Bootstrap Prediction Intervals for Fractionally Integrated Generalized Autoregressive Conditionally Heteroscedastic (FIGARCH) Models

Rukman Ekanayake[1], V.A. Samaranayake[1]
[1]Department of Mathematics & Statistics, Missouri University of Science and Technology, Rolla Building, 400 W. 12th Street, Rolla, MO 65401

Abstract
The Generalized Autoregressive Conditional Heteroscedastic (GARCH) formulations are inadequate to model the persistent volatility found in certain financial assets. The integrated version of the GARCH formulation, namely the IGARCH model, was developed to handle such situations. Fractionally Integrated Generalized Autoregressive Conditionally Heteroscedastic (FIGARCH) models, however, provide a more flexible alternative to modeling long-term dependence of volatility, providing a leptokurtic unconditional distribution for returns having long memory behavior. We propose a method based on the residual bootstrap to obtain prediction intervals for the returns of FIGARCH processes. A Monte-Carlo simulation study, conducted using a variety of distributions for the error terms, show that the proposed intervals have good coverage probabilities in most cases.

Key Words: Fractional Integration, Volatility Modeling, residual-based bootstrap, long memory

1. Introduction

Time series literature is replete with many formulations developed to model the volatility of financial time series. Engle (1982) introduced the well-known Autoregressive Conditional Heteroscedastic (ARCH) model and Bollerslev (1986) extended the ARCH model to the Generalized ARCH (GARCH) model, which accommodate long-term dependence of volatility with a limited number of lag terms compared to the ARCH formulation. Since the introduction of the ARCH and GARCH models, several variations were introduced. For example, the exponential GARCH or the EGARCH model (Nelson, 1991) was developed to allow asymmetric response to positive and negative shocks. A generally known fact about GARCH type models is their ability to models volatility clustering. Volatility clustering refers to the phenomenon where large returns tend to follow large returns and small returns tend to follow small returns. Highly persistent volatility, however, cannot be modeled well using the GARCH model or its alternatives such as the EGARCH. The Integrated GARCH (Engle and Bollerslev, 1986) formulation was developed to model time series with persistent volatility. Fractionally Integrated Generalized Autoregressive Conditional Heteroscedastic (FIGARCH) was introduced by Baillie et al. in 1996 as an alternative to the IGARCH model, allowing the ability to model the long-memory nature of the conditional variance found in many financial time series, but without the assumption of a unit root in the model. Here in we introduced a residual bootstrap-based method of obtaining prediction intervals for the conditional volatility of FIGARCH processes.

The conditional variance of a GARCH process can be written as infinite sum of exponentially decaying terms containing squared past innovations. On the other hand, the conditional variance of FIGARCH model can be expressed as a sum whose terms have a slower hyperbolic rate of decay. This provides the FIGARCH formulation the ability to model squared return processes having long memory. Thus, in the FIGARCH formulation, the effect of a past shock (squared innovation) decays slowly to zero unlike in the GARCH case where such effects decays at a faster exponential rate. In the IGARCH formulation the effect of such a shock persists without decaying. Thus the FIGARCH, while allowing for a past shock to persist for a long period, assumes that eventually its effects become negligible, which is a more reasonable assumption.

There are only few published papers that discuss the construction of prediction intervals for ARCH and GARCH type models. Compared to point forecasts, the prediction intervals give extra assessment about the uncertainty associated with the forecast and is therefore more desirable. In general, the underlying distribution of the point predictor or that of a pivotal statistic is needed to derive prediction intervals. But this is not feasible in some situations and in many instances the asymptotic distribution of such statistics is used. An alternative is the distribution free resampling approach, where a bootstrap-based technique is utilized. Reeves (2005) constructed prediction intervals for ARCH models using a bootstrap method and contrasted it with the traditional asymptotic prediction intervals. Reeves report that the bootstrap-based method improves the coverage accuracy. Pascual *et al*. (2006) developed a bootstrap-based prediction intervals for both returns and volatilities for the GARCH(*1, 1*) model. Their bootstrap method incorporated the uncertainty of parameter estimation when building the prediction intervals, which certainly improved the coverage. However, one drawback of this method is the time-consuming calculation of prediction intervals. Since GARCH model can be re-written as a linear ARMA type model, Chen *et al*. (2010) proposed computationally low cost sieve bootstrap-based prediction intervals for returns and volatiles. Trucíos and Hotta (2016) constructed prediction intervals for returns and volatilities for EGARCH and GJR-GARCH models by adapting the method used by Pascual *et al*. (2006). They found that volatility prediction could be poor when an additive outlier is present near the forecasting origin. Although there are published literature on bootstrap prediction intervals for the conventional volatility models, there are no such work available for long memory volatility models. On the other hand, there is ample literature on the prediction intervals for long memory conditional mean models. For example, Bisaglia and Grigoletto (2001) introduced bootstrap-based prediction intervals for Fractionally Integrated Autoregressive Moving Average (FARIMA) processes. Although this method performs quite well, it is computationally much slower. Rupasinghe and Samaranayake (2013) established a computationally much faster, sieve-bootstrap-based procedure to calculate prediction intervals for FARIMA processes. This latter method yields better results even if the innovation distribution is non-normal. The main objective in this current paper is to introduce a bootstrapped-based prediction interval procedure for the FIGARCH model.

The sections of the paper are organized as follows. First, we introduced the FIGARCH model and then its properties in the Section 2. Section 3 describes the residual based resampling technique and then in Section 4, Monte Carlo simulation results are reported. Section 5 presents an application of the proposed bootstrap-based prediction intervals for the FIGARCH model and conclusions are presented in Section 6.

2. The FIGARCH Model

A real valued discrete time stochastic process $\{\varepsilon_t : t \in \mathbb{Z}\}$ is said to be an ARCH (q) process, if

$$\varepsilon_t = z_t \sqrt{h_t}, \tag{1}$$

with

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{i-1}^2,$$

where $\omega > 0$ and $\alpha_i \geq 0, i = 1,...,q$. In expression (1) it is assumed that, $E(z_t) = 0$, $\text{var}(z_t) = \sigma_z^2 = 1$ and z_t's are uncorrelated. Thus, by the definition $\{\varepsilon_t\}$ is an uncorrelated with mean zero process with conditional variance h_t, where the conditioning is done with respect to the $\sigma-$field $\Im_{t-1}$ generated by the set of random variables $\{z_k : k \leq t-1\}$. The conditional variance is a linear function of squared residuals up to q lags implying a Markovian dependence. The generalized version of ARCH (GARCH), introduced by Bollerslev (1986) gives more flexible structure compared to (1), with the conditional volatility h_t given by,

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{i-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)h_t, \tag{2}$$

where $p > 0, q > 0, \omega > 0, \alpha_i \geq 0, i = 1,...,q, \beta_j \geq 0, j = 1,..., p,$ $\alpha(L)$ and $\beta(L)$ are such that $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + ... + \alpha_q L^q$ and $\beta(L) = \beta_1 L + \beta_2 L^2 + ... + \beta_p L^p$, with L signifying the lag (or backshift) operator. The process defined in (2) is a stationary process and can be written as an ARMA (m, p) formulation in ε_t^2:

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t, \tag{3}$$

where $m = \max(p,\ q)$ and $v_t = \varepsilon_t^2 - h_t$. The process $\{v_t\}$ can be shown to be uncorrelated and is interpreted as the innovations associated with the ARMA process. The formulation in (3) is said to be an IGARCH model if the autoregressive polynomial contains a unit root. Therefore, autoregressive representation of IGARCH can be given as:

$$\phi(L)(1-L)\varepsilon_t^2 = \omega + [1 - \beta(L)]v_t,$$

where $\phi(L) = [1 - \alpha(L) - \beta(L)](1-L)^{-1}$ is of order $m-1$.

Several studies have reported the presence of long memory in the autocorrelations of squared returns in financial asset prices. Thus, Baillie *et al*. (1996) adapted the idea of fractional integration in conditional mean models (FARIMA) in order to develop a FIGARCH process. The class of FARIMA $(k,\ d,\ l)$ models for the discrete time real-valued process $\{y_t\}$ is defined as:

$$a(L)(1-L)^d y_t = b(L)z_t \tag{4}$$

where $a(L)$ and $b(L)$ are polynomials in the lag operators of orders k and l respectively. Here, $\{z_t\}$ is an uncorrelated process with mean zero. The fractional integration parameter, d, lies between -0.5 and 0.5 for the stationary FARIMA model. The fractional differencing operator $(1-L)^d$ has an infinite binomial expansion and can be written in terms of the hypergeometric function,

$$(1-L)^d = F(-d,1,1;L) = \sum_{k=0}^{\infty} \Gamma(k-d)\Gamma(k+1)^{-1}\Gamma(-d)^{-1}L^k,$$

where $\Gamma(.)$ denotes the Gamma function. Analogues to FARIMA $(k,\ d,\ l)$ model for the mean process given in (4), Baillie et al. (1996) defined the FIGARCH model in the following manner:

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1-\beta(L)]v_t\,, \tag{5}$$

where $0<d<1$, and all the roots of $\phi(L)$ and $[1-\beta(L)]$ lie outside the unit circle. Rearranging the terms in (5), an alternative representation for FIGARCH $(p,\ d,\ q)$ can be obtained as

$$[1-\beta(L)]h_t = \omega + [1-\beta(L)-\phi(L)(1-L)^d]\varepsilon_t^2\,. \tag{6}$$

From (6), conditional variance of the $\{\varepsilon_t\}$ is obtained as:

$$\begin{aligned} h_t &= \omega[1-\beta(1)]^{-1} + \{1-[1-\beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t^2 \\ &= \omega[1-\beta(1)]^{-1} + \lambda(L)\varepsilon_t^2 \end{aligned}\,, \tag{7}$$

where $\lambda(L) = \lambda\sum_{k=1}^{\infty}\lambda_k L^k$. For the FIGARCH (*p, d, q*) process given in equation (5) to be well-defined and the conditional variance in the ARCH(∞) representation in (7) to be positive, all the coefficient of ARCH representation in (7), must be non-negative. That is, each $\lambda_k \geq 0$ for $k \in \mathbb{N}$.

From equation (7) the conditional variance of FIGARCH (*1, d, 1*) can be written as follows.

$$h_t = \omega(1-\beta_1)^{-1} + [1-(1-\beta_1)^{-1}(1-\phi_1 L)(1-L)^d]\varepsilon_t^2\,, \tag{8}$$

where,

$$\lambda(L) = \sum_{k=1}^{\infty}\lambda_k L^k = 1-[1-(1-\beta_1)^{-1}(1-\phi_1 L)(1-L)^d]\,.$$

Therefore, coefficients of the infinite ARCH model can be obtained by equating the coefficients of lag operator, thus obtaining:

$$\lambda_1 = \phi_1 - \beta_1 + d\,,$$

$$\lambda_2 = (d - \beta_1)(\beta_1 - \phi_1) + d(1-d)/2,$$

$$\lambda_3 = \beta_1\left[d\beta_1 - d\phi_1 - \beta_1^2 + \beta_1\phi_1 + d(1-d)/2\right] + d(1-d)/2\left[(2-d)/3 - \phi_1\right],$$

$$\vdots$$

$$\lambda_k = \beta_1\lambda_{k-1} + \left[(k-1-d)/d - \phi_1\right]\delta_{d,k-1},\ k \in \mathbb{N},$$

where $\delta_{d,k} = \delta_{d,k-1}(k-1-d)k^{-1},\ k \in \mathbb{N}$ refer to the coefficients in the series expansion of $(1-L)^d$, with $\delta_{d,0} = 1$ *and* $\delta_{d,1} = d$.

The FIGARCH formulation enables us to model a wide range of a conditional volatility models. When $d = 0$ it becomes a GARCH $(p,\ m)$ process where $m = \max(p,\ q)$. Similarly, when $d = 1$ with $\beta(L) \neq 0$ and $\phi(L) = 1$ FIGARCH becomes a regular IGARCH model.

2.1 Non-negativity of the Conditional Variance

For the non-negativity of the conditional variance of the FIGARCH, all λ_k's should be positive. Baillie et al. (1996) derived set of sufficient conditions for the conditional variance to be non-negative. They are $0 \le \beta_1 \le \phi_1 + d$ and $0 \le d \le 1 - 2\phi_1$. We used these set of conditions in our study. Alternatively, Bollerslev and Mikkelsen (1996) state another set of sufficient inequality constraints $\beta_1 - d \le \phi_1 \le (2-d)/3$ and $d\left(\phi_1 - (1-d)/2\right) \le \beta_1(\phi_1 - \beta_1 + d)$. The latter conditions introduced by Bollerslev and MIkkelsen (1996) are less restrictive than the former conditions introduced by Baillie et el. (1996). Chung (1999) suggest another set of sufficient constraints given by $0 \le \phi_1 \le \beta_1 \le d < 1$. Finally Conrad and Haag (2006) derived necessary and sufficient conditions for the non-negativity of the variance for the FIGARCH(*p, d, q*) for $p \le 2$. According to their findings, conditional variance can be negative almost surely, even if all the original parameters of FIGARCH are positive and similarly conditional variance can be non-negative even if all the parameters are negative except d. They also derived sufficient conditions for non-negativity of variance for $p > 2$.

2.2 Asymptotic Normality of the Parameters and the Stationarity of the Process

Baillie et al. (1996) used a dominance type argument by extending the results available for IGARCH(*1, 1*), to claim the asymptotic normality of Q-MLEs of FIGARCH(*1, d, 0*). They did not proved it theoretically, but their empirical study, however, suggests that parameter estimates are asymptotically normal. Robinson and Zaffaroni (2006) established conditions for consistency and asymptotic normality of Q-MLEs for class of ARCH (∞) under some general conditions, which also covers the FIGARCH type processes. According to them strong consistency requires $0 < d < 1$ and asymptotic normality requires $d > 0.5$.

By construction, FIGARCH with $\{\varepsilon_t\}$ defined as in equation (1), has the properties that $\text{cov}(\varepsilon_t, \varepsilon_{t-h}) = 0$ for $h > 0$ and $E(\varepsilon_t) = 0$. The hypergeometric function, $F(-d,\ 1,\ 1;\ u)$ evaluated at $u = 1$ is 0 for $0 < d \le 1$ and thus $\lambda(1) = 1$. Therefore, for $\omega > 0$ the second moment of the $\{\varepsilon_t\}$ does not exist. The implication is that the FIGARCH process is not covariance stationary. Giraitis et el. (2018) established the necessary and sufficient conditions for the FIGARCH to be covariance stationary with $\omega = 0$. Conrad and Haag,

(2006) suggest a way to obtain the covariance stationarity of $\{\varepsilon_t\}$ with $0 < d < 1$ by assuming $\sigma_z^2 = \text{var}(z_t) < 1$ in (1). However, it rules out long memory in ε_t^2, by indicating the absolute summability of auto-covariance function of ε_t^2, as shown in Zaffaroni (2004).

3. Bootstrap Prediction Intervals

In this section, we adopt the procedure proposed by Pascual et al. (2006) for the GARCH case to obtain prediction intervals for future values of returns generated by a FIGARCH process.

1. Let $\{\varepsilon_t\}_{t=1}^n$ be a sequence of realizations of a FIGARCH(*1, d, 1*) process. Then estimate the parameters of the model $\hat{\theta} = (\hat{\omega}_1, \hat{\phi}_1, \hat{d}, \hat{\beta}_1)$ by using Quasi-Maximum Likelihood Estimation (Q-MLE) method.
2. Compute the residuals $\hat{z}_t = \varepsilon_t \big/ \sqrt{\hat{h}_t}\,, t = 1,\ldots,n$ where
$$\begin{aligned}\hat{h}_t &= \hat{\omega}(1-\hat{\beta}_1)^{-1} + [1-(1-\hat{\beta}_1)^{-1}(1-\hat{\phi}_1 L)(1-L)^{\hat{d}}]\varepsilon_t^2 \\ &\approx \hat{\omega}(1-\hat{\beta}_1)^{-1} + \hat{\lambda}_1\varepsilon_{t-1}^2 + \hat{\lambda}_2\varepsilon_{t-2}^2 + \ldots + \hat{\lambda}_k\varepsilon_{t-k}^2\end{aligned},$$
and setting $\varepsilon_t^2 = n^{-1}\sum_{i=1}^n \varepsilon_i^2$, for $t = -k+1, \ldots, -1, 0$. Note that k is a suitably chosen truncation lag of the polynomial $\lambda(L)$.
3. Compute the centered residuals $\tilde{z}_t = \hat{z}_t - \bar{\hat{z}}_t$, where $\bar{\hat{z}}_t = n^{-1}\sum_{i=1}^n \hat{z}_i$.
4. Denote the empirical distribution function of the centered residuals by $\hat{F}_{\tilde{z}}(x) = n^{-1}\sum_{t=1}^n I_{(-\infty,x]}(\tilde{z}_t)$.
5. Draw a bootstrap sample with replacement from the above distribution and denote it by z_t^*, where $t = -m+1, \ldots -1, 0, 1, \ldots, n$.
6. Generate the bootstrapped FIGARCH series $\varepsilon_t^*, t = -m+1, \ldots, -1, 0, 1, \ldots, n$ by first computing a bootstrapped conditional variance series, h_t^* using the FIGARCH parameters estimated in Step 1. Then use $\varepsilon_t^* = z_t^*\sqrt{h_t^*}, t = -m+1, \ldots, -1, 0, 1, \ldots, n$ to generate ε_t^*. The non-positive lags represent 'burn-in' observations that are dropped to mitigate effects due to initial conditions.
7. Estimate the FIGARCH parameters $\theta^* = (\omega^*, \phi^*, d^*, \beta^*)$ for the bootstrapped series $\{\varepsilon_t^*\}$ using the Q-MLE method.
8. Use the new coefficients $\theta^* = (\omega^*, \phi^*, d^*, \beta^*)$ obtained in the previous step, compute the *h*-step ahead bootstrap forecasts of future values based on the following recursions:
$$\begin{aligned}h_{t+h}^* &= \omega^*(1-\beta_1^*)^{-1} + [1-(1-\beta_1^*)^{-1}(1-\phi_1^*)(1-L)^{d^*}]\varepsilon_{t+h}^{2*} \\ &\approx \omega^*(1-\beta_1^*)^{-1} + \lambda_1^*\varepsilon_{t+h-1}^{2*} + \ldots + \lambda_k^*\varepsilon_{t+h-k}^{2*},\end{aligned},$$
$\varepsilon_{t+h}^* = z_{t+h}^*\sqrt{h_{t+h}^*}$, for $h > 0$ and $\varepsilon_t^* = \varepsilon_t$ for $t \le n$.
9. Obtain the estimated bootstrap distribution of ε_{t+h}, denoted by $\hat{F}_{\varepsilon_{n+h}^*}^*(.)$, by repeating steps 5-8 B times $(B = 1000)$ in the simulation study. $\hat{F}_{\varepsilon_{n+h}^*}^*(.)$ is the

estimate of the $F^*_{\varepsilon^*_{n+h}}(.)$, the bootstrap distribution function of ε^*_{t+h}, which is used to approximate unknown distribution of ε_{t+h} given the observed sample.

10. The $100(1-\alpha)\%$ bootstrap prediction interval for ε_{t+h} is then computed by $[Q^*(\alpha/2), Q^*(1-\alpha/2)]$, where $Q^*(.) = \hat{F}^{*-1}_{\varepsilon^*_{t+h}}$ are the percentiles of the estimated bootstrap distribution.

4. The Simulation Study

To investigate the finite sample performance of the proposed bootstrap prediction intervals of the FIGARCH model a Monte-Carlo simulation was carried out. The representations of $\{\varepsilon_t\}$ given in Equations (1) and (8) were used to simulate the FIGARCH process. This method become feasible due to the truncation of the infinite lag polynomial. The effect of the pre-sample values might have a higher impact than regular GARCH due to the long memory nature and the hyperbolic rate of decay of the response to a lagged squared innovation. Thus as suggested by Baillie et al. (1996), truncating lag was selected at k = 1,000 to incorporate the long-run dependencies.

The Monte-Carlo simulation study was carried out for different error distributions, namely standard normal and t with 7 degrees of freedom. Centered exponential distribution was also considered to investigate the effect due to non-symmetric error distributions. Series of lengths 500 and 1500 were used. The t-distributed errors were generated as $z_t = 5^{1/2} z_{1,t}(z^2_{2,t} + z^2_{3,t} + ... + z^2_{8,t})^{-1/2}$ by drawing independent and identically distributed standard normal $z_{i,t}$'s for $i = 1, 2, ..., 8$ as employed in Baillie et al. (1996). Here t-distributed errors also have a unit standard deviation. When generating the realizations, the first 6,000 were dropped to avoid the effects due to initial values.

We considered FIGARCH(*1, d, 0*) and FIGARCH(*1, d, 1*) models to simulate the data with $\omega = 0.1$, $d \in \{0.25, 0.50, 0.75, 0.95\}$, $\phi \in \{0, 0.2\}$ and $\beta \in \{0.10, 0.20, 0.45, 0.70, 0.90\}$. Note that out of these sets of parameter combinations, we only used the combinations which satisfied the sufficient conditions for non-negativity of the variance suggested by Baillie et al. (1996). For each combination of the model, sample size, nominal coverage probability and error distributions, N = 500 independent time series were generated. Then bootstrap steps 1 through 10 were implemented. In each simulation R = 1,000 future values, $\{\varepsilon_{t+h}\}$, $h = 1,\ 10,\ 20$ were generated. We estimated the coverage probabilities by calculating the proportion of those ε_{t+h} values falling between the lower and upper bounds of the bootstrap intervals. Therefore, the coverage for the ith simulation run is given by $C(i) = R^{-1}\sum_{r=1}^{R} I_A[\varepsilon^r_{n+h}(i)]$ where $A = [Q^*(\alpha/2), Q^*(1-\alpha/2)]$ is the $100(1-\alpha)th$ bootstrapped prediction interval. $I_A(.)$ is the indicator function of the set A and $\varepsilon^r_{n+h}(i), r = 1,\ 2, ..., 1000$ are the R future values generated at i^{th} simulation run. The theoretical and bootstrap lengths are obtained by using $L_T(i) = \varepsilon^r_{n+h}(1-\alpha/2) - \varepsilon^r_{n+h}(\alpha/2)$ and $L_B(i) = Q^*(1-\alpha/2) - Q^*(\alpha/2)$ respectively. $L_T(i)$ is the difference between $100(1-\alpha)th$ and the $100(\alpha/2)th$ percentiles generated from R future values of the underling model with known order and the coefficients. The mean coverage, mean bootstrapped prediction intervals, mean theoretical intervals and their standard errors are calculated as follows:

Mean coverage $\bar{C} = N^{-1}\sum_{i=1}^{N} C(i)$,

Standard error of mean coverage $SE_{\bar{C}} = \left\{[N(N-1)]^{-1}\sum_{i=1}^{N}[C(i)-\bar{C}]^2\right\}^{1/2}$,

Mean length (bootstrap) $\bar{L}_B = N^{-1}\sum_{i=1}^{N} L_B(i)$,

Standard error of mean length $SE_{\bar{L}_B} = \{[N(N-1)]^{-1}\sum_{i=1}^{B}[L_B(i)-\bar{L}_B]\}^{1/2}$,

Mean theoretical length $\bar{L}_T = N^{-1}\sum_{i=1}^{N} L_T(i)$.

We investigated the type of model, nominal coverage probability, effect of the bootstrap truncation lag on coverage probabilities and error distribution in this simulation study. We report the mean coverage, mean bootstrap length, mean theoretical length, standard error of mean coverage and standard error of mean bootstrap interval length in Tables 1-6 for standard normal, centered exponential and *t*-distributed innovations. Due to the space limitation we only report the behavior of 95% intervals. The minimum value, percentiles (25th, 50th, 75th), and maximum value of the (a) coverage probabilities, (b) the bootstrap interval bounds (upper and lower), and (c) the theoretical interval bounds (upper and lower), were computed for further investigation and results are available upon request.

Simulation results shows that the coverage probabilities are close to the nominal value for the normal and the *t* error distributions. The maximum and minimum coverage probabilities obtained using the centered and skewed exponential error distribution is 0.9323 and 0.9152 for FIGARCH(*1, d, 1*) with parameters $\omega = 0.1,\ \phi = 0.2,\ d = 0.5,\ \beta = 0.45$ and $\omega = 0.1,\ \phi = 0,\ d = 0.90,\ \beta = 0.20$ respectively. In most cases, the bootstrap lengths are less than the theoretical lengths when using the exponential error distribution as the distribution of the innovations. Note that the coverage probabilities get closer to the nominal value with increasing sample size n. However, the coverage probabilities decrease as sample size n increases for the first lag ahead prediction intervals for FIGARCH with exponentially distributed errors. In some cases, the coverage probabilities exceed 0.95 with for normal error distributions but stays close to 0.95.

Table 1: Coverage of 95% intervals for returns of FIGARCH (*1, d, 1*) with parameters $\omega = 0.1,\ \phi = 0.2,\ d = 0.5,\ \beta = 0.45$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	8.1905	0.9461(0.0009)	8.2818(0.2939)
		1500	8.1903	0.9497(0.0006)	8.2944(0.2811)
	10	500	8.6422	0.9469(0.0010)	9.0641(0.3183)
		1500	8.6663	0.9512(0.0007)	9.0090(0.2768)
	20	500	8.8115	0.9470(0.0011)	9.4244(0.3359)
		1500	8.8575	0.9505(0.0008)	9.2691(0.2779)
t-distr.	1	500	7.6244	0.9474(0.0009)	7.8025(0.4926)
		1500	7.5827	0.9498(0.0006)	7.7746(0.5022)
	10	500	7.9347	0.9469(0.0010)	8.3669(0.4723)
		1500	7.9246	0.9509(0.0007)	8.3218(0.4889)
	20	500	7.9947	0.9463(0.0010)	8.4857(0.4054)
		1500	8.0138	0.9507(0.0007)	8.4643(0.4441)
Exponential	1	500	5.8465	0.9323(0.0017)	5.9865(0.3607)
		1500	5.7658	0.9233(0.0013)	5.8585(0.3405)
	10	500	6.6978	0.9288(0.0015)	6.7201(0.3967)
		1500	6.5921	0.9323(0.0010)	6.5628(0.3796)
	20	500	7.1692	0.9201(0.0017)	6.9116(0.4066)
		1500	7.1493	0.9260(0.0012)	6.8082(0.4021)

Table 2: Coverage of 95% intervals for returns of FIGARCH (*1, d, 0*) with parameters $\omega = 0.1,\ \phi = 0,\ d = 0.95,\ \beta = 0.90$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	25.9854	0.9534(0.0008)	27.6606(1.3980)
		1500	26.1125	0.9501(0.0006)	26.5303(1.3412)
	10	500	26.5898	0.9536(0.0008)	28.5965(1.4415)
		1500	26.6335	0.9531(0.0006)	27.5410(1.3812)
	20	500	27.1725	0.9532(0.0010)	29.7758(1.5132)
		1500	27.0818	0.9524(0.0007)	28.3302(1.4279)
t-distr.	1	500	20.6721	0.9486(0.0010)	21.4025(1.2036)
		1500	20.7467	0.9508(0.0008)	21.3077(1.2124)
	10	500	21.3153	0.9487(0.0012)	22.5039(1.2251)
		1500	21.2706	0.9532(0.0008)	22.3004(1.2043)
	20	500	21.8511	0.9457(0.0013)	23.3292(1.2928)
		1500	21.8165	0.9526(0.0009)	22.9700(1.2512)
Exponential	1	500	15.8917	0.9293(0.0026)	17.6120(1.5387)
		1500	15.8112	0.9161(0.0019)	15.9343(1.0250)
	10	500	16.8619	0.9303(0.0026)	19.9510(2.1806)
		1500	16.7916	0.9281(0.0019)	17.0831(1.0305)
	20	500	17.9403	0.9231(0.0030)	20.3512(2.0408)
		1500	17.9274	0.9239(0.0019)	17.9405(1.0916)

Table 3: Coverage of 95% intervals for returns of FIGARCH (*1, d, 0*) with parameters $\omega = 0.1,\ \phi = 0,\ d = 0.75,\ \beta = 0.70$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	13.2671	0.9491(0.0007)	13.8925(0.7439)
		1500	13.2930	0.9501(0.0006)	13.4264(0.5524)
	10	500	13.8338	0.9481(0.0008)	15.2424(1.2033)
		1500	13.9054	0.9497(0.0006)	14.2174(0.5706)
	20	500	14.4422	0.9449(0.0009)	15.8803(1.2403)
		1500	14.4339	0.9485(0.0007)	14.8183(0.5847)
t-distr.	1	500	10.8616	0.9485(0.0008)	11.1432(0.5269)
		1500	10.9049	0.9489(0.0006)	10.9673(0.5326)
	10	500	11.4262	0.9471(0.0009)	11.8891(0.5807)
		1500	11.4571	0.9488(0.0006)	11.6543(0.5529)
	20	500	11.8185	0.9439(0.0011)	12.4003(0.6218)
		1500	11.8646	0.9484(0.0008)	12.2426(0.5672)
Exponential	1	500	7.9757	0.9285(0.0017)	8.2022(0.3805)
		1500	8.0198	0.9192(0.0012)	7.9749(0.3611)
	10	500	8.9596	0.9270(0.0016)	9.1761(0.4263)
		1500	8.9336	0.9294(0.0009)	8.8358(0.3836)
	20	500	9.7919	0.9182(0.0017)	9.6466(0.4415)
		1500	9.8457	0.9227(0.0010)	9.3081(0.3911)

Table 4: Coverage of 95% intervals for returns of FIGARCH (*1, d, 0*) with parameters $\omega = 0.1,\ \phi = 0,\ d = 0.50,\ \beta = 0.45$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	7.9073	0.9473(0.0008)	7.9823(0.1876)
		1500	7.8858	0.9502(0.0006)	7.9990(0.1850)
	10	500	8.1986	0.9454(0.0008)	8.2987(0.1891)
		1500	8.2050	0.9500(0.0006)	8.3757(0.1860)
	20	500	8.3429	0.9439(0.0009)	8.4755(0.1935)
		1500	8.3903	0.9499(0.0007)	8.6508(0.1934)
t-distr.	1	500	7.1895	0.9471(0.0008)	7.2403(0.2194)
		1500	7.1786	0.9498(0.0006)	7.2885(0.2232)
	10	500	7.4466	0.9456(0.0008)	7.5572(0.2230)
		1500	7.4454	0.9510(0.0006)	7.7070(0.2236)
	20	500	7.5879	0.9442(0.0009)	7.7584(0.2298)
		1500	7.6092	0.9501(0.0007)	7.9539(0.2382)
Exponential	1	500	6.4551	0.9275(0.0015)	6.6834(0.3427)
		1500	6.4130	0.9218(0.0012)	6.5304(0.3274)
	10	500	7.0325	0.9265(0.0014)	7.2449(0.3888)
		1500	7.0402	0.9298(0.0010)	7.1873(0.3633)
	20	500	7.4848	0.9196(0.0015)	7.5463(0.4163)
		1500	7.4890	0.9266(0.0011)	7.4951(0.3780)

Table 5: Coverage of 95% intervals for returns of FIGARCH (*1, d, 0*) with parameters $\omega = 0.1,\ \phi = 0,\ d = 0.25,\ \beta = 0.10$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	3.1390	0.9430(0.0008)	3.0863(0.0327)
		1500	3.1411	0.9491(0.0005)	3.1521(0.0325)
	10	500	3.2306	0.9419(0.0008)	3.1765(0.0282)
		1500	3.2317	0.9483(0.0005)	3.2347(0.0237)
	20	500	3.2541	0.9414(0.0008)	3.1900(0.0274)
		1500	3.2536	0.9474(0.0006)	3.2511(0.0224)
t-distr.	1	500	3.1927	0.9467(0.0008)	3.2160(0.0483)
		1500	3.1950	0.9499(0.0006)	3.2367(0.0450)
	10	500	3.2907	0.9456(0.0008)	3.3061(0.0399)
		1500	3.2815	0.9499(0.0006)	3.3333(0.0339)
	20	500	3.3009	0.9460(0.0008)	3.3363(0.0385)
		1500	3.3133	0.9492(0.0006)	3.3497(0.0310)
Exponential	1	500	2.8655	0.9220(0.0016)	2.8812(0.0652)
		1500	2.8638	0.9182(0.0011)	2.8692(0.0599)
	10	500	3.0639	0.9213(0.0014)	3.0363(0.0575)
		1500	3.0727	0.9249(0.0010)	3.0160(0.0446)
	20	500	3.1488	0.9169(0.0016)	3.0886(0.0537)
		1500	3.1422	0.9226(0.0011)	3.0755(0.0423)

Table 6: Coverage of 95% intervals for returns of FIGARCH (*1, d, 0*) with parameters $\omega = 0.1,\ \phi = 0,\ d = 0.90,\ \beta = 0.20$.

Error Distribution	*Lead Lag*	*Sample size*	*Theoretical Length*	*Mean Coverage (SE)*	*Mean Length (SE)*
Normal	1	500	3.6046	0.9473(0.0008)	3.6046(0.1541)
		1500	3.6015	0.9485(0.0006)	3.6149(0.1578)
	10	500	4.6455	0.9458(0.0009)	4.6233(0.1121)
		1500	4.6360	0.9482(0.0006)	4.6596(0.1123)
	20	500	4.6782	0.9458(0.0009)	4.7088(0.0777)
		1500	4.6751	0.9483(0.0006)	4.7232(0.0778)
t-distr.	1	500	3.2863	0.9480(0.0008)	3.3085(0.1165)
		1500	3.2829	0.9489(0.0006)	3.2998(0.1162)
	10	500	3.9776	0.9469(0.0008)	4.0044(0.0709)
		1500	3.9719	0.9481(0.0006)	3.9796(0.0615)
	20	500	3.9918	0.9467(0.0009)	4.0618(0.0594)
		1500	4.0188	0.9473(0.0006)	3.9977(0.0417)
Exponential	1	500	2.6625	0.9216(0.0017)	2.6375(0.0999)
		1500	2.6571	0.9152(0.0012)	2.6468(0.0964)
	10	500	3.8349	0.9243(0.0011)	3.3212(0.0814)
		1500	3.8152	0.9297(0.0007)	3.3148(0.0637)
	20	500	4.3817	0.9159(0.0012)	3.3669(0.0666)
		1500	4.3714	0.9196(0.0008)	3.3062(0.0397)

5. Application to a Real Data Set

The proposed method was applied to S&P 500 return data from November 5, 2010 through May 2, 2018, for a total of 2201 observations. Data was obtained from the website https://finance.yahoo.com. Following standard practice, daily percentage returns of closing prices i.e. $r_t = 100.\log(s_t / s_{t-1})$ for $t = 2,3,...,2201$ were used. Here s_t denotes the closing price at day *t*. The following figure shows one-step ahead bootstrap prediction interval (95%) for S&P 500 returns.

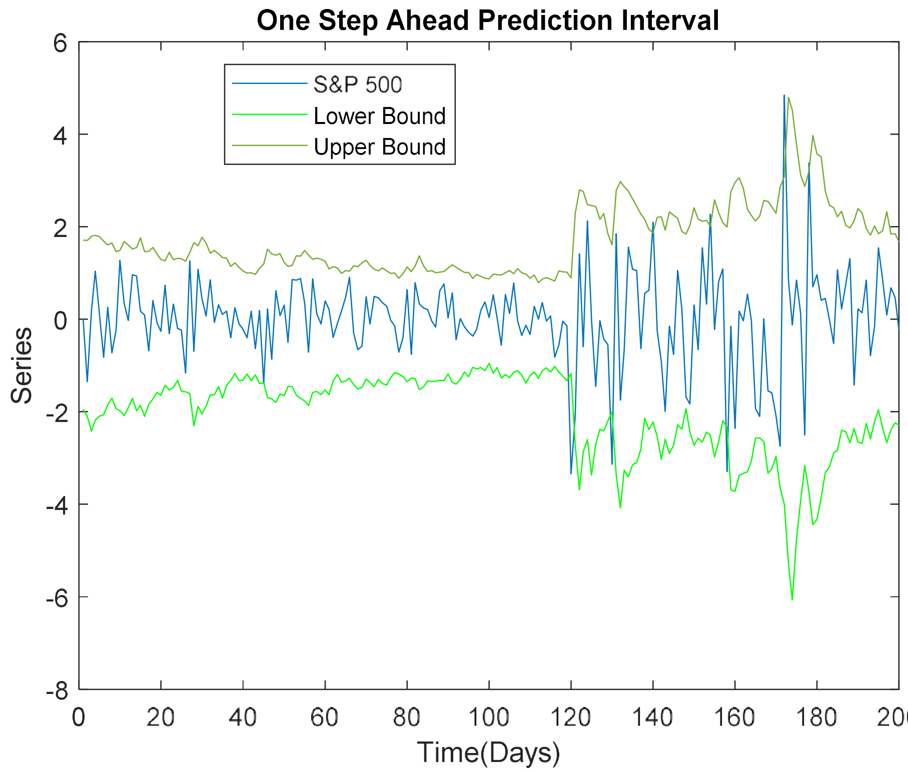

Figure 1: One-step ahead prediction intervals for S&P 500 data

Table 7: Estimated coverage probabilities for future returns

Lag ahead	*Coverage*
1	0.9600
10	0.9424
20	0.9227

6. Conclusion

In this paper we adopt the procedure proposed by Pascual et al. (2006) to construct bootstrap prediction intervals for GARCH realizations. Finite sample properties were investigated using a Monte-Carlo simulation study. In this study it is assumed the order of the FIGARCH process is known. This is not a great limitation because in most empirical modeling situations researchers have found that a GARCH process with orders $p=q=1$ would suffice. Extending this argument, one would assume that FIGARCH (1, d, 1) would suffice in most cases, as was demonstrated in our example with S&P 500 data. Simulation study shows that the proposed bootstrap-based prediction intervals perform well. The coverage probabilities obtained in the simulation study are close to the nominal values for symmetric error distributions, under varying sample sizes and parameter combinations. Further extension of obtaining prediction intervals for models such as Autoregressive-FIGARCH, FARIMA-FIGARCH using sieve bootstrap method is currently ongoing.

References

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *74*(1), 3–30. https://doi.org/10.1016/s0304-4076 (95)01749-6

Bisaglia, L., & Grigoletto, M. (2001). Prediction intervals for FARIMA processes by bootstrap methods. *Journal of Statistical Computation and Simulation*, *68*(2), 185–201. https://doi.org/10.1080/00949650108812065

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076 (86)90063-1

Bollerslev, T., & Ole Mikkelsen, H. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*(1), 151–184. https://doi.org/10.1016/0304-4076 (95)01736-4

Chen, B., Gel, Y. R., Balakrishna, N., & Abraham, B. (2010). Computationally efficient bootstrap prediction intervals for returns and volatilities in ARCH and GARCH processes. *Journal of Forecasting*, *30*(1), 51–71. https://doi.org/10.1002/for.1197

Chung, C.F., 1999. Estimating the fractionally integrated GARCH model. Working Paper, National Taiwan University.

Conrad, C. (2006). Inequality Constraints in the Fractionally Integrated GARCH Model. *Journal of Financial Econometrics*, *4*(3), 413–449. https://doi.org/10.1093/jjfinec/nbj015

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1–50. https://doi.org/10.1080/07474938608800095

Giraitis, L., Surgailis, D., & Škarnulis, A. (2017). Stationary Integrated ARCH (∞) and AR(∞) Processes with Finite Variance. *Econometric Theory*, *34*(6), 1159–1179. https://doi.org/10.1017/s0266466617000391

Nelson, D. B. (1991). Conditional Heteroscedasticity in Asset Returns: A New Approach. *Econometrica*, *59*(2), 347. https://doi.org/10.2307/2938260

Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics & Data Analysis*, *50*(9), 2293–2312. https://doi.org/10.1016/j.csda.2004.12.008

Reeves, J. J. (2005). Bootstrap prediction intervals for ARCH models. *International Journal of Forecasting*, *21*(2), 237–248. https://doi.org/10.1016/j.ijforecast.2004.09.005

Robinson, P. M., & Zaffaroni, P. (2006). Pseudo-maximum likelihood estimation of ARCH(∞) models. *The Annals of Statistics*, *34*(3), 1049–1074. https://doi.org/10.1214/009053606000000245

Rupasinghe, M., Mukhopadhyay, P., & Samaranayake, V. A. (2013). Obtaining prediction intervals for FARIMA processes using the sieve bootstrap. *Journal of Statistical Computation and Simulation*, *84*(9), 2044–2058. https://doi.org/10.1080/00949655.2013.781271

Trucíos, C., & Hotta, L. K. (2016). Bootstrap prediction in univariate volatility models with leverage effect. *Mathematics and Computers in Simulation*, *120*, 91–103. https://doi.org/10.1016/j.matcom.2015.07.001

Zaffaroni, P. (2004). Stationarity and Memory of ARCH (∞) Models. *Econometric Theory*, *20*(01), 147–160. https://doi.org/10.1017/s0266466604201062