Journal of Statistical and Econometric Methods, vol.8, no.4, 2019, 1-24
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2019

A Three-Step Nonparametric Estimation of Conditional Value-At-Risk Admitting a Location-Scale Model

Emmanuel Torsen[1], Peter N. Mwita[2] and Joseph K. Mung'atu[3]

Abstract

Financial institutions owners and regulators are concerned majorly about risk analysis, Value-at-Risk (VaR) is one of the most popular and common measures of risk used in finance, measures the down-side risk and is determined for a given probability level. In this paper, we consider the problem of estimating conditional Value-at-Risk via the nonparametric method and have proposed a three-step nonparametric estimator for conditional Value-at-Risk. The returns are assumed to have a location-scale model where the function of the error innovations is assumed unknown. The asymptotic properties of the proposed estimator were established, a simulation study was also conducted to confirm the properties. Application to real data was carried out, TOTAL stocks quoted on the Nigerian Stock Exchange using daily closing prices for covering the period between January 02, 2008 to December 29, 2017 trading days was used to illustrate the applicability of the estimator.

[1] Department of Mathematics, Pan African University, Institute of Basic Sciences, Technology, and Innovation, Kenya. E-mail: torsen@mautech.edu.ng
[2] Department of Mathematics, Machakos University, Kenya. E-mail:petermwita@mks.ac.ke
[3] Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Kenya. E-mail: j.mungatu@fsc.jkuat.ac.ke

Article Info: *Received* : April 28, 2019. *Revised* : May 31, 2019.
Published online : July 1, 2019

Keywords: Location-Scale Model; Nonparametric Estimation; Three-Step; Conditional Value-at-Risk

1 Introduction

A major concern for financial institutions owners and regulators is the risk analysis, Value-at-Risk (VaR) is one of the most popular and common measures of risk used in finance [4]. It measures the down-side risk and is determined for a given probability level τ. Typically, in measuring losses, VaR is the lowest value which exceeds this level (the quantile of the loss distributions). [10] noted that the volatility in the underlying financial variable and the exposure to this source of risk are the two main drivers of the losses for a financial institution and VaR is the appropriate method to infer the combined effect of the two factors. For a comprehensive description of VaR and its applications in the field of risk management, see [10].

VaR is a well-established risk management practice to measure the potential loss amount due to market risk employed in the financial industry for both the internal control and regulatory reporting. It is a measure which quantifies and controls the risk of a portfolio. Moreover, in many companies the practice is to manage the market risk with a short-term focus, which means that long-term losses are prevented by avoiding losses from one day to the next. On a strategic level, organizations manage market risk by defining and monitoring risk limits in order to reduce the excessive exposure to losses. Within the framework of risk management, VaR is a key value for controlling and complying with external regulations. It provides the basis for the internal risk controlling models proposed by the Basel Committee on Banking Supervision. In particular, financial institutions with activity in trading risky financial assets are required to maintain internally a minimum level of safe capital to counteract unforeseen risk. The level of this capital can be calculated as a function of VaR. Basel II and III require a ten day holding period and a 99% confidence interval. VaR is a statistical risk measure that indicates how much a financial institution can lose on a financial asset (in terms of market value) with a given probability and over a given time horizon. In other terms, is the quantile of the conditional asset return distribution. The VaR measure has the advantage of being a single estimate, which makes it accessible and easy to understand also by the less numerically literate management. It is now obvious that to a risk manager, a good measure of market risk is more than necessary.

There are several ways of calculating VaR for a financial asset, as an estimator. In practice, the most traditional approaches to VaR computation are the variance-covariance method, historical simulation, Monte Carlo simulation and stresstesting. VaR practically the conditional quantile function concerned with the tail behaviour of the conditional distribution function $F(y|x)$. The approaches for constructing quantile estimates namely: historical simulation, which calculates empirical quantiles from past data; fully parametric models, which describe the entire distribution of returns; EVT uses parametric models only for the tails of the return distribution and quantile regression directly models a specific quantile, and not the whole distribution. One of the most established techniques in estimating the conditional quantile function is the Quantile regression, the seminal work of [11] was a major step forward in estimating conditional quantiles [1]. Mostly in the existing risk management literature, VaR estimation has been focused on parametric models and unconditional distributions. For example, one of the most commonly used parametric method is the RiskMetrics model, due to [2] which assumes that returns of a financial asset follow a multivariate normal distribution (the mean change in the value of each variable is assumed to be zero and the variance is expressed as an exponentially weighted moving average of historical squared returns). But the main criticism to this approach is that it does not capture the fat tails property of financial time series. A semiparametric approach is the conditional autoregressive value-at-risk (CAViaR) model of [5], which estimates VaR directly by quantile regression, but with no assumptions on distribution.

It is challenging to find an adequate estimate for VaR which models and incorporates the special characteristics of financial time series. The returns are independent and identical distributed (iid) which means that the returns are assumed to be uncorrelated over successive time intervals, this is the assumption that is heavily relied up on in modeling VaR. [10] related this assumption with the efficient markets concept, which states that the current price includes all relevant information from the financial market (Efficient Market Hypothesis (EMH)). He states that in this context the prices should be uncorrelated and follow a random walk, as prices would only change as a result of news, which cannot be anticipated. However, in practice a series of statistical properties can be observed for financial returns, such as excess kurtosis (fat-tails), time-varying volatility and volatility clustering, indicated by high autocorrelation of the returns (large changes tend to be followed by large changes and small changes tend to be followed by small changes). Moreover, empirical applications consistently show that nonlinearity and changing volatility are very characteristic to financial time series. For instance, [3] showed that stock returns are serially correlated over long time horizons and [8] consider the changing volatility a stylized fact of stock market, when showing the positive relation between expected market risk premiums and the predictable volatility of stock returns. Hence, there is a necessity to find alternative models for VaR prediction, which are not restricted to the independent and identically distributed (iid) case and do not rely on the assumption that financial returns are normally distributed [1].

Nonparametric modeling takes a step further and addresses part of this challenges by constructing estimates without making assumptions on the form of the financial return distribution and allow for more flexibility and nonlinearity.

The conditional quantile estimation which has recently grown rapidly, originates from the seminal work of [11] who introduced the approach in a parametric regression methodology. However, users may instead need nonparametric estimates. For example, in cases where parametric quantile regression model has been rejected by the data, this has led to a growing literature, [13].

Therefore, the focus of this paper is the problem of estimating Conditional Value-at-Risk which admits a location-scale model with the error term assumed to be unknown.

As [13] noted, that most nonparametric methods are based on kernel methods. Let X and Y be random variables having unknown conditional probability density functions (cpdf), $f(y|x)$ and conditional cumulative distribution function (CCDF), $F(y|x)$ belonging to a smooth class of functions. Then the estimators for $f(y|x)$ and $F(y|x)$ are:

$$\hat{f}(y|x) = \frac{\sum_{i=1}^{n} K_b(X_i, x) K_{by}(Y_i, y)}{\sum_{i=1}^{n} K_b(X_i, x)} \tag{1}$$

and

$$\hat{F}(y|x) = \frac{\sum_{i=1}^{n} K_b(X_i, x) G_{by}(Y_i, y)}{\sum_{i=1}^{n} K_b(X_i, x)} \tag{2}$$

where $K_b(X_i, x)$ and $K_{by}(Y_i, y)$ are density kernels and $G_{by}(Y_i, y)$ is cumulative distribution kernel.

A $\tau - th$ quantile associated with conditional distribution function $F(y|x)$ is defined by $(\tau \in (0, 1))$

$$Q_{Y|X}(\tau|x) = inf\{y : F(y|x) \geq \tau\} = F^{-1}(\tau|x) \tag{3}$$

Or equivalently $F(Q_{Y|X}(\tau|x)|x) = \tau$

We can obtain the estimate of the conditional quantile function $Q_{Y|X}(\tau|x)$ by inverting $\hat{F}(y|x)$ at τ. i.e

$$\hat{Q}_{Y|X}(\tau|x) = inf\{y : \hat{F}(y|x) \geq \tau\} \equiv \hat{F}^{-1}(\tau|x) \tag{4}$$

2 Estimation of Conditional Value-at-Risk

We adopt the definitions of Conditional Value-at-Risk (CVaR) of [12]: Let $\{Y_t\}$ denote a stochastic process representing the returns on a given stock, portfolio or market index, where $t \in \mathbb{Z}$ indexes a discrete measure of time, and $F(y|x)$ denote the conditional distribution of Y_t given $X_t = x$. The vector $X_t \in \mathbb{R}^d$ normally includes lag returns $\{Y_t\}$, $1 \leq l \leq p$, for some $p \in \mathbb{N}$, as well as other relevant conditioning variables that reflect economic or market conditions. Then, for $\tau \in (0, 1)$, $CVaR(X)_\tau$ is defined to be the $\tau-$quantile associated with $F(y|x)$.

Here, the estimation of $CVaR(X)_\tau$ for processes Y_t that admit a location-scale representation given as

$$Y_t = m(X_t) + \sqrt{h(X_t)}\epsilon_t \tag{5}$$

where m and $h > 0$ are unknown functions defined on the range of X_t , ϵ_t is independent of X_t, and ϵ_t is an independent and identically distributed (iid) innovation process with $\mathbb{E}(\epsilon_t) = 0$, $\mathbb{V}ar(\epsilon_t) = 1$ and distribution function F_ϵ, this model was also used by [16].

From equation (5), this means that we can write the CCDF of y in terms of the regression function and the CCDF of error term.

$$CVaR(X)_\tau := Q_{Y|X}(\tau|x) = m(X_t) + \sqrt{h(X_t)}q(\tau) \tag{6}$$

where $Q_{Y|X}(\tau|x)$ is the conditional $\tau-$quantile associated with $F(y|x)$ and $q(\tau)$ is the $\tau-$quantile associated with the error innovation F_ϵ.

The problem of estimating $m(.)$ in equation (6) is the same as Local Linear Regression (LLR), estimating the intercept α. Suppose that the second derivative of $m(.)$ exist in a small neighborhood of x, then

$$m(X) \approx m(x) + m'(x)(X - x) \equiv \alpha + \beta(X - x) \tag{7}$$

Now, let us consider a sample $\{X_i, Y_i\}_{i=1}^n$ and LLR: find α and β to minimize

$$\sum_{i=1}^{n}(Y_i - \alpha - \beta(X_i - x))^2 K_1(\frac{x - X_i}{b_1}) \tag{8}$$

Let $\hat{\alpha}$ and $\hat{\beta}$ be the solution to the Weighted Least Square (WLS) problem in equation (8). Then the estimator of $m(.)$ in (6) is equivalent to $\hat{\alpha}$, see [6] and [16] for more details. Which is the first step in our estimation procedure.

The second step follows, for the estimation of $h(.)$ in equation (6), the estimator $\hat{h}(x)$ is given below, a procedure for estimation of variance proposed by [7]

$$\hat{h}(x) := \hat{\Gamma} \tag{9}$$

where

$$(\hat{\Gamma}, \hat{\Gamma}_1) = \arg min \sum_{i=1}^{n}(r_i - \Gamma - \Gamma_1(X_i - x))^2 K_2(\frac{X_i - x}{b_2}) \tag{10}$$

Now, with the estimator in equation (11), we have the sequence of squared residuals $\{r_i = \{Y_i - \hat{m}(x)\}^2\}_{i=1}^n$. For more on this procedure see [7] and [16].

The estimators of mean and variance functions are then used to get a sequence of Standardized Nonparametric Residuals (SNR) $\{\hat{\epsilon}_i\}_{i=1}^n$, where

$$\hat{\epsilon_i} = \begin{cases} \dfrac{Y_i - \hat{m}(X)}{\sqrt{\hat{h}}(X)}, & \text{if } \hat{h}(X) > 0 \\ 0, & \text{if } \hat{h}(X) \leq 0 \end{cases} \tag{11}$$

In the third step, we use these SNR to obtain the cumulative conditional density estimator of F_ϵ, see our previous paper details [16].

With the estimators of the mean function $m(.)$, the variance function $h(.)$ and the unknown error innovation, the three-step estimator for Conditional Value-at-Risk (CVaR) is given as

$$\widehat{CVaR}(x)_\tau := \hat{Q}_{Y|X}(\tau|x) = \hat{m}(x) + \hat{h}^{1/2}(x)\hat{q}(\tau) \tag{12}$$

We will discussed the asymptotic properties of (12), the estimator of (6); to do this we make the following assumptions in the next subsection.

2.1 Assumptions

A: Bandwidth

1. $b \longrightarrow 0$, as $n \longrightarrow \infty$
2. $nb \longrightarrow \infty$, as $n \longrightarrow \infty$

B: Kernel

1. K has compact support
2. K is symmetric
3. K is Lipschitz continuous
4. K is $\int_{-\infty}^{\infty} K(u)du = 1$ and $\int_{-\infty}^{\infty} uK(u)du = 0$ with $\mu_2(K) = \int_{-\infty}^{\infty} u^2K(u)du$ and $R(K) = \int_{-\infty}^{\infty} K(u)^2du$ being the second moment (Variance) and Roughness of the kernel function respectively.
5. K is bounded and there is $\bar{K} \in \mathbb{R}$, with $K(u) \leq \bar{K} < \infty$ and $K(u) \geq 0, \forall u \in \mathbb{R}$

C: Conditional Distribution function

For fixed $y \in \mathbb{R}$, $\exists$

1. $F''_X(y) = \frac{\partial^2 F_X(y)}{\partial X^2}$ in a neighborhood of x.
2. We further assume that the derivatives of
   $f_X(y) = \frac{\partial F_X(y)}{\partial Y}$
   $F^{(10)}_X(y) = \frac{\partial F_X(y)}{\partial X}$ and
   $F^{(20)}_X(y) = \frac{\partial^2 F_X(y)}{\partial X^2}$ exist in the neighborhood of (x, y)

Definition: ($\alpha - mixing$ **or Strong mixing)** Let $\mathcal{F}_k^l$ be the $\sigma - algebra$ of events generated by $\{Y_t, k \leq t \leq l\}$ for $l > k$. The $\alpha - mixing$ coefficient introduced by Rosenblatt (1956) is defined as

$$\alpha(k) = \sup_{\mathcal{A} \in \mathcal{F}_1^i, \mathcal{B} \in \mathcal{F}_{i+k}^\infty} |\mathbb{P}(AB) - \mathbb{P}(A)\mathbb{P}(B)|.$$

The series is said to be $\alpha - mixing$ if

$$\lim_{k \to \infty} \alpha(k) = 0.$$

The dependence described by the $\alpha - mixing$ is the weakest as it is implied by other types of mixing.

Lemma 1: (Serfling, 1980, [15])

Let F be a distribution function. The function $F^{-1}(\tau)$, $0 < \tau < 1$, is nondecreasing and left continuous, and satisfies

1. $F^{-1}(F(x)) \leq x$, $-\infty < x < \infty$
   and
2. $F(F^{-1}(\tau)) \geq \tau$, $0 < \tau < 1$
   hence
3. $F(x) \geq \tau \iff x \geq F^{-1}(\tau)$

Lemma 2: [9] Let $Y_1, \dots, Y_n$ be independent random variables satisfying $\mathbb{P}(a \leq Y_i \leq b) = 1$, each i, where $a < b$. Then, for $t > 0$,

$$\mathbb{P}\Big(\sum_{i=1}^n Y_i - \sum_{i=1}^n \mathbb{E}[Y_i] \geq nt\Big) \leq \exp\{-\frac{2nt^2}{(a-b)^2}\} \tag{13}$$

Theorem 1: (Borel-Cantelli Lemma, [14])

Let

$$A = \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty A_k$$

be terminal and Let A be the $\limsup A_n$ (infinitely many of the A_n occur), then

1. If $\sum_{i=1}^\infty \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A) = 0$

2. Let $\{A_n\}$ be a sequence of independent events with $\sum_{i=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A) = 1$

Proof, see [14]

Theorem 2: [Convergence in probability sufficiently implies convergence with probability 1 (wp1)]

If $\sum_{i=1}^{\infty} \mathbb{P}\Big(\big|X_n - X\big| > \xi\Big) < \infty, \forall_\xi > 0,$

then

$X_n \xrightarrow{wp1} X.$

Proof Let $\xi > 0$, then

$$\begin{aligned}\mathbb{P}\Big(\big|X_n - X\big| > \xi\Big) &= \mathbb{P}\Big(\bigcap_{m=n}^{\infty} \{\big|X_m - X\big| > \xi\}\Big), for some\ \ m \geq n \\ &\leq \sum_{i=1}^{\infty} \mathbb{P}\Big(\big|X_m - X\big| > \xi\Big) \longrightarrow 0, as\ n \longrightarrow \infty; tail\, of\, a\, convergent\, series\end{aligned} \tag{14}$$

The alternate condition for convergence $wp1$ follows.

3 Asymptotic Properties of the Estimator

In **Theorem 1** of [7], under the assumptions of the aforementioned paper,

$$\sqrt{nb}\big[\hat{h}(x) - h(x) - Bias(\hat{h}(x))\big] \xrightarrow{d} \mathcal{N}\Big(0, f^{-1}(x)h^2(x)\lambda^2(x)\int k^2(u)du\Big) \tag{15}$$

where

$\lambda^2(x) = \mathbb{E}\big[(\epsilon^2 - 1)^2 | X = x\big]$, $\epsilon = \frac{Y-m(X)}{h(X)}$, $\mu_2(k) = \int u^2 K(u)du$ and $Bias(\hat{h}(x)) = \frac{b^2}{2}\mu_2(k)h''^2(X) + o(b_1^2 + b_2^2)$

This means that

$$\hat{h}(x) \xrightarrow{d} h(X) \tag{16}$$

with

$\mathbb{E}(\hat{h}(x)) = h(X) + \frac{b^2}{2}\mu_2(k)h''^2(X) = M_h$ and $Var(\hat{h}(x)) = \frac{1}{nbf(x)}R(k)h^2(x)\lambda^2(x) = V_h$

Hence

$$\hat{h}(x) \sim \mathcal{N}\big(M_h, V_h\big) \tag{17}$$

Also in [6], we have that

$$\hat{m}(x) \xrightarrow{d} \mathcal{N}\big(M_{\hat{m}}, V_{\hat{m}}\big) \tag{18}$$

where
$M_{\hat{m}} = m(x) + \frac{b^2}{2} m''(x) \int_{-\infty}^{\infty} u^2 K(u) du = m(x) + \frac{b^2}{2} m''(x) \mu_2(k)$,
$V_{\hat{m}} = \frac{\sigma^2(x)}{nbf(x)} \int_{-\infty}^{\infty} K^2(u) du = \frac{\sigma^2(x) R(k)}{nbf(x)}$, and $Bias(\hat{m}(x)) = \frac{b^2}{2} m''(x) \mu_2(k)$

We want to show the mean and Variance of our estimator using Slutsky's theorem

$$G \xrightarrow{d} h(X) q(\tau)$$

where $G = \hat{h}(x)\hat{q}(\tau)$ and $\hat{q}(\tau) > 0$

Now

$$\begin{aligned} \mathbb{E}[G] &= \mathbb{E}\big[\hat{h}(x)\hat{q}(\tau)\big] \\ &= \hat{q}(\tau)\mathbb{E}\big[\hat{h}(x)\big] \\ &= \hat{q}(\tau)\Big[h(X) + \frac{b^2}{2}\mu_2(k) h''^2(X)\Big] \\ &= h(X)\hat{q}(\tau) + \frac{b^2 \hat{q}(\tau)}{2}\mu_2(k) h''^2(X) \end{aligned} \tag{19}$$

Therefore,

$$\mathbb{E}[G] \approx h(X) q(\tau) + \underbrace{\frac{b^2 q(\tau)}{2}\mu_2(k) h''^2(X)}_{=\,0,\ \text{Assumption A(1)}} \tag{20}$$

and

$$\begin{aligned} Var[G] &= Var\big[\hat{h}(x)\hat{q}(\tau)\big] \\ &= \hat{q}(\tau)^2 Var\big[\hat{h}(x)\big] \\ &= \hat{q}(\tau)^2 \Big[\frac{1}{nbf(x)} R(k) h^2(x) \lambda^2(x)\Big] \\ &= \frac{\hat{q}(\tau)^2 R(k) h^2(x) \lambda^2(x)}{nbf(x)} \end{aligned} \tag{21}$$

$$Var[G] \approx \underbrace{\frac{q(\tau)^2 R(k) h^2(x) \lambda^2(x)}{nbf(x)}}_{=\,0,\ \text{Assumption A(1)}} \tag{22}$$

Hence by Assumption A(1) and Slutsky's theorem,

$$\hat{h}(x)\hat{q}(\tau) \xrightarrow{d} h(X)q(\tau) \tag{23}$$

with mean and variance as given above.

Hence,

$$\begin{aligned}\hat{Q} &= \hat{m}(x) + \hat{h}(x)\hat{q}(\tau)\\ &= \hat{m}(x) + G\\ &= H + G\end{aligned} \tag{24}$$

where $\hat{Q} = \hat{Q}_{Y|X}(\tau/x)$

Now,

$$\begin{aligned}\mathbb{E}[\hat{Q}] &= \mathbb{E}[H+G]\\ &= \mathbb{E}[H] + \mathbb{E}[G]\\ &= \Big[m(x) + \frac{b^2}{2}m''(x)\mu_2(k)\Big] + \hat{q}(\tau)\Big[h(X) + \frac{b^2}{2}\mu_2(k)h''^2(X)\Big]\\ &= m(x) + \frac{b^2}{2}m''(x)\mu_2(k) + \hat{q}(\tau)h(X) + \frac{b^2\hat{q}(\tau)}{2}\mu_2(k)h''^2(X)\\ &= m(x) + h(X)\hat{q}(\tau) + \frac{b^2}{2}\mu_2(k)\big[m''(x) + h''^2(X)\hat{q}(\tau)\big]\\ &\approx m(x) + h(X)q(\tau) + \underbrace{\frac{b^2}{2}\mu_2(k)\big[m''(x) + h''^2(X)q(\tau)\big]}_{=\text{Bias}}\end{aligned} \tag{25}$$

So that,

$$Bias(\hat{Q}) \approx \frac{b^2}{2}\mu_2(k)\big[m''(x) + h''^2(X)q(\tau)\big] \tag{26}$$

and

$$\begin{aligned}Var(\hat{Q}) &= Var(H+G)\\ &= Var(H) + Var(G) + 2Cov(H,G), \quad Cov(H,G) = 0\\ &= \frac{\sigma^2(x)R(k)}{nbf(x)} + \frac{\hat{q}(\tau)^2R(k)h^2(x)\lambda^2(x)}{nbf(x)}\\ &= \frac{R(k)}{nbf(x)}\Big[\sigma^2(x) + \hat{q}(\tau)^2h^2(x)\lambda^2(x)\Big]\\ &\approx \frac{R(k)}{nbf(x)}\Big[\sigma^2(x) + q(\tau)^2h^2(x)\lambda^2(x)\Big]\end{aligned} \tag{27}$$

$\Longrightarrow\ \hat{Q} \xrightarrow{d} Q$, with mean and variance given above (using Slutsky's theorem and that by assumption A(1), (26) and (27) tends to 0), where

$$\widehat{CVaR}(x)_\tau := \hat{Q}_{Y|X}(\tau|x) = \hat{m}(x) + \hat{h}^{1/2}(x)\hat{q}(\tau) \tag{28}$$

Hence, (28) which is also (12), is the estimator for $CVaR(X)_\tau$.

3.1 Consistency

For simplicity of notation defined;

$$\hat{Q} := \hat{Q}_{Y|X}(\tau|x) = \widehat{CVaR}(x)_\tau \tag{29}$$

Next, consider the following theorem with Assumption B holding

Theorem 3

Let for $\xi > 0$ and $0 < \tau < 1$, $F_x(Q-\xi) \leq \tau \leq F_x(Q+\xi)$, then

$$\hat{Q} \xrightarrow{wp1} Q \tag{30}$$

Proof: Let $\xi > 0$. By definition and uniqueness condition, we have that

$$F_x(Q-\xi) < \tau < F_x(Q+\xi)$$

It can be easily verified that

$$F_n(Q-\xi) \xrightarrow{wp1} F_x(Q-\xi)$$

and

$$F_n(Q+\xi) \xrightarrow{wp1} F_x(Q+\xi)$$

this is convergence with probability 1 ($wp1$).

Hence,

$$\mathbb{P}\Big(F_m(Q-\xi) < \tau < F_m(Q+\xi)\Big) \longrightarrow 1\ \forall_{m\geq n}\ \ as\ \ n \longrightarrow \infty$$

Now, using Lemma 1(3), we have that

$$\mathbb{P}\Big((Q-\xi) < \hat{Q}_m < (Q+\xi)\Big) \longrightarrow 1\ \forall_{m\geq n}\ \ as\ \ n \longrightarrow \infty$$

Thus

$$\mathbb{P}\Big(\sup_{m\geq n}\Big|\hat{Q}_m - Q\Big| > \xi\Big) \longrightarrow 0\ \forall_{m\geq n}\ \ as\ \ n \longrightarrow \infty \tag{31}$$

Hence, $\hat{Q}$ is strongly consistent for the estimation of Q.

Theorem 4

Let the conditions of Theorem 1 be satisfied and let for $\xi > 0$ and $0 < \tau < 1$, $F_x(Q-\xi) < \tau < F_x(Q+\xi)$ and also assume that $nb^2 \longrightarrow \infty$, as $n \longrightarrow \infty$ then

$$\hat{Q} \xrightarrow{p} Q \tag{32}$$

also, if we let

$$\sum_{i=1}^{n} e^{-2\xi nb^2} < \infty \quad hold$$

then,

$$\hat{Q} \xrightarrow{a.e} Q \tag{33}$$

Proof:

Let $\xi > 0$, then it holds that

$$\begin{aligned}\mathbb{P}\Big(\Big|\hat{Q} - Q\Big| > \xi\Big) &= \mathbb{P}\Big(\hat{Q} > Q + \xi\Big) + \mathbb{P}\Big(\hat{Q} < Q - \xi\Big) \\ &= T_1 + T_2\end{aligned}$$

We used Lemma 1 to evaluate T_1 and T_2, and hence applied Lemma 2 on the results:

$$\begin{aligned}T_1 &= \mathbb{P}\Big(\hat{Q} > Q + \xi\Big) \\ &= \mathbb{P}\Big(\tau > F_{n,x}(Q+\xi)\Big) \\ &= \mathbb{P}\Big(1 - \tau > 1 - F_{n,x}(Q+\xi)\Big) \\ &= \mathbb{P}\left(\frac{\sum_{i=1}^{n} 1(Y_i < Q+\xi) K\Big(\frac{x-X_i}{b}\Big)}{\sum_{i=1}^{n} K\Big(\frac{x-X_i}{b}\Big)} - \frac{\sum_{i=1}^{n} \mathbb{E}\Big[1(Y_i < Q+\xi)\Big] K\Big(\frac{x-X_i}{b}\Big)}{\sum_{i=1}^{n} K\Big(\frac{x-X_i}{b}\Big)} > \right. \\ &\qquad \left. 1 - \tau - \Big(1 - F_{n,x}(Q+\xi) - o(b)\Big)\right) \\ &= \mathbb{P}\Big(\sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \mathbb{E}[V_i] > n\Big[\frac{1}{n}\Big(F_{n,x}(Q+\xi)\Big) - \tau + o(b)\Big]\Big) \\ &= \mathbb{P}\Big(\sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \mathbb{E}[V_i] > \eta_1\Big)\end{aligned}$$

where $\eta_1 = (F_{n,x}(Q + \xi) - \tau + o(b)$

and $V_i = \dfrac{\sum_{i=1}^{n} 1(Y_i < Q + \xi) K\Big(\dfrac{x - X_i}{b}\Big)}{\sum_{i=1}^{n} K\Big(\dfrac{x - X_i}{b}\Big)}$

$\implies T_1 = \mathbb{P}\Big(\hat{Q} > V_\tau + \xi\Big) \leq e^{-2nb^2\eta_1^2}$

Similarly,

$$\begin{aligned} T_2 &= \mathbb{P}\Big(\hat{Q} < Q - \xi\Big) \\ &= \mathbb{P}\Big(\tau < F_{n,x}(Q - \xi)\Big) \\ &= \mathbb{P}\Big(\sum_{i=1}^{n} w_i - \sum_{i=1}^{n} \mathbb{E}[w_i] > \eta_2\Big) \end{aligned}$$

$\implies T_2 = \mathbb{P}\Big(\hat{Q} < Q - \xi\Big) \leq e^{-2nb^2\eta_2^2}$, as $n \longrightarrow \infty$

Therefore, putting $\eta_3 = min\{\eta_1, \eta_2\}$ we have that

$$\mathbb{P}\Big(\Big|\hat{Q} - Q\Big| > \xi\Big) \leq e^{-2nb^2\eta_3^2}$$

which completes the proof.

Hence, $\hat{Q}\Big(= \hat{Q}_{Y|X}(\tau/x)\Big)$ converges in probability to $Q\Big(= Q_{Y|X}(\tau/x)\Big)$.

Thus,

$$\mathbb{P}\Big(\Big|\hat{Q} - Q\Big| > \xi\Big) \longrightarrow 0$$

exponentially fast, which implies that $\hat{Q}$ converges completely to Q (Theorem 2).

Furthermore, using the second part of Theorem 1; $\sum_{n=m}^{\infty} e^{-2\gamma m b^2}$, for $\gamma > 0$ as $n \longrightarrow \infty$, we see that $\hat{Q}$ converges almost everywhere to Q.

i.e

$$\hat{Q} \xrightarrow{a.e} Q$$

Then by central limit theorem, the asymptotic normal distribution for the three-step LLR in (12) is given as

$$\sqrt{nb}\Big(\hat{Q} - Q - Bias(\hat{Q})\Big) \xrightarrow{d} \mathcal{N}\Big(0, Var(\hat{Q})\Big) \tag{34}$$

with mean, bias, and variance given in (25), (26), and (27) respectively.

The nonparametric prediction intervals for the proposed estimator (28) was constructed, it was found to perform very well, see [17] for more details.

4 Smoothing Parameter (Bandwidth) Selection

In Nonparametric methods, the choice of an optimal smoothing parameter can not be over emphasized. We choose the smoothing parameter that minimizes the Asymptotic Mean Square Error (AMSE) below:

$$\begin{aligned}
AMSE(\hat{Q}) &= \mathbb{E}\big[(\hat{Q}-Q)^2\big] \\
&= \mathbb{E}\Big[\Big(\hat{Q}-\mathbb{E}[\hat{Q}]+Bias(\hat{Q})\Big)^2\Big] \\
&= \mathbb{E}\Big[\Big(\hat{Q}-\mathbb{E}[\hat{Q}]\Big)^2\Big]+Bias(\hat{Q})\times\mathbb{E}\Big[\hat{Q}-\mathbb{E}[\hat{Q}]\Big]+Bias^2(\hat{Q}) \\
&= Var(\hat{Q})+Bias^2(\hat{Q}) \\
&= \frac{R(k)}{nbf(x)}\Big[\sigma^2(x)+\hat{q}(\tau)^2h^2(x)\lambda^2(x)\Big]+\left[\frac{b^2}{2}\mu_2(k)\big[m''(x)+\hat{q}(\tau)h''^2(x)\big]\right]^2 \\
&= \frac{R(k)}{nbf(x)}\Big[\sigma^2(x)+\hat{q}(\tau)^2h^2(x)\lambda^2(x)\Big]+\frac{b^4}{4}\mu_2^2(k)\Big[m''(x)+\hat{q}(\tau)h''^2(x)\Big]^2 \\
&\approx \frac{R(k)}{nbf(x)}\Big[\sigma^2(x)+q(\tau)^2h^2(x)\lambda^2(x)\Big]+\frac{b^4}{4}\mu_2^2(k)\Big[m''(x)+q(\tau)h''^2(x)\Big]^2
\end{aligned} \tag{35}$$

Therefore,

$$b_{opt}=\underset{b>0}{argmin}AMSE(\hat{Q})$$

and hence,

$$\frac{d}{db}AMSE(\hat{Q})=0$$

which gives

$$b_{opt}=\left\{\frac{R(k)\Big[\sigma^2(x)+q(\tau)^2h^2(x)\lambda^2(x)\Big]}{\mu_2^2(k)f(x)\Big[m''(x)+q(\tau)h''^2(X)\Big]^2}\right\}^{\frac{1}{5}}\times n^{-\frac{1}{5}} \tag{36}$$

5 Results

5.1 Simulation Study

We conducted a simulation study to examine the performance of our estimator in (12) and (28), considering the following data generating location-scale model

$$Y_t = m(Y_{t-1}) + h(t)^{1/2}\epsilon_t, \quad t = 1, 2, ..., n \tag{37}$$

where $m(Y_{t-1}) = sin(0.5Y_{t-1}),\quad \epsilon_t \sim t(\nu = 3),\ h(t) = h_i(Y_{t-1}) + \theta h(Y_{t-1})$, $i = 1, 2$ and $h_1(Y_{t-1}) = 1 + 0.01Y_{t-1}^2 + 0.5sin(Y_{t-1})$, $h_1(Y_{t-1}) = 1 - 0.9exp\big(-2Y_{t-1}^2\big)$

Y_t and $h(t)$ are set to zero (0) initially, then Y_t is generated recursively from (37) above. To reduce the effect of the choice of our initial values on the samples, the first 1000 observations are discarded, the above data generating process was also considered by [12] and used by [16].

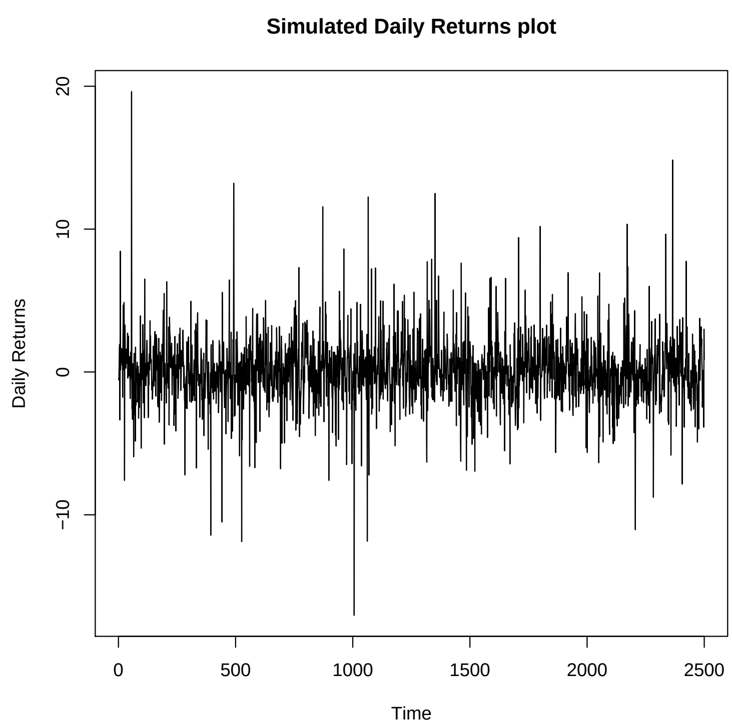

Figure 1: Plot of the simulated returns showing its evolution.

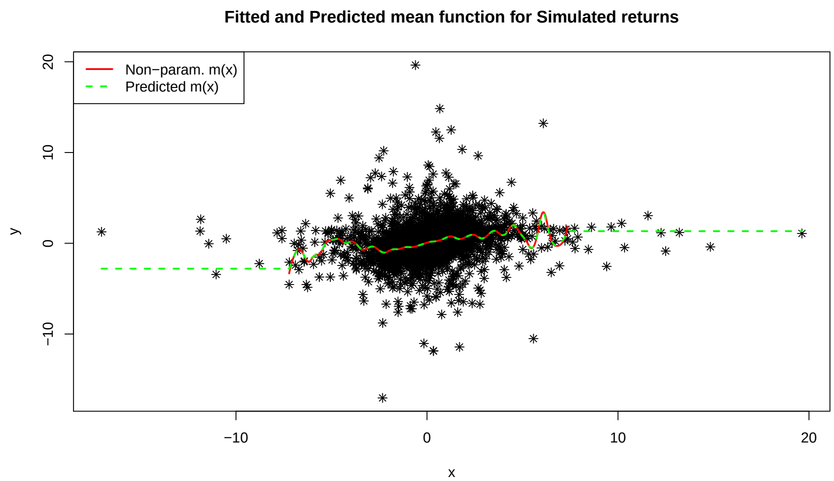

Figure 2: Plot of the mean $m(.)$ function and the predicted mean function of the simulated returns.

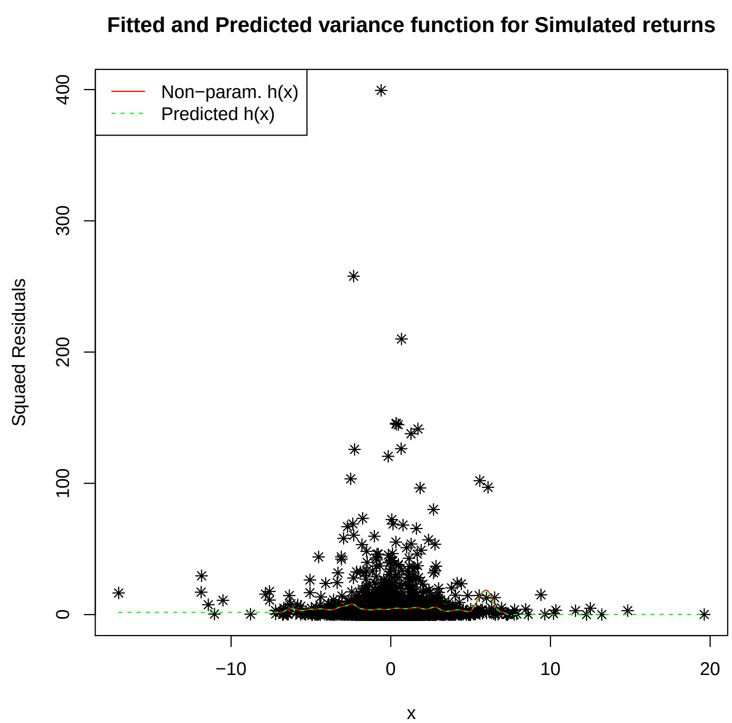

Figure 3: Graph of the variance $h(.)$ function and the predicted mean function of the simulated returns.

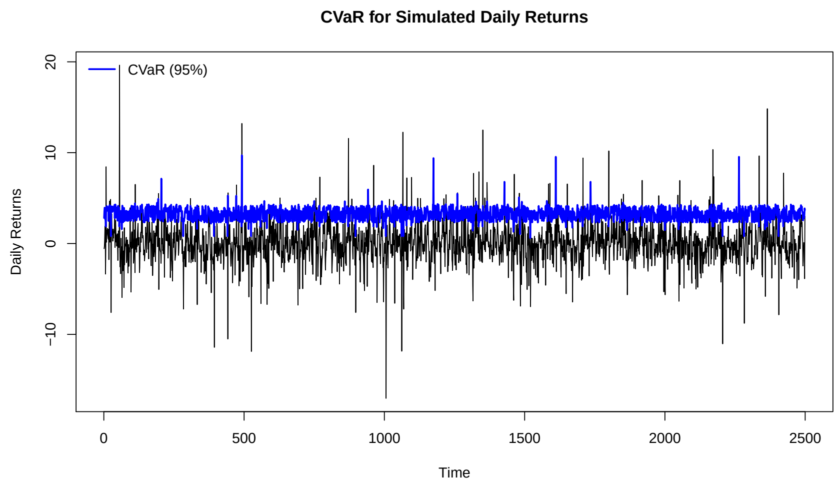

Figure 4: Graph showing the 95% Conditional Value-at-Risk for the simulated returns.

5.2 Application to financial market data

To see how the proposed three-step nonparametric estimate for CVaR performs on a real data set, the closing prices for the period between January 02, 2008 to December 29, 2017 trading days of TOTAL company quotd on the Nigerian Stock Exchange was used, which gave 2 475 observations. Thel daily closing prices series $\{P_t\}$ where used to obtain the log returns as defined below

$$y_t = \ln P_t - \ln P_{t-1} = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{38}$$

where $\{P_t\}$ is the the asset closing price on day t and ln is the natural logarithmic function.

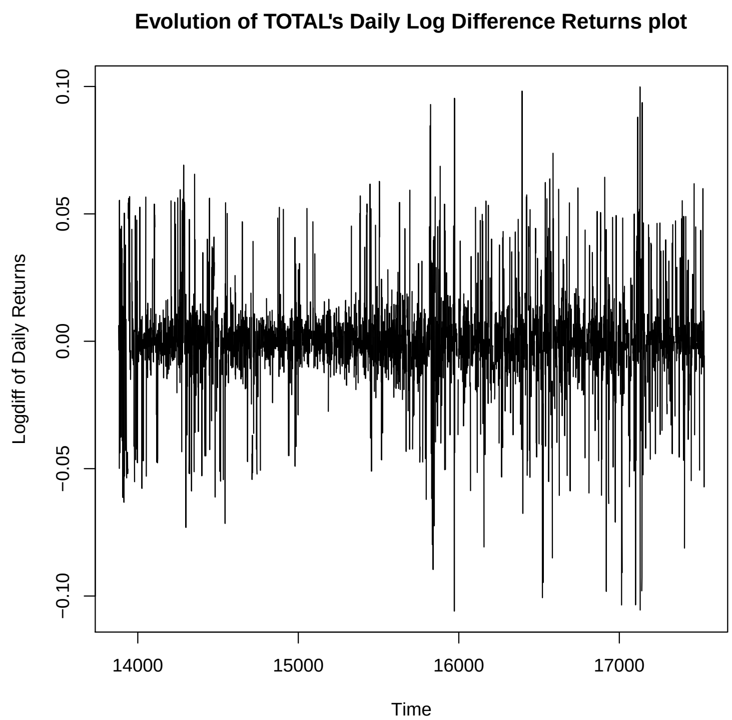

Figure 5: Graph showing the evolution of TOTAL's daily log difference of returns.

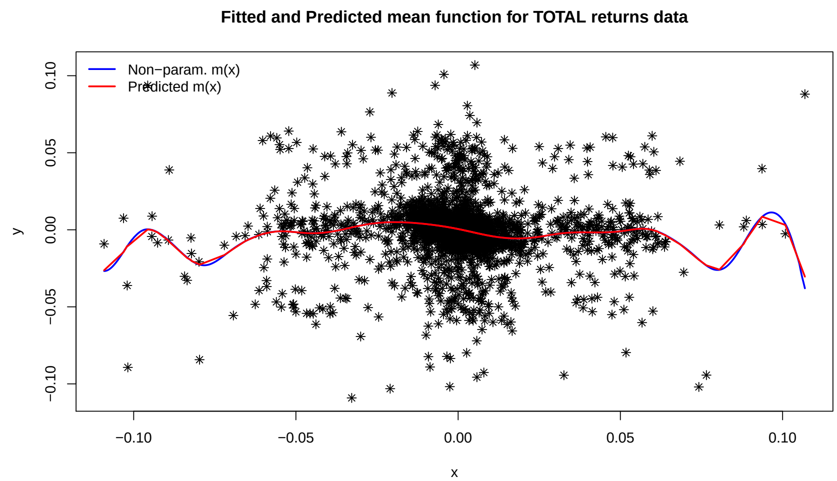

Figure 6: Graph showing the nonparametric mean $m(.)$ function and the predicted nonparametric mean $m(.)$ function of TOTAL's daily log difference of returns.

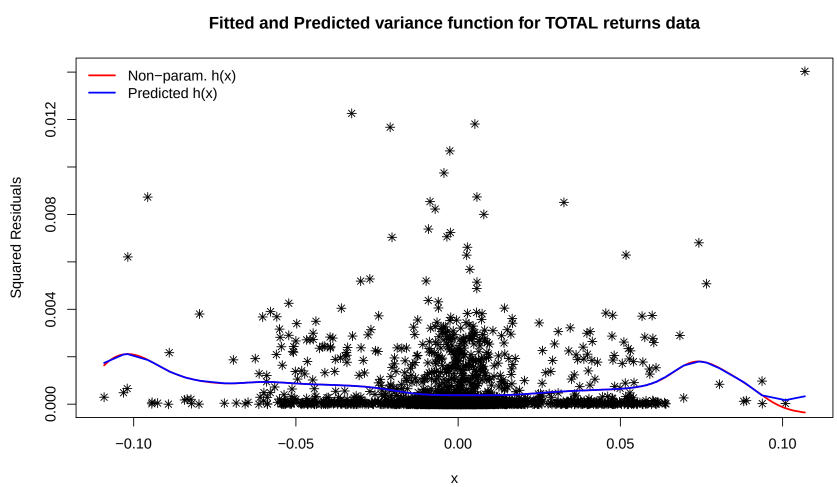

Figure 7: Graph showing the nonparametric variance $h(.)$ function and the predicted nonparametric variance $h(.)$ function of TOTAL's daily log difference of returns.

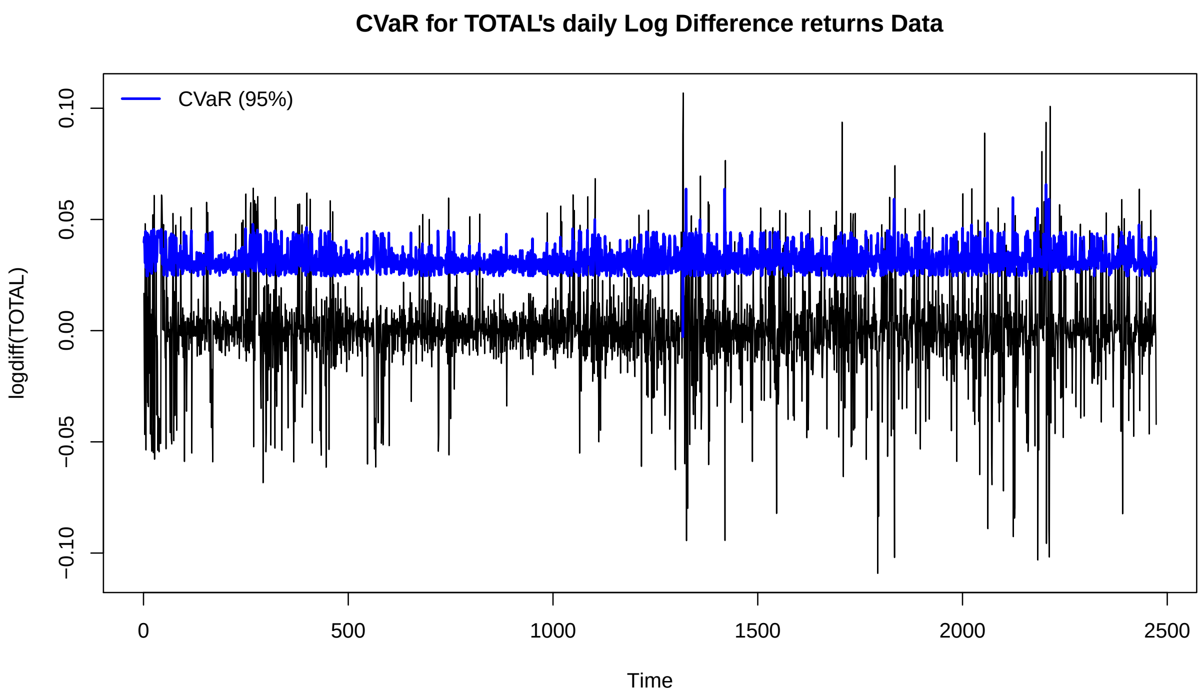

Figure 8: Graph showing the 95% Conditional Value-at-Risk for TOTAL's daily log difference of returns.

Table 1: Conditional Value-at-Risk Results at different quantiles

Tau	**q_tau**	**CVaR**	**se_CVaR**
0.9	0.8889	0.0189	0.0131
0.91	1.0858	0.0231	0.0131
0.92	1.3023	0.0277	0.0131
0.93	1.516	0.0323	0.0131
0.94	1.7187	0.0366	0.0131
0.95	1.8593	0.0396	0.0131
0.96	2.0915	0.0446	0.0131
0.97	2.2931	0.0489	0.0131
0.98	2.492	0.0531	0.0131
0.99	2.7399	0.0584	0.0131

6 Discussion

We have presented in Figure 1 the evolution of the simulated returns, characterized by high volatility which is known of financial time series dat. Using the same simulated returns, in Figures 2 and 3, we presented the smoothed estimation of the mean function $m(.)$ and the variance $h(.)$ function by LLR as proposed by [6] and [7] respectively. Hence, the 95% Conditional Value-at-Risk is shown in Figure 4 for the daily simulated returns data. The three-step nonparametric estimator for CVaR was applied to real data from the Nigerian Stock Exchange; Figure 5 showed the time series plot of the TOTAL's daily returns for the period under consideration, in Figures 6 and 7 we have shown graphically the smooth estimation of the mean function $m(.)$ and the variance $h(.)$ function using LLR as mentioned earlier on, and Figure 8 presented the 95% Conditional Value-at-Risk for the TOTAL data. The three-step estimator is consistent as given in Theorem 3, converges almost everywhere to the true as shown in Theorem 4, and is asymptotically normal.

The results of Conditional Value-at-Risk at different quantiles can be seen as shown in Table 1 above, it can be observed that CVaR increases as the quantile increased while their standard error remained unchanged.

7 Conclusions

A Three-Step Nonparametric Estimator for Conditional Value-at-Risk is proposed in this paper, we assumed that the returns on a portfolio or investment admit the location-scale model with heteroscedasticity. The distribution of the innovation was considered unknown. We examined the asymptotic properties of the estimator, the three-step estimator is consistent, converges almost everywhere to the true, and is asymptotically normal. A Simulation study was conducted and application to real data was also done.

Acknowledgements

The first author wishes to thank African Union, Pan African University, Institute for Basic Sciences Technology and Innovation Nairobi, Kenya and Japan International Cooperation (JICA) for supporting this research.

Disclosure statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] I. Balcau, Nonparametric estimate for conditional quantiles of time series, *Master's thesis, Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät*, (2012).

[2] J. M. Bank, TRiskmetrics technical manual, *Biometrika,JP Morgan Bank. New York* (1995).

[3] W. F. Bondt and R. Thaler, Does the stock market overreact?, *The Journal of finance*, *40*(3), (1985), 793–805.

[4] P. Embrechts, S. I. Resnick and G. Samorodnitsky, Extreme value theory as a risk management tool, *North American Actuarial Journal*, **3**(2), (1999), 30–41.

[5] R. F. Engle and S. Manganelli, CAViaR: Conditional autoregressive value at risk by regression quantiles, *Journal of Business and Economic Statistics*, **22**(4), (2004), 367–381.

[6] J. Fan, Design-adaptive nonparametric regression, *Journal of the American statistical Association*, **87**(420), (1992), 998–1004.

[7] J. Fan and Q. Yao, Efficient estimation of conditional variance functions in stochastic regression, *Biometrika*, **85**(3), (1998), 645–660.

[8] K. R. French, G. W. Schwert and R. F.Stambaugh, Expected stock returns and volatility, *Journal of financial Economics*, **19**(1), (1987), 3–29.

[9] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American statistical association*, **58**(301), (1963), 13–30.

[10] P.Jorion, Value at risk: The new benchmark for managing market risk, *McGraw Hill, 2nd.*, (2000).

[11] R. Koenker and G. Bassett Jr, Regression quantiles, *Econometrica: journal of the Econometric Society*, (2008), *pages*, 33–50.

[12] C. Martins-Filho, F. Yao and M. Torero, Nonparametric estimation of conditional value-at-risk and expected shortfall based on extreme value theory, *Econometric Theory*, **34**(1), (2018), 23-67.

[13] J. S. Racine and K. Li, Nonparametric conditional quantile estimation: A locally weighted quantile kernel approach, *Journal of Econometrics*, **201**(1), (2017), 72–94.

[14] V. K. Rohatgi and A. M. Ehsanes Saleh, An Introduction to Probability and Statistics, *2nd ed; Wiley Series in Probability and Statistics: New York, NY, USA: Wiley*, (2001).

[15] R. J. Serfling, Approximation theorems of mathematical statistics, *volume 162. John Wiley & Sons.*, (2009).

[16] E. Torsen, P. N. Mwita and J. K. Mung'atu, Nonparametric estimation of the error functional of a location-scale model, *Journal of Statistical and Econometric Methods*, **7**(4), (2018), 1–18.

[17] E. Torsen and L. L. Seknewna, Bootstrapping nonparametric prediction intervals for conditional value-at-risk with heteroscedasticity, *Journal of Probability and Statistics, Hindawi*, **2019**, (2019).