Computational Statistics manuscript No.
(will be inserted by the editor)

Empirical wavelet analysis of tail and memory properties of LARCH and FIGARCH models

Agnieszka Jach · Piotr Kokoszka

Received: date / Accepted: date

Abstract Using computationally efficient wavelet methods, we study two nonlinear models of financial returns $\{r_t\}$: linear ARCH (LARCH) and fractionally integrated GARCH (FIGARCH). We estimate the tail index α and the long memory parameter d of the squared returns $X_t = r_t^2$ of LARCH, and of the powers $X_t = |r_t|^p$ of FIGARCH. We find that the X_t have infinite variance and long memory, and show how the estimates of α and d depend on the model parameters. These relationships are determined empirically, as the setting is quite complex, and no suitable theory has been developed so far. In particular, we provide empirical relationships between the estimates $\hat{d}$ and the difference parameters in LARCH and FIGARCH. Our computational work uncovers tail and memory properties of LARCH and FIGARCH for practically relevant parameter ranges, and provides some guidance on modeling returns on speculative assets including FX rates, stocks and market indices.

Keywords Heavy tails · Long memory · Volatility · Wavelets

Mathematics Subject Classification (2000) 62M10 · 42C40

Research supported by NSF grants DMS-0413653 and DMS-0804165.

Agnieszka Jach
Departamento de Estadística
Universidad Carlos III de Madrid
28903 Getafe-Madrid, Spain
Tel.: +34-91-6249893
Fax: +34-91-6248749
E-mail: ajach@est-econ.uc3m.es

Piotr Kokoszka
Department of Mathematics and Statistics
Utah State University
Logan UT, 84322, USA
Tel.: +1-435-7970746
Fax: +1-435-7971822
E-mail: Piotr.Kokoszka@usu.edu

1 Introduction

We present a computationally intensive study of two nonlinear models for financial returns $\{r_t\}$, the LARCH process proposed by Robinson (1991) and developed by Giraitis et al (2000b), and the FIGARCH introduced by Baillie et al (1996), see Baillie et al (2007), Beine et al (2002), Giraitis et al (2003), Giraitis et al (2004), Lardic and Mignon (2004), Lee (2005) for a few recent applications. These two processes were introduced to model observed properties of long series of returns on speculative assets which include lack of autocorrelation in r_t, but significant autocorrelations of the squares r_t^2 at large lags, and possibly infinite fourth moments of the r_t's. Even though the FIGARCH and LARCH processes were introduced to model long memory in volatility and non-Gaussian heavy-tailed marginal distributions, these two fundamental properties are only partially theoretically understood. It is known that the FIGARCH process has infinite variance, and that the LARCH process has long memory in squares if it has finite fourth moment. The empirically most relevant case of infinite fourth moment LARCH has not been studied, either theoretically or through simulations. Using new computationally efficient tools based on the wavelet analysis, we investigate the tail and memory properties of the LARCH and FIGARCH models for broad practically relevant ranges of their parameters. We also discuss the implications of our computational analysis for modeling daily returns on financial assets.

We now recall the definitions of the two models under consideration. A series $\{r_t\}$ is said to be LARCH if it satisfies

$$r_t = \sigma_t \varepsilon_t \tag{1}$$

and

$$\sigma_t = b_0 + \sum_{j=1}^{\infty} b_j r_{t-j}, \qquad b_0 \neq 0. \tag{2}$$

We assume that ε_t's are iid standard normal. Theorem 2.1 of Giraitis et al (2000b) shows that, if

$$B_2 = \sum_{j=1}^{\infty} b_j^2 < 1, \tag{3}$$

then equations (1), (2) have a strictly stationary solution with $E[r_t^2] < \infty$. If

$$b_j \sim c j^{d-1} \qquad \text{for some } d \in (0, 1/2),\ c > 0 \tag{4}$$

and

$$7\sqrt{3} \sum_{j=1}^{\infty} b_j^2 < 1 \qquad (B_2 < 0.082479), \tag{5}$$

then, by Theorems 2.2 and 2.3 of Giraitis et al (2000b), $E[X_t^2] < \infty$, where $X_t = r_t^2$,

$$\mathrm{Cov}(X_0, X_k) \sim C k^{2d-1}, \qquad C > 0 \tag{6}$$

and the normalized partial sums of $\{X_t\}$ tend to a fractional Brownian motion (fBm) with Hurst parameter $H = d + 1/2$.

A series $\{r_t\}$ is said to be a FIGARCH process, if it satisfies (1) and

$$\sigma_t^2 = a_0 + \sum_{j=1}^{\infty} a_j r_{t-j}^2 \tag{7}$$

with

$$\sum_{j=1}^{\infty} a_j = 1, \quad a_j \sim cj^{-\delta-1} \text{ for some } \delta \in (0,1). \tag{8}$$

The moments $|r_t|^p$, $p < 2$ are finite, but since r_t has infinite variance (Baillie et al (1996) and Douc et al (2008)), we cannot compute the covariances of $X_t = r_t^2$ as in (6), and we even have $E|X_t| = \infty$.

One of the most important characteristics of financial returns is excess kurtosis or heavy tails, see e.g. Chapter 1 of Campbell et al (1997). We say that $\{X_t\}$ is heavy-tailed with index $\alpha > 0$ if

$$P(|X_t| > x) \sim cx^{-\alpha}, \quad x \to \infty,$$

where $\sim$ indicates that the ratio of the left- and right-hand sides tends to 1. Our goal is to determine the tail indices of $X_t = r_t^2$, when $\{r_t\}$ is modeled by LARCH, and of $X_t = |r_t|^p$, $p \in (0, 2]$, when $\{r_t\}$ follows the FIGARCH model. In the case of the LARCH nothing is known about the tail index if condition (5) is violated. Douc et al (2008) conjecture that for FIGARCH, $X_t = r_t^2$ have heavy tails with index $\alpha = 1$ and rescaled partial sums of $X_t = |r_t|^p$, $p < 1$ may converge to fBm with $H \in (0.5, 1)$. Known theoretical results on the tails of the marginal distributions of nonlinear processes include those for GARCH, ARCH and the stochastic volatility model. Basrak et al (2002) showed that GARCH(p, q) processes have regularly varying marginal distributions. They used the embedding of the GARCH process into a vector process satisfying a stochastic recurrence equation (SRE). In GARCH(1,1) and ARCH(1) models it is possible to compute α from the model parameters. In these two cases, the tail index α is the (numerical) solution of the equation $h(\alpha) = 1$, where the function $h(\cdot)$ depends on the GARCH (Section 8.4 of Embrechts et al (1997), Mikosch and Stărică (2000)) or ARCH (de Hann et al (1989)) parameters. Davis and Mikosch (2009) characterized the tails of the marginals of stochastic volatility model by analyzing logarithm-transformed squares of the process and using an inheritance principle. These approaches cannot be readily extended to LARCH and FIGARCH models because it is not apparent how to express the infinite series in (2) and (7) via a SRE, and, in addition, the squares of LARCH and powers of FIGARCH are likely to have distributional long memory. We therefore resort to a computational methodology of estimating tail index based on the discrete wavelet transform. The advantage of working with the wavelet coefficients of $\{X_t\}$ rather than the process itself is that the former are approximately uncorrelated, whilst $\{X_t\}$ has a complex dependence structure which makes estimation of the tail index very difficult. Abry et al (2000) used this principle to estimate α in the linear fractional stable motion, Jach and Kokoszka (2008) in α-stable FARIMA and Kokoszka et al (2006) to study probability tails of magnetometer records.

Turning now to the long memory property, recall that a weakly stationary process $\{X_t\}$ is typically said to have long memory or to be long-range dependent (LRD) if

$$\mathrm{Cov}(X_0, X_k) \sim ck^{2d-1}, \quad k \to \infty, \ (c > 0),$$

where $d \in (0, 1/2)$ is the memory parameter. This is often referred to as the covariance long memory, whilst the convergence of appropriately normalized partial sums $S_N(t) = \sum_{k=1}^{[Nt]}(X_k - E[X_k])$ to a random process with LRD increments, as distributional long memory, see Giraitis et al (2008).

We have seen that under assumption (5), the squares of the LARCH process have both covariance and distributional long memory. However, nothing is known about its memory properties if (5) fails to hold. In particular the relationship between the parameter d in (4) and the estimated long memory parameter of the $X_t = r_t^2$ is unknown. Even though the FIGARCH process was defined already in Baillie et al (1996), the proof of the existence of a stationary solution to the FIGARCH equations was delivered only by Douc et al (2008). While the definition of covariance long memory for this infinite-variance model is irrelevant, the distributional long memory of its powers $|r_t|^p$ is of interest as the intent of Baillie et al (1996) was to model long memory in volatility. There are no theoretical results on the limits of the partial sums of the abovementioned processes. The relationship between the memory parameter d of the $|r_t|^p$'s and δ is unknown, thus our goal is to establish a connection between the estimates of d and δ.

Teyssière and Abry (2006) used the wavelet regression estimator of Abry et al (2002), to estimate d in nonlinear processes including squares of LARCH model, if condition (5) holds. The wavelet estimator of Stoev et al (2002), which we apply here, is an extension to the case of infinite-variance data. Another estimation method is the wavelet pseudo maximum likelihood approach of Jach and Kokoszka (2008), which builds on the work of Craigmile et al (2005). Moulines et al (2008) studied this estimator under the assumption of linearity and finite fourth moment. None of the above wavelet schemes was applied to nonlinear infinite-variance processes.

The LARCH and FIGARCH models introduced above are defined in terms of infinite sums over $1 \le j < \infty$. In practical modeling settings, these sums are replaced by finite sums over $1 \le j \le j_{\text{trunc}}$. In this paper, we use $j_{\text{trunc}} = 10^4$ (B_2 is thus $\sum_{j=1}^{j_{\text{trunc}}} b_j$). All our findings pertain to this truncation level, but we note that the conclusions remain largely unchanged if it is replaced by truncation levels as large as 10^5. In the case of the FIGARCH model, using the same j_{trunc} for all values of δ means that the values of the sum $\sum_{j=1}^{\infty} a_j$ change with δ. An alternative approach would therefore be to vary j_{trunc} with δ, and keep the value of $1 - \sum_{j=1}^{\infty} a_j$ equal to the same small number for all δ's. We however prefer to keep j_{trunc} fixed, as has been done in empirical work (Baillie et al (1996)). For all models, we generate $R = 100$ replications of length $N = 10^4$ (after discarding a presample of the same length).

The paper is organized as follows. In Section 2 we describe the methods and results of tail index estimation in the squares of LARCH and powers of FIGARCH; Section 3 deals with their memory properties. In Section 4 we explore the tail and memory properties of the return data from a large collection of financial time series, and connect them to the computational analysis of the previous sections. We outline the conclusions of our study in Section 5.

2 Tail index estimation

In this section, we explain a wavelet-based methodology of tail index estimation of LRD time series, and apply it to the squares of the LARCH and the powers of the FIGARCH process.

Let $X_0, X_1, \ldots, X_{N-1}$ be a long-memory sequence of nonnegative heavy-tailed observations with index α. We apply the discrete wavelet transform (DWT) to X_t's and denote by $W_{j,k}$ the wavelet coefficients at level (octave) j and time $2^j k$. The level index j takes values $j = 1, 2, \ldots, J$, $J = \log_2\lfloor N \rfloor$ ($\lfloor \cdot \rfloor$ denotes the integer part), while for fixed j, $k = 0, 1, \ldots, N_j - 1$, $N_j = 2^{J-j}$. The vectors $\boldsymbol{W}_j = [W_{j,0}, W_{j,1}, \ldots, W_{j,N_j-1}]^T$ extract changes in $\{X_t\}$ over time scales 2^j, and are obtained by filtering $\{X_t\}$ with a level j filter derived from a (mother) wavelet filter of length L. We refer the reader to Percival and Walden (2000) for a detailed background. The $\boldsymbol{W}_j$'s are approximately uncorrelated and within each of these vectors the $W_{j,k}$'s are approximately uncorrelated. The longer the filter, the better the decorrelation, but the use of very long filters reduces the number of available wavelet coefficients, due to the exclusion of the boundary ones. The idea of tail index estimation is that although the process $\{X_t\}$ exhibits (or may exhibit) long memory, the $W_{j,k}$'s are approximately uncorrelated, and tail index estimation is easier for such nearly independent variates. Due to the abundance of the coefficients at level $j = 1$, we consider the vector $\boldsymbol{W}_1$ as an idealized iid heavy-tailed sequence with tail index α. This approach was first exploited by Abry et al (2000) to estimate tail index α in a linear fractional stable motion.

Once the problem of nontrivial dependence structure in the observations X_t has been mitigated, one can focus on the tail index estimation based on $\boldsymbol{W}_1$. Below we describe an approach introduced by Abry et al (2000) and justified theoretically by Stoev and Taqqu (2005) in the case when a heavy-tailed distribution with parameter α is an exact α-stable distribution. This approach uses the concept of self-similarity. Stoev et al (2006) recently developed a method based on the idea of max self-similarity. However, the estimator of Stoev et al (2006), although superior to the well-known Hill estimator, tends to overestimate the tail index, when the observations come from the α-stable distribution with α close to 2. Because of the importance of this case in our analysis, we chose the method of Abry et al (2000).

Consider then a sequence $\{X_t^{'}\}$ of iid heavy-tailed observations with index α. In our considerations we set $X_k^{'} = W_{1,k}$, $k = 0, 1, \ldots, n-1$, $n = 2^{J-1}$, where $\boldsymbol{W}_1$ are DWT coefficients at level $j = 1$ of a realization $X_0, X_1, \ldots, X_{N-1}$ of a long-memory process $\{X_t\}$ with tail index α. The partial sums of iid heavy-tailed random variables with index $0 < \alpha < 2$ converge to the stable Lévy motion (Gnedenko and Kolmogorov (1954)), which has the self-similarity index $H^{'} = 1/\alpha$. An estimator of α is thus obtained as

$$\hat{\alpha} = 1/\hat{H}^{'},$$

where $\hat{H}^{'}$ is an estimator of the self-similarity parameter of the partial sum process. A wavelet-based generalized least-squares regression estimator is defined as $\hat{H}^{'} = \sum_{j=j_1}^{j_2} w_j Y_j^{'}$. The weights $\{w_j\}$ are such that $\sum_{j=j_1}^{j_2} w_j = 0$ and $\sum_{j=j_1}^{j_2} j w_j = 1$, and the $Y_j^{'}$ are $Y_j^{'} = \frac{1}{n_j} \sum_{k=0}^{n_j-1} \log_2 |W_{j,k}^{'}|$, where $W_{j,k}^{'}$ is the k-th wavelet coefficient at level j obtained from the sequence $\{X_t^{'}\}$, and n_j is the number of nonboundary wavelet coefficients at level j, $j = 1, \ldots, I$, $I = \lfloor \log_2(n) \rfloor$.

Table 1 LARCH A, B, C model specifications of $\{r_t\}$ for selected d's and the corresponding values of B_2.

	LARCH A ($b_0 = 0.1,\ \theta = 0,\ \phi = 0$)								
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
B_2	0.0044	0.0195	0.0488	0.0984	0.1788	0.3094	0.5276	0.9090	1.6147
	LARCH B ($b_0 = 0.1,\ \theta = 0,\ \phi = -0.2$)								
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
B_2	0.0264	0.0196	0.0237	0.0430	0.0847	0.1621	0.3012	0.5545	1.0337
	LARCH C ($b_0 = 0.1,\ \theta = -0.2,\ \phi = 0$)								
d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
B_2	0.0235	0.0150	0.0167	0.0326	0.0692	0.1389	0.2656	0.4978	0.9383

Combining the results of Abry et al (2000) and Pipiras et al (2007) (with suitably chosen mother wavelet) we have

$$\sqrt{n}(\hat{H}^{'} - H^{'}) \xrightarrow{d} N\left(0, (\log_2(e))^2\pi^2(1 + 2(H^{'})^2)\left(\sum_{i,j=j_1}^{j_2} w_i w_j 2^j\right)/12\right), \quad (9)$$

so it is possible to select the lower scale j_1 automatically, following, for example, procedure in Section 5.2 of Stoev et al (2006), as well as calculate asymptotic confidence bounds for α. For $J = \lfloor \log_2(10^4) \rfloor = 13$ the upper cut-off j_2 is set to 8, which corresponds to the penultimate level unaffected by the boundary coefficients.

In large sample, it is also possible to construct $100(1-\beta)\%$ confidence intervals for $H^{'}$ based on a normal approximation

$$\left(\bar{\hat{H}}' \pm z_{\beta/2} s/\sqrt{R}\right), \quad (10)$$

where $\bar{\hat{H}}'$ and s are the sample mean and sample standard deviation based on R estimates $\hat{H}^{'}$ and $z_{\beta/2}$ is the $\beta/2$-th quantile of the standard normal distribution. Corresponding confidence intervals for α result from inverting the end-points of (10).

In Sections 2.1 and 2.2 we also assess the assumption of independence of the wavelet coefficients at level $j = 1$. This is done by calculating the empirical size of a nonparametric difference-sign test (Chapter 1 of Brockwell and Davis (2002)) for a randomly permuted sequence $\boldsymbol{W}_1$. Random permutation is introduced to mitigate the residual effects of only approximate decorrelation of the DWT and to such permuted sequences we apply the tail estimator.

2.1 Tail properties of the LARCH process

We now present the results of tail index estimation for $X_t = r_t^2$, where $\{r_t\}$ follows the LARCH model. We use the least asymmetric filter with 4 vanishing moments (Percival and Walden (2000)). We generate replications of three LARCH models, denoted A, B, C, corresponding to three FARIMA filter parametrizations of the b_j's (Table 1) for $d \in \{0.05, 0.06, \ldots, 0.45\}$. Explicitly, we set $b_0 = 0.1$, and for $j \geq 1$ define the b_j's via

$$\sum_{j=0}^{\infty} b_j z^j = (1-z)^{-d}\frac{1+\theta z}{1-\phi z}.$$

Table 2 Empirical size of the independence test (nominal level 5%) for permuted level $j = 1$ wavelet coefficients of squares of LARCH A, B, C models based on $R = 1000$ replications.

	Empirical size								
Model \ d	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
LARCH A	0.047	0.052	0.058	0.056	0.055	0.046	0.038	0.048	0.047
LARCH B	0.051	0.058	0.048	0.047	0.057	0.053	0.042	0.049	0.053
LARCH C	0.057	0.041	0.041	0.053	0.053	0.049	0.065	0.053	0.057

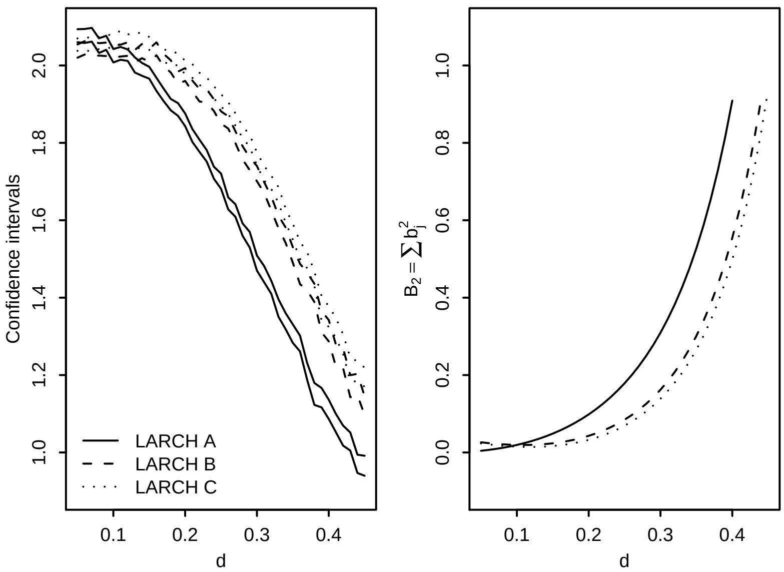

Fig. 1 *Left:* 95% confidence intervals (10) for the tail index of $X_t = r_t^2$, where $\{r_t\}$ follows the three LARCH models (Table 1) as a function of memory parameter d, $d \in \{0.05, 0.06, \ldots, 0.45\}$, based on $R = 100$ replications of length $N = 10^4$. *Right:* Values of B_2 for the same LARCH models, as a function of d.

When calculating tail estimates, we check the assumption of independence of $\boldsymbol{W}_1$ of the squares of LARCH models. The empirical sizes of the difference-sign test (Table 2) are very close to the nominal level of $\beta = 0.05$ with the standard error of $\sqrt{\beta(1-\beta)/R} \approx 0.0069$, suggesting the acceptance of the null hypothesis for all models and all values of the memory parameter d. In Figure 1 we present 95% confidence intervals (10) for the tail of squares of the three LARCH models under consideration as a function of d (left panel). For small d's, the estimates and the bounds should be treated with caution as the models have finite-variance. For larger d's, when condition (5) is violated, the estimates decrease with d and those for models B and C are larger compared with model A. In the right panel of Figure 1 we show the relationship between B_2 and d. There is a striking resemblance in the behavior of the curves between the two panels, where in the left $\hat{\alpha}$ decreases with d in the same manner as B_2 increases with d in the right. Estimates of α as a function of B_2, $B_2 \in (0, 1)$ for the same replications as in Figure 1, are displayed in Figure 2 (left panel). When B_2 is not too large the relationship between $\hat{\alpha}$ and B_2 can be well described by a power law. For large values of B_2 the

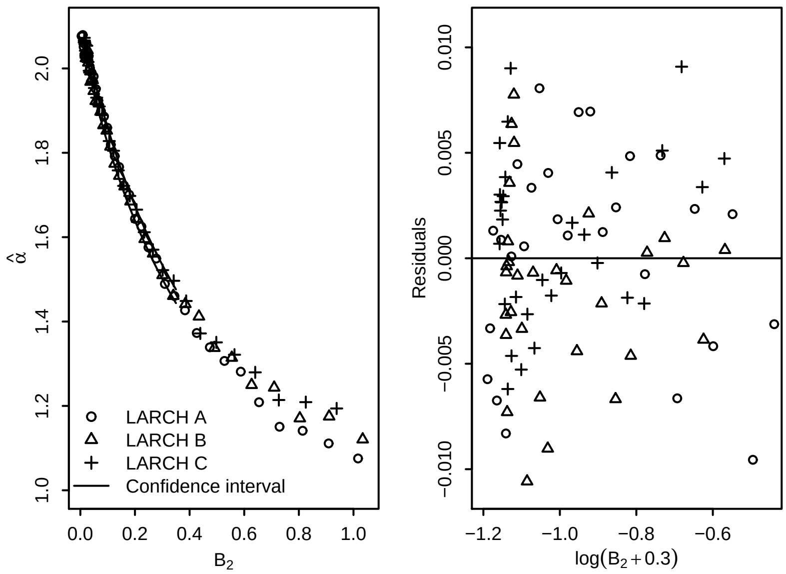

Fig. 2 *Left:* Estimates of the tail index based on the same $R = 100$ replications of $X_t = r_t^2$ as in Figure 1, where $\{r_t\}$ follows the three LARCH models (Table 1) as a function of B_2, $B_2 \in (0, 0.35)$, with 95% confidence bounds for the individual predictions derived from (11). *Right:* Residuals of the least-squares linear regression for $\log(\hat{\alpha} + 0.3)$ on $\log(B_2 + 0.3)$.

sum of d and $1/\hat{\alpha}$ yields values greater than 1, thus if this combination approximates Hurst index $H \in (0, 1)$ of a self-similar process with dependent increments, the range of admissible values of B_2 needs to be restricted. To explain, suppose α depends only on B_2 and is substantially less than 2. This indicates that the normalized partial sums of $\{X_t\}$ are attracted by a non-Gaussian self-similar process possibly with infinite second moment. A natural choice of the limiting process is a linear fractional stable motion (lfsm) with Hurst index $H = d+1/\alpha$, $\alpha \in (0, 2)$, $H \in (0, 1)$, $H \neq 1/\alpha$. We say, following Samorodnitsky and Taqqu (1994), that the infinite-variance increments of a lfsm, the linear fractional stable noise, have long memory if $H > 1/\alpha$, i.e., when $d \in (0, 1/\alpha)$. The upper bound on the self-similarity parameter H (and on its estimates) admits only those values of B_2 for which $d + 1/\hat{\alpha} < 1$, resulting in the condition $B_2 \in (0, 0.35)$, what agrees with our empirical findings. With this restriction in mind, we apply various transformations to $\hat{\alpha}$ or/and B_2 based on 88 pairs of observations, intending to establish closed-form relationship between these two variables. Fitting a least-squares regression line for $\log(\hat{\alpha}+0.3)$ on $\log(B_2+0.3)$ yields the most satisfactory result, with R-squared $= 0.9969$, residual plot with no pattern (right panel of Figure 2), residual standard deviation of 0.004437, and two estimates of the intercept and the slope with their standard errors, $a = 0.3913$ (0.002416) and $b = -0.4037$ (0.002423), respectively. In terms of the untransformed variables, we can quantify the relationship between the tail estimates and B_2 as

$$\hat{\alpha} = \overbrace{\mathrm{e}^{0.4}}^{1.5} (B_2 + 0.3)^{-0.4} - 0.3, \qquad B_2 \in (0, 0.35). \tag{11}$$

Table 3 Empirical size of the independence test (nominal level 5%) for permuted level $j = 1$ wavelet coefficients of $X_t = |r_t|^p$, where $\{r_t\}$ follows FIGARCH(0,δ,0) model based on $R = 1000$ replications. The last row corresponds to FIGARCH, $X_t = r_t$.

					Empirical size					
$p \setminus \delta$	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
2.0	0.054	0.047	0.061	0.057	0.049	0.051	0.050	0.041	0.059	0.052
1.8	0.051	0.056	0.041	0.043	0.050	0.057	0.062	0.053	0.057	0.057
1.6	0.041	0.053	0.051	0.047	0.053	0.056	0.052	0.045	0.059	0.058
1.4	0.051	0.053	0.043	0.051	0.059	0.054	0.052	0.057	0.047	0.064
1.2	0.050	0.041	0.053	0.051	0.043	0.063	0.059	0.044	0.047	0.057
1.0	0.049	0.064	0.052	0.045	0.065	0.042	0.038	0.034	0.040	0.060
0.8	0.046	0.052	0.048	0.052	0.055	0.064	0.049	0.052	0.061	0.063
0.6	0.047	0.052	0.038	0.045	0.060	0.053	0.060	0.054	0.043	0.044
0.4	0.055	0.041	0.038	0.053	0.053	0.059	0.041	0.044	0.064	0.061
0.2	0.055	0.066	0.056	0.042	0.053	0.064	0.052	0.052	0.050	0.048
	0.042	0.050	0.043	0.042	0.055	0.046	0.054	0.054	0.053	0.064

This empirically-derived formula for $\hat{\alpha}$ as a function of B_2, depicted through the 95% confidence intervals for the individual predictions in the left panel of Figure 2, calls for the theoretically-justified connection between α and B_2.

2.2 Tail properties of the FIGARCH process

We now assume that the $\{r_t\}$ follows a FIGARCH model, and consider time series of the form $X_t = |r_t|^p$, $p \in \{0.2, 0.4, \ldots, 2\}$ and $X_t = r_t$ (11 processes in total). As in the LARCH analysis, we apply the DWT with the least asymmetric filter with 4 vanishing moments and estimate heavy tail index α from the reshuffled first level wavelet coefficients of $\{X_t\}$. We generate replications of 11 FIGARCH models each defined for $\delta \in \{0.5, 0.55, \ldots, 0.95\}$. The filter $\{a_j\}$ is defined by $a_0 = 0.1$, and for $j \geq 1$ via

$$\sum_{j=0}^{\infty} a_j z^j = 1 - (1-z)^{\delta}.$$

As in Section 2.1, we assess the assumption of independence of the $\boldsymbol{W}_1$ for these models by examining the empirical sizes of the difference-sign (Table 3). In almost all cases this assumption is accepted at 5% level. In Figure 3 we show 95% confidence intervals for the tail index as a function of δ when $p < 1$ (left panel) and $p \geq 1$ (right panel). In the former case tail estimates are very similar for different δ's with values well above 2, suggesting that the tail estimation scheme might not be valid here. In the latter, tail estimates for δ's close to 0.5 are larger than those for δ close to 1. This might stem from the truncation effect that leads to the processes with lighter tails for larger differences $1 - \sum_{j=1}^{\infty} a_j$. In Douc et al (2008), the authors give a sufficient condition on the existence of a stationary solution to FIGARCH equations (formula (5) there). This condition yields approximately $\delta > 0.87$. If we were to consider only $\delta \in (0.87, 1)$, then the tail index estimates in the right panel, would not depend on δ. Absolute values of FIGARCH yield confidence bounds close to 2, which are larger than those of FIGARCH. The tails get heavier as p increases, and approach $\hat{\alpha} = 1$ when $p = 2$. When $p \in [1, 2]$, tail estimates of $X_t = |r_t|^p$ are between 1 and 2. These findings, suggesting

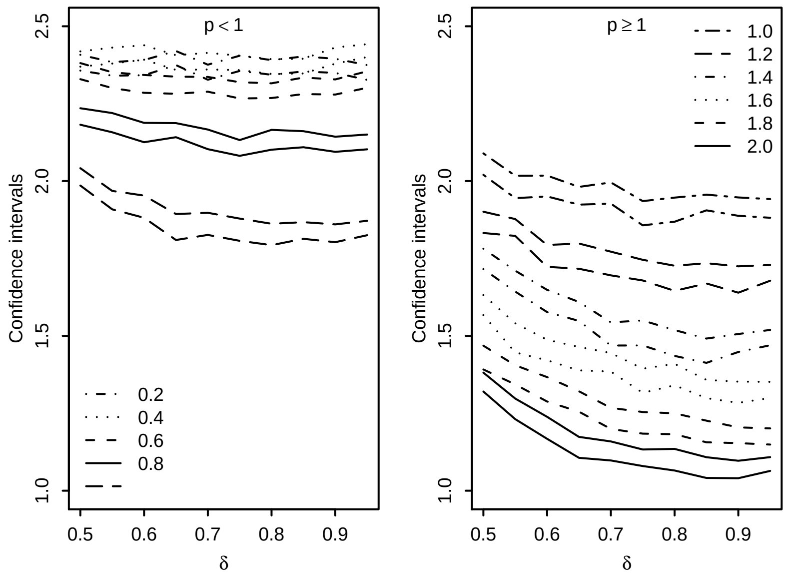

Fig. 3 *Left:* 95% confidence intervals (10) for the tail index of $X_t = |r_t|^p$, $p \in \{0.2, 0.4, \ldots, 0.8\}$ as a function of $\delta \in \{0.5, 0.55, \ldots, 0.95\}$, where $\{r_t\}$ follows the FIGARCH model. The unlabeled line corresponds to FIGARCH, $X_t = r_t$. *Right:* 95% confidence intervals when $p \in \{1, 1.2, \ldots, 2\}$.

that the marginal distributions of $\{|r_t|^p\}$, $p \in [1,2]$ have regularly varying tails with $\alpha \in [1,2]$, need to be verified theoretically.

3 Estimation of the memory parameter d

In this section, we estimate the memory parameter d of the squares of the LARCH process, and attempt to establish a connection between memory estimates of the powers of the FIGARCH process and its parameter δ.

The relationship between the memory parameter d, the self-similarity parameter H of the limiting distribution of the partial sums of $\{X_t\}$, and the tail index α of the observations is known only for linear processes, e.g., FARIMA. Partial sums of a long-memory finite-variance Gaussian FARIMA process with parameter $d \in (0, 1/2)$ converge to a fBm with Hurst index $H = d + 1/2$. However, partial sums of infinite-variance α-stable FARIMA, $\alpha \in (1,2)$ with memory parameter $d \in (0, 1/\alpha)$ are attracted to a lfsm with self-similarity parameter $H = d + 1/\alpha$. No corresponding results exist for the LARCH and FIGARCH processes, but it is of interest to know how the well known memory estimators relate to the memory parameters d and δ in LARCH and FIGARCH, respectively. This question is of particular interest if condition (5) fails, as no results at all, except the existence under (3), are available in this case. Below we list the estimators we focus on. Except perhaps the wavelet estimators, these are well-known estimators, so we do not describe them, but rather give references to detailed descriptions (see also Chapter 4 of Palma (2007) and Chapter 9 of Percival and Walden (2000)).

Time domain: a parametric estimator that assumes the FARIMA(1,d,1) model for $\{X_t\}$ (thus we call it FARIMA) and uses an approximate time-domain Gaussian likelihood. Estimates are obtained from the `R` routine `fracdiff`.

Frequency domain: semiparametric maximum-likelihood estimator studied by Robinson (1995b), local Whittle (LW) based on the frequency range of Henry (2001); semiparametric regression estimator of Geweke and Porter-Hudak (1983) (GPH) over frequencies specified by Hurvich et al (1998). Both estimators are designed for second-order Gaussian stationary time series.

Wavelet domain: semiparametric wavelet maximum likelihood estimator of Jach and Kokoszka (2008) (WMLE); semiparametric regression estimator of self-similarity parameter $H = d + 1/\alpha$ of a lfsm of Stoev et al (2002) (LOG). It can be used to estimate d only if α is known or can be estimated. Following recommendation of Teyssière and Abry (2006) we use the lower scale that corresponds to the frequency bandwidth of Henry (2001) and the upper as the penultimate boundary-effect-free. Wavelet analysis is performed with the least asymmetric filter with 4 vanishing moments. Theoretical justification for both estimators is developed only for heavy-tailed, linear time series.

3.1 Memory properties of the LARCH process

In our computational analysis, we allow the sufficient condition (5) to be violated. In addition, the restriction for B_2 following our earlier tail index estimation in squares of LARCH, originating from the condition $d + 1/\hat{\alpha} < 1$, coincidices almost exactly with the range of values of B_2 for which the self-similarity estimates in the LOG scheme are less than 1. This is another evidence suggesting that the partial sums of the squares of LARCH in the infinite-variance setting may converge to a lfsm with $H = d + 1/\alpha$. We present values of the RMSE of memory parameter estimators in the models with self-similarity estimates less than 1. If this condition is violated, all estimators fail. In general, RMSEs displayed in Figure 4 are small and rarely exceed 0.1. The largest RMSEs correspond to large values of B_2 and d, however B_2 seems to control RMSEs more than d (compare LARCH A with LARCH B and C). To be able to relate self-similarity estimator with memory estimators we subtract 1/2, which in finite-variance case gives us the memory parameter estimator, but is likely to overestimate d when $\alpha < 2$. When condition (5) holds (to the left of vertical lines in Figure 4), LOG-1/2 estimator is the best with RMSE around 0.05, then FARIMA, followed by LW and GPH very close together, and finally WMLE. The situation changes in the infinite-variance scenario, LOG-1/2 is the worse. It would have improved if we subtracted $1/\hat{\alpha}$ rather than 1/2, but then it would get much worse in the finite-variance case. WMLE is the second worse, whilst frequency and time estimators perform very well (in LARCH A, FARIMA is closer to WMLE rather than to LW and GPH) with RMSEs of about 0.05. Standard deviations in all models consistently increase with B_2 and d (values of B_2 have greater impact than d) from about 0.025 to 0.075, with LW and GPH having smaller values than time and wavelet estimators. Biases of all estimators except for LOG-1/2 are all negative, between -0.05 and 0 for LARCH A and between -0.1 and 0 in LARCH B and C, with WMLE exhibiting highest underestimation, then frequency estimators followed by FARIMA. Squared biases show roughly similar pattern to those of the RMSE curves. LOG-1/2 has the smallest bias (in absolute terms), close to 0

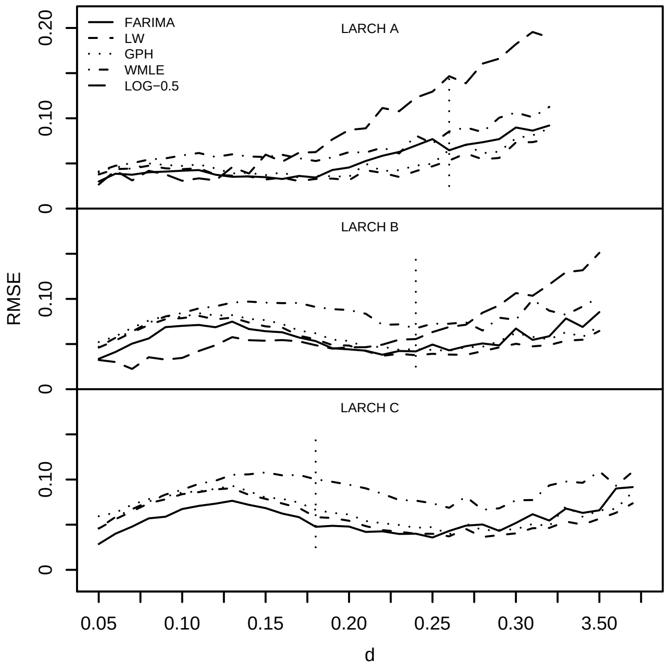

Fig. 4 RMSE of the memory parameter estimators of $X_t = r_t^2$, where $\{r_t\}$ follows three LARCH models (Table 1) as a function of memory parameter d, $d \in \{0.05, 0.06, \dots, 0.45\}$. Vertical lines indicate the cut-off from condition (5). Values are reported only for d's with self-similarity estimates less than 1.

for finite-variance LARCH A and negative for finite-variance LARCH B and C. In infinite-variance scenario its bias becomes positive and increases with d, substantially contributing to the RMSE. The shape of RMSE curve for model B resembles that of model C, and the two are different from the RMSE curve of model A. This suggests the influence of the filter b_j on RMSE. When B_2 is close to zero (right panel of Figure 1), the values of the RMSE prior to the vertical lines are the largest.

3.2 Memory properties of the FIGARCH process

Memory estimates $\hat{d}$ in time, frequency (we show only LW, results for GPH are very similar) and wavelet domains of the 11 models described in Section 2 are displayed in Figures 5-6. Note that the vertical range in all panels is 0.45 except the last one where it is twice as big.

All series (besides $X_t = r_t$ that has short memory) yield relatively large memory estimates that consistently increase with increasing power when $p < 1$, but which consistently decrease with power when $p \geq 1$, for all δ's. Curves based on the memory estimates follow similar inverted-parabola-type pattern, although time and frequency estimates cover wider range of values compared to WMLE. Time estimates are overall smaller than frequency estimates, which in turn are smaller than those of WMLE.

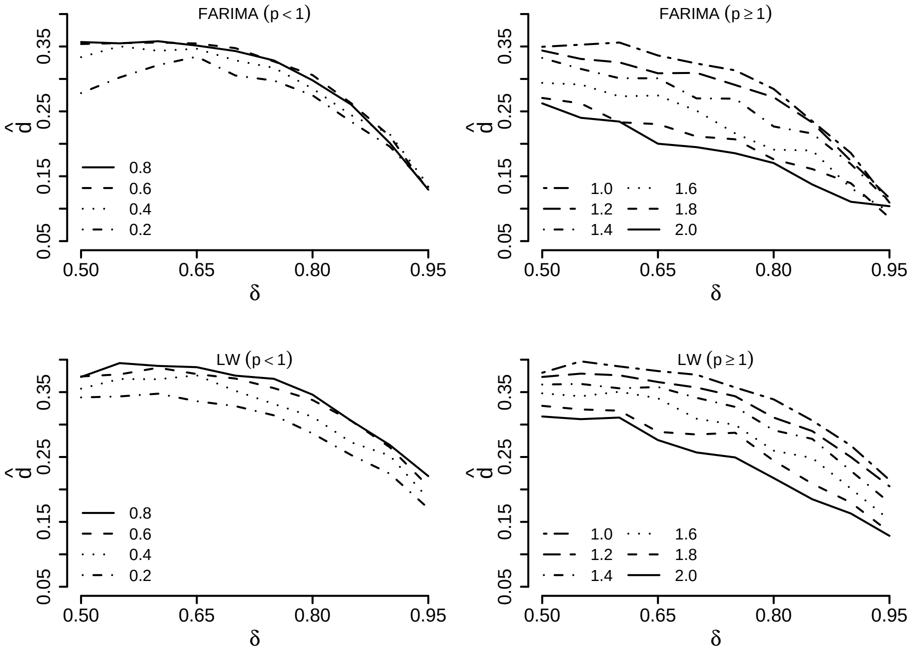

Fig. 5 *Top:* Memory parameter estimates for FARIMA method of $X_t = |r_t|^p$ (left: $p < 1$, right: $p \geq 1$), where $\{r_t\}$ follows FIGARCH model, as a function of the parameter $\delta \in \{0.5, 0.55, \ldots, 0.95\}$. *Bottom:* Memory parameter estimates for LW method.

When $p < 1$, self-similarity estimates differ from memory estimates by 0.5 and fall in $(0.5, 1)$ suggesting that the limiting process of the partial sums is a fBm with $H \in (0.5, 1)$. Although memory estimates for $p \geq 1$ depend on δ in a similar fashion as for $p < 1$, this is not the case with self-similarity estimates. When p is large, majority of the self-similarity estimates are greater than 1, indicating that the limiting process is likely to be non-Gaussian. We are unable to compute RMSEs of the estimators because, unlike in LARCH model, in FIGARCH framework we do know exactly how d and δ are connected. The relationship between the memory parameter and δ seems to be quite complex and it needs to be understood theoretically.

4 Application to return data

In this section we study a broad selection of financial time series. We consider assets such that: 1) their daily returns (first differences of log-prices) have negligible autocorrelations; 2) the autocorrelations of their squared returns decay very slowly remaining significant at lags of several hundred days, and are positive. Our goal is to estimate the tail index α and the memory parameter d of the squared returns using techniques from Sections 2 and 3. We want to see if these estimates are consistent with those obtained for the LARCH and FIGARCH models.

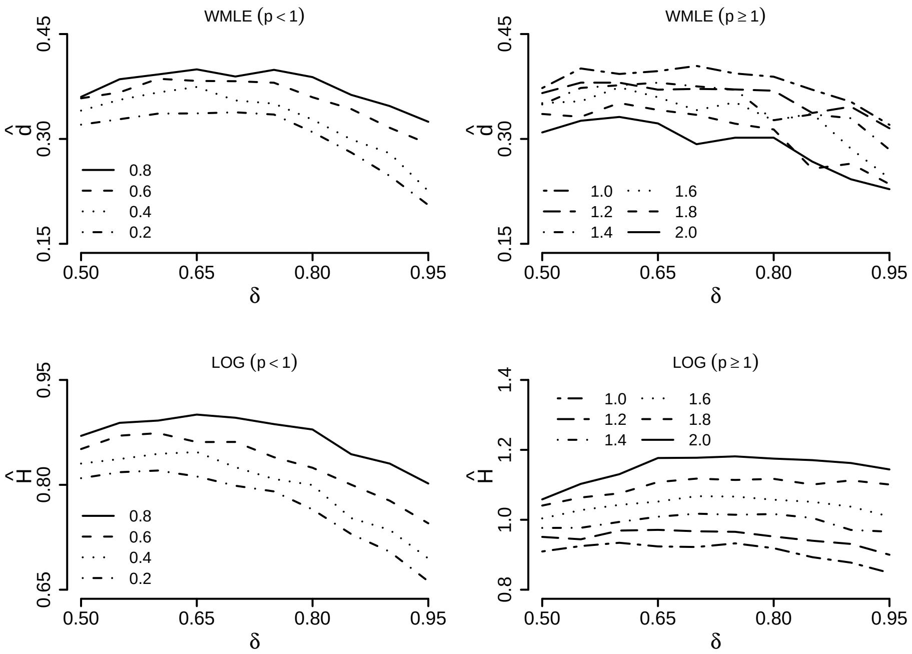

Fig. 6 *Top:* Memory parameter estimates for WMLE method of $X_t = |r_t|^p$ (left: $p < 1$, right: $p \geq 1$), where $\{r_t\}$ follows FIGARCH model, as a function of the parameter $\delta \in \{0.55, 0.6, \ldots, 0.95\}$. *Bottom:* Self-similarity parameter estimates for LOG method.

Twenty data sets described in Table 4 include foreign exchange rates obtained from the Federal Reserve Statistical Releases[1], as well as stock market indices and individual stock prices collected from Wessa.net[2]. First we perform wavelet-based tail index estimation of the squares of returns (in all cases, the hypothesis of independence of permuted wavelet coefficients at level $j = 1$ was accepted). In Table 5 we show asymptotic and permutation bootstrap confidence intervals for α. Asymptotic confidence intervals are derived from (9). Permutation bootstrap confidence limits are obtained by: 1) permuting elements of $\boldsymbol{W}_1$ and calculating an estimate $\hat{H}^{'(*)}$ of the self-similarity parameter $H^{'}$ of $\boldsymbol{W}_1$; 2) repeating this procedure $P = 1000$ times to obtain empirical distribution of $H^{'}$ based on P resampling estimates $\hat{H}^{'(*)}$; 3) inverting $\beta/2$-th and $(1-\beta/2)$-th percentiles of this distribution to obtain $100(1-\beta/2)\%$ permutation bootstrap cofidence intervals for α. In addition we calculate permutation bootstrap tail estimate $\hat{\alpha}^*$ by inverting the median of the empirical distribution of $H^{'}$. Tail index estimates are less than 2 suggesting that the returns have infinite fourth moment. Infinite-variance LARCH could be a reasonable model for the data, or even FIGARCH, as many estimates, especially for FX returns, are smaller than 1.5, c.f. Figure 3. Asymptotic and resampling approaches yield overall similar estimates, however confidence bounds based on the former one are narrower.

[1] http://www.federalreserve.gov/releases/h10/Hist/
[2] http://www.wessa.net/db.wasp/

Table 4 Description of the financial time series used in the analysis, one observation per day.

TS	Abbreviation	Period
BELGIUM FRANCS/US $	be/us	4 Jan 1971 - 31 Dec 1998
CANADA $/US $	ca/us	4 Jan 1971 - 31 Dec 2007
ITALY LIRA/US $	it/us	4 Jan 1971 - 31 Dec 1998
SOUTH AFRICA RAND/US $	sf/us	4 Jan 1971 - 31 Dec 2007
UK POUND STERLING/ US $	uk/us	4 Jan 1971 - 31 Dec 2007
10-year note	10YEARNOTE	02 Jan 1962 - 31 Dec 2007
30-year bond	30YEARBOND	15 Feb 1977 - 31 Dec 2007
5-year note	5YEARNOTE	02 Jan 1962 - 31 Dec 2007
DAX	DAX	26 Nov 1990 - 31 Dec 2007
Dow Jones Industrial Average	DJINDAVE	01 Oct 1928 - 31 Dec 2007
Dow Jones Transportation	DJTRANS	01 Oct 1928 - 31 Dec 2007
Dow Jones Utilities	DJUTIL	02 Jan 1929 - 31 Dec 2007
FTSE 100	FTSE100	02 Apr 1984 - 31 Dec 2007
Goldman Sachs Group	GOLDMAN	04 May 1999 - 31 Dec 2007
KFX	KFX	26 Jan 1993 - 31 Dec 2007
NASDAQ Composite	NASDAQCOMP	05 Feb 1971 - 31 Dec 2007
NASDAQ Transportation	NASDAQTRANS	25 Oct 1990 - 31 Dec 2007
RUSSELL 1000	RUSSELL1000	10 Dec 1992 - 31 Dec 2007
S&P 400 Midcap	SP400	20 Aug 1991 - 31 Dec 2007
S&P 500	SP500	03 Mar 1950 - 31 Dec 2007

Table 5 Tail index estimates and corresponding 95% confidence intervals for α of the squared returns of time series listed in Table 4. For asymptotic confidence bounds see formula (9), for permutation bootstrap, see discussion prior to this table.

		Asymptotic	Permutation bootstrap	
TS	$\hat{\alpha}$	95% confidence interval	$\hat{\alpha}^*$	95% confidence interval
be/us	1.4108	(1.3284,1.5041)	1.4545	(1.3433,1.5979)
ca/us	1.6268	(1.5562,1.7041)	1.5958	(1.5112,1.7057)
it/us	1.3619	(1.2836,1.4503)	1.3852	(1.2818,1.5271)
sf/us	1.0522	(1.0145,1.0927)	1.0626	(1.0154,1.1212)
uk/us	1.5299	(1.4656,1.6001)	1.6041	(1.5077,1.7134)
10YEARNOTE	1.5341	(1.4696,1.6046)	1.5324	(1.4481,1.6333)
30YEARBOND	1.6556	(1.5507,1.7757)	1.6222	(1.4917,1.8109)
5YEARNOTE	1.6171	(1.5472,1.6937)	1.5558	(1.4614,1.6632)
DAX	1.5679	(1.4714,1.6779)	1.5772	(1.4482,1.7394)
DJINDAVE	1.3060	(1.2706,1.3435)	1.2945	(1.2464,1.3471)
DJTRANS	1.3726	(1.3345,1.4129)	1.3283	(1.2819,1.3785)
DJUTIL	1.1787	(1.1481,1.2110)	1.1736	(1.1341,1.2145)
FTSE100	1.5960	(1.4969,1.7093)	1.6572	(1.5037,1.8603)
GOLDMAN	1.5207	(1.3855,1.6851)	1.5564	(1.3750,1.8580)
KFX	2.0522	(1.8371,2.3244)	1.7413	(1.5139,2.0875)
NASDAQCOMP	1.3490	(1.2957,1.4069)	1.3435	(1.2773,1.4252)
NADAQTRANS	1.6940	(1.5853,1.8187)	1.6427	(1.4863,1.8530)
RUSSELL1000	1.7034	(1.5428,1.9011)	1.6136	(1.4109,1.9225)
SP400	1.6601	(1.5548,1.7808)	1.6605	(1.5202,1.8506)
SP500	1.5556	(1.4897,1.6277)	1.6451	(1.5474,1.7728)

In Table 6 we present memory estimates of the squared returns. Most of the considered assets exhibit long memory in volatility. FARIMA estimator typically yields smaller values compared to two frequency schemes, that overall produce very similar, relatively large estimates. WMLE and adjusted LOG estimates are somewhere between their time and frequency counterparts, although WMLE seems to vary less. The range

Table 6 Memory estimates of the squared returns of time series listed in Table 4.

TS	FARIMA	LW	GPH	WMLE	LOG-$1/\hat{\alpha}$
be/us	0.0977	0.2088	0.2225	0.0629	0.2029
ca/us	0.0848	0.3629	0.3459	0.1731	0.3317
it/us	0.1398	0.2444	0.2426	0.1576	0.3862
sf/us	0.1184	0.1834	0.1715	0.0881	0.4275
uk/us	0.0542	0.2940	0.3188	0.2316	0.3197
10YEARNOTE	0.0000	0.3403	0.3122	0.2136	0.3536
30YEARBOND	0.0001	0.2746	0.2502	0.2524	0.2745
5YEARNOTE	0.1844	0.3470	0.2806	0.2244	0.4095
DAX	0.0000	0.4462	0.4493	0.3663	0.2835
DJINDAVE	0.1502	0.2483	0.2309	0.0057	0.1730
DJTRANS	0.3001	0.3379	0.3097	0.1851	0.1612
DJUTIL	0.2713	0.2879	0.3093	0.0509	0.0534
FTSE100	0.2928	0.2610	0.1982	0.1687	0.1589
GOLDMAN	0.0001	0.2908	0.2499	0.1339	0.0258
KFX	0.1824	0.3072	0.2867	0.2860	0.3032
NASDAQCOMP	0.1810	0.3512	0.3619	0.2229	0.2501
NADAQTRANS	0.0585	0.2208	0.2058	0.2313	0.2086
RUSSELL1000	0.0959	0.3110	0.2999	0.1988	0.1779
SP400	0.0000	0.3547	0.3153	0.3195	0.3482
SP500	0.1665	0.1573	0.1283	0.1484	0.3356

of frequency bands and scales in LW approach is the same as the range of scales in WMLE and LOG schemes, thus it would be interesting to see why this frequency procedure gives slightly larger values than the wavelet methods.

The pattern of smaller $\hat{d}$ values for the FARIMA method compared to the other methods is consistent with the results shown in Figures 5-6, suggesting that FIGARCH might be a suitable model, at least for those assets with estimated α less than 1.5 (like FX returns). This pattern is not observed for the LARCH models we considered, but the estimates of α from the LARCH model agree better with those in Table 6 for returns on individual stocks and indices. These findings suggest that the FIGARCH model is more appropriate for FX returns while the LARCH model for the returns on stocks and indices. Model selection cannot however be based exclusively on an approximate agreement of some statistical parameters. Typically, its predictive and explanatory performance plays a crucial role.

5 Conclusions

In our empirical analysis of squares of LARCH and powers of FIGARCH we exploited decorrelation property of the DWT to estimate tail indices of the corresponding marginal distributions. We brought together time-, frequency-, and wavelet-domain estimators to analyze memory properties of these models. Tail and memory estimation techniques were applied to twenty asset returns. Our empirical findings, that we hope will stimulate theoretical continued study of LARCH and FIGARCH models, can be summarized as follows:

- The estimates of the tail index α of the marginal distributions of the squares of infinite-variance LARCH are smaller than 2, depend on $B_2 = \sum_{j\geq 1} b_j^2$, but not on the memory parameter d, and decrease as B_2 increases. If (5) holds, these estimates exceed 2 and should be treated with caution. If B_2 is not too large ($d + 1/\hat{\alpha} < 1$), the relationship between $\hat{\alpha}$ and B_2 can be well described by a power law.

- The relationship between estimated H, d and α suggests that the partial sums of the squares of LARCH may converge to a lfsm with $H = d + 1/\alpha$, if $H \in (0, 1)$.
- Tail index estimates of the marginals of $X_t = |r_t|^p$, $p \in \{0.2, 0.4, \dots, 0.8\}$, where $\{r_t\}$ follows FIGARCH model exceed 2. For $\in \{1, 1.2, \dots, 2\}$ they decrease with δ for relatively small δ's (truncation effect) and fall into $[1, 2]$; if $p = 2$, these estimates are less than 1.5.
- Memory estimators in three different domains, perform well when applied to the squares of LARCH with RMSEs around 0.1, given that unrescaled filter $\{b_j\}$ is used. Estimating d by $\hat{H}_{\text{LOG}} - 0.5$ is recommended if it is reasonable to assume the finite fourth moment of the returns.
- For FIGARCH, memory estimates of $|r_t|^p$ decrease as δ increases in a "parabolic" manner, and are mostly in the range $[0.2, 0.4]$. When $p < 1$, the limiting process is likely to be fBm with $H \in (0.5, 1)$, but when $p \geq 1$, the process is likely to be non-Gaussian and only relatively small powers p yield models with $H \approx \hat{d} + 1/\hat{\alpha} < 1$, where H is the Hurst parameter of the limiting process for the partial sums.
- Analysis of return data indicates that FIGARCH might be suitable for modeling long series of FX returns, while LARCH may be a reasonable approximation to returns on stock and market indices.

None of the above conclusions has been reported, and there is at present no theoretical understanding to underpin them. We hope that our computational analysis will provide some guidance on models and estimators for important classes of financial data.

References

Abry P, Taqqu MS, Pesquet-Popescu B (2000) Wavelet based estimators for self-similar α-stable processes. In: Int. Conf. on Signal Processing, 16th World Computer Congress, Beijing, pp 369–375

Abry P, Flandrin P, Taqqu MS, Veitch D (2002) Self-similarity and long-range dependence through the wavelet lens. In: Doukhan P, Oppenheim G, Taqqu MS (eds) Theory and Applications of Long Range Dependence, Birkhäuser, Boston, pp 527–554

Baillie RT, Bollerslev T, Mikkelsen HO (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74:3–30

Baillie RT, Han YW, Myers RJ, Song J (2007) Long memory and FIGARCH models for daily and high frequency commodity prices. Journal of Futures Markets 27:643–668

Basrak B, Davis R, Mikosch T (2002) Regular variation of GARCH processes. Stochastic Processes and Their Applications 99:95–115

Beine M, Laurent S, Lecourt C (2002) Accounting for conditional leptokurtosis and closing days effects in FIGARCH models of daily exchange rates. Applied Financial Economics 12:589–600

Brockwell P, Davis R (2002) Introduction to Time Series and Forecasting, 2nd edn. Springer-Verlag, New York

Campbell JY, Lo AW, MacKinlay AC (1997) The Econometrics of Financial Markets. Princeton University Press, New Jersey

Craigmile P, Guttorp P, Percival DB (2005) Wavelet-Based Parameter Estimation for Polynomial Contaminated Fractionally Differenced Process. IEEE Transactions on Signal Processing 53:3151–3161

Davis R, Mikosch T (2009) Extremes of stochastic volatility models. In: Andersen T, Davis R, Kreiß J, Mikosch T (eds) Handbook of Financial Time Series, Springer, New York, pp 355–364

Douc R, Roueff F, Soulier P (2008) On the existence of some ARCH(∞) processes. Stochastic Processes and their Applications 118:755–761

Embrechts P, Klüppelberg C, Mikosch T (1997) Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin

Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series models. Journal of Time Series Analysis 4:221–238

Giraitis L, Robinson P, Surgailis D (2000b) A model for long memory conditional heteroskedasticity. The Annals of Applied Probability 10:1002–1024

Giraitis L, Kokoszka P, Leipus R, Teyssière G (2003) Rescaled variance and related tests for long memory in volatility and levels. Journal of Econometrics 112:265–294

Giraitis L, Leipus R, Robinson P, Surgailis D (2004) LARCH, leverage and long memory. Journal of Financial Econometrics 2:177–210

Giraitis L, Leipus R, Surgailis D (2008) ARCH(∞) models and long memory properties. In: Andersen T, Davis R, Kreiß J, Mikosch T (eds) Handbook of Financial Time Series, Springer, New York, pp 71–84

Gnedenko BV, Kolmogorov AN (1954) Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Reading

de Hann L, Resnick S, Rootzén H, de Vries CG (1989) Extremal behaviour of solutions to a stochastic difference equation with application to ARCH processes. Stochastic Processes and Their Applications 32:213–224

Henry M (2001) Robust automatic bandwidth for long memory. Journal of Time Series Analysis 22:293–316

Hurvich CM, Deo R, Brodsky J (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long memory time series. Journal of Time Series Analysis 19:19–46

Jach A, Kokoszka P (2008) Robust wavelet-domain estimation of the fractional difference parameter in heavy-tailed time series: an empirical study. Methodology and Computing in Applied Probability Doi:10.1007/s11009-008-9105-3

Kokoszka P, Maslova I, Sojka J, Zhu L (2006) Probability tails of wavelet coefficients of magnetometer records. Journal of Geophysical Research Doi:10.1029/2005JA011321

Lardic S, Mignon V (2004) Term premium and long-range dependence in volatility: A FIGARCH-M estimation on some asian countries. Journal of Emerging Market Finance 3:1–19

Lee SH (2005) Forecasting conditional volatility of returns by using the relationship among returns, trading volume, and open interest in commodity futures markets. PhD thesis, Purdue University

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

Moulines E, Roueff E, Taqqu MS (2008) A wavelet Whittle estimator of the memory parameter of a nonstationary Gaussian time series. The Annals of Statistics 36:1925–1956

Palma W (2007) Long-Memory Time Series: Theory and Methods. Wiley-Interscience, New Jersey

Percival DB, Walden AT (2000) Wavelet Methods for Time Series Analysis. Cambridge University Press, Cambridge

Pipiras V, Taqqu MS, Abry P (2007) Bounds for the covariance of functions of infinite variance stable random variables with applications to central limit theorems and wavelet-based estimation. Bernoulli 13:1091–1123

Robinson P (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity. Journal of Econometrics 47:67–84

Robinson P (1995b) Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 23:1630–1661

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

Stoev S, Taqqu MS (2005) Asymptotic self-similarity and wavelet estimation for long-range dependent FARIMA time series with stable innovations. Journal of Time Series Analysis 26:211–249

Stoev S, Pipiras V, Taqqu MS (2002) Estimation of the self-similarity parameter in linear fractional stable motion. Signal Processing 80:1873–1901

Stoev S, Michailidis G, Taqqu MS (2006) Estimating heavy-tail exponents through max self-similarity. Tech. Rep. TR445, University of Michigan

Teyssière G, Abry P (2006) Wavelet analysis of nonlinear long-range dependent processes. Applications to financial time series. In: Teyssière G, Kirman A (eds) Long memory in Economics, Springer, New York, pp 173–238