Hindawi
Journal of Probability and Statistics
Volume 2021, Article ID 6697120, 10 pages
https://doi.org/10.1155/2021/6697120

Research Article

Adjusted Extreme Conditional Quantile Autoregression with Application to Risk Measurement

Martin M. Kithinji,[1] **Peter N. Mwita**,[2] **and Ananda O. Kube**[3]

[1]*Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation, P.O. Box 62000, Nairobi 00200, Kenya*
[2]*Department of Mathematics, Machakos University, P.O. Box 136, Machakos 90100, Kenya*
[3]*Department of Statistics and Actuarial Sciences, Kenyatta University, P.O. Box 43844, Nairobi 00100, Kenya*

Correspondence should be addressed to Martin M. Kithinji; mkithinji@kyu.ac.ke

Received 16 November 2020; Revised 6 March 2021; Accepted 22 March 2021; Published 8 April 2021

Academic Editor: Aera Thavaneswaran

In this paper, we propose an extreme conditional quantile estimator. Derivation of the estimator is based on extreme quantile autoregression. A noncrossing restriction is added during estimation to avert possible quantile crossing. Consistency of the estimator is derived, and simulation results to support its validity are also presented. Using Average Root Mean Squared Error (ARMSE), we compare the performance of our estimator with the performances of two existing extreme conditional quantile estimators. Backtest results of the one-day-ahead conditional Value at Risk forecasts are also given.

1. Introduction

Correct specification of a loss/returns distribution is key to the accuracy of a risk measure such as Value at Risk. As noted in [9], the major difference among many estimators of Value at Risk lies in estimation of the distribution of returns. Complexity in modeling financial data is due to its failure to exhibit standard statistical properties such as normality, independence, and identical distribution [18]. Statistical tests have revealed that returns exhibit fat-tails, time-varying volatility, and volatility clustering. Moreover, [7] showed that returns exhibit serial correlation over long time horizons. Models based on mean autoregression coupled with results from extreme value theory such as the AR(1)-GARCH(1,1) model in [30] incorporated most of the aforementioned characteristics of financial data but suffer from lack of robustness due to the effect of extreme observations on the mean. Extreme quantile autoregression in [2, 24, 25] among others leads to a more robust model. This is because they combine regression quantiles introduced by [17] in an autoregressive fashion while using extreme value techniques on the resulting residuals to capture the tail behaviour. A major challenge of this approach is possible quantile crossing.

The challenge of quantile crossing has been addressed by smoothing suggestions in [5, 6, 16] among others in a nonparametric setting. Equally, the conditional location-scale model used in obtaining Restricted Regression Quantiles (RRQ) in [11] averts possible crossing in extreme quantiles but can suffer from the same when estimating the median. To avert quantile crossing even at the middle, [1] added a forced ordering constraint in the estimation of multiple quantiles. Simulation results revealed that, based on standard error of the estimates, noncrossing quantile regression in [1] produces better estimates in the middle. However, the RRQ estimator produced better estimates at the tails, especially when the sample size was large.

The lack of monotonicity in estimation of conditional quantiles is addressed in [3] through sorting of originally estimated nonmonotone quantile curves using a functional delta method. The monotonic quantile functions obtained were found to be closer to the true quantile than the nonmonotonic quantiles. Function limit theory for the rearranged estimators was also derived. The resulting

monotonic quantile functions were then used in estimating economic functions using Vietnam data. However, the model was not extended to extreme cases to cover heavy tails beyond the sample.

Parametric quantile regression is used in [27] to estimate percentiles in positive valued datasets. Specifically, a linear quantile regression model was used with the error term assumed to follow a generalized gamma distribution. The idea of quantile regression was achieved by allowing parameters for the error distribution to depend on the univariate covariate leading to a location-scale model. The four-, five-, and six-parameter generalized conditional gamma distributions were considered and likelihood ratio test was used in selecting the best-fit model for each dataset. Asymptotics for the three resulting models were also derived. However, the use of generalized gamma distribution limits the model to applications where the covariate is greater than zero. This together with the fact that some financial datasets have heavier tails than the Gamma distribution limits the application of the model in finance.

We seek to improve the extreme conditional quantile estimator in [25] using an interquantile dispersion from the central conditional quantile.

2. Methods and Estimation

Let $\{S_t, t \in \mathbb{R}^+ U\{0\}\}$ be a real valued financial time series on a complete probability space $(\Omega, \mathcal{F}, P)$. We assume that $S_t \in \mathbb{R}^+$ and it is $\mathcal{F}_t$-measurable where $\{\mathcal{F}_t, t \in \mathbb{R}^+\}$ is an increasing sequence of σ-algebras representing information available up to time t. In particular, let S_t be the value of a portfolio at trading time t. The return on the portfolio at time t, used to quantify the gain in value of the portfolio from trading time $t - \Delta t$ to trading time t, is given by

$$R_t = \frac{S_t - S_{t-\Delta t}}{S_{t-\Delta t}}, \tag{1}$$

so that

$$X_t = -R_t, \tag{2}$$

is the corresponding loss return of S_t.

Definition 1 (Risk Measure). A risk measure is a function ρ from a set $\mathcal{L}$ of risks in a financial position (in this case, the loss distribution) to $\mathbb{R}$; that is, ρ: $\mathcal{L} \longrightarrow \mathbb{R}$.

We assume that X_t can be expressed using a linear heteroscedastic model of the following form:

$$X_t = \mu_t + e_t, \tag{3}$$

where $\mu_t \equiv f$: $\mathbb{R}^d \longrightarrow \mathbb{R}$ is the conditional mean function of X_t given $\mathcal{F}_{t-1}$ and it is defined as $\mu_t = Y_t'\beta$. e_t are errors and Y_t is a d-dimensional process which is $\mathcal{F}_{t-1}$-measurable. In particular, Y_t has 1 as the first element and a collection of the last observed returns up to time $t-1$; that is, $Y_t = (1, X_{t-1}, \ldots, X_{t-d})$. To ensure that the model is smooth and obeys some of the financial norms such as clustering of shocks, we further assume that e_t can be decomposed into

$$e_t = \epsilon_t \sigma_t, \tag{4}$$

where ϵ_t are independent and identically distributed random variables and $\sigma_t > 0$ is the conditional volatility. In this case, X_t is said to assume a location-scale model of the form

$$X_t = \mu_t + \sigma_t \epsilon_t. \tag{5}$$

The corresponding α quantile of X_t under this formulation is given by

$$\mu_{t,\alpha}^m = \mu_t + \sigma_t q_\alpha^\epsilon, \tag{6}$$

where q_α^ϵ is the α-quantile of ϵ_t. Let us now define a conditional quantile autoregressive model on X_t of the form

$$X_t = \mu_{t,\theta} + \varepsilon_t, \tag{7}$$

where $\mu_{t,\theta}$ is the central conditional θ-quantile of X_t and ε_t are errors with zero θ-quantile. Let $\varepsilon_t = \sigma_{t,\theta} Z_t$, where $\sigma_{t,\theta}$ is the central conditional scale of X_t and Z_t are i.i.d. residuals. Using an approach similar to Points Over Threshold (POT), we propose an extreme conditional quantile of the form given in equation (8) and refer to it as the adjusted extreme conditional quantile. That is, suppose that we are interested in an extreme quantile, $\mu_{t,\theta,\alpha}$, for some $\alpha \approx 1$ or 0; the idea is to estimate the central quantile, $\mu_{t,\theta}$, and scale $\sigma_{t,\theta}$, for some level θ in the middle and approximate the extreme quantile as

$$\mu_{t,\theta,\alpha} = \mu_{t,\theta} + \sigma_{t,\theta}\left[q_\alpha^z - q_\theta^z\right], \tag{8}$$

where q_α^z and q_θ^z are the α, θ-quantile of Z_t, respectively, for $\alpha, \theta \in (0, 1)$. If the parametric distribution of Z_t is known, then q_α^z and q_θ^z are easily determined as the inverse of the cumulative distribution of Z_t at probability levels α and θ, respectively; otherwise, appropriate estimates are determined. Note that $\mu_{t,\theta,\alpha}$ is $\mu_{t,\alpha}^m$ in conditional quantile autoregressive form. Observe that when $\alpha = \theta$, equation (8) reduces to

$$\mu_{t,\theta,\theta} = \mu_{t,\theta}, \tag{9}$$

which is the central conditional quantile of X_t given $Y_t = y$. This confirms that indeed ε_t has a zero θ-quantile. From equation (8), we obtain the following estimator for the extreme conditional quantile:

$$\widehat{\mu}_{t,\theta,\alpha} = \widehat{\mu}_{t,\theta} + \widehat{\sigma}_{t,\theta}\left[\widehat{q}_\alpha^z - \widehat{q}_\theta^z\right], \tag{10}$$

where $\widehat{q}_\theta^z$ and $\widehat{q}_\alpha^z$ are appropriate estimates of the θ, α-quantiles of the i.i.d. residuals, respectively. We compare this estimator with

$$\widehat{\mu}_{t,\theta,\alpha}^m = \widehat{\mu}_{t,\theta} + \widehat{\sigma}_{t,\theta}\widehat{q}_\alpha^z, \tag{11}$$

proposed in [25], and

$$\widehat{\mu}_{t,\theta,\alpha}^h = \widehat{\mu}_{t,\theta} + \widehat{\sigma}_{t,\theta} c_\alpha, \tag{12}$$

where c_α is the resulting coefficient from quantile regression of ε_t against $\widehat{\varepsilon_t}$ at $100\alpha\%$. Note that $\widehat{\mu}_{t,\theta,\alpha}^h$ is the estimator proposed in [11].

2.1. Estimation of Central Quantiles. Let μ_t: $\mathbb{R}^d \longrightarrow \mathbb{R}$ be an unknown smooth function and define the loss function M_θ:

$$\begin{aligned} M_\theta(x,\mu_t) &= \theta\left|x-\mu_t\right|^+ + (1-\theta)\left|x-\mu_t\right|^- \\ &= (x-\mu_t)\left(\theta - \mathbb{I}_{x-\mu_t\le 0}\right), \end{aligned} \tag{13}$$

where $|x-\mu_t|^+$ and $|x-\mu_t|^-$ represent absolute positive and negative values, respectively, and $\mathbb{I}_{x-\mu_t\le 0}$ is the indicator function. Assuming that the conditional quantile process is well defined, we expect

$$E\left[M_\theta\left(X_t,\mu_{t,\theta}\right)|Y_t = y\right] = 0. \tag{14}$$

So the θ-conditional quantile of X_t is given by

$$\mu_{t,\theta} = \arg\min_{\mu_t} E\left[M_\theta(X_t,\mu_t)|Y_t = y\right]. \tag{15}$$

Note that equation (14) can be used to check whether the conditional quantile process is correctly specified or not. We impose the following regularity assumptions to ensure consistency of the conditional quantiles.

Assumption 1. $\{X_t, Y_t\}$ are identically distributed with the joint probability density $f_{X,Y}(x,y)$ and a continuous conditional probability density $f_{X|Y}(x|y)$ of X_t given $Y_t = y$.

Assumption 2. There exists $\delta > 1$ such that $E(\|Y_t\|^\delta) < \infty$.

Assumption 3. $\mu_{t,\theta} \equiv f: B \longrightarrow \mathbb{R}$, where B is a compact subset of $\mathbb{R}^d$.

Assumption 4. $f_{\varepsilon|Y}(\varepsilon|x) > 0 \forall x$, where $f_{\varepsilon|Y}(\varepsilon|x)$ is the conditional probability density of $\varepsilon_t = X_t - \mu_{t,\theta}$ given $Y_t = y$.

From the sample analog of equation (15), we obtain

$$\widehat{\mu}_{t,\theta} = \operatorname*{argmin}_{\mu_t} \frac{1}{n}\sum_{t=1}^{n} M_\theta\left[(X_t,\mu_t)|Y_t = y\right], \tag{16}$$

which is the θ-conditional quantile estimator for a sample of size n. To overcome the limitation of quantile crossing, we used the approach in [1] where required quantiles were estimated simultaneously with a noncrossing constraint using the optimization problem:

$$\begin{aligned} &\min_{\mu_t} \sum_{i=1}^{k} w(\theta_i) \sum_{t=1}^{n} M_{\theta_i}\left[\left(X_t,\mu_{t_i}\right)|Y_t = y\right] \\ &\text{subject to} \quad \mu_{t_i} \ge \mu_{t_{i-1}}, Y_t \in B, i = 1,\ldots,k, \end{aligned} \tag{17}$$

for some weight function $w(\theta_i) > 0$. A conveniently used practical choice of the weight function which was also adopted for this study is $w(\theta_i) = 1$, $\forall i = 1,\ldots,k$.

Lemma 1 (consistency central quantiles' estimator). *Given that Assumptions 1, 2, 3, and 4 hold,* $\widehat{\mu}_{t,\theta} - \mu_{t,\theta} = o_p(1)$.

2.1.1. The Scale Function. To still maintain dependence and ensure positivity of the scale function, this study incorporated a scale function in the form of a quantile autoregressive (QAR) function on the absolute of the nonstandardized residuals. This was achieved by replacing $\mu_{t,\theta}$ with its corresponding estimate in equation (7) so that $\varepsilon_t = X_t - \widehat{\mu}_{t,\theta}$ and

$$\left|\varepsilon_t\right| = \sigma_{t,\theta} + \vartheta_{t,\theta}\eta_t, \tag{18}$$

where $\sigma_{t,\theta}$ is the central conditional θ-quantile of $|\varepsilon_t|$, $\vartheta_{t,\theta}$ is the conditional scale of $|\varepsilon_t|$, η_t are i.i.d. residuals, and $\varpi_t = \vartheta_{t,\theta}\eta_t$ are errors with zero θ-quantile. Similarly, as is the case in Section 2.1, we let σ_t: $\mathbb{R}^d \longrightarrow \mathbb{R}^+$ be an unknown smooth function and define the loss function $M_\theta(x,\sigma_t)$. We assume that the QAR process in equation (18) obeys the four regularity assumptions given earlier so that, using noncrossing quantile regression approach, we obtain the following estimate of the scale:

$$\widehat{\sigma}_{t,\theta} = \operatorname*{argmin}_{\sigma_t} \frac{1}{n}\sum_{t=1}^{n} M_\theta\left[\left(\left|\varepsilon_t\right|,\sigma_t\right)|Y_t = y\right], \tag{19}$$

where $M_\theta[(|\varepsilon_t|,\sigma_t)|Y_t = y]$ is a loss function of the form given in equation (13).

Lemma 2 (consistency of the scale estimator). *Given that the QAR process defined in equation (18) satisfies Assumptions 1, 2, 3, and 4,* $\widehat{\sigma}_{t,\theta} - \sigma_{t,\theta} = o_p(1)$.

2.2. Extreme Value Theory (EVT). Most of the financial datasets are heavy-tailed [12]. Therefore, it is fundamentally important to incorporate extreme value theory in the estimation of extreme quantiles. A basic requirement for the application of EVT is independence in the particular distribution. The study in [25] observed that it is appropriate to assume that, at high (low) levels of α, the standardized residuals in equation (7) given by

$$Z_t = \frac{\varepsilon_t}{\widehat{\sigma}_{t,\theta}}, \tag{20}$$

where $\widehat{\sigma}_{t,\theta}$ is the estimate of the scale, are approximately independent, which allows us to apply EVT. Let $Z_1, Z_2, \ldots$, follow a common distribution function F. Consider a sample $(Z_i)_{i=1,\ldots,n}$ from which $M_n = \max\{Z_1,\ldots,Z_n\}$ and is such that $M_n \longrightarrow z_F$ (a.s) where $z_F \cong \sup\{z \in \mathbb{R} | F(z) < 1\}$. Pursuant to Fisher–Tippett's theorem in [8], the random variable Z (or alternatively the distribution F of Z) is said to belong to the Maximum Domain of Attraction (MDA) of the extreme value distribution H if there exist norming constants $c_n > 0$ and $d_n \in \mathbb{R}$ such that

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H. \tag{21}$$

Consequently, H is referred to as the Generalized Extreme Value (GEV) distribution.

This study applied Points Over Threshold (POT) method because it uses more data leading to better estimates compared to the Block Maxima method. POT models the distribution

$$F_u(z) = P(Z - u \le z | Z > u) = \frac{F(z+u) - F(u)}{1 - F(u)}. \quad (22)$$

of all excesses above a particular threshold u, where $0 \le z \le z_F - u$.

Theorem 1 (Pickands–Balkema–de Haan) (see [22]). *We can find (positive-measurable function) $\beta(u)$ such that*

$$\lim_{u \longrightarrow z_F} \sup_{0 \le z < z_F - u} \left| F_u(z) - G_{\lambda,\beta(u)}(z) \right| = 0, \quad (23)$$

if and only if $F \in MDA(H_\lambda), \lambda \in \mathbb{R}$, and $G_{\lambda,\beta(u)}(z)$ is the Generalized Pareto Distribution (GPD) given by

$$G_{\lambda,\beta(u)}(z) = G_{\lambda,\beta}(z-u) = \begin{cases} 1 - \left(1 + \dfrac{\lambda(z-u)}{\beta}\right)^{-(1/\lambda)}, & \lambda \neq 0, \\ 1 - \exp - \left(\dfrac{z-u}{\beta}\right), & \lambda = 0, \end{cases} \quad (24)$$

with $z \ge u$ when $\lambda > 0$ and $u \le z \le u - (\beta/\lambda)$ when $\lambda < 0$. λ and β are the shape and scale parameters, respectively.

A major challenge in POT is accuracy in choosing a threshold to separate extreme observations from the center of the distribution [30]. Among the methods discussed in [32], most authors such as those of [21, 28, 30] prefer the conventional method in which a threshold that ensures that between 5% and 10% of the sample data is classified as extreme observations is chosen. Although conventional method is subjective, the choice of the threshold can be checked for appropriateness using a mean excess plot. The mean excess function for the GPD is given by

$$E[Z - u | Z > u] = \frac{\beta - \lambda u}{1 + \lambda}, \quad (25)$$

where $1 + \lambda > 0$. This implies that an optimal threshold corresponds to start of approximate linearity of the mean excess plot with the sign of the slope, $\lambda/1 - \lambda$, indicating the specific family of the GPD. A positive sign corresponds to the Frechet family, while a negative sign implies the Weibull family [8].

For convenience in making inferences on variability of the estimated quantiles, a recommendation in [13] on the use of Probability Weighted Moments (PWM) method in estimating parameters of the GPD was adopted. Using the first and second PWMs, we obtain the corresponding parameters estimates as

$$\lambda = \frac{M_0}{M_0 - 2M_1} - 2, \qquad \beta = \frac{2M_0M_1}{M_0 - 2M_1}, \quad (26)$$

where M_0 and M_1 are obtained by replacing for $k = 0$ and $k = 1$ in

$$M_{1,0,k} = M_k = E\left[X\{1 - F(X)\}^k\right] = \frac{\beta}{(k+1)(k+1+\lambda)}, \quad (27)$$

which is the PWM of GPD with $\lambda > -1$. See [10, 23] for details on PWM method. For a sample of size n, the corresponding PWM estimates are given by

$$\widehat{M}_k = \frac{1}{n} \sum_{i=1}^{n} (1 - P_{i:n})^k x_{i:n}, \quad (28)$$

where $x_{1:n} \le \cdots \le x_{n:n}$ is the ordered sample and $P_{j:n} = i + \gamma/n + \delta$ for suitable constants γ and δ. As recommended in [20], $\gamma = -0.35$ and $\delta = 0$.

The overall distribution of the standardized residuals was obtained by splicing the GPD with the empirical bulk distribution at the threshold using the approach in [21, 29, 30] among others to obtain

$$F(z) = (1 - F(u)) G_{\lambda,\beta}(z-u) + F(u), \quad (29)$$

for $z > u$ and $G_{\lambda,\beta}(z-u)$ the Generalized Pareto Distribution. When $F(u)$ is approximated empirically, we obtain the following estimate of $F(z)$:

$$\widehat{F}(z) = 1 - \frac{m}{N}\left(1 + \frac{\widehat{\lambda}(z-u)}{\widehat{\beta}}\right)^{(-1/\widehat{\lambda})}, \quad (30)$$

where N is the sample of size, m is the number of exceedances above the threshold u, and $\widehat{\beta}$ together with $\widehat{\lambda}$ are the estimated GPD parameters.

Lemma 3 (consistency of Probability Density Function (PDF) estimator). *Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. random variables from a Cumulative Distribution Function (CDF) $F(z)$ belonging to the MDA of $H(\lambda)$. Suppose that $F(z)$ has a right endpoint at $z_F \le \infty$; then $\sup_z |\widehat{F}(z) - F(z)| \longrightarrow 0$ as $n \longrightarrow \infty$.*

From equation (30), we get the following estimate for the quantile of the standardized residuals at level α:

$$\widehat{q}_\alpha^z = u + \frac{\widehat{\beta}}{\widehat{\lambda}}\left[\left(\frac{N}{m}(1-\alpha)\right)^{-\widehat{\lambda}} - 1\right]. \quad (31)$$

Lemma 4 (consistency of error quantiles). *Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. random variables from a CDF F belonging to the MDA of $H(\lambda)$ satisfying $\alpha < F(q_\alpha^z + \epsilon)$ for any $\epsilon > 0$. Then, for every $\epsilon > 0$ and $n = 1, 2, \ldots,$*

$$\lim_{n \longrightarrow \infty} P\left(\left|\widehat{q}_\alpha^z - q_\alpha^z\right| < \epsilon\right) = 1. \quad (32)$$

Using equation (10), we obtained the one-step VaR predictions as

$$\widehat{VaR}_{t+1}^{\alpha} = \widehat{\mu}_{t+1,\theta,\alpha} = \widehat{\mu}_{t+1,\theta} + \widehat{\sigma}_{t+1,\theta}\left[\widehat{q}_\alpha^z - \widehat{q}_\theta^z\right], \quad (33)$$

where $\widehat{\mu}_{t+1,\theta}$ and $\widehat{\sigma}_{t+1,\theta}$ are the corresponding one-step θ-quantile and scale estimates, respectively, from the linear conditional quantile process.

Theorem 2 (consistency of extreme quantile estimator). *Given that the QAR processes defined satisfy Assumptions 1, 2, 3, and 4 and $F(z)$ belongs to the MDA of $H(\lambda)$, $\widehat{\mu}_{t,\theta,\alpha} - \mu_{t,\theta,\alpha} = o_p(1)$.*

3. Simulations

To evaluate the accuracy of our estimators, a sample of size $T = 4250$ was generated using the model $X_t = 0.5 + 0.3X_{t-1} + \sqrt{1 + 0.35X_{t-1}^2}Z_t$, where Z_t follows Student's t-distribution with 4 degrees of freedom. The sample was partitioned into design data of size n and test data of size $T - n$. Figure 1 represents a sample path of the model superimposed with the median. Note that, for simulation purposes, $\theta = 0.5$ was used in estimating the central quantile.

Clearly, from the sample path, there is some level of volatility clustering, which is common in most financial data. An ACF plot of the resulting standardized residuals in Figure 2 confirms that indeed they are independent.

$z = 2.3283$ was chosen as the threshold to ensure that 10% of the resulting ordered standardized errors were classified as extremes. This was confirmed by approximate linearity of the mean excess plot after the threshold as shown in Figure 3.

The corresponding shape and scale parameter estimates from the GPD fit were 0.1156182 and 1.272937, respectively. Table 1 outlines the sample statistics of the estimates of the various quantiles together with the data.

Figure 4 shows the corresponding quantile estimates at different levels of α.

The accuracy of our extreme quantile estimator was evaluated using the Average Root Mean Squared Error (ARMSE). The RMSE seeks to return the Mean Squared Error (MSE) to the original scale of the sample. For k sample paths of simulations of size n of the extreme quantiles, the average RMSE is given by

$$A.\text{R.M.S.E} = \frac{1}{m}\sum_{i=1}^{k}\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\widehat{\mu}_{t,\theta,\alpha} - \mu_{t,\alpha}^{m}\right)^2}, \tag{34}$$

where $\widehat{\mu}_{t,\theta,\alpha}$ is replaced by $\widehat{\mu}_{t,\theta,\alpha}^{m}$ or $\widehat{\mu}_{t,\theta,\alpha}^{h}$ when considering extreme conditional quantiles or restricted regression quantiles, respectively. To check how our model behaves under different choices of central quantile, we computed ARMSE for the extreme conditional quantile at $\alpha = 0.95$, where $\theta = 0.25, 0.5$ and 0.75 were considered. 1000 sample paths, each of sizes 250, 500, 1000, 2000, and 4000, were used in the computation of ARMSE. Table 2 reports the obtained ARMSE values.

We note that, for a large enough sample (2000 observations and above), ARMSE is lowest when $\theta = 0.5$ and thus this choice of theta is maintained in investigating the accuracy of our estimator and forecasting the one-day-ahead VaR. Table 3 outlines the obtained ARMSE for the extreme conditional quantile at $\alpha = 0.95$ under three different models. The sample sizes and number or replications are still maintained.

Based on ARMSE, both RRQ and ECQ perform better than AECQ for small samples. However, as the sample size increases, AECQ outperforms both RRQ and ECQ. The decreasing ARMSE with increase in the sample size for AECQ and ECQ confirms that both are consistent estimators of the extreme conditional quantile. Also, for sample size above 2000, the rate of convergence of the AECQ estimator is higher than that of the ECQ estimator. It was not possible to comment on the consistency of the RRQ estimator, since its ARMSE fluctuated with increase in sample size. The consistent reduction in ARMSE when noncrossing constraint is added during estimation, confirming that indeed this constraint increases accuracy of resulting estimators.

4. Evaluating VaR Forecasts

In Section 3, we evaluated accuracy of our in-sample quantile estimates. In this section, we extend this by evaluating the out-of-sample VaR forecasts from our quantile estimator. To achieve this, we carry out backtests on 250 one-day-ahead VaR forecasts (as recommended in the Basel Accord) using coverage tests in [4, 31]. Coverage tests were adopted due to their popularity in literature and practice [26]. Consider the failure process $I_t^{\alpha} = I\{X_t > \text{VaR}_t^{\alpha}\}$, where I is the indicator function such that

$$I_t^{\alpha} = \begin{cases} 1, & \text{if } X_t > \text{VaR}_t^{\alpha}, \\ 0, & \text{otherwise}, \end{cases} \tag{35}$$

and $t = n+1, n+2, \ldots, T$. By Lemma 1 in [4], $I_t^{\alpha} \overset{iid}{\sim} \text{Binom}(\alpha)$ which is tested using the conditional coverage test that combines both the unconditional coverage test in [19] and the test for independence in [4] to under the null hypothesis $E[I_t^{\alpha}|I_{t-1}^{\alpha}] = \alpha$. The likelihood under the null hypothesis is

$$L\left(\alpha; I_{n+1}^{\alpha}, I_{n+2}^{\alpha}, \ldots, I_T^{\alpha}\right) = (1-\alpha)^{n_0}\alpha^{n_1}, \tag{36}$$

where n_1 is the number of VaR exceedances and $n_0 = T - n - n_1$. Now consider the first-order Markov chain generated by the transition probabilities of I_t^{α}:

$$\Pi_1 = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix} = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}, \tag{37}$$

where $\pi_{ij} = \Pr(I_t^{\alpha} = j | I_{t-1}^{\alpha} = i)$. This has an approximate likelihood function:

$$L\left(\Pi_1; I_{n+1}^{\alpha}, I_{n+2}^{\alpha}, \ldots, I_T^{\alpha}\right) = \pi_{00}^{n_{00}}\pi_{01}^{n_{01}}\pi_{10}^{n_{10}}\pi_{11}^{n_{11}}, \tag{38}$$

where n_{ij} is the number of times observation i is followed by j in the failure process I_t^{α}. From equation (38), we obtain the maximum likelihood estimate of π_{ij} as

$$\widehat{\pi}_{ij} = \frac{n_{ij}}{\sum_{j=0}^{1} n_{ij}}. \tag{39}$$

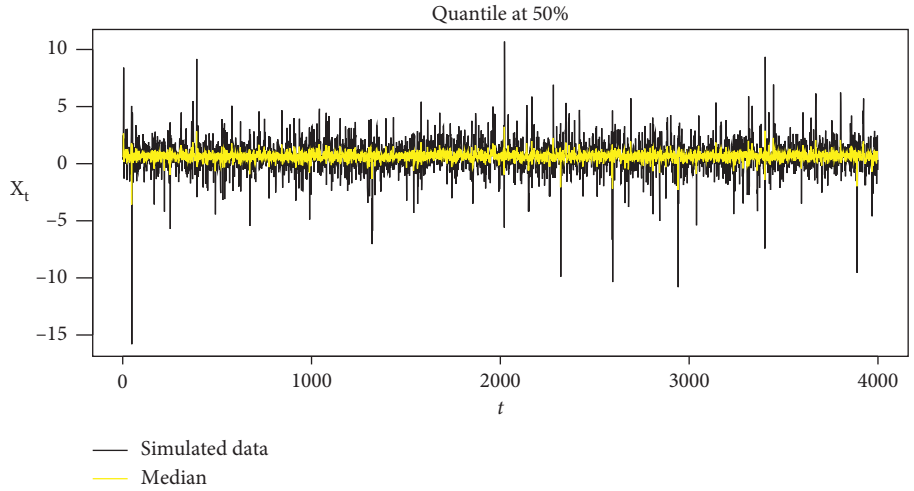

Figure 1: Sample path superimposed with median.

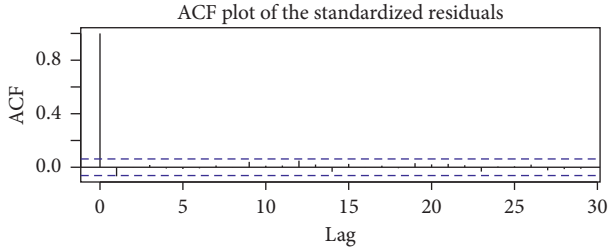

Figure 2: Autocorrelation plot of resulting residuals.

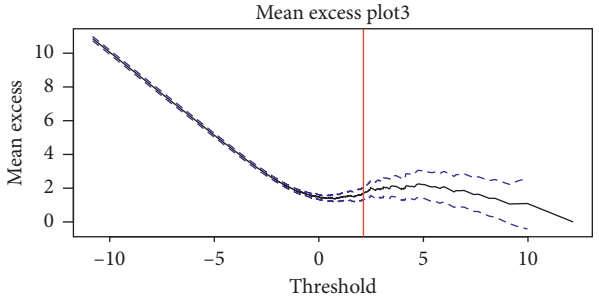

Figure 3: Mean excess plot.

Table 1: Sample statistics.

α	0	0.50	0.75	0.95	0.99
Minimum	−15.81	−3.58	−2.80	−0.73	1.76
Median (Q_1, Q_3)	0.67 (0.01, 1.38)	0.65 (0.48, 0.83)	1.15 (0.99, 1.32)	2.47 (2.34, 2.61)	4.07 (3.97, 4.17)
Mean ± SD	0.69 ± 1.41	0.66 ± 0.36	1.16 ± 0.34	2.48 ± 0.27	4.07 ± 0.20
Maximum	10.69	3.22	3.55	4.42	5.47

Q_1, lower quartile; Q_3, upper quartile; SD, standard deviation.

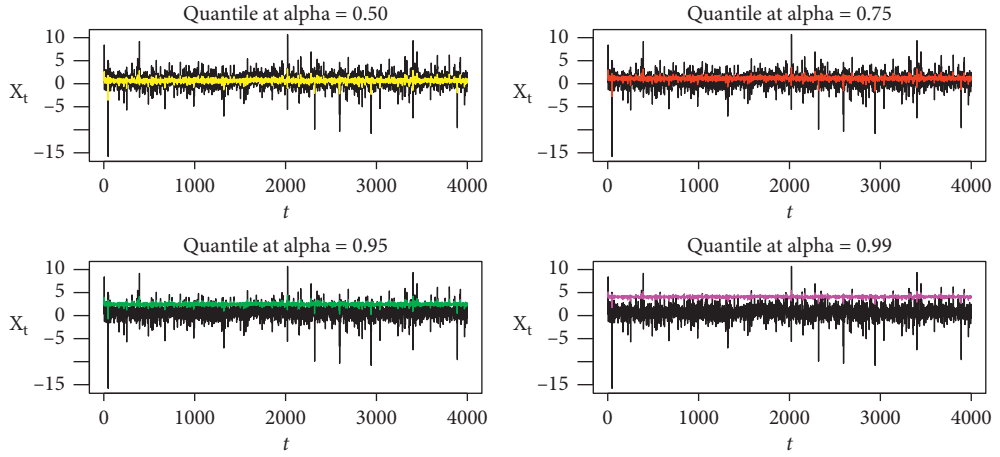

Figure 4: Quantile estimates at various α-levels for $\theta = 50\%$.

Table 2: ARMSE for $\widehat{\mu}_{t,\theta,\alpha}$ for $\alpha = 95\%$.

Sample size (n)	$\theta = 0.25$	$\theta = 0.5$	$\theta = 0.75$
250	0.96844	0.74493	0.62015
500	0.79632	0.60788	0.56179
1000	0.71974	0.54056	0.53254
2000	0.71416	0.49585	0.50037
3000	0.69751	0.47862	0.49713
4000	0.69608	0.46439	0.48776

Table 3: ARMSE for $\widehat{\mu}_{t,\theta,\alpha}$ under different models for $\theta = 50\%$ and $\alpha = 95\%$.

Sample size (n)	Without noncrossing constraint			With noncrossing restriction		
	RRQ	ECQ	AECQ	RRQ	ECQ	AECQ
250	0.58236	0.64552	0.73307	0.58200	0.64230	0.72728
500	0.56989	0.61336	0.60766	0.56816	0.61083	0.60341
1000	0.53386	0.58654	0.53386	0.55780	0.58529	0.53186
2000	0.56051	0.58204	0.49562	0.55800	0.57880	0.49360
3000	0.55478	0.57663	0.48235	0.55235	0.57369	0.48073
4000	0.55271	0.57619	0.47032	0.55013	0.57295	0.46832

RRQ, restricted regression quantiles in [11]. ECQ, extreme conditional quantiles in [25]. AECQ, proposed adjusted extreme conditional quantiles.

Table 4: P values for the likelihood ratio tests.

Sample size (n)	RRQ model			ECQ model			AECQ model		
	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}
250	0.000	0.001	0.000	0.014	0.013	0.002	0.001	0.006	0.000
500	0.000	0.001	0.000	0.036	0.010	0.003	0.009	0.005	0.000
1000	0.000	0.000	0.000	0.042	0.008	0.003	0.019	0.004	0.000
2000	0.000	0.001	0.000	**0.055**	0.010	0.005	**0.051**	0.005	0.002
3000	0.000	0.001	0.000	**0.062**	0.012	0.006	**0.062**	0.005	0.003
4000	0.000	0.000	0.000	**0.067**	0.009	0.005	**0.066**	0.004	0.002

Therefore, the conditional coverage hypothesis can be assessed using the likelihood ratio:

$$\mathrm{LR}_{\mathrm{cc}} = -2\log\left[\frac{L\left(\alpha; I_{n+1}^{\alpha}, I_{n+2}^{\alpha}, \ldots, I_{T}^{\alpha}\right)}{L\left(\widehat{\Pi}_1; I_{n+1}^{\alpha}, I_{n+2}^{\alpha}, \ldots, I_{T}^{\alpha}\right)}\right] \overset{\text{asy}}{\sim} \chi^2(2). \quad (40)$$

Table 4 reports the obtained P values for the likelihood ratios of the three tests in [4]. The tests were conducted on 250 one-day-ahead 5% VaR forecasts from the three considered models.

$\mathrm{LR}_{\mathrm{uc}}$ denotes likelihood ratio for the unconditional coverage test. $\mathrm{LR}_{\mathrm{ind}}$ denotes likelihood ratio for the independence test. $\mathrm{LR}_{\mathrm{cc}}$ denotes likelihood ratio for the conditional coverage test. Models accepted at 5% level of significance are highlighted in bold. Note that n is size of the sample used in estimation, while testing was done using a sample of size 250 for all n.

Observe that, as a consequence of consistency, the accuracy of ECQ and AECQ forecasts improves with increase in sample size. It can also be seen that all the three models perform poorly under $\mathrm{LR}_{\mathrm{ind}}$ and $\mathrm{LR}_{\mathrm{cc}}$ due to dependence in autoregression. Based on $\mathrm{LR}_{\mathrm{uc}}$, the RRQ model performed poorly when it comes to forecasting. This can be attributed to the failure to incorporate extreme value theory in estimating residual quantiles in the RRQ model.

5. Conclusions and Recommendations

We have derived the extreme conditional quantile estimator and used it to obtain the one-step-ahead conditional Value at Risk forecast for a simulated financial distribution. Consistency of our estimators has been proved and illustrated through Monte Carlo simulations. We noticed that adding the noncrossing restriction during estimation improves accuracy of the resulting extreme conditional quantile estimator. Backtesting results from the one-step-ahead conditional Value at Risk forecasts indicate that independence and conditional coverage tests in [4] are not appropriate for our estimators due to dependence in autoregressive models.

6. Proofs

Proof of Lemma 1. Let $\zeta(\mu_{t,\theta}) = M_\theta(X_t, \mu_t) - M_\theta(X_t, \mu_{t,\theta})$, $\zeta_n(\mu_{t,\theta}) = (1/n)\sum_{\mu_{t,\theta}}^{n}\zeta(\mu_{t,\theta})$, and $\overline{\zeta}(\mu_{t,\theta}) = E(\zeta_n(\mu_{t,\theta}))$. Note that, by Assumption 1, $\overline{\zeta}_n(\mu_{t,\theta})$ does not depend on n. Since $M_\theta(X_t, \mu_t)$ does not depend on $\mu_{t,\theta}$, then $\widehat{\mu}_{t,\theta} \in \operatorname{argmin}_{\mu_t}(1/n)\sum_{t=1}^{n}\zeta(\mu_{t,\theta})$. We need to show that the objective function $\zeta(\mu_{t,\theta})$ satisfies the following conditions for application of Theorem 12.2 in [34]:

(1) $\zeta(\mu_{t,\theta})$ is measurable for each $\beta \in B$

(2) $\zeta(\mu_{t,\theta})$ is continuous on B almost surely

(3) $\exists$ a measurable function m: $\mathbb{R}^d \longrightarrow \mathbb{R}$ such that

(i) $|\zeta(\mu_{t,\theta})| \le m(Y_t)\ \forall \beta \in B$

(ii) $\exists \delta > 0$: $E|m(Y_t)|^{1+\delta} \le M \le \infty$

(4) $\overline{\zeta}(\mu_{t,\theta})$ has a unique minimum at $\mu_{t,\theta}$

The functional form of $\zeta(\mu_{t,\theta})$ and Assumption 3 guarantees measurability of $M_\theta(X_t, \mu_{t,\theta})$. To prove condition (2), we first show that $\zeta(\mu_{t,\theta}$ is Lipschitz continuous. By definition,

$$\begin{aligned}\zeta\left(\mu_{t,\theta}\right) &= \left(\left(X_t - \mu_t\right)\left[\theta - I\left(X_t - \mu_t \le 0\right)\right]\right) \\ &\quad -\left(\left(X_t - \mu_{t,\theta}\right)\left[\theta - I\left(X_t - \mu_{t,\theta} \le 0\right)\right]\right) \\ &= \theta\left(\mu_{t,\theta} - \mu_t\right) - \left[\left(X_t - \mu_t\right)I\left(X_t - \mu_t \le 0\right)\right. \\ &\quad \left. -\left(X_t - \mu_{t,\theta}\right)I\left(X_t - \mu_{t,\theta} \le 0\right)\right].\end{aligned} \quad (41)$$

Considering the possible range X_t, we have the following:

(i) For $\mu_t < X_t < \mu_{t,\theta}$, it follows that $I(X_t - \mu_t \le 0) = 0$ and $I(X_t - \mu_{t,\theta} \le 0) = 1$; hence, equation (41) reduces to

$$\begin{aligned}\zeta\left(\mu_{t,\theta}\right) &= \theta\left(\mu_{t,\theta} - \mu_t\right) + X_t - \mu_{t,\theta} \\ &= \theta\left(\mu_{t,\theta} - \mu_t\right) + X_t - \mu_{t,\theta} + \left(\mu_t - \mu_t\right) \\ &= \left(X_t - \mu_t\right) - (1-\theta)\left(\mu_{t,\theta} - \mu_t\right).\end{aligned} \quad (42)$$

Since $(X_t - \mu_t) > 0$ and $(X_t - \mu_{t,\theta}) < 0$, we have

$$-(1-\theta)\left(\mu_{t,\theta} - \mu_t\right) \le \zeta\left(\mu_{t,\theta}\right) \le \theta\left(\mu_{t,\theta} - \mu_t\right), \quad (43)$$

and so $|\zeta(\mu_{t,\theta})|$ is bounded above by either $(1-\theta)(\mu_{t,\theta} - \mu_t)$ or $\theta(\mu_{t,\theta} - \mu_t)$.

(ii) Similarly, when $\mu_t \le \mu_{t,\theta} < X_t$, $I(X_t - \mu_t \le 0) = 0$ and $I(X_t - \mu_{t,\theta} \le 0) = 0$; hence,

$$\zeta\left(\mu_{t,\theta}\right) = \theta\left(\mu_{t,\theta} - \mu_t\right), \quad (44)$$

(iii) $X_t < \mu_t \le \mu_{t,\theta}$; then $I(X_t - \mu_{t,\theta} \le 0) = 1$ and $I(X_t - \mu_{t,\theta} \le 1)$; hence,

$$\zeta\left(\mu_{t,\theta}\right) = (\theta - 1)\left(\mu_{t,\theta} - \mu_t\right). \quad (45)$$

Combining equations (43)–(45), we have

$$\begin{aligned}\left|\zeta\left(\mu_{t,\theta}\right)\right| &\le \max(\theta, 1-\theta)\left|\mu_{t,\theta} - \mu_t\right| \\ &\le \left|\mu_t - \mu_{t,\theta}\right|.\end{aligned} \quad (46)$$

Thus, $\zeta(\mu_{t,\theta})$ is Lipschitz continuous for $K = 1$ and hence is differentiable almost everywhere by Rademacher's theorem in [14] which implies that $\zeta(\mu_{t,\theta})$ is continuous everywhere.

To prove Condition 3(i), let $m(Y_t) = \|Y_t\|(L\|\overline{\beta}\|^2)(\theta + 1)$, where $\overline{\beta} = \max_{\beta \in B}\|\beta\|$. Existence of $\overline{\beta}$ is guaranteed by Assumption 3. Clearly, $m(Y_t)$ is measurable and $|\zeta(\mu_{t,\theta}| \le m(Y_t)$. Assumption 2 ensures that Condition 3(ii) is satisfied.

To verify Condition 4, let $\delta = \mu_t - \mu_{t,\theta}$; we need to show that $E(\zeta(\mu_{t,\theta})) > 0$ for any $\delta > 0$. By Knight's identity in [15], we have

$$\begin{aligned} E\left[\zeta\left(\mu_{t,\theta}\right)\right] &= E\left[M_\theta(X_t,\mu_t) - M_\theta\left(X_t,\mu_{t,\theta}\right)\right] \\ &= E\left[\left(\mu_t - \mu_{t,\theta}\right)\left(I\left\{X_t - \mu_{t,\theta} \le 0\right\} - \theta\right) + \int_0^{\mu_t-\mu_{t,\theta}} \left[I\left\{X_t - \mu_{t,\theta} \le t\right\} - I\left\{X_t - \mu_{t,\theta} \le 0\right\}\right]\mathrm{d}t\right] \\ &= \int_0^{\mu_t-\mu_{t,\theta}} \left[F_{\varepsilon|Y}(t) - F_{\varepsilon|Y}(0)\right]\mathrm{d}t. \end{aligned} \tag{47}$$

Thus, $E[\zeta(\mu_{t,\theta})] > 0$ for all $\delta > 0$ by monotonicity of the CDF. Therefore, $\widehat{\mu}_{t,\theta} - \mu_{t,\theta} = o_p(1)$ by Theorem 12.2 in [34].

Proof of Lemma 2. The proof proceeds in a similar way to the proof of Lemma 1.

Proof of Lemma 3. Observe that

$$\begin{aligned} \sup_z |\widehat{F}(z) - F(z)| &= \sup_z \left|[1 - \widehat{F}(u)]\widehat{G}_{\lambda,\beta}(z-u) + \widehat{F}(u) - [1 - F(u)]G_{\lambda,\beta}(z-u) - F(u)\right| \\ &= \sup_z \Big|[\widehat{F}(u) - F(u)] + \left[\widehat{G}_{\lambda,\beta}(z-u) - G_{\lambda,\beta}(z-u)\right] - \widehat{F}(u)\widehat{G}_{\lambda,\beta}(z-u) + F(u)G_{\lambda,\beta}(z-u) \\ &\quad + \widehat{F}(u)G_{\lambda,\beta}(z-u) - \widehat{F}(u)G_{\lambda,\beta}(z-u)\Big| \\ &\le \sup_u |\widehat{F}(u) - F(u)| + \sup_z \left|\widehat{G}_{\lambda,\beta}(z-u) - G_{\lambda,\beta}(z-u)\right| \\ &\quad - G_{\lambda,\beta}(z-u)\sup_u |\widehat{F}(u) - F(u)| - \widehat{F}(u)\sup_z \left|\widehat{G}_{\lambda,\beta}(z-u) - G_{\lambda,\beta}(z-u)\right|. \end{aligned} \tag{48}$$

Note that $\sup_u|\widehat{F}(u) - F(u)| \longrightarrow 0$ as $n \longrightarrow \infty$ by Glivenko-Cantelli theorem in [33] and $\sup_z|\widehat{G}_{\lambda,\beta}(z-u) - G_{\lambda,\beta}(z-u)| \longrightarrow 0$ as $n \longrightarrow \infty$ heuristically from consistency of GPD parameters, Lemma 4.1 in [25], and asymptotic normality of PWM estimators in [13]. Therefore, since $0 \le G_{\lambda,\beta}(z-u) \le 1$ and $0 \le \widehat{F}(u) \le 1$, we have the result.

Proof of Lemma 4. Let $\epsilon > 0$ and $\delta = \min\{F(q_\alpha^z + \epsilon) - \alpha, \alpha - F(q_\alpha^z - \epsilon)\}$ and note that, for any CDF F defined on $\mathbb{R}$ $F(z) \ge k$, if $z \ge F^{-1}(k)$, then

$$\begin{aligned} P\left(\widehat{q}_\tau^z - q_\tau^z < \varepsilon\right) &= P\left(\widehat{q}_\tau^z > q_\tau^z + \varepsilon\right) \\ &= P\left(\widehat{F}\left(q_\tau^z\right) > \widehat{F}\left(q_\tau^z + \varepsilon\right)\right) \\ &= P\left(\alpha > \widehat{F}\left(q_\tau^z + \varepsilon\right)\right) \\ &= P\left(F\left(q_\tau^z + \varepsilon\right) - \widehat{F}\left(q_\tau^z + \varepsilon\right) > F\left(q_\tau^z + \varepsilon\right) - \alpha\right) \\ &\le P\left(\sup_z |F(z) - \widehat{F}(z)| > \delta\right), \end{aligned} \tag{49}$$

which tends to 0 as $n \longrightarrow \infty$ by Lemma 3.

Proof of Theorem 2. Combining Lemma 1, Lemma 2, and Lemma 4, we have the desired result.

Data Availability

The data used in the article were simulated, and the data generating process (DGP) is included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the Pan-African University of Basic Sciences, Technology and Innovation (PAUSTI) for funding this research.

References

[1] H. D. Bondell, B. J. Reich, and H. Wang, "Noncrossing quantile regression curve estimation," *Biometrika*, vol. 97, no. 4, pp. 825–838, 2010.

[2] V. Chernozhukov, "Extremal quantile regression," *The Annals of Statistics*, vol. 33, no. 2, pp. 806–839, 2005.

[3] V. Chernozhukov, I. Fernández-Val, and A. Galichon, "Quantile and probability curves without crossing," *Econometrica*, vol. 78, no. 3, pp. 1093–1125, 2010.

[4] P. F. Christoffersen, "Evaluating interval forecasts," *International Economic Review*, vol. 39, no. 4, pp. 841–862, 1998.

[5] T. J. Cole, "Fitting smoothed centile curves to reference data," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 151, no. 3, pp. 385–406, 1988.

[6] T. J. Cole and P. J. Green, "Smoothing reference centile curves: the lms method and penalized likelihood," *Statistics in Medicine*, vol. 11, no. 10, pp. 1305–1319, 1992.

[7] W. F. M. De Bondt and R. Thaler, "Does the stock market overreact?" *The Journal of Finance*, vol. 40, no. 3, pp. 793–805, 1985.

[8] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: For Insurance and Finance*, Springer Science & Business Media, Berlin, Germany, 2013.

[9] R. F. Engle and S. Manganelli, "CAViaR," *Journal of Business & Economic Statistics*, vol. 22, no. 4, pp. 367–381, 2004.

[10] J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis, "Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form," *Water Resources Research*, vol. 15, no. 5, pp. 1049–1054, 1979.

[11] X. He, "Quantile curves without crossing," *The American Statistician*, vol. 51, no. 2, pp. 186–192, 1997.

[12] J. B. Hill, "Robust estimation and inference for heavy tailed garch," *Bernoulli*, vol. 21, no. 3, pp. 1629–1669, 2015.

[13] J. R. M. Hosking and J. R. Wallis, "Parameter and quantile estimation for the generalized pareto distribution," *Technometrics*, vol. 29, no. 3, pp. 339–349, 1987.

[14] R. Howard, Alexandrov's theorem on the second derivatives of convex functions via rademacher's theorem on the first derivatives of lipschitz functions. http://people.math.sc.edu/howard/Notes/alex.pdf, 1998.

[15] K. Knight, "Limiting distributions for l1 regression estimators under general conditions," *Annals of Statistics*, pp. 755–770, 1998.

[16] R. Koenker, *Quantile Regression*, Cambridge University Press, Cambridge, UK, 2005.

[17] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[18] K. Kuester, S. Mittnik, and M. S. Paolella, "Value-at-risk prediction: a comparison of alternative strategies," *Journal of Financial Econometrics*, vol. 4, no. 1, pp. 53–89, 2006.

[19] P. Kupiec, "Techniques for verifying the accuracy of risk measurement models," *Journal of Financial Econometrics*, vol. 4, no. 1, pp. 53–89, 1995.

[20] J. M. Landwehr, N. C. Matalas, and J. R. Wallis, "Estimation of parameters and quantiles of wakeby distributions: 1. known lower bounds," *Water Resources Research*, vol. 15, no. 6, pp. 1361–1372, 1979.

[21] A. J. McNeil and R. Frey, "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach," *Journal of Empirical Finance*, vol. 7, no. 3-4, pp. 271–300, 2000.

[22] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools-Revised Edition*, Princeton University Press, Princeton, NJ, USA, 2015.

[23] S. H. Moharram, A. K. Gosain, and P. N. Kapoor, "A comparative study for the estimators of the generalized pareto distribution," *Journal of Hydrology*, vol. 150, no. 1, pp. 169–185, 1993.

[24] D. N. Mutunga, "Estimation of market risk volatility using interquantile autoregressive range," Ph.D. thesis, JOMO Kenyatta University of Agriculture and Technology, Juja, Kenya, 2018.

[25] N. P. Mwita, "Semiparametric estimation of conditional quantiles for time series, with applications in finance," Ph. D thesis, University of Kaiserslautern (Technische Universität Kaiserslautern), Kaiserslautern, Germany, 2003.

[26] M. R. Nieto and E. Ruiz, "Frontiers in var forecasting and backtesting," *International Journal of Forecasting*, vol. 32, no. 2, pp. 475–501, 2016.

[27] A. Noufaily and M. Jones, "Parametric quantile regression based on the gener-alized gamma distribution," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 62, no. 5, pp. 723–740, 2013.

[28] K. Nystrom and J. Skoglund, *Univariate Extreme Value Theory, Garch and Measures of Risk*, 2002.

[29] C. O. Omari, P. N. Mwita, and A. G. Waititu, "Using conditional extreme value theory to estimate value-at-risk for daily currency exchange rates," *Journal of Mathematical Finance*, vol. 7, no. 4, pp. 846–870, 2017.

[30] S. Rydell, *The Use of Extreme Value Theory and Time Series Analysis to Estimate Risk Measures for Extreme Events*, UMEA University, Umeå, Sweden, 2013.

[31] P. A. Santos, "Interval forecasts evaluation: R programs for a new independence test," *Notas e Comunicações CEAUL*, vol. 17, no. 10, 2010.

[32] C. Scarrott and A. MacDonald, "A review of extreme value thresholdes-timation and uncertainty quantification," *REVSTAT–Statistical Journal*, vol. 10, no. 1, pp. 33–60, 2012.

[33] H. Tucker, *A Graduate Course in Probability*, Academic Press, New York, NY, USA, 1967.

[34] J. M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data Mit Press*, p. 108, Cambridge, MA, USA, 2002.