 1998). The regular variation of the PGARCH process (cf. Corollary 1) yields the vague convergence of $n\,\mathbb{P}[a_n^{-1}\vec{W}_0(m) \in \cdot]$ to a measure τ on $\overline{\mathbb{R}}^{d(m+1)} \backslash \{0\}$, the Lévy measure of a κ-stable random vector with $\kappa \in (0,2)$ (and such that κ/ν is not an even integer). Then we have the definition

$$V_k = \lim_{\delta \to 0} \left(\sum_{i,j \geq 1} P_i^2 Q_{ij}^{(0,q+1)} Q_{ij}^{(k,q+1)} 1_{\{P_i^2 |Q_{ij}^{(0,q+1)} Q_{ij}^{(k,q+1)}| \geq \delta\}} - \int_{B_{\delta,k}} \mathbf{x}^{(0,q+1)} \mathbf{x}^{(k,q+1)} \, \rho(dx) \right) \quad (15)$$

for $B_{\delta,k} = \{\mathbf{x} \in \mathbb{R}^{d(m+1)} : \delta < |\mathbf{x}^{(0,q+1)} \mathbf{x}^{(k,q+1)}|\}$, where $k \geq 0$. The superscript notation is explained in (13). The points $Q_{ij}^{(0,q+1)}$ and $Q_{ij}^{(k,q+1)}$ correspond to the $(0, q+1)$ and $(k, q+1)$ components of $\vec{W}_t(m)$, that is, X_t and X_{t+k}. In addition, all the theorems contain the random variables

$$W_r = \sum_{i,j \geq 1} |P_i|^r |Q_{ij}^{(0,q+1)}|^r,$$

for $r > 0$. This first result extends the work of BDM, which is concerned with the GARCH(p,q) process, in three ways: the process is a PGARCH (which nest the GARCH), we consider convergence jointly with sample moments, and we also derive the result for acs.

Theorem 2: *Let $\{X_t\}$ be a PGARCH(p,q) process (2) satisfying Assumptions A, B, and D; also assume that Z_0 is symmetric about zero. Then:*

(i): in the case $\kappa \in (0,2)$:

$$\left(\left\{na_n^{-2}\,\hat{\gamma}_{X,X}(k)\right\}_{k=0}^{m}, na_n^{-4}\,\hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\{V_k\}_{k=0}^{m}, W_4\right)$$

$$\left(\left\{\hat{\rho}_{X,X}(k)\right\}_{k=1}^{m}, na_n^{-4}\,\hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\{V_k V_0^{-1}\}_{k=1}^{m}, W_4\right),$$

where $V_k = \sum_{i,j\geq 1} P_i^2 Q_{ij}^{(0,q+1)} Q_{ij}^{(k,q+1)}$.

(ii): in the case $\kappa \in (2,4)$:

$$\left(\left\{na_n^{-2}\left[\hat{\gamma}_{X,X}(k)-\gamma_{X,X}(k)\right]\right\}_{k=0}^{m}, na_n^{-4}\hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\{V_k\}_{k=0}^{m}, W_4\right)$$

$$\left(\left\{na_n^{-2}\left[\hat{\rho}_{X,X}(k)-\rho_{X,X}(k)\right]\right\}_{k=1}^{m}, na_n^{-4}\hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\left\{\gamma_{X,X}^{-1}(0)\left[V_k-\rho_{X,X}(k)V_0\right]\right\}_{k=1}^{m}, W_4\right),$$

where V_k is defined by (15).

(iii): in the case $\kappa > 4$:

$$\left(\left\{n^{1/2}\left[\hat{\gamma}_{X,X}(k)-\gamma_{X,X}(k)\right]\right\}_{k=0}^{m}, \hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\{G_k\}_{k=0}^{m}, \mathbb{E}X^4\right)$$

$$\left(\left\{n^{1/2}\left[\hat{\rho}_{X,X}(k)-\rho_{X,X}(k)\right]\right\}_{k=1}^{m}, \hat{\mu}_{X^4}\right) \xrightarrow{\mathcal{L}} \left(\left\{\gamma_{X,X}^{-1}(0)\left[G_k-\rho_{X,X}(k)G_0\right]\right\}_{k=1}^{m}, \mathbb{E}X^4\right),$$

where the G_k satisfy $\mathrm{Cov}(G_j,G_k)=\sum_\ell \mathrm{Cov}(X_0X_j, X_\ell X_{\ell+k})$ and are multivariate normal.

Remark 3: In cases (ii) and (iii) of Theorem 2, $\gamma_{X,X}(k)=0$ if $k>0$, and in particular $\rho_{X,X}(k)=0$ in the results for $\hat{\rho}_{X,X}(k)$. Therefore, these results could be stated as $na_n^{-2}\,\hat{\rho}_{X,X}(k) \xrightarrow{\mathcal{L}} \gamma_{X,X}^{-1}(0)V_k$ and $\sqrt{n}\,\hat{\rho}_{X,X}(k) \xrightarrow{\mathcal{L}} \gamma_{X,X}^{-1}(0)G_k$ for cases (ii) and (iii), respectively. Note that these limiting distributions are nondegenerate, which is important for our subsampling applications discussed in the next section.

The next result is concerned with autocovariances and acs for the powered process $\{Y_t\}$. The middle case (ii) involves a somewhat complicated limit distribution, which is described in Proposition 1 of Appendix A (Supplementary Data), which requires the additional Assumption M discussed therein. The result for the autocovariances is stated without proof in BDM (and for the GARCH(1,1) case, a full proof is given in Davis and Mikosch, 1998), and our proof here relies on the recursive relations of the previous subsection. We also describe results jointly with the sample moments, and provide the autocorrelation results.

Theorem 3: *Let $\{X_t\}$ be a PGARCH(p,q) process (2) satisfying Assumptions A, B, and D; also assume that Z_0 is symmetric about zero. Then:*

(i): in the case $\kappa \in (0,2\nu)$:

$$\left(\left\{na_n^{-2\nu}\,\hat{\gamma}_{Y,Y}(k)\right\}_{k=0}^{m}, na_n^{-4\nu}\,\hat{\mu}_{Y^4}\right) \xrightarrow{\mathcal{L}} \left(\{V_k\}_{k=0}^{m}, W_{4\nu}\right)$$

$$\left(\left\{\hat{\rho}_{Y,Y}(k)\right\}_{k=1}^{m}, na_n^{-4\nu}\,\hat{\mu}_{Y^4}\right) \xrightarrow{\mathcal{L}} \left(\{V_k V_0^{-1}\}_{k=1}^{m}, W_{4\nu}\right),$$

where $V_k = \sum_{i,j\geq 1} |P_i|^{2\nu} |Q_{ij}^{(0,q+1)} Q_{ij}^{(k,q+1)}|^{\nu}$.

(ii): in the case $\kappa \in (2\nu, 4\nu)$, assume Assumption M as well:

$$\left(\left\{na_n^{-2\nu}\left[\hat{\gamma}_{Y,Y}(k)-\gamma_{Y,Y}(k)\right]\right\}_{k=0}^{m}, na_n^{-4\nu}\hat{\mu}_{Y^4}\right) \stackrel{\mathcal{L}}{\Rightarrow} \left(\{U_k\}_{k=0}^{m}, W_{4\nu}\right)$$

$$\left(\left\{na_n^{-2\nu}\left[\hat{\rho}_{Y,Y}(k)-\rho_{Y,Y}(k)\right]\right\}_{k=1}^{m}, na_n^{-4\nu}\hat{\mu}_{Y^4}\right) \stackrel{\mathcal{L}}{\Rightarrow} (\{_{\gamma_{Y,Y}^{-1}}^{(0)\left[U_k-\rho_{Y,Y}(k)U_0\right]\}} k = 1m, W_{4\nu}),$$

with U_k defined as in Proposition 1.

(iii): in the case $\kappa > 4\nu$:

$$\left(\left\{n^{1/2}\left[\hat{\gamma}_{Y,Y}(k)-\gamma_{Y,Y}(k)\right]\right\}_{k=0}^{m}, \hat{\mu}_{Y^4}\right) \stackrel{\mathcal{L}}{\Rightarrow} \left(\{G_k\}_{k=0}^{m}, \mathbb{E}|X|^{4\nu}\right)$$

$$\left(\left\{n^{1/2}\left[\hat{\rho}_{Y,Y}(k)-\rho_{Y,Y}(k)\right]\right\}_{k=1}^{m}, \hat{\mu}_{Y^4}\right) \stackrel{\mathcal{L}}{\Rightarrow} \left(\left\{\gamma_{Y,Y}^{-1}(0)\left[G_k-\rho_{Y,Y}(k)G_0\right]\right\}_{k=1}^{m}, \mathbb{E}|X|^{4\nu}\right),$$

where the G_k satisfy $\mathrm{Cov}(G_j, G_k) = \sum_\ell \mathrm{Cov}(Y_0 Y_j, Y_\ell Y_{\ell+k})$ *and are multivariate normal.*

The final result of this section utilizes some of the results of Theorems 2 and 3, because the normalization for the cross-correlations involves both $\gamma_{X,X}(0)$ and $\gamma_{Y,Y}(0)$. It turns out that symmetry in the cross-covariance variables makes the results resemble those of Theorem 2, though the limits depend on whether ν is less than or greater than one. Also, when $\kappa > 1+\nu$, the limit involves the random variables

$$V_k = \lim_{\delta\to 0}\left(\sum_{i,j\geq 1} P_i|P_i|^{\nu} Q_{ij}^{(0,q+1)}|Q_{ij}^{(k,q+1)}|^{\nu}\mathbf{1}_{\{|P_i|^{\nu+1}|Q_{ij}^{(0,q+1)}||Q_{ij}^{(k,q+1)}|^{\nu}\geq\delta\}} - \int_{B_{\delta,k}} \mathbf{x}^{(0,q+1)}|\mathbf{x}^{(k,q+1)}|^{\nu}\rho(dx)\right) \quad (16)$$

for $B_{\delta,k}=\{\mathbf{x}\in\mathbb{R}^{d(m+1)}:\delta<|\mathbf{x}^{(0,q+1)}[\mathbf{x}^{(k,q+1)}]^{\nu}|\}$, where $k\geq 0$. The points $Q_{ij}^{(0,q+1)}$ and $[Q_{ij}^{(k,q+1)}]^{\nu}$ correspond to the $(0,q+1)$ component and the powered $(k,q+1)$ component, respectively, of $\vec{W}_t(m)$, that is, X_t and Y_{t+k}. These types of statistics have not been mathematically studied in prior literature, to our knowledge.

Theorem 4: *Let $\{X_t\}$ be a PGARCH(p,q) process (2) satisfying Assumptions A, B, and D; also assume that Z_0 is symmetric about zero. If $\kappa > 1+\nu$, then $\gamma_{X,Y}(k)$ exists and equals zero for all k. Supposing that $\nu \in (0,1)$, then:*

(i) in the case $\kappa \in (0, 2\nu)$:

$$\left(\left\{na_n^{-(1+\nu)}\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, na_n^{-2}\hat{\gamma}_{X,X}(0), na_n^{-2\nu}\hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\hat{\mu}_{X^2Y^2}\right)$$
$$\stackrel{\mathcal{L}}{\Rightarrow} \left(\{V_k\}_{k=0}^{m}, W_2, W_{2\nu}, W_{2+2\nu}\right)$$
$$\left(\left\{\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\hat{\mu}_{X^2Y^2}\right) \stackrel{\mathcal{L}}{\Rightarrow} \left(\{V_k/\sqrt{W_2\,W_{2\nu}}\}_{k=1}^{m}, W_{2+2\nu}\right),$$

where $V_k = \sum_{i,j\geq 1} P_i|P_i|^{\nu}Q_{ij}^{(0,q+1)}|Q_{ij}^{(k,q+1)}|^{\nu}$.

(ii) in the case $\kappa \in (2\nu, 1+\nu)\cup(1+\nu, 2)$, assuming Assumption M:

$$\left(\left\{na_n^{-(1+\nu)}\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, na_n^{-2}\hat{\gamma}_{X,X}(0), \hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\hat{\mu}_{X^2Y^2}\right) \stackrel{\mathcal{L}}{\Rightarrow}$$
$$\left(\{V_k\}_{k=0}^{m}, W_2, \gamma_{Y,Y}(0), W_{2+2\nu}\right)$$
$$\left(\left\{n^{1/2}a_n^{-\nu}\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\hat{\mu}_{X^2Y^2}\right) \stackrel{\mathcal{L}}{\Rightarrow}$$
$$\left(\left\{V_k/\sqrt{W_2\,\gamma_{Y,Y}(0)}\right\}_{k=1}^{m}, W_{2+2\nu}\right),$$

with V_k given by case (i) if $\kappa \in (2\nu, 1+\nu)$ and by (16) if $\kappa \in (1+\nu, 2)$.

(iii) in the case $\kappa \in (2, 2+2\nu)$, assuming Assumption M:

$$\left(\left\{na_n^{-(1+\nu)}\,\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, \hat{\gamma}_{X,X}(0), \hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\{V_k\}_{k=0}^{m}, \gamma_{X,X}(0), \gamma_{Y,Y}(0), W_{2+2\nu}\right)$$
$$\left(\left\{na_n^{-(1+\nu)}\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\left\{V_k/\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)}\right\}_{k=1}^{m}, W_{2+2\nu}\right),$$

with V_k given by (16).

(iv) in the case $\kappa > 2+2\nu$, assuming Assumption M:

$$\left(\left\{n^{1/2}\,\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, \hat{\gamma}_{X,X}(0), \hat{\gamma}_{Y,Y}(0), \hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\{G_k\}_{k=0}^{m}, \gamma_{X,X}(0), \gamma_{Y,Y}(0), \gamma_{X^2,Y^2}(0)\right)$$
$$\left(\left\{n^{1/2}\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, \hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow} \left(\left\{G_k/\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)}\right\}_{k=1}^{m}, \mathbb{E}|X|^{2+2\nu}\right),$$

where the G_k satisfy $\operatorname{Cov}(G_j, G_k) = \sum_{\ell} \operatorname{Cov}(X_0Y_j, X_\ell Y_{\ell+k})$ and are multivariate normal. Supposing that $\nu \geq 1$, then:

(i) in the case $\kappa \in (0, 2)$:

$$\left(\left\{na_n^{-(1+\nu)}\,\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, na_n^{-2}\hat{\gamma}_{X,X}(0), na_n^{-2\nu}\,\hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right)$$
$$\overset{\mathcal{L}}{\Rightarrow} \left(\{V_k\}_{k=0}^{m}, W_2, W_{2\nu}, W_{2+2\nu}\right)$$
$$\left(\left\{\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow} \left(\left\{V_k/\sqrt{W_2\,W_{2\nu}}\right\}_{k=1}^{m}, W_{2+2\nu}\right),$$

where $V_k = \sum_{i,j\geq 1} P_i|P_i|^{\nu} Q_{ij}^{(0,q+1)}|Q_{ij}^{(k,q+1)}|^{\nu}$.

(ii) in the case $\kappa \in (2, 1+\nu) \cup (1+\nu, 2\nu)$, assuming Assumption M:

$$\left(\left\{na_n^{-(1+\nu)}\,\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, \hat{\gamma}_{X,X}(0), na_n^{-2\nu}\,\hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\{V_k\}_{k=0}^{m}, \gamma_{X,X}(0), W_{2\nu}, W_{2+2\nu}\right)$$
$$\left(\left\{n^{1/2}a_n^{-1}\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\left\{V_k/\sqrt{\gamma_{X,X}(0)\,W_{2\nu}}\right\}_{k=1}^{m}, W_{2+2\nu}\right),$$

with V_k given by case (i) if $\kappa \in (2, 1+\nu)$ and by (16) if $\kappa \in (1+\nu, 2\nu)$.

(iii) in the case $\kappa \in (2\nu, 2+2\nu)$, assuming Assumption M:

$$\left(\left\{na_n^{-(1+\nu)}\,\hat{\gamma}_{X,Y}(k)\right\}_{k=0}^{m}, \hat{\gamma}_{X,X}(0), \hat{\gamma}_{Y,Y}(0), na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\{V_k\}_{k=0}^{m}, \gamma_{X,X}(0), \gamma_{Y,Y}(0), W_{2+2\nu}\right)$$
$$\left(\left\{na_n^{-(1+\nu)}\hat{\rho}_{X,Y}(k)\right\}_{k=1}^{m}, na_n^{-(2+2\nu)}\,\hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$\left(\left\{V_k/\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)}\right\}_{k=1}^{m}, W_{2+2\nu}\right),$$

with V_k given by (16).

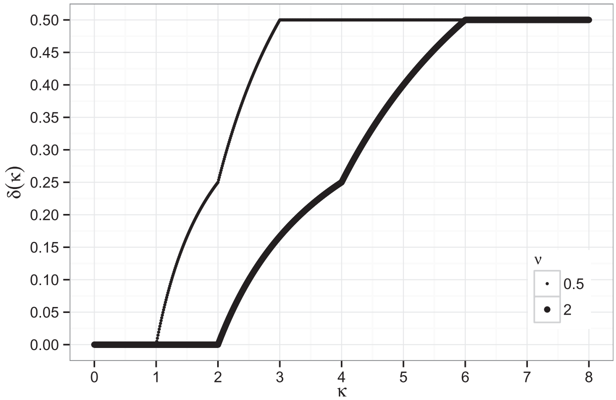

Figure 1. $\delta(\kappa)$ versus κ, where $\delta(\kappa)$ determines the order in probability of the sample cross-correlation, $n^{\delta(\kappa)}\hat{\rho}_{X,Y}(k) = O_P(1)$.

(iv) *in the case $\kappa > 2 + 2\nu$, assuming Assumption M:*

$$\left(\{n^{1/2}\,\hat{\gamma}_{X,Y}(k)\}_{k=0}^{m}, \hat{\gamma}_{X,X}(0), \hat{\gamma}_{Y,Y}(0), \hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow}$$
$$(\{G_k\}_{k=0}^{m}, \gamma_{X,X}(0), \gamma_{Y,Y}(0), \gamma_{X^2,Y^2}(0))$$
$$\left(\{n^{1/2}\hat{\rho}_{X,Y}(k)\}_{k=1}^{m}, \hat{\mu}_{X^2Y^2}\right) \overset{\mathcal{L}}{\Rightarrow} \left(\{G_k/\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)}\}_{k=1}^{m}, \mathbb{E}|X|^{2+2\nu}\right),$$

where the G_k satisfy $\mathrm{Cov}(G_j, G_k) = \sum_{\ell} \mathrm{Cov}(X_0 Y_j, X_\ell Y_{\ell+k})$ *and are multivariate normal.*

Remark 4: The rates of convergence for the sample cross-correlation, and its root, vary considerably depending upon κ. To summarize (recall that $a_n \sim n^{1/\kappa}$ up to a constant), when $\nu \in (0,1)$ the rate is n^0 for $\kappa < 2\nu$; $n^{1/2-\nu/\kappa}$ for $\kappa \in (2\nu, 2)$; $n^{1-(1+\nu)/\kappa}$ for $\kappa \in (2, 2+2\nu)$; $n^{1/2}$ for $\kappa > 2+2\nu$. Whereas for $\nu \geq 1$, the rates are n^0 for $\kappa < 2$; $n^{1/2-1/\kappa}$ for $\kappa \in (2, 2\nu)$; $n^{1-(1+\nu)/\kappa}$ for $\kappa \in (2\nu, 2+2\nu)$; $n^{1/2}$ for $\kappa > 2+2\nu$. See Figure 1 for a depiction of this rate in the case of $\nu \in \{0.5, 2\}$; the odd behavior is due to the unusual normalization involving both $\hat{\gamma}_{X,X}(0)$ and $\hat{\gamma}_{Y,Y}(0)$.

2.3 Self-normalization for Correlation Statistics

We next provide the asymptotic results for self-normalized roots.

Theorem 5: *Let $\{X_t\}$ be a PGARCH(p,q) process (2) satisfying Assumptions A, B, and D; also assume that Z_0 is symmetric about zero.*

(i) *With quantities defined in Theorem 2 according to the value of κ,*

$$n\left\{\frac{\hat{\rho}_{X,X}(k)}{\hat{\sigma}_{X,X}}\right\}_{k=1}^{m} \overset{\mathcal{L}}{\Rightarrow} \begin{cases} \left\{\dfrac{V_k}{V_0}\right\}_{k=1}^{m} & \text{if } \kappa \in (0,2) \\ \left\{\dfrac{V_k}{\gamma_{X,X}(0)\sqrt{W_4}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2,4) \\ \left\{\dfrac{G_k}{\gamma_{X,X}(0)\sqrt{\mathbb{E}X^4}}\right\}_{k=1}^{m} & \text{if } \kappa > 4. \end{cases} \tag{17}$$

(ii) With quantities defined in Theorem 3 according to the value of κ, and assuming Assumption M when $\kappa > 2\nu$,

$$n\left\{\frac{\hat{\rho}_{Y,Y}(k) - \rho_{Y,Y}(k)}{\hat{\sigma}_{Y,Y}}\right\}_{k=1}^{m} \overset{\mathcal{L}}{\Rightarrow} \begin{cases} \left\{\dfrac{V_k}{V_0}\right\}_{k=1}^{m} & \text{if } \kappa \in (0, 2\nu) \\ \left\{\dfrac{U_k - \rho_{Y,Y}(k)U_0}{\gamma_{Y,Y}(0)\sqrt{W_{4\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2\nu, 4\nu) \\ \left\{\dfrac{G_k - \rho_{Y,Y}(k)G_0}{\gamma_{Y,Y}(0)\sqrt{\mathbb{E}[|X|^{4\nu}]}}\right\}_{k=1}^{m} & \text{if } \kappa > 4\nu. \end{cases} \tag{18}$$

(iii) If $\nu \in (0, 1)$, with quantities defined in Theorem 4 according to the value of κ

$$n\left\{\frac{\hat{\rho}_{X,Y}(k)}{\hat{\sigma}_{X,Y}}\right\}_{k=1}^{m} \overset{\mathcal{L}}{\Rightarrow} \begin{cases} \left\{\dfrac{V_k}{\sqrt{W_2\, W_{2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (0, 2\nu) \\ \left\{\dfrac{V_k}{\sqrt{\gamma_{Y,Y}(0)\, W_2^{1+\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2\nu, 2) \\ \left\{\dfrac{V_k}{\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)\, W_{2+2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2, 2+2\nu) \\ \left\{\dfrac{G_k}{\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)\,\mathbb{E}|X|^{2+2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa > 2+2\nu. \end{cases} \tag{19}$$

If $\nu \geq 1$, with quantities defined in Theorem 4 according to the value of κ

$$n\left\{\frac{\hat{\rho}_{X,Y}(k)}{\hat{\sigma}_{X,Y}}\right\}_{k=1}^{m} \overset{\mathcal{L}}{\Rightarrow} \begin{cases} \left\{\dfrac{V_k}{\sqrt{W_2\, W_{2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (0, 2) \\ \left\{\dfrac{V_k}{\sqrt{\gamma_{X,X}(0)\, W_{2\nu}^{1+1/\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2, 2\nu) \\ \left\{\dfrac{V_k}{\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)\, W_{2+2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa \in (2\nu, 2+2\nu) \\ \left\{\dfrac{G_k}{\sqrt{\gamma_{X,X}(0)\,\gamma_{Y,Y}(0)\,\mathbb{E}|X|^{2+2\nu}}}\right\}_{k=1}^{m} & \text{if } \kappa > 2+2\nu. \end{cases}$$

In each case, the studentized correlations have a nondegenerate limit distribution, and the rate of convergence does not depend on κ. Although the limit distributions are not pivotal, they can be approximated via subsampling.

Remark 5: Although our normalization by the sample moment matches the growth rate of its corresponding root, it is not scale-invariant. Hence, to obtain a scale-invariant statistic, we propose to divide each sample kth moment by the exponential of the log moment

$$\hat{\mu}_{\log\,(|X|^k)} = n^{-1}\sum_{t=1}^{n} \log\,(|X_t|^k).$$

This quantity converges to $\mathbb{E}\log\,(|X|^k)$ no matter the value of κ, and it also scales such that $\hat{\mu}_{\log\,(|aX|^k)} = \log\,(|a|^k) + \hat{\mu}_{\log\,(|X|^k)}$. Since $\hat{\mu}_{\log\,(|X|^k)}$ converges in probability to its expectation, we can use its exponential to correct for scale; $\exp\,\hat{\mu}_{\log\,(|aX|^k)} = |a|^k \exp\,\hat{\mu}_{\log\,(|X|^k)}$.

For notational transparency the estimators of Subsection 1.3 are expressed in terms of the original normalizations, but the simulations in Appendix C (Supplementary Data) contain results with the scale-invariant normalizations introduced in Remark 5.

2.4 Consistency of the Subsampling Estimators

To establish the consistency of the various subsampling estimators introduced in Section 1, we employ the mixing properties of the corresponding processes and Theorem 11.3.1 of Politis, Romano, and Wolf (1999) combined with the convergence results derived in Section 2. By consistency, we mean that the subsampling distribution estimators converge in probability to the respective cdf of the limiting distributions.

Corollary 2: *Let $\{X_t\}$ be a PGARCH(p,q) process (2) satisfying Assumptions A, B, and D; also assume that Z_0 is symmetric about zero. Assume that $b/n + 1/b \to 0$ as $n \to \infty$.*

(i) With quantities defined in Theorem 2 according to the value of κ

$$\hat{L}_{X,X,k}(x) \xrightarrow{P} L_{\infty,X,X,k}(x)$$

for each x that is a continuity point of the limit distribution of (17). If this distribution is continuous, the convergence is also uniform. Moreover, the asymptotic coverage of the intervals (4) and (5) is the nominal level $1-p$.

(ii) With quantities defined in Theorem 3 according to the value of κ, and assuming Assumption M when $\kappa > 2\nu$,

$$\hat{L}_{Y,Y,k}(x) \xrightarrow{P} L_{\infty,Y,Y,k}(x)$$

for each x that is a continuity point of the limit distribution of (18). If this distribution is continuous, the convergence is also uniform. Moreover, the asymptotic coverage of the intervals (6) and (7) is the nominal level $1-p$.

(iii) With quantities defined in Theorem 4 according to the value of κ

$$\hat{L}_{X,Y,k}(x) \xrightarrow{P} L_{\infty,X,Y,k}(x)$$

for each x that is a continuity point of the limit distribution of (19). If this distribution is continuous, the convergence is also uniform. Moreover, the asymptotic coverage of the intervals (8) and (9) is the nominal level $1-p$.

3 Application to Returns on Stock Indices

We consider two datasets analyzed in Section 5.5 of Francq and Zakoian (2010). One corresponds to the daily log returns of CAC 40 stock index from 02/03/1990 to 29/12/2006 (4244 observations) and other to those of FTSE 100 index from 04/04/1984 to 03/04/2007 (5811 observations). According to Francq and Zakoian (2010), a GARCH(1,1) model is a reasonable guess for the first series, but higher order GARCH or ARCH is more likely to fit the second series.

In Figure 2, we show Hill plots of the returns. For both series, the assumption about the finiteness of the second moment does not seem to be too unreasonable; however, the assumption regarding the finiteness of the fourth moment might be questionable.

The 95% confidence intervals for the acf of the returns are plotted in Figure 3. PLUGIN intervals were calculated by fitting a GARCH(1,1) model with normal innovations (`R` routine `garchFit` of package `fGarch`) to the returns and plugging in the estimated α_1 and β_1 values into the exact normal asymptotic confidence bounds (under normal errors) formula. The estimated values for CAC 40 returns were $\hat{\alpha}_1 = 0.076907$ and $\hat{\beta}_1 = 0.907047$

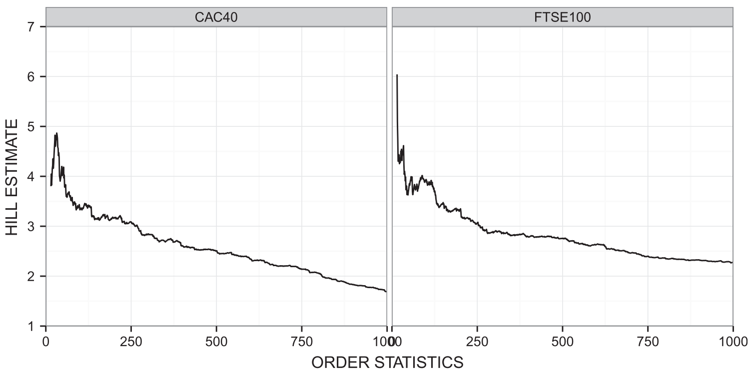

Figure 2. Hill plot of CAC 40 and FTSE 100 returns.

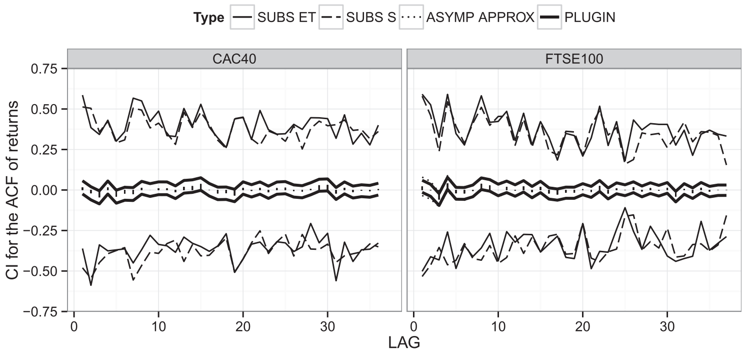

Figure 3. 95% confidence intervals for the acs of returns of CAC 40 and FTSE 100 indices based on several methods.

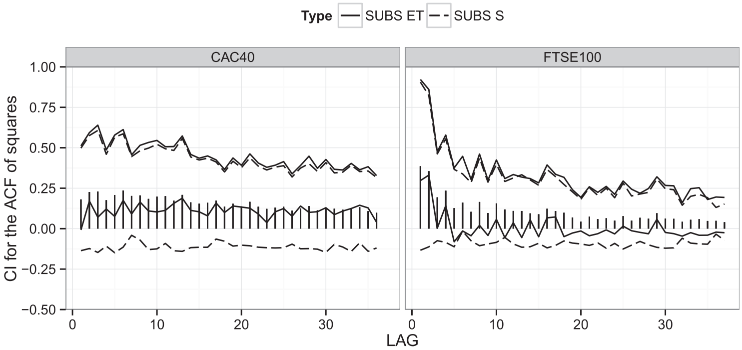

Figure 4. 95% subsampling confidence intervals for the acs of squared returns of CAC 40 and FTSE 100 indices.

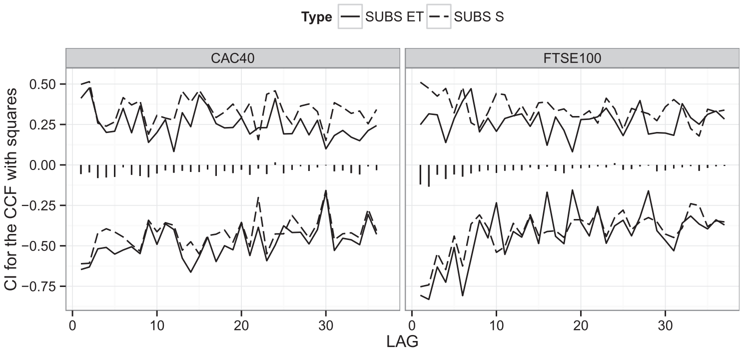

Figure 5. 95% subsampling confidence intervals for the cross-correlations of the returns and squared returns of CAC 40 and FTSE 100 indices.

($\hat{\alpha}_0 = 0.000003$) while those for FTSE 100 returns were $\hat{\alpha}_1 = 0.089941$ and $\hat{\beta}_1 = 0.891497$ ($\hat{\alpha}_0 = 0.000002$). The remaining confidence intervals are those employed in the simulation study of Appendix C (Supplementary Data).

For CAC 40 returns, the confidence intervals for $\rho_{X,X}(k)$ based on the subsampling method are much wider than the other two types and oscillate around ± 0.4. For both series, and all interval types, the acs seem to be not significantly different from zero.

The 95% confidence intervals for $\rho_{X^2,X^2}(k)$ are depicted in Figure 4 (these intervals have to be treated with caution if we are unsure about the finiteness of the fourth moment, because in such case $\rho_{X^2,X^2}(k)$ is undefined). As pointed out in Appendix C (Supplementary Data), the symmetric confidence intervals are too wide for the acs of the squares, hence we focus our attention on the equal-tailed intervals. For CAC 40 and FTSE 100 squared returns, the width of the confidence intervals for the acs is about 0.25. Bounds for CAC 40 indicate slower decay of $\rho_{X^2,X^2}(k)$ compared to bounds for FTSE 100.

The 95% confidence intervals for $\rho_{X,X^2}(k)$ are shown in Figure 5. The two subsampling methods yield similar bounds, which are slightly asymmetric. Both types of intervals indicate that the process' cross-correlations are non-significant.

4 Summary

In this article, we study the asymptotic properties of sample autocovariances and acs of PGARCH processes. Autocorrelations of a time series and its powers have been used in the literature to determine a nonlinear process' serial structure, but the inference for PGARCH is complicated because the rate of convergence of sample autocovariance and autocorrelation estimators depends upon the tail index. This tail index is typically unknown, unless it has been previously estimated. However, through an appropriate studentization of autocovariance/autocorrelation roots (i.e., the estimator minus its estimand) it is possible to avoid the necessity of knowing the tail index; this strategy was adopted for heavy-tailed moving average time series in Davis and Resnick (1985) and McElroy and Politis (2002). The latter paper approximated the limiting quantiles of the studentized root by the subsampling methodology. Here we study the related inference problem for PGARCH processes: we derive the limiting distributions, provide effective studentizations, and examine subsampling methods for estimating the limiting quantiles.

To obtain nondegenerate weak limits of all studentized roots, we derive the recursive relationships for the PGARCH and thus substantially extend existing theoretical results. A simulation study, which can be found in Appendix C (Supplementary Data), indicates that the subsampling confidence intervals for the acs of GARCH processes with a finite fourth moment are generally wider than the asymptotic confidence intervals (approximate or exact). The subsampling approach can still be employed when no asymptotic formula is available (providing intervals for the acs when the fourth moment is infinite, and for the acs of squares and ccs of the process with its squares). Nevertheless, the subsampling-based empirical coverage probabilities tend to be higher than the nominal level. Equal-tailed subsampling confidence intervals are preferable over the symmetric ones.

Supplementary Data

Supplementary data are available at *Journal of Financial Econometrics* online.

References

Baek, C., V. Pipiras, H. Wendt, and P. Abry. 2009. Second Order Properties of Distribution Tails and Estimation of Tail Exponents in Random Difference Equations. *Extremes* 12: 361–400.

Basrak, B., R. Davis, and T. Mikosch. 2002. Regular Variation of GARCH Processes. *Stochastic Processes and their Applications* 99: 95–115.

Bollerslev, T. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31: 307–327.

Bougerol, P. and N. Picard. 1992. Stationarity of GARCH Processes and of Some Nonnegative Time Series. *Journal of Econometrics* 52: 115–127.

Brockwell, P. J. and R. A. Davis. 1991. *Time Series: Theory and Methods*. New York: Springer.

Carrasco, M. and X. Chen. 2002. Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models. *Econometric Theory* 18: 17–39.

Davis, R. 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 S. I. Resnick. 1985. Limit Theory for Moving Averages of Random Variables with Regularly Varying Tail Probabilities. *The Annals of Probability* 13 (1): 179–195.

Francq, C. and J.-M. Zakoian. 2010. *GARCH Models. Structure, Statistical Inference and Finanial Applications*. UK: Wiley.

Fryzlewicz, P. and S. Subba Rao. 2011. Mixing Properties of ARCH and Time-Varying ARCH Processes. *Bernoulli* 1: 320–346.

Hall, P. and Q. Yao. 2003. Inference in ARCH and GARCH Models. *Econometrica* 71: 285–317.

Huang, D., H. Wang, and Q. Yao. 2008. Estimating GARCH Models: When to Use What? *Econometrics Journal* 11: 27–38.

Jach, A., T. McElroy, and D. Politis. 2012. Subsampling Inference for the Mean of Heavy-Tailed Long-Memory Time Series. *Journal of Time Series Analysis* 33: 96–111.

Kokoszka, P. and D. Politis. 2011. Nonlinearity of ARCH and Stochastic Volatility Models and Bartlett's Formula. *Probability and Mathematical Statistics* 31: 47–59.

Kokoszka, P., G. Teyssière, and A. Zhang. 2004. "Confidence Intervals for the Autocorrelations of the Squares of GARCH Sequences". In Computational Science - ICCS 2004*, Volume 3039 of* Lecture Notes in Computer Science. Springer, pp. 837–844.

Lindner, A. 2009. Stationarity, Mixing, Distributional Properties and Moments of GARCH(p, q)-Processes. In J.-P. K. T. Andersen, R. Davis and T. Mikosch (eds.), *Handbook of Financial Time Series*. Berlin, Germany: Springer, pp. 481–496.

McElroy, T. and D. N. Politis. 2002. Robust Inference for the Mean in the Presence of Serial Correlation and Heavy-Tailed Distributions. *Econometric Theory* 18: 1019–1039.

McElroy, T. and D. N. Politis. 2007. Self-Normalization for Heavy-Tailed Time Series with Long Memory. Statistica Sinica 17 (1): 199–220.

Mikosch, T. and C. Stărică. 2000. Limit Theory for the Sample Autocorrelation and Extremes of a GARCH(1, 1) Process. *The Annals of Statistics* 28: 1427–1451.

Mittnik, S., M. Paolella, and S. Rachev. 2002. Stationarity of Stable Power-GARCH Processes. *Journal of Econometrics* 106: 97–107.

Peng, L. and Q. Yao. 2003. Least Absolute Deviations Estimation for ARCH and GARCH Models. *Biometrika* 90: 967–975.

Politis, D. N., J. P. Romano, and M. Wolf. 1999. *Subsampling*. New York: Springer.

Tully, E. and B. Lusey. 2007. A Power GARCH Examination of the Gold Market. *Research in International Business and Finance* 21: 316–325.

Wagner, N. and T. Marsh. 2004. Measuring Tail Thickness under GARCH and an Application to Extreme Exchange Rate Changes. *Journal of Empirical Finance* 12: 165–185.