Smoothed Conditional Scale Function Estimation in AR(1)-ARCH(1) Processes

Lema Logamou Seknewna, Peter Mwita Nyamuhanga, and Benjamin Kyalo Muema

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

Correspondence should be addressed to Lema Logamou Seknewna; logamou.lema@students.jkuat.ac.ke

Received 27 November 2017; Revised 11 January 2018; Accepted 24 January 2018; Published 12 March 2018

Abstract

The estimation of the Smoothed Conditional Scale Function for time series was taken out under the conditional heteroscedastic innovations by imitating the kernel smoothing in nonparametric QAR-QARCH scheme. The estimation was taken out based on the quantile regression methodology proposed by Koenker and Bassett. And the proof of the asymptotic properties of the Conditional Scale Function estimator for this type of process was given and its consistency was shown.

1. Introduction

Consider a Quantile Autoregressive model,

\[ X_t = \alpha_\tau(Z_t) + u_t, \quad t = 1, 2, \ldots, \]  

(1)

where \( \alpha_\tau(Z_t) \) is the \( \tau \)th Conditional Quantile Function of \( X_t \) given \( Z_t \) and the innovation \( u_t \) is assumed to be independent and identically distributed with zero \( \tau \)th quantile and constant scale function; see [1]. A kernel estimator of \( \alpha_\tau(Z_t) \) has been determined and its consistency is shown [2]. A bootstrap kernel estimator of \( \alpha_\tau(Z_t) \) was determined and shown to be consistent [3]. This research will extend [3] by assuming that the innovations follow Quantile Autoregressive Conditional Heteroscedastic process similar to Autoregressive-Quantile Autoregressive Conditional Heteroscedastic process proposed in [1]:

\[ X_t = \alpha_\tau(Z_t) + \omega_\tau(Z_t) \epsilon_t, \quad t = 1, 2, \ldots, \]  

(2)

where \( \alpha_\tau(Z_t) \) is the conditional \( \theta \)-quantile function of \( X_t \) given \( Z_t \); \( \omega_\tau(Z_t) \) is a conditional scale function at \( \tau \)-level, and \( \epsilon_t \) is independent and identically distributed (i.i.d.) error with zero \( \tau \)-quantile and unit scale. The function \( \omega_\tau(Z_t) \) can be expressed as

\[ \omega_\tau(Z_t) = \lambda \omega(Z_t), \]  

(3)

where \( \omega(Z_t) \) is the so-called volatility found in [4, 5] which are papers of reference on Engle’s ARCH models among many others and \( \lambda \) is a positive constant depending on \( \tau \) [see [6]]. An example of this kind of function is Autoregressive-Generalized Autoregressive Conditional Heteroscedastic AR(1)-GARCH(1,1),

\[ X_t = \alpha_t + \omega_t \epsilon_t, \quad t = 1, 2, \ldots, \]  

(4)

where \( \omega_t \) is the so-called volatility found in [4, 5] which are papers of reference on Engle’s ARCH models among many others and \( \lambda \) is a positive constant depending on \( \tau \) [see [6]]. An example of this kind of function is Autoregressive-Generalized Autoregressive Conditional Heteroscedastic AR(1)-GARCH(1,1),

\[ X_t = \alpha_t + \omega_t \epsilon_t, \quad t = 1, 2, \ldots, \]  

where \( \alpha_t = \mu + \delta X_{t-1}, \omega_t = \sqrt{\lambda + \alpha X_{t-1}^2 + \beta \omega_{t-1}^2}, \mu \in (-\infty, \infty), \delta < 1, \beta > 0, \alpha > 0, \omega > 0, \alpha + \beta < 1, \text{ and } \epsilon_t \sim \text{i.i.d. with 0 mean and variance 1. Note that } \alpha_t \text{ may also be an ARMA (see [7]). The specifications for model (4) are given in Section 4.2.} \]
Considering other financial time series models, the model (1) can be seen as a robust generalization of AR-ARCH-models, introduced in [7], and their nonparametric generalizations reviewed by [8]. For instance, consider a financial time series model of AR(\(p\))-ARCH(\(q\))-type, and show its consistency using standard estimate of Nadaraya-Watson [10] type. This estimate is obtained from the estimate of the conditional scale function (KDE), bias reduction [13], asymptotic properties, and normality proof) are necessary (see Table 1).

**Assumption 1.** 
(i) \(f(x, z)\) exist.
(ii) For fixed \((x, z)\), \(0 < f(x \mid z) < 1\) and \(f(z) > 0\) are continuous in the neighborhood of \(z\) where the estimator is to be estimated.
(iii) The derivatives \(F^{(j)}(x) = d^j F(x \mid z) / dz^j\) and \(f^{(j)}(z) = df(x \mid z) / dz\) for \(j = 1, 2\) exist.
(iv) \(F(x \mid z)\) is a convex function in \(x\) for fixed \(z\).
(v) The conditional density \(f(x \mid z) = dF(x \mid z) / dx\) exists and is continuous in the neighborhood of \(x\).
(vi) \(f(\bar{\omega}(z) \mid z) > 0\).

**Assumption 2.** The kernel function \(K : \mathbb{R}^d \rightarrow \mathbb{R}\) is
(i) Symmetrical: \(K(s) = K(-s)\) with \(s \in \mathbb{R}^d\)
(ii) Nonnegative and bounded: for \(\Gamma < \infty, 0 < K(s) \leq \Gamma, s \in \mathbb{R}^d\)

<table>
<thead>
<tr>
<th>Kernel functions</th>
<th>Expressions (K(u))</th>
<th>(r)</th>
<th>(R(K))</th>
<th>(\mu_2(K))</th>
<th>(\text{Eff}(K))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian</td>
<td>(\frac{1}{\sqrt{2}} \exp \left(-\frac{u^2}{2}\right) I_R)</td>
<td>(\infty)</td>
<td>(\frac{1}{2\sqrt{2}})</td>
<td>1</td>
<td>0.2821</td>
</tr>
<tr>
<td>Epanechnikov</td>
<td>(\frac{3}{4} \left(1 - u^2\right) I_{{</td>
<td>u</td>
<td>\leq1}})</td>
<td>2</td>
<td>(\frac{3}{5})</td>
</tr>
<tr>
<td>Uniform</td>
<td>(\frac{1}{2} I_{{</td>
<td>u</td>
<td>\leq1}})</td>
<td>0</td>
<td>(\frac{1}{2})</td>
</tr>
<tr>
<td>Triangular</td>
<td>(\left(1 -</td>
<td>u</td>
<td>\right) I_{{</td>
<td>u</td>
<td>\leq1}})</td>
</tr>
<tr>
<td>Tricube</td>
<td>(\frac{35}{32} \left(1 - u^2\right)^3 I_{{</td>
<td>u</td>
<td>\leq1}})</td>
<td>6</td>
<td>(\frac{2}{3})</td>
</tr>
<tr>
<td>Biweight</td>
<td>(\frac{70}{81} \left(1 -</td>
<td>u</td>
<td>\right)^3 I_{{</td>
<td>u</td>
<td>\leq1}})</td>
</tr>
<tr>
<td>Cosine</td>
<td>(\frac{\pi}{4} \cos \left(\frac{\pi}{2} u\right))</td>
<td>(\infty)</td>
<td>(\frac{\pi^2}{16})</td>
<td>(-8 + \pi^2)</td>
<td>0.2685</td>
</tr>
</tbody>
</table>

**2. Methods and Estimations**

Let \(f(z)\) and \(f(x, z)\) denote the probability density function (pdf) of \(X_t\) and the joint pdf of \((X_t, Z_t)\). The dependence between the exogenous \(X_t\) and the endogenous variables is described by the following conditional probability density function (CPDF):

\[
f(x \mid z) = \frac{f(x, z)}{f(x)} \tag{6}
\]

and the conditional cumulative distribution function (CCDF)

\[
F(x \mid z) = \int_{-\infty}^{x} f(s \mid z) ds = P(X \leq x \mid Z_t = z) \tag{7}
\]

The estimation of the conditional scale function is derived through the CCDF. However, the following assumptions and definitions (these assumptions are commonly used for kernel density estimation (KDE), bias reduction [13], asymptotic properties, and normality proof) are necessary (see Table 1).
(iii) Lipschitz: \( \exists \lambda > 0, m_k < \infty \) such that \( |K(s) - K(t)| \leq m_k |s - t| \) for all \( s, t \in \mathbb{R}^d \).

(iv) A pdf: \( \int K(s)ds = 1 \) with \( \int_{\mathbb{R}^d} sK(s) = 0 \).

Assumption 3. The process \( \{X_t, Z_t\}, \ t = 1, 2, \ldots \) is strong mixing with \( \alpha(s) = o(s^{2-\delta}), \delta > 0 \); see [14, Theorem 1.7].

Assumption 4. The sequence \( \{b_n\}_{n \in \mathbb{N}} \) of the smoothing parameters is such that \( b\to 0, nb^n \to \infty \) as \( n \to \infty \) and \( b_n > 0 \).

Definition 5 (strong mixing). Let \( X_t = \{X_{t-1}, X_{t+1}, \ldots\} \) be a stationary time series endowed with \( \sigma \)-algebras \( \mathcal{A}_t = \{X_j, \ -\infty < j < t \} \) and \( \mathcal{A}_t = \{X_j, \ t < j < \infty \} \). Define \( \alpha(s) \) as

\[
\alpha(s) = \sup_{\mathcal{A}_t, \mathcal{B}_t} \left| \left| \mathbb{E}(A \mathbb{I}_{B}) - \mathbb{E}(A) \mathbb{E}(B) \right| \right|. \tag{8}
\]

If \( \alpha(s) \to 0 \) as \( s \to \infty \), then the process is strong mixing.

The results in this section are about the case when the Autoregressive part of the model (4) \( \alpha_{\tau, \tau} = \alpha_{\tau}(z) = 0 \) for any \( \tau \in (0, 1) \). We therefore consider the model

\[
X_t = \alpha_{\tau}(Z_t) Z_t, \quad t = 1, 2, \ldots \tag{9}
\]

Define the check-function as

\[
\gamma_t(X, \mu) = \gamma_t(X - \mu) = \left( \tau - I_{\{X < \mu \}} \right)(X - \mu). \tag{10}
\]

Here, \( I_{\{\cdot\}} \) is the indicator function. Therefore, \( \gamma_t \) is a piecewise monotone increasing function. \( \gamma_t \) is a function of any real random variable \( X \) with distribution function \( F_X(x) = P(X \leq x) = E[I_{\{X \leq x\}}] \), and a real value, \( \mu \in \mathbb{R} \), is the asymmetric absolute value function whose amount of asymmetry depends on \( \tau \); see [15]. In case where \( X_t \) is symmetric and \( \tau = 1/2 \), then we have the fact that \( 2\gamma_t(X_t, \mu) \) is an absolute value function and \( \alpha_{\tau, \tau}(Z_t) \) is the conditional median absolute deviation (CMAD) of \( X_t \). When \( \alpha \) became 0 in model (5), we have a purely heteroscedastic ARCH model introduced in [16] and \( \alpha_{\tau}(Z_t) \) for \( \tau > 0.5 \), which, in this particular case, can be seen as a conditional scale function at \( \tau \)-level.

The check-function in (10) is Lipschitz continuous by the following theorem.

**Theorem 6.** Let \( \gamma_t \) be defined as in (10) and \( (x, \sigma) \in \mathbb{R}^2 \). Then, \( \gamma_t \) satisfies the Lipschitz continuity condition:

\[
\left| \gamma_t(x, \sigma) - \gamma_t(x, \sigma') \right| \leq M \left| \sigma - \sigma' \right| \tag{11}
\]

with the Lipschitz constant \( M = 1 \) and for all \( \sigma, \sigma' \).

**Proof of Theorem 6.** See the proof of Lemma 3.1 in [1, p. 74-75].

By the next theorem we show clearly why the errors \( \{\epsilon_t\} \) in model (2) are assumed to be zero \( \tau \)-quantile and unit scale.

**Theorem 7.** Consider model (5) and the so-called check-function in (10); then, for \( \alpha_{\tau}(Z_t) \in \mathbb{R}^+ \),

\[
\epsilon_t = \frac{X_t - \alpha_{\tau}(Z_t)}{\alpha_{\tau}(Z_t)} \tag{12}
\]

is zero \( \tau \)-quantile and unit scale. And the following equations are verifiable:

\[
P(X_t \leq \alpha_{\tau}(Z_t) | Z_t) = \tau, \tag{13}
\]

\[
P(\gamma_t(X_t, \alpha_{\tau}(Z_t)) \leq \alpha_{\tau}(Z_t) | Z_t) = \tau. \tag{14}
\]

**Proof of Theorem 7.** The \( \tau \)-th quantile operator is

\[
\tau_t(Y_t) = \inf \{ \mu \in \mathbb{R} : P(Y_t \leq \mu | Z_t) \geq \tau \} \tag{15}
\]

with well-defined properties in [1, p. 9-10]. From model (5), the conditional \( \tau \)-quantile of \( X_t \) is

\[
\gamma_t(Z_t) = \omega(Z_t), \tag{16}
\]

where \( \omega(t) \) is the \( \tau \)-quantiles of \( \epsilon_t \). Then, using model (5) and (16), we get

\[
X_t - \gamma_t(Z_t) = \omega(Z_t) (\epsilon_t - q_t^e), \tag{17}
\]

\[
\gamma_t(X_t, \gamma_t(Z_t)) = \omega(Z_t) \gamma_t(\epsilon_t, q_t^e). \tag{18}
\]

And the \( \tau \)-quantile of (18) is

\[
Q_t(\gamma_t(X_t, \gamma_t(Z_t))) = \omega(Z_t) Q_t(\gamma_t(\epsilon_t, q_t^e)) \tag{19}
\]

\[
= \omega(Z_t) Q_t^e, \tag{20}
\]

where \( Q_t^e \) is the \( \tau \)-quantile of \( \gamma_t(\epsilon_t, q_t^e) \). Note that, from (17), \( Q_t(X_t - \gamma_t(Z_t)) = 0 \). The quotient

\[
\frac{X_t - \alpha_{\tau}(Z_t)}{Q_t(\gamma_t(X_t, \alpha_{\tau}(Z_t)))} = \frac{e_t - q_t^e}{Q_t^e} \tag{21}
\]

is zero \( \tau \)-quantile and unit scale and can be seen as model (2) if \( \epsilon_t = (e_t - q_t^e)/Q_t^e, \alpha_{\tau}(Z_t) = q_t(Z_t), \) and \( \alpha_{\tau}(Z_t) = Q_t(\gamma_t(X_t, \alpha_{\tau}(Z_t))). \)

Now, assuming that \( \epsilon_t \) (independent of \( Z_t \)) in model (2) is zero \( \tau \)-quantile, it is equivalent to write

\[
Pr(\epsilon_t < 0) = Pr(\epsilon_t < 0 | Z_t) = \tau \implies \Pr\left(\frac{X_t - \alpha_{\tau}(Z_t)}{\alpha_{\tau}(Z_t)} \leq 0 \right) \rightarrow Z_t = \tau. \tag{21}
\]

This proves (13) for \( \alpha_{\tau}(z) > 0 \). Also, \( \epsilon_t \) is unit scale, which means

\[
Pr(\gamma_t(\epsilon_t) \leq 1) = \tau \implies \Pr\left(\gamma_t\left(\frac{X_t - \alpha_{\tau}(Z_t)}{\alpha_{\tau}(Z_t)}\right) \leq 1 \right) < Z_t = \tau. \tag{22}
\]

\[
Pr(\gamma_t(X_t - \alpha_{\tau}(Z_t)) \leq \alpha_{\tau}(Z_t) | Z_t) = \tau. \tag{22}
\]
Assuming $\alpha_r(Z_t) = 0$, the estimator, $\hat{\theta}_s(Z_t)$, of the conditional scale function, $\omega_s(Z_t)$, is obtained through the minimization of the objective function

$$
\varphi(z, \omega) = E \left[ Y_t \left( Y_t(\mathbf{X}_t), \omega \right) \mid Z_t = z \right].
$$

Thus, the conditional scale function may be obtained by minimizing $\varphi(z, \omega)$ with respect to $\omega$; that is,

$$
\omega_s(z) = \arg\min_{\omega \in \mathbb{R}_+} \varphi(z, \omega),
$$

$$
\omega_s(z) = \inf \left\{ \mu \in \mathbb{R}_+^* : F(\mu \mid z) \geq \tau \right\}
$$

$$
= F^{-1}(\tau \mid z).
$$

The kernel estimator of (24) at $Z_t = z$ is given by

$$
\hat{\omega}_s(z) = \arg\min_{\omega \in \mathbb{R}_+} \hat{\varphi}_s(z, \omega).
$$

We can express the estimate of $\varphi(z, \omega)$ in the random design as it was developed in [17]. Let $Y^*_t = Y_t(Y_t(X_t), \omega)$ be a nonnegative function of $X_t$, $Y^* = (Y^*_1, Y^*_2, \ldots, Y^*_n)$ a random vector in $\mathbb{R}^n_+$, $\tau = 1, 2, \ldots, n$. In the random design, the conditional expectation (23) can be rewritten as follows:

$$
\varphi(z, \omega) = E[Y^* \mid Z_t = z] = \int y^* f(y^* \mid z) dy^*
$$

$$
= \int y^* \frac{f(y^*, z)}{f(z)} dy^*,
$$

where $f(y^* \mid z)$ represents the conditional pdf of $Y^*_t = y^*$ given $Z_t = z$, $f(y^*, z)$ is the joint pdf of the two random variables $Y^*$ and $Z_t$, and $f(z)$ is the pdf of $Z_t = z$. Using [9, 10] with $K_0(\mu) = b^{-1}K(\mu b^{-1})$, a 1-dimensional rescaled kernel with bandwidth $b > 0$, we have the following estimates of $f(y^*, z)$ and $f(z)$ [18]:

$$
\tilde{f}(y^*, z) = \frac{1}{n} \sum_{t=1}^{n} K_b(z_t - z) K_b(y^* - y^*_t),
$$

$$
\tilde{f}(z) = \frac{1}{n} \sum_{t=1}^{n} K_b(z_t - z).
$$

From the estimations above, $\hat{\varphi}(z, \omega)$, the estimate of $\varphi(z, \omega)$, is

$$
\hat{\varphi}_s(z, \omega) = \int \frac{y^* \sum_{t=1}^{n} K_b(z_t - z) K_b(y^* - y^*_t)}{\sum_{t=1}^{n} K_b(z_t - z)} dy^*
$$

$$
= \frac{\sum_{t=1}^{n} K_b(z_t - z) \int y^* K_b(y^* - y^*_t) dy^*}{\sum_{t=1}^{n} K_b(z_t - z)}
$$

$$
= \frac{\sum_{t=1}^{n} K_b(z_t - z) \int (y^* - y^*_t)^2 K_b(y^* - y^*_t) dy^*}{\sum_{t=1}^{n} K_b(z_t - z)}
$$

and considering the regularity conditions of $K_b$ in Assumption 2 and also the fact that $d(y^* - Y^*_t) = dy^*$, $Y^*_t \in \mathbb{R}_+$, we have

$$
\hat{\varphi}_s(z, \omega) = \frac{\sum_{t=1}^{n} K_b(z_t - z) Y^*_t}{\sum_{t=1}^{n} K_b(z_t - z)}
$$

$$
= \frac{n^{-1} \sum_{t=1}^{n} K_b(z_t - z) Y^*_t}{\hat{f}(z)},
$$

where $\hat{f}(z)$ is the estimate of the marginal pdf of $Z_t$ at point $z$ and $Y^*$ can be rewritten as

$$
Y^*_t = [X_t (r - I_{X_t(z)}} \omega)] (r - I_{X_t(z)} \omega)\tau)
$$

and the derivative of $\varphi_s(z, \omega)$ with respect to $\omega$ is

$$
\frac{d\hat{\varphi}_s(z, \omega)}{d\omega} = (n \tilde{f}(z))^{-1} \sum_{t=1}^{n} K_b(z_t - z) \left( I_{X_t(z)} \omega)\tau) - \tau \right).
$$

The minimizer of (30) is obtained from $d\hat{\varphi}_s(z, \omega)/d\omega = 0$. This leads to the following equation:

$$
\left( n \tilde{f}(z) \right)^{-1} \sum_{t=1}^{n} K_b(z_t - z) \left( I_{X_t(z)} \omega)\tau) - \tau \right) = \tau
$$

$$
\text{for all } X_t \in \mathbb{R}_+, t = 1, 2, \ldots.
$$

Note that $Y^*_t = I_{X_t(z)} \omega)\tau)$ in (27). The left part of (33) is a (unsmoothed) conditional cumulative distribution function (CCDF),

$$
\tilde{F}(x^* \mid z) = (n \tilde{f}(z))^{-1} \sum_{t=1}^{n} K_b(z_t - z) \left( I_{X_t(z)} \omega)\tau) \right),
$$

that needs to be estimated and our estimator is therefore

$$
\tilde{\omega}_s(z) = \inf \left\{ x^* \in \mathbb{R}_+ : \tilde{F}(x^* \mid z) \geq \tau \right\} \equiv \tilde{F}^{-1}(\tau \mid z)
$$

which is equivalent to $\tilde{F}(\tilde{\omega}_s(z) \mid z) = \tau$.

An algorithm for estimating $\tilde{F}(x^* \mid z)$ is proposed in the following section. This estimator suffers from the problem of boundary effects as we can see it on Figure 2 due to outliers. We obtain unsmoothed curves of the CCDF because the smoothness is only in the $Z$ direction. A method is proposed by [19] to smooth it in the $y$. The form of Smoothed Conditional Distribution Estimator is

$$
\tilde{F}(x^* \mid z) = (n \tilde{f}(z))^{-1} \sum_{t=1}^{n} K_h(x - Z_t) G \left( \frac{x^* - X_t^*}{h_0} \right),
$$

where $G(\cdot)$ is an integrated kernel with the smoothing parameter $h_0$ in the $X^*$ direction. This estimate is smooth rather
than the NW which is a jump function in $y$. To deal with boundary effects, one may think of the Weighted Nadaraya-Watson (WNW) estimate of the CDF discussed in [12, 20], [21, p. 3–18] among others. The WNW estimator’s expression is

$$\hat{F}_{\text{	iny WNW}}(x^* \mid z) = \frac{\sum_{i=1}^{n} p_i(z, \lambda) K_b(z_i - z) I_{|z_i^* \leq x^*|}}{\sum_{i=1}^{n} p_i(z, \lambda) K_b(z_i - z)}$$

(38)

with conditions $\sum_{i=1}^{n} p_i(z, \lambda) = 1$ and $\Lambda$ is determined using the Newton-Raphson iteration. Smoothing the CDF does not smooth the estimator in (36).

2.1. Algorithm. This algorithm estimates the empirical CCDF, $\hat{F}(x^* \mid z)$, and its inverse $\hat{F}^{-1}(\tau \mid z)$. Starting with the estimation of the former, the denominator is easy to compute as the estimator of the probability density function of $Z$ as vector of points $z$.

1. Obtain $X_{\tau} = (X_{\tau})_{i} \in \mathbb{R}^{n}$, for all $\tau \in (0, 1)$.

2. Check if each $X_{\tau}^*$ is less than or equal to each observation of the whole sequence $x^* = (x_{1}^*, x_{2}^*, ..., x_{n}^*) \in \mathbb{R}^{n}$. The result determines $I_{|x_{i}^* \leq x^*|}$ which can be expressed in (0, 1)-matrix of order $n \times n$.

3. Construct $z_{i}^* = \min(Z) < z_{1}^* < \cdots < z_{n}^* = \max(Z)$ from the sequence of random variable $Z = (Z_{1}, Z_{2}, ..., Z_{n})$ with observation $z = (z_{1}, z_{2}, ..., z_{n})$. $N$ is the number of $z_{i}^*$ from which the probability density function (pdf) of $Z_{i}$ is to be estimated.

4. Determine the matrix of kernels $K$ which is

$$K_{b}(z_{i}^* - Z_{1}) K_{b}(z_{i}^* - Z_{2}) \cdots K_{b}(z_{i}^* - Z_{n})$$

$$\vdots$$

$$K_{b}(z_{n}^* - Z_{1}) K_{b}(z_{n}^* - Z_{2}) \cdots K_{b}(z_{n}^* - Z_{n})$$

(39)

The row sums of $K$ over $n$ give the estimator of the pdf of $Z_{i}$ at $z_{i}^*$, $\tilde{g}(z_{i}^*)$, $i = 1, 2, ..., N$. We obtain the matrix of weights $W$ by the ration of $K$ and $K_{b}$ (element-wise), where $L_{n}$ is a matrix of $n \times n$ ones. Note that the row sums of $W$ are 1.

Let $M$ be the $(0, 1)$-matrix from 2. The estimator of the Conditional Cumulative Distribution Function (CCDF) is

$$\hat{F}(x^* \mid z^*) = W M = K M / (K L_{n})$$

(40)

2.2. Nadaraya-Watson Smoothing Method. We can make $\hat{\alpha}_{\tau}(z)$ smooth by using NW regression (one can also use LOWESS (LOcally WEighted Scatter plot Smoother) regression introduced by [22] to smooth the estimator in (36) and it solves the problem of boundary effects). This will provide a smoothed curve at each level $\tau \in (0, 1)$. We write the regression equation as

$$Y_{i} = \hat{\alpha}_{\tau_{i}}(Z_{i}) + \eta_{i}$$

(41)

with $Y_{i} = \hat{\alpha}_{\tau_{i}}(Z_{i})$ and $\hat{\alpha}_{\tau_{i}}(x) = E[\hat{\alpha}_{\tau}(z) \mid Z_{i} = z]$ and the errors $\eta_{i}$ satisfy $E[\eta_{i}] = 0, V(\eta_{i}) = \sigma_{\eta}^{2}$, and $\text{cov}(\eta_{i}) = 0$ for $i \neq j$. Note that $\hat{\alpha}_{\tau_{i}}(x)$ can be derived using joint pdf $f(y, z)$ as

$$\hat{\alpha}_{\tau_{i}}(z) = E[Y \mid Z = z] = \int y f(y, z) \frac{dv}{f(z)}$$

(42)

where $f(y, z)$ and $f(z)$ are estimated as in (28).

We can perform some transformations on (42) in order to show that it is actually better than the unsmoothed one. By Assumption 1 (iv) and the fact that $F(\hat{\alpha}_{\tau}(z) \mid z) = \tau$, we have

$$F(\hat{\alpha}_{\tau_{i}}(Z_{i}) \mid z) = F(E[\hat{\alpha}_{\tau}(z) \mid Z_{i} = z] \mid z) \leq E[F(\hat{\alpha}_{\tau}(z) \mid z) \mid Z_{i} = z]$$

$$= F(\hat{\alpha}_{\tau}(Z_{i}) \mid z) = \tau$$

(43)

We have used Jensen’s theorem for conditional expectation found in [23] and stated as follows.

Theorem 8 (Jensen’s inequality). For any convex function $l, E[l(X)] \geq l(E[X])$.

(44)

Proof of Theorem 8. Suppose that $l$ is differentiable. The function $l$ is convex if

$$l(x) \geq l(y) + (x - y) \frac{dl}{dx} (y), \text{ for any } x, y.$$ (45)

Let $x = X$ and $y = E[X]$. The inequality $l(X) \geq l(E[X])$ + $(X - E[X])l'(X)$ is true for all $X$ and taking its expectation on both sides proves the theorem.

This inequality is applicable when $f$ is a conditional convex function and when $E[l]$ is a conditional expectation. The estimator $\hat{\alpha}_{\tau_{i}}(Z_{i})$ is also element of the set to which the unsmoothed estimator belongs. This means that $F(\hat{\alpha}_{\tau_{i}}(Z_{i}) \mid z) \geq \tau$. The estimator is empirically given by

$$\hat{\alpha}_{\tau_{i}}(z) = \frac{\sum_{i=1}^{n} K_{b}(Z_{i} - z) y_{i}}{\sum_{i=1}^{n} K_{b}(Z_{i} - z)} \frac{\sum_{i=1}^{n} K_{b}(Z_{i} - z) \hat{\alpha}_{\tau}(Z_{i})}{\sum_{i=1}^{n} K_{b}(Z_{i} - z)}.$$ (46)

2.2.1. Asymptotic Properties. To show the asymptotic properties of our estimator, we compute its expectation and variance. Assuming the data $(Y, Z)$ is i.i.d, the expectation of the numerator is given by

$$E \left[ K_{b}(Z_{i} - z) Y_{i} \right] = \iint_{y = b}^{v} K \left( \frac{u - z}{b} \right) f(u, v) \, du \, dv$$

$$= \int v K(s) f(v | z + sb) f(z + sb) \, ds \, dv$$

$$= \int K(s) f(z + sb) \left( \int v f(v | z = sb) \, dv \right) ds$$

$$= \int K(s) f(z + sb) \hat{\alpha}_{\tau_{i}}(z + sb) ds.$$ (47)
We assume that the first and the second derivatives of \( \omega_{r,s}(z) \) at point \( Z = z \) exist. That is, by Taylor's expansion of \( f(z+sb) \) and \( \omega_{r,s}(z+sh) \) given by

\[
f(z + sh) = f(z) + \frac{f^{(1)}(z)}{1!} sb + \frac{f^{(2)}(z)}{2!} (sb)^2 + o\left(b^2\right),
\]

\[
\omega_{r,s}(z + sh) = \omega_{r,s}(z) + \frac{\omega^{(1)}(z)}{1!} sb + \frac{\omega^{(2)}(z)}{2!} (sb)^2 + o\left(b^2\right),
\]

we get

\[
E\left[K_b (Z_1 - z) Y_1\right] = \omega_{r,s}(z) f(z) + \frac{1}{2} b^2 \mu_2(K) \cdot f^{(2)}(z)
\]

\[
+ f^{(2)}(z) \omega_{r,s}(z) + o\left(b^2\right).
\]

Similarly, the expectation of the numerator is

\[
E\left[K_b (Z_1 - z)\right] = f(z) + \frac{1}{2} b^2 \mu_2(K) \cdot f^{(2)}(z)
\]

\[
+ o\left(b^2\right).
\]

For \( b^2 \) small enough, \( 1 + (1/2) b^2 \mu_2(K) (f^{(2)}(z)/f(z))^{-1} \approx 1 - (1/2) b^2 \mu_2(K) (f^{(2)}(z)/f(z)) \). Thus,

\[
E\left[\tilde{\omega}_{r,s}(z)\right] = \omega_{r,s}(z) + \frac{1}{2} b^2 \mu_2(K) \left(\omega^{(2)}(z) + 2 f^{(1)}(z) \omega^{(1)}(z)\right).
\]

The variance of the numerator, say \( V(N) \), is

\[
V \left(\sum_{t=1}^{n_b} K_b (Z_1 - z) Y_t\right) = \frac{1}{nb^2} \left[\sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) \right] Y_t^2
\]

\[
- \left(\sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) \right)^2 \frac{1}{nb}
\]

\[
\cdot \int v^2 K^2(s) f(v | z + sb) f(z + sb) ds \, dv
\]

\[
- o\left(\frac{1}{n}\right) = \frac{1}{nb} R(K) f(z) \left[\sigma_\eta^2 + \omega_{r,s}^2(z)\right].
\]

Note that \( \int v^2 f(v | z + sb) ds \approx E[Y_t^2 | Z_t = z] \). Similarly, the variance of the denominator, \( V(D) \), is \( V((-1/n) \sum_{t=1}^{n_b} K_b (Z_t - z)) = (1/nb) f(z) R(K) \).

The covariance of the numerator and the denominator of the estimator in (46) are given by

\[
\text{cov}\left(N, D\right)
\]

\[
= \text{cov}\left(\frac{1}{nb} \sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) Y_t, \frac{1}{nb} \sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right)\right)
\]

\[
= \frac{1}{nb^2} \left(\sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) \right) Y_t, K\left(z_t - \frac{z}{b}\right)\right)
\]

\[
- \frac{1}{nb^2} \left(\sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) \right) Y_t, \left(\sum_{t=1}^{n_b} K\left(z_t - \frac{z}{b}\right) \right)\right)
\]

\[
\cdot R(K) f(z) \omega_{r,s}(z) - o\left(\frac{1}{n}\right).
\]

The variance of the estimator in (46) is the variance of a ratio of correlated variables that can be calculated using the approximation found in [24]:

\[
V\left(\frac{N}{D}\right)
\]

\[
= \left(\frac{E[N]}{E[D]}\right)^2 \left[\frac{V(N)}{E[N]^2} + \frac{V(D)}{E[D]^2} - 2 \text{cov}(N, D)\right]
\]

\[
= \frac{R(K) \sigma_\eta^2}{nbf(z)}.
\]

If Assumption 3 for strong mixing processes holds, then from the Central Limit Theorem (CLT) we have

\[
\sqrt{nb} \left(\tilde{\omega}_{r,s}(z) - \omega_{r,s}(z) - \text{Bias}(\tilde{\omega}_{r,s}(z))\right)
\]

\[
\xrightarrow{D} \mathcal{N}\left(0, \frac{R(K) \sigma_\eta^2}{f(z)}\right).
\]

2.3. Asymptotic Normality of QARCH. The CCDF in (35) can be written in the form of an arithmetic mean of a random variable \( L \):

\[
\tilde{F}(x^* | z) = \sum_{t=1}^{n_b} L_t
\]

\[
\xrightarrow{n_b} \mathcal{N}\left(\frac{1}{n_b} \sum_{t=1}^{n_b} K_b (Z_t - z) \right), \frac{1}{n_b} \sum_{t=1}^{n_b} K_b (Z_t - z)\right)
\]

and the approximation of the expectation of \( L \) is

\[
E[L_t] = \frac{E\left[K_b (Z_t - z) I_{[x^* < x^*]}\right]}{E\left(\sum_{t=1}^{n_b} K_b (Z_t - z)\right)} = \frac{E[N]}{E[D]}.
\]
[see [24]]. Using the i.i.d assumption over the data, the numerator is

\[
E[N] = \frac{1}{b_z} E \left[ K \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right]
\]

\[
= \frac{1}{b_z} \int \int K \left( \frac{u - z}{b_z} \right) f(u, v) du dv
\]

\[
= \int F(x^* \mid z + sh) K(s) f(z + sh) ds.
\]

We have used the change of variables \( s = (u - z) / b_z \), the definition of the conditional density function turned into \( f(z + s b_z, v) = f(v \mid z + s h) f(z + s b_z) \), and Fubini's theorem for multiple integrals. Taylor series expansions of \( F(v \mid z + sh) \) and \( f(z + sh) \) yield

\[
E[N] = f(z) F(x^* \mid z) + b_z^2 \mu_z(K)
\]

\[
\cdot \left( \frac{f^{(1)}(z)}{f(z)} F^{(1)}(x^* \mid z) + \frac{1}{2} f^{(2)}(z) F(x^* \mid z) \right)
\]

and, for the denominator, we have

\[
E[D] = f(z) + \frac{1}{2} b_z^2 \mu_z(K) f^{(2)}(z) + o \left( b_z^2 \right).
\]

Thus,

\[
E[L_t] \approx f(z) \left[ F(x^* \mid z) + b_z^2 \mu_z(K) \left( \frac{f^{(1)}(z)}{f(z)} F^{(1)}(x^* \mid z) + \frac{1}{2} f^{(2)}(z) F(x^* \mid z) \right) \right]
\]

\[
= F(x^* \mid z) + \frac{1}{2} b_z^2 \mu_z(K) \left( 2 \frac{f^{(1)}(z)}{f(z)} F^{(1)}(x^* \mid z) + F^{(2)}(x^* \mid z) \right) + o \left( b_z^2 \right).
\]

From the assumption that \( b_z \to 0 \), the denominator is approximated to \( 1 - b_z^2 \mu_z(K) f^{(2)}(z) / 2 f(z) \). Hence,

\[
\text{Bias} \left( \hat{F}(x^* \mid z) \right) \approx \frac{1}{2} b_z^2 \mu_z(K)
\]

\[
\cdot \left( 2 \frac{f^{(1)}(z)}{f(z)} F^{(1)}(x^* \mid z) + F^{(2)}(x^* \mid z) \right).
\]

Some authors assumed that, in this case, the first derivative of the true pdf of \( Z \) at point \( z \) can be zero [19] as the one for the fixed design and, therefore, the bias can be given by

\[
\text{Bias} \left( \tilde{F}(x^* \mid z) \right) \approx \frac{1}{2} b_z^2 \mu_z(K) \left( F^{(2)}(x^* \mid z) \right).
\]

We have

\[
V(N) = V \left( \frac{1}{b_z} K \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right) = \frac{1}{b_z^2}
\]

\[
\cdot V \left( K \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right)
\]

\[
= \frac{1}{b_z^2} \left( E \left[ K^2 \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right] \right)^2
\]

\[
- \left( E \left[ K \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right] \right)^2
\]

\[
\approx \frac{F(x^* \mid z) f(z) R(K)}{b_z} - o(1),
\]

\[
V(D) = V \left( \frac{1}{n} \sum_{i=1}^{n} K_{b_z} (Z_i - z) \right) = \frac{1}{n b_z^2}
\]

\[
\cdot V \left( K \left( \frac{Z_i - z}{b_z} \right) \right) + \frac{1}{n b_z^2} \left( E \left[ K^2 \left( \frac{Z_i - z}{b_z} \right) \right] \right)
\]

\[
- \left( E \left[ K \left( \frac{Z_i - z}{b_z} \right) \right] \right)^2 \approx \frac{f(z) R(K)}{n b_z} - o \left( \frac{1}{n} \right).
\]

\[
\text{cov}(N, D) = \frac{1}{n b_z^2}
\]

\[
\cdot \text{cov} \left( K \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)}, K \left( \frac{Z_t - z}{b_z} \right) \right) \approx \frac{1}{n b_z^2}
\]

\[
\cdot E \left[ K^2 \left( \frac{Z_t - z}{b_z} \right) I_{(X^*_t \leq x^*_t)} \right] - o \left( \frac{1}{n} \right) \approx \frac{1}{n b_z^2}
\]

\[
\cdot F(x^* \mid z) f(z) R(K).
\]

Using the same approximation in (54), the variance of \( \tilde{F}(x^* \mid z) \) is

\[
V(L_t) \approx F(x^* \mid z) \left[ \frac{R(K) (1 - F(x^* \mid z))}{b_z f(z)} \right].
\]
and by the Central Limit Theorem, using Assumption 3 for 
\{X_t^*, Z_t\}, \ t = 1, 2, \ldots, 
\[\sqrt{n}(\hat{F}(x^* | z) - F(x^* | z) - \text{Bias}(F(x^* | z))) \xrightarrow{D} \mathcal{N}(0, V(L_z)).\] (67)

Notice that the expectation of \(\hat{F}(x^* | z)\) is the same as the one of \(L\) and the variance is \(V(L_z)/n\). To show the asymptotic normality of \(\hat{t}_z\), we use the following theorem.

**Theorem 9** (delta method). Suppose \(\hat{F}(x^* | z)\) has the asymptotic normal distribution as in (67). Suppose \(g(\cdot)\) is a continuous function that has a derivative \(g^{(1)}(\cdot)\) at \(\mu = E[\hat{F}(x^* | z)]\). Then

\[\sqrt{n}b_z(g(\hat{F}(x^* | z)) - g(\mu)) \xrightarrow{D} \mathcal{N}(0, [g^{(1)}(\mu)]^2 R(K)(1 - F(x^* | z))/f(z)).\] (68)

**Proof of Theorem 9.** The first-order Taylor expansion of \(g(\cdot)\) about the point \(\mu\), and evaluated at the random variable \(\hat{F}(x^* | z)\), is

\[g(\hat{F}(x^* | z)) = g(\mu) + g^{(1)}(\mu)(\hat{F}(x^* | z) - \mu)\] (69)

and subtracting \(g(\mu)\) from both sides and multiplying by \(\sqrt{n}b_z\), we get

\[\sqrt{n}b_z(g(\hat{F}(x^* | z)) - g(\mu)) = \sqrt{n}b_z [g^{(1)}(\mu)(\hat{F}(x^* | z) - \mu)]\] (70)

which tends to \(\mathcal{N}(0, [g^{(1)}(\mu)]^2 R(K)(1 - F(x^* | z))/f(z))\) in distribution. \(\square\)

For \(g(\mu) = F^{-1}(\mu | z)\), thus, \(g^{(1)}(\mu) = 1/f(F^{-1}(\mu | z) | z)\).

In the next section, it is shown that the AMSE (Asymptotic Mean Squared Error) of \(\hat{F}(x^* | z)\) is equal to \(o(b^4) + o(1/nb)\) which tends to 0 as \(n \to \infty\) and \(b \to 0\). This shows the consistency of the CCDF estimate, that is \(\hat{F}(x^* | z) \to^p F(x^* | z)\), and we have

\[\frac{1}{f(F^{-1}(\mu | z) | z)} = \frac{1}{f(\hat{F}(x^* | z) | z)} = \frac{1}{f(\hat{t}_z(z) | z)}.\] (71)

at points \(x^*\)'s that satisfy (36). Using again the first-order Taylor expansion, we also have

\[g(\mu) = g(F(x^* | z) + \text{Bias}(\hat{F}(x^* | z))) \approx g(F(x^* | z)) + \text{Bias}(\hat{F}(x^* | z)) \times g^{(1)}(F(x^* | z)) = x^* + \text{Bias}(\hat{F}(x^* | z)) / f(x^* | z)\] (72)

for \(x^*\) satisfying (36) and replacing \(\hat{F}(\omega_z(z) | z)\) by \(F(\hat{t}_z(z) | z)\) using the uniqueness assumption of \(\omega_z(z)\), (68) becomes

\[\sqrt{n}b_z(\hat{b}_z(z) - \hat{t}_z(z) - \text{Bias}(\hat{t}_z(z))) \xrightarrow{D} \mathcal{N}(0, R(K) \{\tau(1 - \tau)|f(\omega_z(z) | z)|^2\}).\] (73)

with \(\text{Bias}(\hat{t}_z(z)) = \text{Bias}(\hat{F}(\omega_z(z) | z))/f(\omega_z(z) | z) = (1/2)\sqrt{\mu_z^2 \{f(\omega_z(z) | z)\} | f(\omega_z(z) | z) |} \).\)

This result can be used to calculate the optimal bandwidth to compute the good estimation of the CSF.

### 3. Bandwidth Selections

#### 3.1. Optimal Bandwidth for Density Estimations.

In nonparametric estimations, specially in Kernel Density Estimations, computing a curve of an arbitrary function from the data without guessing the shape in advance requires an adequate choice of the smoothing parameter. The most used method is the "plug-in" method which consists of assigning a pilot bandwidth in order to estimate the derivatives of \(\hat{f}(z)\). We choose the bandwidth that minimizes the AMISE (Asymmetric Mean Integrated Squared Error) below.

\[\text{AMISE}(\hat{f}(z)) = \int E[(\hat{f}(z) - f(z))^2] dz\]

\[= \int E[(\hat{f}(z) - E[\hat{f}(z)])^2] dz\]

\[= \int \{E[(\hat{f}(z) - E[\hat{f}(z)])^2] + \text{Bias}^2(\hat{f}(z))\} dz\] (74)

\[= \int \{V(\hat{f}(z)) + \text{Bias}^2(\hat{f}(z))\} dz\]

\[= \int \{\frac{R(K)f(z)}{nb} + \frac{1}{4}b^4 \mu_z^2(K) f(z)^2\} dz\]

\[= \frac{R(K)}{nb} + \frac{1}{4}b^4 \mu_z^2(K) R(f(z)).\]

The general form of the \(r\)th derivatives of the AMISE with respect to \(b\) was studied in [25], considering that the unknown functions in (74) are also functions of the smoothing parameter.

\[\frac{d}{dz} \text{AMISE}(\hat{f}(z)) = \frac{R(K^{(r)})}{nb^{2r+1}}\]

\[+ \frac{1}{4}b^4 \mu_z^2(K) R(f^{2r}(z)).\] (75)

The optimal smoothing parameter minimizing (75) is

\[b^* = \left[\frac{(2r + 1) R(K^{(r)})}{4 \mu_z^2(K) R(f^{2r}(z))}\right]^{1/(2r+5)} \times n^{-1/(2r+5)}.\] (76)
AMISE in (74) is to get the optimal bandwidth for each \( f^{(r)} \) directly. As an example, we consider the Epanechnikov Kernel function in order to compute \( R(K), \mu_2(K) \), and the efficiency of the kernel function given by \( \sqrt{\mu_2(K)}R(K) \). Epanechnikov’s kernel function is

\[
K(u) = \frac{3}{4} \left(1 - u^2\right) I_{|u|\leq1} \implies R(K) = \frac{3}{4}
\]

\[
\mu_2(K) = \int_{-1}^{1} u^2 K(u) du = \int_{-1}^{1} \left(u^2 - u^4\right) du = \frac{3}{5}
\]

and its efficiency is measured by

\[
\text{Eff}(K) = R(K) \sqrt{\mu_2(K)} = \frac{3}{4} \sqrt{\frac{5}{7}} = 0.268
\]

which is the smallest of all the other kernel functions.

3.2. Optimal Bandwidth for CCDF. The optimal bandwidth for the CCDF estimate is the one that minimizes the AMSE. It is shown below that the AMSE is actually the summation of the variance and the bias of the CCDF estimate. This is useful because the two are linked. When the variance is big, the bias also is big and when the variance is small, the bias is small.

\[
\text{AMSE} \left( \hat{F}(x^*|z) \right) = E \left[ \left( \hat{F}(x^*|z) - F(x^*|z) \right)^2 \right] \\
= E \left[ \left( \hat{F}(x^*|z) - \hat{F}(x^*|z) \right)^2 \right] + \text{Bias} \left( \hat{F}(x^*|z) \right)^2 + E \left[ \hat{F}(x^*|z) \right] - E \left[ \hat{F}(x^*|z) \right] \\
+ E \left[ \hat{F}(x^*|z) \right] - E \left[ \hat{F}(x^*|z) \right] \\
+ \text{Bias} \left( \hat{F}(x^*|z) \right) V \left( \hat{F}(x^*|z) \right) \\
+ \text{Bias} \left( \hat{F}(x^*|z) \right) = \frac{R(K)}{nb_2f(z)} F(x^*|z) \left( 1 - F(x^*|z) \right) + \frac{b^4}{4} \mu_2(K) \left( F^{(2)}(x^*|z) \right)^2
\]

which is given by (66) and (64). Therefore,

\[
b^* = \arg \min_{b>0} \text{AMSE} \left( \hat{F}(x^*|z) \right)
\]

and \((d/db)\text{AMSE}(\hat{F}(x^*|z)) = 0\) leads to

\[
b^* = \left\{ \frac{R(K)F(x^*|z)(1 - F(x^*|z))}{\mu_2^2(K)f(z)(F^{(2)}(x^*|z))^2} \right\}^{1/5} \times n^{-1/5}.
\]

This result is practically possible by estimating the unknown functions which are dependent on the smoothing parameter.

\( \hat{F}^{(2)}(x^*|z) \) is the second derivative of the CCDF from (35) at point \( Z_t = z \). The estimator of \( r \)th derivatives of (35) is

\[
\hat{F}^{(r)}(x^*|z) = \frac{d^r}{dz^r} \sum_{t=1}^{n} W_t(z) X_{X_t^* < x^*}
\]

\[
= \sum_{t=1}^{n} W_t^{(r)}(z) X_{X_t^* < x^*}
\]

with

\[
W_t(z) = \frac{K((Z_t - z)/b)}{\sum_{i=1}^{n} K((Z_t - z)/b)} = \frac{K((Z_t - z)/b)}{nb_2f(z)},
\]

the function of weights. Thus, the first derivative is given by

\[
W_t^{(1)}(z) = \frac{1}{nb_2^2} \frac{K((Z_t - z)/b)}{f(z)} \hat{f}(z) - bK((Z_t - z)/b) \hat{f}(1)(z)
\]

\[
= \frac{1}{nb_2^2} A
\]

and the second derivative is also

\[
W_t^{(2)}(z) = \frac{1}{nb_2^2} A^{(1)} - b^{(1)} A
\]

with \( A^{(1)} = (1/b)K^{(2)}((Z_t - z)/b) \hat{f}(z) - bK((Z_t - z)/b) \hat{f}^{(2)}(z) \) and \( b^{(1)} = 2\hat{f}^{(1)}(z) \hat{f}(z) \). Note that the estimation of the CCDF is function of the estimation of the empirical pdf of \( z \). An optimal bandwidth that minimizes the AMISE of \( \hat{f}(z) \) can also be the one that is optimal for the estimation of the CCDF.

Recent findings on the estimation of an optimal band- width for KDE (Kernel Density Estimation) are numerous ([25–27]) but the estimation of an optimal smoothing parameter remains irksome due to computation issue and time consuming routines. To do so, we adopt what had been done by [27] to estimate the \( r \)th derivatives of the pdf of \( Z_t \) with respect to \( z \). We extend the idea to estimate the first and the second derivative of the CCDF with respect to \( z \).

4. Simulation Study

4.1. Model Specification. The ARCH(q) models introduced by [16] are widely used in financial applications. An AR(1)-ARCH(1) is a mixed model from an AR(d) and GARCH(p, q) for \( d = 1, p = 1, \) and \( q = 0 \). In time series, an observation at one time can be correlated with the observations in the previous time. That is,

Note that the operator \( : \cdot : \) means the element-wise division between matrices.

(5) For each row of \( \hat{F}(\cdot|\cdot) \), find the smallest \( x^* \) such that \( \hat{F}(x^*|z^*) \geq \tau, \tau \in (0, 1) \).

(6) The quantiles \( \hat{q}_i(z) \) are the \( x^* \)'s which satisfy (36).

This gives an unsmoothed estimator curve with bad shape at boundaries (see Figure 2).
The data to be simulated is given by \( X_t = \mu + \delta X_{t-1} + (w + \alpha \varepsilon_i^2)^{1/2} \varepsilon_i, t = 1, 2, \ldots \).

(i) autoregressive process of order \( p = 1, 2, \ldots \),

\[
\text{AR}(p): X_t = \mu + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \cdots + \delta_p X_{t-p} + \varepsilon_t, \quad \text{with } \varepsilon_t \text{ i.i.d,}
\]

(ii) autoregressive \((p)\)-General Autoregressive Conditional Heteroscedastic process of order \((d = 1, 2, \ldots; \; p = 1, 2, \ldots; \; q = 1, 2, \ldots)\),

\[
\text{AR}(d) - \text{GARCH}(p, q): X_t = \sum_{i=1}^{p} a_i X_{t-i} + \alpha_i \varepsilon_t,
\]

with \( \varepsilon_t \) i.i.d. and \( \alpha_i \) given by

\[
\alpha_i = \left( w + \sum_{i=1}^{p} a_i \mu_i^2 + \sum_{i=1}^{q} \beta_i \sigma_i^2 \right)^{1/2}.
\]

4.2. Specifications for AR(1)-GARCH(1,1)

4.2.1. Unconditional Expectation. The unconditional expectation is

\[
E[X_t] = \mu + \delta E[X_{t-1}] + E[\alpha_i \varepsilon_t]
\]

\[= \mu + \delta E[X_t] + E[\alpha_i] E[\varepsilon_t].
\]

Note that \( E[X_t] = E[X_{t-1}] \) is used to ensure the stationarity of the process. That is, the expectation is therefore given by

\[
E[X_t] = \frac{\mu}{1 - \delta}.
\]

4.2.2. Unconditional Variance. The unconditional variance of the model is given by the law of total variance

\[
V(X_t) = E[V(X_t | X_{t-1})] + V(E[X_t | X_{t-1}])
\]

\[= E[\alpha_t^2] + V[\alpha_t].
\]

We have

\[
E[\alpha_t^2] = \omega + \alpha E[X_{t-1}^2] + \beta E[\alpha_{t-1}^2].
\]

Using the i.i.d. assumption on the sequence of random variables \( X_1, X_2, \ldots, X_n \), the expected value of \( X_t^2 \) can be calculated as follows:

\[
E[X_t^2] = E[\mu X_t + \delta X_{t-1} X_t + \alpha \varepsilon_i X_t]
\]

\[= \mu E[X_t] + \delta E[X_t^2] + \alpha \mu^2.
\]

which is independent of time. In another way,

\[
E[X_t^2] = \omega + \alpha \varepsilon_i^2 + \alpha \mu^2
\]

\[= E[\alpha_t^2] + E[\alpha_t^2] \quad \text{equation (94)}
\]

Equation (92) becomes

\[
E[\alpha_t^2] = \omega + \alpha (E[\alpha_t^2] + E[\alpha_{t-1}^2]) + \beta E[\alpha_{t-1}^2]
\]

\[= \omega + \alpha E[\alpha_t^2] + (\alpha + \beta) E[\alpha_{t-1}^2] \quad \text{stationarity}.
\]

We obtain

\[
E[\alpha_t^2] = \frac{\omega + \alpha \mu^2}{1 - \alpha - \beta}.
\]

The expectation of \( \alpha_t^2 \) is given by

\[
E[X_t^2] = (\mu + \delta X_{t-1})^2
\]

\[= \mu^2 + 2 \delta \mu^2 E[X_t] + \delta^2 E[X_t^2]
\]

\[= \mu^2 + 2 \frac{\delta \mu^2}{1 - \delta} + \frac{\delta^2 \mu^2}{(1 - \delta)^2}
\]

\[= \frac{\mu^2}{(1 - \delta)^2}.
\]

It follows that

\[
E[\alpha_t^2] = \frac{\omega (1 - \delta)^2 + \alpha \mu^2}{(1 - \alpha - \beta) (1 - \delta)^2}
\]

and the variance in (91) becomes

\[
V(X_t) = \frac{\omega (1 - \delta)^2 + \alpha \mu^2}{(1 - \alpha - \beta) (1 - \delta)^2} + V(\mu + \delta X_{t-1})
\]

\[= \frac{\omega (1 - \delta)^2 + \alpha \mu^2}{(1 - \alpha - \beta) (1 - \delta)^2} + \delta^2 V(X_t)
\]

\[= \frac{\omega (1 - \delta)^2 + \alpha \mu^2}{(1 - \alpha - \beta) (1 - \delta)^2}.
\]

This variance is positive and finite for \( \mu \in \mathbb{R}, |\delta| < 1, \omega > 0, \alpha > 0, \beta > 0, \) and \( \alpha + \beta < 1 \).

4.3. Model Simulation. We simulated the data from (1) with \( \mu = 0.5, \delta = 0.3 \), for the AR(1) part and \( w = 0.1, \alpha = 0.35 \), for the ARCH(1) and \( \varepsilon_i \sim \text{i.i.d. } \mathcal{N}(0, 1) \). The data plot is represented by Figure 1.

Our algorithm gives the estimation of the conditional scale function which suffers from boundary effects as it is seen from Figure 2. This issue is recurrent while performing Kernel Density Estimations. The reason is that, at the boundaries, \( g(z) \) is underestimated because of the minimal number of points \([28] \). The consistency of our estimator is dependent on this problem of big variations at the boundaries. This increases the Average Squared Error between two different estimations from the same model.
4.4. Boundary Correction. To correct the boundary effects, we use the method of box-plot fences proposed by [29] to detect the extreme values that make the estimation too rough at the extremities of the CCDF estimations' curves. Our estimator, being the inverse of the CCDF, is naturally rough at extremities. Among the Kernel functions, only the Gaussian can handle the sparseness of points at boundaries because its domain is $\mathbb{R}$. The other kernel functions can bring zero at extremities and make the estimation of the CCDF wrong. What we do is to omit the points that are extremely far from the others by the box-plot fences method. The method consists of determining the first and the third quantiles from the $Z_i$'s. Outliers are the points that are located outside the interval

$$[Q1 - 3 \times (Q3 - Q1), Q3 + 3 \times (Q3 - Q1)],$$  \hspace{1cm} (100)

where $Q1$ and $Q3$ are the first and the third quartiles. Figure 3 is the representation of $Z_i$ and the transformed response variable $X^*_i$ defined in (34) at level $\tau = 0.75$.

The gray points are outliers from (100). We lose some information by deleting them but we get the possibility of performing the estimation of a continuous curve of the CSF. Figure 4 is the estimations of the CSF at levels 0.25, 0.5 (median), 0.75, and 0.9. As we can see on the graphic, despite the optimal bandwidth for the empirical pdf of $Z_i$ at point $z$, we get unsmoothed curves at high level $\tau > 0.5$.

![Figure 1: Plot of the simulated AR(1)-ARCH(1).](image1)

![Figure 2: Conditional scale function estimate at $\tau = 0.75$.](image2)

The curves represent the estimations of the CSF at $\tau = 0.9, 0.75, 0.50, 0.25$ from up to down. As it is seen in Figure 4, some curves are not smooth; that is why the NW method is discussed in Section 2.2 which requires that unsmoothed estimator and the bins $z^*_1, z^*_2, \ldots, z^*_N$. We obtain Figure 5 which combines the two estimations.

The next section discusses how precise is our estimation with the optimal bandwidth selection with the calculation of the MASE (Mean Average Squared Errors).

4.5. Consistency. The consistency of the estimator can be shown with the calculation of the Mean Average Squared Error providing the quantitative assessment of the accuracy of our estimator. This is a kind of bootstrap method to calculate the average gap between $m$ estimated CSFs. The formula is

$$\text{MASE}(\hat{\alpha}_r(z)) = \frac{1}{n} \sum_{j=1}^{n} \left[ \frac{1}{m} \sum_{i=1}^{m} (\hat{\alpha}_{r,i}(z_j) - \hat{\alpha}_{r,j}(z_i))^2 \right].$$  \hspace{1cm} (101)

Table 2 shows that the estimator of the CSF is more precise at level $\tau \leq 0.55$ for both the smoothed and the LOWESS versions.
Table 2: Mean average squared errors (MASE).

<table>
<thead>
<tr>
<th>Kern. func.</th>
<th>n</th>
<th>$m = 10$</th>
<th>$m = 50$</th>
<th>$m = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\text{mase}(.25)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>250</td>
<td>0.0017</td>
<td>0.0013</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0011</td>
<td>0.0008</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0006</td>
<td>0.0004</td>
<td>0.0005</td>
</tr>
<tr>
<td>Epanechn</td>
<td>200</td>
<td>0.0012</td>
<td>0.0009</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0006</td>
<td>0.0004</td>
<td>0.0007</td>
</tr>
<tr>
<td>Triweight</td>
<td>200</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0006</td>
<td>0.0005</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0003</td>
<td>0.0002</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{mase}(.50)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>250</td>
<td>0.0054</td>
<td>0.0045</td>
<td>0.0063</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0036</td>
<td>0.0029</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0028</td>
<td>0.0022</td>
<td>0.0025</td>
</tr>
<tr>
<td>Epanechn</td>
<td>200</td>
<td>0.0067</td>
<td>0.0057</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0045</td>
<td>0.0036</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0031</td>
<td>0.0026</td>
<td>0.0041</td>
</tr>
<tr>
<td>Triweight</td>
<td>200</td>
<td>0.0008</td>
<td>0.0007</td>
<td>0.0030</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0023</td>
<td>0.0020</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0019</td>
<td>0.0016</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{mase}(.75)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>250</td>
<td>0.0234</td>
<td>0.0183</td>
<td>0.0237</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0227</td>
<td>0.0178</td>
<td>0.0223</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0099</td>
<td>0.0079</td>
<td>0.0138</td>
</tr>
<tr>
<td>Epanechn</td>
<td>200</td>
<td>0.0156</td>
<td>0.0123</td>
<td>0.0266</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0184</td>
<td>0.0152</td>
<td>0.0235</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0162</td>
<td>0.0130</td>
<td>0.0102</td>
</tr>
<tr>
<td>Triweight</td>
<td>200</td>
<td>0.0190</td>
<td>0.0176</td>
<td>0.0145</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0112</td>
<td>0.0099</td>
<td>0.0131</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0075</td>
<td>0.0064</td>
<td>0.0073</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{mase}(.90)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian</td>
<td>250</td>
<td>0.0880</td>
<td>0.0692</td>
<td>0.1180</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0468</td>
<td>0.0377</td>
<td>0.0890</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0932</td>
<td>0.0690</td>
<td>0.0491</td>
</tr>
<tr>
<td>Epanechn</td>
<td>200</td>
<td>0.0664</td>
<td>0.0577</td>
<td>0.1074</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0816</td>
<td>0.0515</td>
<td>0.0827</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0740</td>
<td>0.0510</td>
<td>0.0449</td>
</tr>
<tr>
<td>Triweight</td>
<td>200</td>
<td>0.0510</td>
<td>0.0434</td>
<td>0.0452</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.0453</td>
<td>0.0337</td>
<td>0.0390</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>0.0172</td>
<td>0.0133</td>
<td>0.0268</td>
</tr>
</tbody>
</table>

5. Conclusion

We have derived an estimator for the conditional scale function in an AR(1)-GARCH(1) and despite the heavy-tailed data, we could deal with the boundary effect and were able to show the consistency of the estimator through a Monte Carlo study. We assumed that the QAR(1) is known and is zero and, along with the regularity assumptions, we derived the estimator which can be improved in some next papers. The very next paper will focus on the estimation when the QAR(1) is unknown.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Acknowledgments

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

References


