Fitting a pth order parametric generalized linear autoregressive multiplicative error model[1]

N. Balakrishna H. L. Koul M. Ossiander L. Sakhanenko[2]

Cochin University of Science and Technology, Oregon State University and Michigan State University

Abstract

This paper is concerned with the problem of fitting a generalized linear model to the conditional mean function of multiplicative error time series models. These models are particularly suited to model nonnegative time series such as the duration between trades at a stock exchange and volume transactions. The proposed test, based on a marked residual empirical process whose marks are suitably defined residuals and which jumps at the estimated indices, is shown to be asymptotically distribution free.

1 Introduction

In the last two decades the family of multiplicative error models has attracted considerable attention for modeling nonnegative time series. Engle and Russell (1998) and Engle (2002) used these models for analyzing financial durations and trading volume of orders, respectively. For many other applications see Pacarur (2008), Hautsch (2012), and the references therein.

To proceed further, let Y_i, $i \in \mathbb{Z} := \{0, \pm 1, \pm 2, \cdots\}$, be a discrete time nonnegative stationary process with $EY_0^2 < \infty$, and let $\mathcal{H}_{i-1}$ denote the information available up to time $i-1$ for forecasting Y_i. A multiplicative error model (MEM) takes the form

$$Y_i = E[Y_i \mid \mathcal{H}_{i-1}]\varepsilon_i, \quad i \in \mathbb{Z}, \tag{1.1}$$

where $\varepsilon_i, i \in \mathbb{Z}$ are independent and identically distributed (i.i.d.) non-negative error random variables (r.v.'s) with $E(\varepsilon_0) = 1$, $E(\varepsilon_0^2) < \infty$. Moreover, ε_i is assumed to be independent of the past information $\mathcal{H}_{i-1}$, for all $i \in \mathbb{Z}$.

Pacarur (2008) and Hautsch (2012) discuss many parametric specifications for the conditional mean function $E(Y_i \mid \mathcal{H}_{i-1})$. The problem of fitting a given parametric MEM model is of practical importance since knowing the right parametric model can lead to optimal inference. Koul, Perera and Sillvaphule (2013) (KIS) proposed a lack of fit test for fitting a parametric model to this conditional mean function when the underlying time series is

[1] *Keywords & phrases: Martingale transform, AR conditional duration models.*

[2] This author's research was supported in part by the NSF grant DMS–1612867

Markovian, i.e., when $E[Y_i \mid \mathcal{H}_{i-1}] = \tau(Y_{i-1})$, $i \in \mathbb{Z}$, a.s., for some positive measurable function τ defined on $R^+ := [0, \infty)$. To describe the KIS test a bit more precisely, let $q \geq 1$ be a given positive integer, $\Omega \subset \mathbb{R}^q$, and $\{\psi(y, \omega),\ y \geq 0, \omega \in \Omega\}$ be a given family of positive functions. Consider the problem of testing

$$\begin{aligned} &H_0 : \tau(y) = \psi(y, \omega), \quad \text{for all } y \geq 0, \text{ and some } \omega \in \Omega, \text{ versus} \\ &H_1 : H_0 \text{ is not true.} \end{aligned}$$

Let

$$\mathcal{U}_n(y, \omega) = n^{-1/2} \sum_{i=1}^{n} \Big(\frac{Y_i}{\psi(Y_{i-1}, \omega)} - 1 \Big) I(Y_{i-1} \leq y), \quad y \geq 0,\ \omega \in \Omega,$$

The test of H_0 proposed in KIS is based on the process $\mathcal{U}_n(\cdot, \hat{\omega})$, where $\hat{\omega}$ is a $n^{1/2}$-consistent estimator of ω under H_0. Asymptotic null distribution of this process depends on the null model parameter ω_0 under H_0 and its estimator in a complicated fashion. In KIS, an analog of the martingale transform of Stute, Thies and Zhu (1998) (STZ) of this process was shown to converge weakly, under H_0, to a time transform of Brownian motion. In particular the test based on the supremum of the absolute value of this transform is asymptotically distribution free, consistent against a large class of nonparametric alternatives, and has nontrivial asymptotic power against a large class of $n^{-1/2}$-local alternatives. In a finite sample simulation study this test was shown to have very desirable empirical level property and much larger empirical power, compared to several other model specification tests including the Ljung-Box Q test and Lagrange multiplier type test of Meitz and Teräsvirta (2006).

Admittedly restricting the dependence of the conditional mean function to one lag variable has limited applications. Consider the situation where

$$E[Y_i \mid \mathcal{H}_{i-1}] = \gamma(Y_{i-1}, \cdots, Y_{i-p}), \quad i \in \mathbb{Z}, \tag{1.2}$$

for a known positive integer p and a positive function γ defined on $[0, \infty)^p$. In practice there are several time series where the conditional mean function γ depends on the previous p lags via a linear combination. Then the question of interest is which one of these models fit the given time series.

More precisely, let $\Theta_1 \subset \mathbb{R}^p$, $\Theta := [0, \infty) \times \Theta_1$, and φ be a known positive link function defined on $\mathbb{R}$. Let x' denote the transpose of an Euclidean vector x. The problem of interest is to test the hypotheses

$$\begin{aligned} &\mathcal{H}_0 : \gamma(y) = \theta_0 + \varphi(\theta' y), \quad \text{for all } y \in [0, \infty)^p, \text{ and for some } (\theta_0, \theta')' \in \Theta, \quad \text{versus} \\ &\mathcal{H}_1 : \mathcal{H}_0 \text{ is not true.} \end{aligned}$$

A version of the $\mathcal{U}_n$-process suitable for this testing problem is

$$\mathcal{V}_n(y, \delta) = n^{-1/2} \sum_{i=1}^{n} \Big(\frac{Y_i}{\vartheta_0 + \varphi(\vartheta' \mathbf{Y}_{i-1})} - 1 \Big) I\{\vartheta' \mathbf{Y}_{i-1} \le y\}, \quad y \ge 0,$$

where $\vartheta := (\vartheta_1, \cdots, \vartheta_p)'$, $\delta := (\vartheta_0, \vartheta')' \in \Theta$. Let δ_0 denote the parameter vector for which H_0 holds, i.e., $\delta_0 := (\theta_0, \theta')'$. The proposed tests of $\mathcal{H}_0$ will be based on the process $\mathcal{V}_n(\cdot, \hat{\delta})$, where $\hat{\delta}$ is a $n^{1/2}$-consistent estimator of δ_0, under $\mathcal{H}_0$. The linear combinations $\hat{\theta}' \mathbf{Y}_{i-1}$ are known as estimated indices and $\mathcal{V}_n(\cdot, \hat{\delta})$ is a marked residual empirical process with the marks $\hat{\eta}_i := Y_i / [\hat{\theta}_0 + \varphi(\hat{\theta}' \mathbf{Y}_{i-1})] - 1$ jumping at the indices $\hat{\theta}' \mathbf{Y}_{i-1}$.

The asymptotic null distribution of this process under $\mathcal{H}_0$ is discussed in the sub-section 2.1. An analog of the martingale transform of STZ test is given in the sub-section 2.2. Several proofs are relegated to section 3. In the sequel, $\|\cdot\|$ denotes the Euclidean norm, all limits are taken as $n \to \infty$, and $u_p(1)$ denotes a sequence of stochastic processes that converges to zero uniformly over its time domain, in probability.

2 Main Results

In this section we derive the asymptotic expansion of the process $\mathcal{V}_n(\cdot, \hat{\delta})$ under $\mathcal{H}_0$ and describe a martingale type transformation that converges weakly to the standard Brownian motion on $[0, \infty)$.

2.1 Asymptotic linearity of $\mathcal{V}$ under H_0

Here we derive an asymptotic linearity result for the process $\mathcal{V}_n(x, \hat{\delta})$. To begin with we shall state the needed assumptions.

(F) The error variable ε_0 is positive, $E\varepsilon_0 = 1$, $E\varepsilon_0^2 < \infty$.

(G) The distribution function (d.f.) G of $\theta' \mathbf{Y}_0$ is continuous, where θ is as in $\mathcal{H}_0$.

(φ) $E\varphi^2(\theta' \mathbf{Y}_0) < \infty$, and φ is continuously differentiable with the derivative $\dot{\varphi}$ satisfying $E\|\mathbf{Y}_0 \dot{\varphi}(\theta' \mathbf{Y}_0)\|^2 < \infty$, and for every $0 < c < \infty$,

$$\sup_{1 \le i \le n,\ \sqrt{n}\|\vartheta - \theta\| \le c} \sqrt{n} |\varphi(\vartheta' \mathbf{Y}_{i-1}) - \varphi(\theta' \mathbf{Y}_{i-1}) - (\vartheta - \theta)' \mathbf{Y}_{i-1} \dot{\varphi}(\theta' \mathbf{Y}_{i-1})| = o_p(1).$$

(C) There exist estimators $\hat{\theta}_0 > 0$ and $\hat{\theta} \in \mathbb{R}^p$ such that under $\mathcal{H}_0$,

$$n^{1/2} \Big(|\hat{\theta}_0 - \theta_0| + \|\hat{\theta} - \theta\| \Big) = O_p(1).$$

Let $\hat{\delta} := (\hat{\theta}_0, \hat{\theta}')'$. The derivation of the asymptotic linearity of the process $\mathcal{V}_n(\cdot, \hat{\delta})$ is facilitated by the following lemma of a general interest. Consider the following assumptions.

(E) The nonnegative process $Y_j, j \in \mathbb{Z}$ is stationary and ergodic. The r.v.'s η_i, $i \in \mathbb{Z}$ are i.i.d., and η_i is independent of Y_j, $j \le i-1$, for all $i \in \mathbb{Z}$.

(g) $g(x,z)$ is a measurable function of $(x,z) \in [0,\infty)^p \times \mathbb{R}$ such that $Eg^2(\mathbf{Y}_0, \eta_1) < \infty$.

Let $\mathcal{F}_i := \sigma$-field$\{\mathbf{Y}_j, j \le i\}$, $j \in \mathbb{Z}$, and define, for $s \in \mathbb{R}^p$ and $y \ge 0$,

$$V_n(y,s) \;:=\; \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \big[g(\mathbf{Y}_{i-1}, \eta_i) - E\{g(\mathbf{Y}_{i-1}, \eta_i)|\mathcal{F}_{i-1}\}\big] I((\theta + n^{-1/2}s)'\mathbf{Y}_{i-1} \le y).$$

Lemma 2.1 *Assume that the processes η_j, $Y_j, j \in \mathbb{Z}$ satisfies assumption* (E) *with the d.f. G of $\theta'\mathbf{Y}_0$ satisfying condition* (G), *and that* (g) *holds. Then for every $c \in [0,\infty)$,*

$$\sup_{y \ge 0, \|s\| \le c} |V_n(y,s) - V_n(y,0)| = o_P(1).$$

The proof of this lemma is given in the last section. We shall now use this to derive the limiting distribution of the process $\mathcal{V}_n(y,\hat{\delta})$. To begin with consider the process

$$T_n(y,\vartheta) \;:=\; n^{-1/2} \sum_{i=1}^{n} \Big[\frac{Y_i}{\theta_0 + \varphi(\theta'\mathbf{Y}_{i-1})} - 1\Big] I\big(\vartheta'\mathbf{Y}_{i-1} \le y\big), \quad y \ge 0,\, \vartheta \in \mathbb{R}^p.$$

Note that $T_n(y,\theta) = \mathcal{V}_n(y,\delta_0)$. We have the following corollary.

Corollary 2.1 *Assume that the model (1.1), (1.2), and $\mathcal{H}_0$ hold, and that the given time series $Y_j, j \in \mathbb{Z}$ is stationary and ergodic. In addition, suppose that the assumptions (F), (G) and (C) hold. Then*

$$\sup_{y \ge 0} \big|T_n(y,\hat{\theta}) - \mathcal{V}_n(y,\delta_0)\big| = o_p(1). \tag{2.1}$$

Proof. In view of the assumption (C), it suffices to prove that for every $0 < c < \infty$,

$$\sup_{y \ge 0, \|s\| \le c} \big|T_n(y, \theta + n^{-1/2}s) - T_n(y,\theta)\big| = o_p(1).$$

But this follows from Lemma 2.1 upon taking

$$\eta_i \equiv [Y_i/\big(\theta_0 + \varphi(\theta'\mathbf{Y}_{i-1})\big)] - 1 \equiv \varepsilon_i - 1, \tag{2.2}$$

and $g(x,z) \equiv z$ in there, because, in this case η_i is independent of $\mathcal{F}_{i-1}$ for every $i \in \mathbb{Z}$, and $E\{g(\mathbf{Y}_{i-1}, \eta_i)|\mathcal{F}_{i-1}\} \equiv E(\eta_i) \equiv 0$, and $V_n(y,s) \equiv T_n(y, \theta + n^{-1/2}s)$.

Next, we state a weak convergence result for

$$\mathcal{V}_n(y,\delta_0) = n^{-1/2} \sum_{i=1}^{n} \eta_i I\big(\theta'\mathbf{Y}_{i-1} \le y\big),$$

where η_i's are as in (2.2). By the classical CLT, $\mathcal{V}_n(\infty, \delta_0) = n^{-1/2} \sum_{i=1}^n \eta_i = O_p(1)$.

Recall G is the d.f. of $\theta' \mathbf{Y}_0$. Let η be a copy of η_1 and $\sigma^2 = \mathrm{Var}(\eta)$. Under $\mathcal{H}_0$,

$$\text{(2.3)} \qquad \mathrm{Cov}\big(\mathcal{V}_n(y, \delta_0), \mathcal{V}_n(z, \delta_0)\big) = \sigma^2 G(y \wedge z) = \mathrm{Cov}\big(B(\sigma G(y)), B(\sigma G(z)\big), \quad y, z \geq 0,$$

where B is the standard Brownian motion on $[0, \infty)$. The following lemma proves the weak convergence of the $\mathcal{V}_n(\cdot, \delta_0)$ to the time transformed Brownian motion $B(\sigma G)$. Its proof appears in the last section.

Lemma 2.2 *Suppose (1.1), (1.2), $\mathcal{H}_0$,* (G) *hold. Then*

$$\text{(2.4)} \qquad \mathcal{V}_n(\cdot, \delta_0) \rightarrow_D \sigma B \circ G \ \textit{ in } D[0, \infty] \textit{ and uniform metric.}$$

Next, consider $\mathcal{V}_n(\cdot, \hat{\delta})$. To simplify the exposition, we let

$$h(\delta, y) := \vartheta_0 + \varphi(\vartheta' y), \quad \delta = (\vartheta_0, \vartheta')' \in \Theta, \ y \in \mathbb{R}^p.$$

Note that $\dot{h}(\delta, y) := \partial h(\delta, y)/\partial \delta = (1, y' \dot{\varphi}(\vartheta' y)'$. To state the main result about $\mathcal{V}_n(\cdot, \hat{\delta})$ we need to define

$$\text{(2.5)} \qquad \nu(y, \delta) := E\Big\{\frac{\dot{h}(\delta, \mathbf{Y}_0)}{h(\delta, \mathbf{Y}_0)} I(\vartheta' \mathbf{Y}_0 \leq y)\Big\}, \quad \delta = (\vartheta_0, \vartheta')' \in \Theta, \ y \geq 0.$$

Condition (φ) implies that $E\|\dot{h}(\delta_0, \mathbf{Y}_0)\|^2 < \infty$. This and the fact that $\inf_y h(\delta_0, y) \geq \theta_0 > 0$ imply that

$$\sup_{y \geq 0} \|\nu(y, \delta_0)\| \leq E\big\{\|\dot{h}(\delta_0, \mathbf{Y}_0)\|/h(\delta_0, \mathbf{Y}_0)\big\} < \infty.$$

We have the following theorem.

Theorem 2.1 *Suppose (1.1), (1.2), $\mathcal{H}_0$,* (F), (G), (φ) *and* (C) *hold. Then*

$$\mathcal{V}_n(y, \hat{\delta}) = \mathcal{V}_n(y, \delta_0) - n^{1/2}(\hat{\delta} - \delta_0)' \nu(y, \delta_0) + u_p(1).$$

Proof. Recall that $\eta_i \equiv [Y_i/h(\delta_0, \mathbf{Y}_{i-1})] - 1$. Consider the following decomposition.

$$\begin{aligned}
\text{(2.6)} \ \ & \mathcal{V}_n(y, \hat{\delta}) \\
&= n^{-1/2} \sum_{i=1}^n \Big[\frac{Y_i}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - 1\Big] I\Big(\hat{\theta}' \mathbf{Y}_{i-1} \leq y\Big) \\
&= T_n(y, \hat{\theta}) + n^{-1/2} \sum_{i=1}^n \eta_i h(\delta_0, \mathbf{Y}_{i-1}) \Big[\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}\Big] I\Big(\hat{\theta}' \mathbf{Y}_{i-1} \leq y\Big) \\
&\qquad + n^{-1/2} \sum_{i=1}^n h(\delta_0, \mathbf{Y}_{i-1}) \Big[\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}\Big] I\Big(\hat{\theta}' \mathbf{Y}_{i-1} \leq y\Big) \\
&= T_n(y, \hat{\theta}) + \hat{S}_n(y) + \hat{R}_n(y), \quad \text{say.} \\
&= \mathcal{V}_n(y, \delta_0) + u_p(1) + \hat{S}_n(y) + \hat{R}_n(y), \qquad \text{by (2.1).}
\end{aligned}$$

We shall now analyze the remaining two terms $\hat{R}_n$ and $\hat{S}_n$. Let $\Delta := n^{1/2}(\hat{\delta}-\delta_0)$, and

$$\xi_{ni} := h(\hat{\delta}, \mathbf{Y}_{i-1}) - h(\delta_0, \mathbf{Y}_{i-1}) - (\hat{\delta}-\delta_0)'\dot{h}(\delta_0, \mathbf{Y}_{i-1}),\ 1 \le i \le n.$$

By conditions (φ), (C) and Ergodic Theorem, $\theta_0^{-1} < \infty$, and

$$\begin{aligned}(2.7)\qquad n^{1/2}\max_{1\le i\le n}|\xi_{ni}| &= o_p(1), \qquad n^{-1/2}\max_{1\le i\le n}\|\mathbf{Y}_{i-1}\dot{\varphi}(\theta'\mathbf{Y}_{i-1})\| = o_p(1),\\
n^{-1/2}\max_{1\le i\le n}\|\dot{h}(\delta_0, \mathbf{Y}_{i-1})\| &\le n^{-1/2}\big\{1+\max_{1\le i\le n}\|\mathbf{Y}_{i-1}\dot{\varphi}(\theta'\mathbf{Y}_{i-1})\|\big\} = o_p(1),\\
\max_{1\le i\le n}\Big|\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}\Big| &\le \theta_0^{-2}\max_{1\le i\le n}|h(\hat{\delta}, \mathbf{Y}_{i-1}) - h(\delta_0, \mathbf{Y}_{i-1})|\\
&\le \theta_0^{-2}\Big[\max_{1\le i\le n}|\xi_{ni}| + \|\Delta\|n^{-1/2}\max_{1\le i\le n}\|\dot{h}(\delta_0, \mathbf{Y}_{i-1})\|\Big]\\
&= o_p(1),\\
n^{-1}\sum_{i=1}^{n}\|\mathbf{Y}_{i-1}\dot{\varphi}(\theta'\mathbf{Y}_{i-1})\| &= E\|\mathbf{Y}_0\dot{\varphi}(\theta'\mathbf{Y}_0)\| + o_p(1).\end{aligned}$$

Now consider the $\hat{R}_n$ term. Rewrite

$$\begin{aligned}(2.8)\quad -\hat{R}_n(y) \;:=\; & n^{-1/2}\sum_{i=1}^{n}\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})}\xi_{ni}I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
&+\Delta' n^{-1}\sum_{i=1}^{n}\Big[\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}\Big]\dot{h}(\delta_0, \mathbf{Y}_{i-1})I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
&+\Delta' n^{-1}\sum_{i=1}^{n}\frac{\dot{h}(\delta_0, \mathbf{Y}_{i-1})}{h(\delta_0, \mathbf{Y}_{i-1})}I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
=\; & \hat{R}_{n1}(y) + \Delta'\hat{R}_{n2}(y) + \Delta'\hat{R}_{n3}(y), \qquad \text{say.}\end{aligned}$$

Then by (C) and because $\theta_0 > 0$, for sufficiently large n, $P(2\hat{\theta}_0 > \theta_0) \to 1$. This fact and $\varphi \ge 0$ imply that on the event $A := \{\hat{\theta}_0 > \theta_0/2\}$, $\min_{1\le i\le n}h(\hat{\delta}, \mathbf{Y}_{i-1}) > 2^{-1}\theta_0$. Hence, in view of (2.7), we obtain that on the event A,

$$(2.9)\qquad \sup_{y\ge 0}|\hat{R}_{n1}(y)| \le 2\theta_0^{-1}n^{1/2}\max_{1\le i\le n}|\xi_{ni}| = o_p(1).$$

Similarly, by (2.7),

$$\begin{aligned}(2.10)\quad \sup_{y\ge 0}|\hat{R}_{n2}(y)| &\le \max_{1\le i\le n}\Big|\frac{1}{h(\hat{\delta}, \mathbf{Y}_{i-1})} - \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}\Big| n^{-1}\sum_{i=1}^{n}\{1+\|\mathbf{Y}_{i-1}\dot{\varphi}(\theta'\mathbf{Y}_{i-1})\|\}\\
&= o_p(1).\end{aligned}$$

Upon combining (2.8), (2.9), (2.10) with (3.4) of Section 3 below, we obtain

$$(2.11)\qquad \hat{R}_n(y) = -n^{1/2}(\hat{\delta}-\delta_0)'\nu(y, \delta_0) + u_p(1).$$

Next, consider the term $\hat{S}_n$. Rewrite

$$\begin{aligned}
-\hat{S}_n(y) \;&:=\; n^{-1/2}\sum_{i=1}^{n}\frac{\eta_i}{h(\hat{\delta},\mathbf{Y}_{i-1})}\xi_{ni}I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
&\quad+\Delta'\frac{1}{n}\sum_{i=1}^{n}\Big[\frac{1}{h(\hat{\delta},\mathbf{Y}_{i-1})}-\frac{1}{h(\delta_0,\mathbf{Y}_{i-1})}\Big]\eta_i\,\dot{h}(\delta_0,\mathbf{Y}_{i-1})I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
&\quad+\Delta'\frac{1}{n}\sum_{i=1}^{n}\eta_i\frac{\dot{h}(\delta_0,\mathbf{Y}_{i-1})}{h(\delta_0,\mathbf{Y}_{i-1})}I\Big(\hat{\theta}'\mathbf{Y}_{i-1}\le y\Big)\\
&=\; \hat{S}_{n1}(y)+\Delta'\hat{S}_{n2}(y)+\Delta'\hat{S}_{n3}(y), \qquad \text{say.}
\end{aligned}$$

Because $E|\eta_0|<\infty$, $n^{-1}\sum_{i=1}^{n}|\eta_i|=O_p(1)$, and by (2.7), on A,

$$\sup_{y\ge 0}\big|\hat{S}_{n1}(y)\big|\le 2\theta_0^{-1}n^{1/2}\max_{1\le i\le n}|\xi_{ni}|\,n^{-1}\sum_{i=1}^{n}|\eta_i|=o_p(1).$$

An argument similar to the one used to prove (2.10) yields $\sup_{y\ge0}|\hat{S}_{n2}(y)|=o_p(1)$. Upon combining these facts with (3.5) below we obtain $\sup_{y\ge0}|\hat{S}_n(y)|=o_p(1)$. This fact together with (2.11), (2.6) and (2.1) completes the proof of the theorem.

2.2 Asymptotically distribution free tests

A consequence of Theorem 2.1 is that the asymptotic null distribution of the process $\mathcal{V}_n(\cdot,\hat{\delta})$ depends on the null model and estimator, and hence in general is unknown. We shall now describe a transformation of this process which will yield a process with known asymptotic null distribution. This transformation is an analog of the one in Stute, Thies and Zhu (1998), Koul and Stute (1999) and Stute and Zhu (2002).

Let G_ϑ denote the d.f. of $\vartheta'\mathbf{Y}_0$. Note that $G_\theta=G$. Recall that $h(\delta_0,y):=\theta_0+\varphi(\theta'y)$, $\dot{h}(\delta_0,y)=\big(1,y'\dot{\varphi}(\theta'y)\big)'$ and $\nu(y,\delta)$ given at (2.5). Let $\xi_\vartheta(x)=E\big(\mathbf{Y}_0|\vartheta'\mathbf{Y}_0=x\big)$, $x\ge0$ and

$$L(x,\delta)\;:=\;E\Big(\frac{\dot{h}(\delta,\mathbf{Y}_0)}{h(\delta_0,\mathbf{Y}_0)}\Big|\vartheta'\mathbf{Y}_0=x\Big)=\frac{1}{\vartheta_0+\varphi(x)}\begin{pmatrix}1\\ \xi_\vartheta(x)\dot{\varphi}(x)\end{pmatrix}$$

Then $E\{\mathbf{Y}_0\varphi(\vartheta'\mathbf{Y}_0)|\vartheta'\mathbf{Y}_0=x)\}=\xi_\vartheta(x)\varphi(x)$, and

$$\nu(y,\delta)\;=\;E\Big(\frac{\dot{h}(\delta,\mathbf{Y}_{i-1})}{h(\delta,\mathbf{Y}_{i-1})}\Big)I(\theta'\mathbf{Y}_{i-1}\le y)=\int_{x\le y}L(x,\delta)dG_\vartheta(x).$$

Define

$$A(x,\delta):=\int_x^\infty \nu(y,\delta)\nu(y,\delta)'dG_\vartheta(y).$$

Assume

$$A(x, \delta_0) \text{ is positive definite for all } 0 \le x < \infty. \tag{2.12}$$

Write $A_\delta^{-1}(x)$ for $(A(x,\delta))^{-1}$. Define, for $y \ge 0$, $\delta \in \Theta$,

$$\begin{aligned} J_n(y,\delta) &:= \int_0^y \nu(z,\delta)' A_\delta^{-1}(z) \int_z^\infty \nu(s,\delta) d\mathcal{V}_n(s,\delta) dG_\vartheta(z), \\ \mathcal{W}_n(y,\delta) &:= \mathcal{V}_n(y,\delta) - J_n(y,\delta). \end{aligned}$$

We note here that by arguing as in Khmaladze and Koul (2009), one can verify that the transformation $\mathcal{W}_n(\cdot,\delta_0)$ is well defined even when (2.12) does not hold.

Next, suppose $\mathcal{H}_0$ holds. Then clearly, $E\mathcal{W}_n(y,\delta_0) \equiv 0$, and more importantly,

$$\mathrm{Cov}\Big(\mathcal{W}_n(y_1,\delta_0), \mathcal{W}_n(y_2,\delta_0)\Big) = \sigma^2 G(y_1 \wedge y_2), \quad y_1, y_2 \ge 0. \tag{2.13}$$

The proof of this claim appears in the last section. In fact, arguing as in say Koul and Stute (1999) and using Lemma 2.4, one can show that $\mathcal{W}_n(\cdot,\delta_0)$ converges weakly to $\sigma^2 B \circ G$, in $D[0,\infty)$ and uniform metric. However, this result is of little use from the practical point of view, since this process depends on the unknown parameters θ_0, θ and G. We need the analog of the above process when all parameters are estimated. The needed estimated entities are defined below.

Analogous to the regression set up of Stute and Zhu (2002), an estimate of ξ_θ here can be taken to be a kernel estimator. Because here observations are nonnegative, it is desirable to take kernels that are supported on $[0,\infty)$. One such example is that of the inverse gamma kernel of Mnatsakanov and Sarkisian (2012) defined as

$$K_\alpha(y,u) := \frac{1}{u\Gamma(\alpha+1)} (\frac{\alpha y}{u})^{\alpha+1} \exp\{-(\frac{\alpha y}{u})\}, \quad \alpha > 0, y > 0, u > 0.$$

The corresponding estimator of ξ_θ then is

$$\tilde{\xi}_n(y) := \frac{\sum_{i=1}^n K_\alpha(y, \hat{\theta}'\mathbf{Y}_{i-1}) Y_i}{\sum_{i=1}^n K_\alpha(y, \hat{\theta}'\mathbf{Y}_{i-1})}, \quad y > 0.$$

Now we are ready to define estimates of the other entities needed in the transformation. Let

$$\begin{aligned} G_n(y) &:= n^{-1} \sum_{i=1}^n I(\hat{\theta}'\mathbf{Y}_{i-1} \le y), \quad L_n(y) := \frac{1}{\hat{\theta}_0 + \varphi(y)} \begin{pmatrix} 1 \\ \tilde{\xi}_n(y)\dot{\varphi}(y) \end{pmatrix}, \\ \nu_n(y) &:= \int_{x \le y} L_n(x) dG_n(x), \qquad A_n(y) := \int_y^\infty \nu_n(z)\nu_n'(z) dG_n(z), \\ \widetilde{J}_n(y) &:= \int_0^y \nu_n(z)' A_n^{-1}(z) \int_z^\infty \nu_n(s) d\mathcal{V}_n(s,\hat{\theta}) dG_n(z), \\ \widetilde{\mathcal{W}}_n(y) &:= \mathcal{V}_n(y,\hat{\theta}) - \widetilde{J}_n(y), \qquad y > 0. \end{aligned} \tag{2.14}$$

Again, using the arguments similar to those used in Koul and Stute (1999) or Khmaladze and Koul (2004, 2009) and the tightness of the process $\mathcal{V}_n(\cdot,\theta)$ proved in Lemma 2.4 here, we can prove that $\widetilde{\mathcal{W}}_n(y,\hat{\delta}_n) \to_D \sigma^2 B \circ G$, in $D[0,\infty)$ and uniform metric. Consequently, tests based on continuous functionals of the process $\widetilde{\mathcal{W}}_n(y,\hat{\delta}_n), y \geq 0$ will be asymptotically distribution free. In particular, the test that rejects $\mathcal{H}_0$ whenever

$$\hat{\sigma}_n^{-1} \sup_{0 \leq y < \infty} |\widetilde{\mathcal{W}}_n(y,\hat{\theta})| > b_\alpha$$

is of the asymptotic size α, where b_α be such that $P(\sup_{0\leq t\leq 1} |B(t)| > b_\alpha) = \alpha$ and

$$\hat{\sigma}_n^2 := n^{-1} \sum_{i=1}^n \Big(\frac{Y_i}{\hat{\theta}_0 + \varphi(\hat{\theta}'\mathbf{Y}_{i-1})} - 1 \Big)^2.$$

A computational formula for the above process is as follows. Let $X_{i-1} = \hat{\theta}'\mathbf{Y}_{i-1}$, $\hat{\eta}_i = \big\{Y_i/[\hat{\theta}_0 + \varphi(X_{i-1})]\big\} - 1$. Then

$$\begin{aligned}
\widetilde{J}_n(y) &= n^{-3/2} \sum_{j=1}^n \hat{\eta}_j \nu_n(X_{j-1})' \sum_{i=1}^n A_n^{-1}(X_{i-1}) \nu_n(X_{i-1}) I(X_{j-1} \leq X_{i-1} \wedge y), \\
\widetilde{\mathcal{W}}_n(y) &= n^{-1/2} \sum_{j=1}^n \hat{\eta}_j \Big\{ I(X_{j-1} \leq y) \\
&\qquad -n^{-1} \sum_{i=1}^n \nu_n(X_{j-1})' A_n^{-1}(X_{i-1}) \nu_n(X_{i-1}) I(X_{j-1} \leq X_{i-1} \wedge y) \Big\}.
\end{aligned}$$

As an example consider the problem of fitting an autoregressive conditional duration (ACD$(p,0)$) model given by

$$\begin{aligned}
Y_i &= \psi(\mathbf{Y}_{i-1},\delta_0)\varepsilon_i, \qquad \psi(\mathbf{Y}_{i-1},\delta_0) = \theta_0 + \varphi(\theta'\mathbf{Y}_{i-1}), \\
\varphi(\theta'\mathbf{Y}_{i-1}) &:= \theta_1 Y_{i-1} + \theta_2 Y_{i-2} + \cdots + \theta_p Y_{i-p},
\end{aligned}$$

where $\{\varepsilon_i\}$ is a sequence of iid non-negative r.v.'s with unit mean and finite variance. The resulting sequence $\{Y_i\}$ defined by the above ACD$(p,0)$ is a p^{th} order Markov sequence, which is weakly stationary if $\theta_1 + \theta_2 + \cdots + \theta_p < 1$. The unconditional mean is given by $E(Y_i) = \theta_0/[1 - \theta_1 - \theta_2 - \ldots - \theta_p]$, which requires θ_0 to be positive.

In general the rv ε_i could follow any continuous distribution on the non-negative support with required moment conditions. If, for example, we assume that ε_i follows a unit exponential distribution, then the conditional pdf of Y_i, given $\mathbf{Y}_{i-1}$ is

$$f(y_i|\mathbf{Y}_{i-1}) = \frac{1}{\psi(\mathbf{Y}_{i-1},\delta_0)} e^{-y_i/\psi(\mathbf{Y}_{i-1},\delta_0)}.$$

One can use likelihood based on this density to estimate δ.

In this example $\varphi(x) \equiv x$ and the condition (φ) is *a priori* satisfied with $\dot{\varphi} \equiv 1$. Thus the above testing procedure is applicable for fitting ACD$(p, 0)$ model to the given time series.

Similarly it is also applicable to fit an ARCH(p) model given by

$$X_i = \sigma_i \, \zeta_i, \qquad \sigma_i^2 = \theta_0 + \theta_1 X_{i-1}^2 + \theta_2 X_{i-2}^2 + \cdots + \theta_p X_{i-p}^2,$$

where $\{\zeta_i\}$ is a sequence of iid symmetric rvs with mean zero and finite variance. If we define $Y_i = X_i^2$ and $\zeta_i^2 = \varepsilon_i$ and $\sigma_i^2 = \psi_i$ then the ARCH(p) model becomes an ACD(p,0) model. Hence the conditions for stationarity and other distributional properties will follow similarly. In particular if we assume standard normal distribution for ζ_i then Y_i becomes an ACD model with Chi-square-1 innovations.

3 Proofs

This section contains the proof of Lemma 2.1, which is facilitated by the following preliminary result. Let γ be a measurable function from $[0, \infty)^p \times \mathbb{R}$ to $\mathbb{R}$, $\eta_i := [\varepsilon_i / \varphi(\theta' \mathbf{Y}_{i-1})] - 1$, $i \in \mathbb{Z}$, and define, for $y \in \mathbb{R}^+, s \in \mathbb{R}^p, b \in \mathbb{R}$,

$$\mathcal{H}_n(y, s, b) := n^{-1} \sum_{i=1}^{n} \gamma(\mathbf{Y}_{i-1}, \eta_i) I\big((\theta + n^{-1/2} s)' \mathbf{Y}_{i-1} \leq y + b n^{-1/2} \|\mathbf{Y}_{i-1}\|\big),$$

$$H_n(y) := n^{-1} \sum_{i=1}^{n} \gamma(\mathbf{Y}_{i-1}, \eta_i) I\big(\theta' \mathbf{Y}_{i-1} \leq y\big), \quad H(y) := E\gamma(\mathbf{Y}_0, \eta_1) I\big(\theta' \mathbf{Y}_0 \leq y\big).$$

We are now ready to state the first preliminary result.

Lemma 3.1 *Let $\{Y_i, i \in \mathbb{Z}\}$ be a stationary and ergodic time series satisfying $E\|\mathbf{Y}_0\|^2 < \infty$, (1.1), (1.2), and having continuous stationary distribution. Let γ be a non-negative function on $[0, \infty)^p \times \mathbb{R}$ satisfying $E\gamma(\mathbf{Y}_0, \eta_1) < \infty$. Then, for every $0 < c < \infty, b \in \mathbb{R}$,*

$$\sup_{y \geq 0, \|s\| \leq c} \big|\mathcal{H}_n(y, s, b) - H(y)\big| = o_p(1). \tag{3.1}$$

Proof. By using a Glivenko-Cantelli type argument where $[0, \infty]$ is totally bounded by the nondecreasing bounded function $H(y)$, and by the Ergodic Theorem,

$$\sup_{y \geq 0} |H_n(y) - H(y)| = o(1), \quad \text{a.s.} \tag{3.2}$$

Stationarity and the assumption $E\|Y_0\|^2 < \infty$ imply

$$\max_{1 \leq i \leq n} n^{-1/2} \|\mathbf{Y}_{i-1}\| = o_p(1). \tag{3.3}$$

Thus for every $\epsilon > 0$ there is an N_ϵ such that $P(A_n) > 1 - \epsilon$, for all $n > N_\epsilon$, where $A_n := \{\max_{1 \le i \le n} n^{-1/2} \|\mathbf{Y}_{i-1}\| \le \epsilon\}$. But, on A_n, for all $\|s\| \le c$, $b \in \mathbb{R}$,

$$H_n(y - (c+b)\epsilon) \le \mathcal{H}_n(y, s, b) \le H_n(y + (c+b)\epsilon), \quad \forall\, y \ge 0.$$

The claim (3.1) follows from these inequalities, (3.2), and the continuity of the stationary distribution, thereby completing the proof of the lemma.

Next, consider, for $y \ge 0$,

$$\hat{R}_{n3}(y) := n^{-1} \sum_{i=1}^{n} \frac{\dot{h}(\delta_0, \mathbf{Y}_{i-1})}{h(\delta_0, \mathbf{Y}_{i-1})} I\Big(\hat{\theta}' \mathbf{Y}_{i-1} \le y\Big), \quad \hat{S}_{n3}(y) := \frac{1}{n} \sum_{i=1}^{n} \eta_i \frac{\dot{h}(\delta_0, \mathbf{Y}_{i-1}}{h(\delta_0, \mathbf{Y}_{i-1}} I\{\hat{\theta}' \mathbf{Y}_{i-1} \le y\}.$$

We have

Corollary 3.1 *Under the assumptions of Theorem 2.1,*

$$\sup_{y \ge 0} \Big\|\hat{R}_{n3}(y) - \nu(y, \delta_0)\Big\| = o_p(1). \tag{3.4}$$

$$\sup_{y \ge 0} \Big\|\hat{S}_{n3}(y)\Big\| = o_p(1). \tag{3.5}$$

Proof. *Proof of* (3.4). Let

$$\begin{aligned}
r_0(\delta_0, \mathbf{Y}_{i-1}) &:= \frac{1}{h(\delta_0, \mathbf{Y}_{i-1})}, \qquad r_j(\delta_0, \mathbf{Y}_{i-1}) := Y_{i-j} \frac{\dot{\varphi}(\theta' \mathbf{Y}_{i-1})}{h(\delta_0, \mathbf{Y}_{i-1})}, \quad j = 1, \cdots, p, \\
T_j(y, s) &:= n^{-1} \sum_{i=1}^{n} r_j(\delta_0, \mathbf{Y}_{i-1}) I\{(\theta + n^{-1/2} s)' \mathbf{Y}_{i-1} \le y\}, \quad j = 0, 1, \cdots, p, \\
T(y, s) &:= (T_0(y, s), T_1(y, s), \cdots, T_p(y, s))', \quad y \ge 0,\ s \in \Theta_1.
\end{aligned}$$

Then with $s = n^{1/2}(\hat{\theta} - \theta)$, $\hat{R}_{n3}(y) = T(y, s)$. Thus, in view of the assumption (C), to prove (3.4), it suffices to show that for every $0 < c < \infty$,

$$\sup_{y \ge 0, \|s\| \le c} \Big\|T(y, s) - \nu(y, \delta_0)\Big\| = o_p(1). \tag{3.6}$$

But note that the jth coordinates of $T(y, s)$ and $\nu(y, \delta_0)$ are like the $\mathcal{H}(y, s, 0)$ and $H(y)$ of Lemma 3.1 with $\gamma = r_j$, $j = 0, \cdots, p$, respectively. Thus (3.6) follows from Lemma 3.1 upon writing $r_j = r_j^+ - r_j^-$ and applying that lemma with $\gamma = r_j^\pm$ for each $j = 0, \cdots, p$.

Proof of (3.5). Follows similarly from Lemma 3.1 upon applying it with the gamma functions $\gamma_0(x, z) := z / h(\delta_0, x)$, $\gamma_j(x, z) := (\dot{\varphi}(\theta' x) / h(\delta_0, x))\, z\, x_j$, $j = 1, \cdots, p$. In this case the corresponding H functions are as follows.

$$\begin{aligned}
H_0(y) &= E\gamma_0(\mathbf{Y}_0, \eta_1) I\{\theta' \mathbf{Y}_0 \le y\} = E\Big(\frac{1}{h(\delta_0, \mathbf{Y}_0)} I\{\theta' \mathbf{Y}_0 \le y\}\Big) E(\eta_1) = 0, \\
H_j(y) &= E \frac{\dot{\varphi}(\theta' \mathbf{Y}_0)}{h(\delta_0, \mathbf{Y}_0)} \eta_1 \mathbf{Y}_0 I\{\theta' \mathbf{Y}_0 \le y\} \equiv 0,
\end{aligned}$$

because of the independence of η_1 and $\mathbf{Y}_0$ and because $E\eta_1 = 0$. This completes the proof of the corollary.

Proof of Lemma 2.1. We will follow the structure of the proof of uniform Glivenko-Cantelli theorems in van der Vaart and Wellner (1997) (VW). Accordingly, let $\xi_i, i \in \mathbb{Z}$ be an independent copy of the errors $\eta_i, i \in Z$ and $\zeta, \zeta_i, i \in \mathbb{Z}$, be i.i.d. r.v.'s, independent of $Y_i, \eta_i, i \in \mathbb{Z}$, with $P(\zeta = 1) = P(\zeta_i = -1) = 1/2$. For any r.v. Z, let P_Z (E_Z) denote the conditional probability distribution (expectation) of Z, given all other r.v.'s. Then

$$\begin{aligned}
& E_\eta \sup_{y\geq 0, \|s\|\leq c} |V_n(y,s) - V_n(y,0)| \\
= \; & E_\eta \sup_{y\geq 0, \|s\|\leq c} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^n \Big\{ [g(\mathbf{Y}_{i-1}, \eta_i) - E_\xi\{g(\mathbf{Y}_{i-1}, \xi_i)|\mathcal{F}_{i-1}\}] \\
& \qquad \times [I((\theta + n^{-1/2}s)'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)] \Big\} \Bigg| \\
\leq \; & E_\eta E_\xi \Big\{ \sup_{y\geq 0, \|s\|\leq c} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^n [g(\mathbf{Y}_{i-1}, \eta_i) - g(\mathbf{Y}_{i-1}, \xi_i)] \\
& \qquad \times [I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)] \Bigg| \Big| \mathcal{F}_0 \Big\} \\
= \; & E_\zeta E_\eta E_\xi \Big\{ \sup_{y\geq 0, \|s\|\leq c} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i [g(\mathbf{Y}_{i-1}, \eta_i) - g(\mathbf{Y}_{i-1}, \xi_i)] \\
& \qquad \times [I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)] \Bigg| \Big| \mathcal{F}_0 \Big\} \\
\leq \; & 2 E_\zeta E_\eta \Big\{ \sup_{y\geq 0, \|s\|\leq c} \Bigg| \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i g(\mathbf{Y}_{i-1}, \eta_i) \\
& \qquad \times [I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)] \Bigg| \Big\}.
\end{aligned}$$

Now note that the process

$$W(y,s) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i g(\mathbf{Y}_{i-1}, \eta_i)[I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)]$$

is a Rademacher process X_f indexed by the functions

$$f = \frac{1}{\sqrt{n}} g(\mathbf{Y}_{i-1}, \eta_i)[I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \leq y) - I(\theta'\mathbf{Y}_{i-1} \leq y)], \; i = 1, \ldots, n.$$

We shall follow the proof of 2.4.3 in VW.

Accordingly, let P_n be the empirical measure based on $(\mathbf{Y}_{i-1}, \eta_i), i = 1, ..., n$. For any $\epsilon > 0$, let $\mathcal{G}$ be an ϵ-net in $L_1(P_n)$ over

$$
\begin{aligned}
\mathcal{F} \;=\; & \Big\{ f : f(z,\eta) = g(z,\eta)[I((\theta + \frac{s}{\sqrt{n}})'z \le y) - I(\theta' z \le y)] : \\
& \qquad\qquad\qquad\qquad z \in [0,\infty)^p, s \in \mathbb{R}^p,\ y \ge 0, \|s\| \le c \Big\}.
\end{aligned}
$$

Then

$$
\sup_{y \ge 0, \|s\| \le c} |W(y,s)| = \sup_{f \in \mathcal{F}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i f(\mathbf{Y}_{i-1}, \eta_i) \right|.
$$

Moreover,

$$
E_\delta \sup_{f \in \mathcal{F}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i f(\mathbf{Y}_{i-1}, \eta_i) \right| \le E_\delta \sup_{f \in \mathcal{G}} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \zeta_i f(\mathbf{Y}_{i-1}, \eta_i) \right| + \epsilon.
$$

Using a maximal inequality with the appropriate Orlicz norm for $e^{x^2} - 1$ taken over $\zeta_1, \cdots, \zeta_n$ with the rest of variables fixed and using Hoeffding's inequality for the Rademacher process as on pages 123–124 in VW, we bound the above expectation further by

$$
C(1 + \log N(\epsilon, \mathcal{F}, L_1(P_n)))^{1/2} \sup_{f \in \mathcal{G}} \left(\frac{1}{n} \sum_{i=1}^n f^2(\mathbf{Y}_{i-1}, \eta_i) \right)^{1/2} + \epsilon,
$$

where $\{y \ge 0, \|s\| \le c\}$ is a VC class, so the cardinality of $\mathcal{G}$ is $N(\epsilon, \mathcal{F}, L_1(P_n)) \le C_V \epsilon^{-(V-1)}$ with a finite VC dimension $V > 1$. The last supremum above can be further bounded from the above by

$$
\begin{aligned}
& \sup_{f \in \mathcal{G}} \left(\frac{1}{n} \sum_{i=1}^n f^2(\mathbf{Y}_{i-1}, \eta_i) \right) \\
&= \sup_{y \ge 0, \|s\| \le c} \left(\frac{1}{n} \sum_{i=1}^n g^2(\mathbf{Y}_{i-1}, \eta_i) \Big[I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \le y) - I(\theta'\mathbf{Y}_{i-1} \le y) \Big]^2 \right) \\
&= \sup_{y \ge 0, \|s\| \le c} \left(\frac{1}{n} \sum_{i=1}^n g^2(\mathbf{Y}_{i-1}, \eta_i) | I((\theta + \frac{s}{\sqrt{n}})'\mathbf{Y}_{i-1} \le y) - I(\theta'\mathbf{Y}_{i-1} \le y) | \right) \\
&\le \sup_{y \ge 0, \|s\| \le c} \left(\frac{1}{n} \sum_{i=1}^n g^2(\mathbf{Y}_{i-1}, \eta_i) I(\theta'\mathbf{Y}_{i-1} \text{ is between } y, y - \frac{s}{\sqrt{n}}'\mathbf{Y}_{i-1}) \right) \\
&\le \sup_{y \ge 0} \left(\frac{1}{n} \sum_{i=1}^n g^2(\mathbf{Y}_{i-1}, \eta_i) I(\theta'\mathbf{Y}_{i-1} \in [y - \frac{c\|\mathbf{Y}_{i-1}\|}{\sqrt{n}}, y + \frac{c\|\mathbf{Y}_{i-1}\|}{\sqrt{n}}]) \right) \\
&= o_p(1).
\end{aligned}
$$

The last claim follows from Lemma 3.1 because the process in the last but one bound equals $\mathcal{H}_n(y, 0, c) - \mathcal{H}_n(y, 0, -c)$ of Lemma 3.1 with $\gamma = g^2$.

Proof of Lemma 2.2. Apply the CLT for martingales (Hall and Heyde: 1980, Corollary 3.1) to show that the fidis tend to the right limit.

The proof of tightness is given in the following two lemmas, where $\sigma^2 = E\eta^2$. Since δ_0 is fixed, we shall write $\mathcal{V}_n(y)$ for $\mathcal{V}_n(y, \delta_0)$ in this proof. Let $Z_i := \theta'\mathbf{Y}_{i-1}$, and $\eta'_i, i \in \mathbb{Z}$ be an independent copy of η_i. Let

$$\mathcal{V}'_n(z) = n^{-1/2} \sum_{i=1}^{n} \eta'_i I(Z_{i-1} \leq z), \quad z \geq 0.$$

Recall the definition of G_n from (2.14). Upon taking $\gamma \equiv 1, b = 0, s = 0$ in (3.1), we see that $\mathcal{H}_n(y, 0, 0) \equiv G_n(y)$, $H(y) \equiv G(y)$. Hence from (3.1), we obtain

$$\lim_n \sup_z |G_n(z) - G(z)| = 0, \quad \text{a.s.} \tag{3.7}$$

Let $\mathbb{R}^+ := [0, \infty]$, $\rho(y, z) = |G(y) - G(z)|, y, z \in \mathbb{R}^+$, and, for $f : \mathbb{R}^+ \to \mathbb{R}$,

$$\|f\|_\delta = \sup_{y,z \in \mathbb{R}^+ : \rho(y,z) < \delta} |f(y) - f(z)|.$$

We are now ready to state an inequality useful in proving the tightness of $\mathcal{V}_n$.

Lemma 3.2 (Symmetrization Lemma): *For $\epsilon, \gamma, \delta > 0$ fixed with $\epsilon < \gamma$ and $\delta < \epsilon^2/3\sigma^2$,*

$$P(\|\mathcal{V}_n\|_\delta > \gamma) \leq (1 - 3\delta\sigma^2/\epsilon^2)^{-1} P(\|\mathcal{V}_n - \mathcal{V}'_n\|_\delta > \gamma - \epsilon) + P(\sup_{k \geq n} \sup_{z \in \mathbb{R}^+} |G_k(z) - G(z)| > \delta).$$

Proof: Let $P_{\eta'}$ and $E_{\eta'}$ denote the conditional probability measure and expectation, respectively, given $\eta_i, Z_i, i \in \mathbb{Z}$. Let $\rho_n(y, z) := |G_n(y) - G_n(z)|$, and

$$B_n(\gamma; y, z) := \Big[|\mathcal{V}_n(y) - \mathcal{V}_n(z)| > \gamma\Big], \qquad B_n(\gamma, \delta) := \Big[\|\mathcal{V}_n\|_\delta > \gamma\Big].$$

Note

$$B_n(\gamma, \delta) = \cup_{y,z \in \mathbb{R}^+ : \rho(y,z) < \delta} B_n(\gamma; y, z).$$

For fixed $y, z \in \mathbb{R}^+$ with $\rho(y, z) < \delta$,

$$B_n(\gamma; y, z) \cap \Big[\|\mathcal{V}_n - \mathcal{V}'_n\|_\delta > \gamma - \epsilon\Big] \supset B_n(\gamma; y, z) \cap \Big[|\mathcal{V}'_n(y) - \mathcal{V}'_n(z)| < \epsilon\Big].$$

Since $E_{\eta'}(\mathcal{V}'_n(y) - \mathcal{V}'_n(z)) = 0$ and $E_{\eta'}(\mathcal{V}'_n(y) - \mathcal{V}'_n(z))^2) = \sigma^2 \rho_n(y, z)$, Chebyshev's inequality gives $P_{\eta'}\big(|\mathcal{V}'_n(y) - \mathcal{V}'_n(z)| < \epsilon\big) \geq 1 - \sigma^2 \rho_n(y, z)/\epsilon^2$. Then

$$\begin{aligned} I\big(B_n(\gamma; y, z)\big) P_{\eta'}\Big(\|\mathcal{V}_n - \mathcal{V}'_n\|_\delta > \gamma - \epsilon\Big) &\geq I(B_n(\gamma; y, z)) P_{\eta'}\Big(|\mathcal{V}'_n(y) - \mathcal{V}'_n(z)| < \epsilon\Big) \\ &\geq I(B_n\big(\gamma; y, z)\big)\big(1 - \sigma^2 \rho_n(y, z)/\epsilon^2\big). \end{aligned}$$

Let

$$A_n(\delta) = \cup_{k \geq n} \Big[\sup_{z \in \mathbb{R}^+} |G_k(z) - G(z)| > \delta\Big].$$

Using the triangle inequality, on $A_n^c(\delta)$, $\rho_k(y,z) \le 2\delta + \rho(y,z)$ for all $k \ge n$, $y,z \in \mathbb{R}^+$. Then for $y,z \in \mathbb{R}^+$ with $\rho(y,z) < \delta$,

$$\begin{aligned} & I\big(B_n(\gamma; y,z) \cap A_n^c(\delta)\big) P_{\eta'}\Big(\|\mathcal{V}_n - \mathcal{V}_n'\|_\delta > \gamma - \epsilon\Big) \\ \ge\ & I\big(B_n(\gamma; y,z) \cap A_n^c(\delta)\big)\big(1 - \sigma^2 \rho_n(y,z)/\epsilon^2\big) \\ \ge\ & I\big(B_n(\gamma; y,z) \cap A_n^c(\delta)\big)\big(1 - 3\sigma^2\delta/\epsilon^2\big). \end{aligned}$$

This, in turn, gives

$$\begin{aligned} I\big(B_n(\gamma,\delta) \cap A_n^c(\delta)\big)\big(1 - 3\sigma^2\delta/\epsilon^2\big) &\le I\big(B_n(\gamma,\delta) \cap A_n^c(\delta)\big) P_{\eta'}\Big(\|\mathcal{V}_n - \mathcal{V}_n'\|_\delta > \gamma - \epsilon | \mathcal{F}\Big) \\ &\le P_{\eta'}\Big(\|\mathcal{V}_n - \mathcal{V}_n'\|_\delta > \gamma - \epsilon\Big) \end{aligned}$$

and

$$\begin{aligned} I\big(B_n(\gamma,\delta)\big) &\le I\big(B_n(\gamma,\delta) \cap A_n^c(\delta)\big) + I(A_n(\delta)) \\ &\le \big(1 - 3\sigma^2\delta/\epsilon^2\big)^{-1} P_{\eta'}\Big(\|\mathcal{V}_n - \mathcal{V}_n'\|_\delta > \gamma - \epsilon\Big) + I(A_n(\delta)\Big). \end{aligned}$$

Now take the expectation to obtain

$$P\big(\|\mathcal{V}_n\|_\delta > \gamma\big) \le (1 - 3\delta\sigma^2/\epsilon^2)^{-1} P\big(\|\mathcal{V}_n - \mathcal{V}_n'\|_\delta > \gamma - \epsilon\big) + P(A_n(\delta)).$$

From the above inequality it thus suffices to prove the tightness of the given process when η_i are symmetric around zero, which is done in the next lemma.

Lemma 3.3 *Suppose η is symmetrically distributed around 0 and $E\eta^2 < \infty$. Then for any fixed δ_0 the process $\mathcal{V}_n(\cdot,\delta_0)$ is sub-Gaussian with the semi-metric $d(y,z) = (\sigma^2\rho(y,z))^{1/2}$, $y,z \in [0,\infty]$, that is, for any $y,z,x > 0$,*

$$\limsup_n P\Big(|\mathcal{V}_n(z,\delta_0) - \mathcal{V}_n(y,\delta_0)| > x\Big) \le 2\exp\Big(-\frac{x^2}{2d^2(y,z)}\Big). \tag{3.8}$$

Proof. Without the loss of generality assume $y < z$. As before, let $Z_{i-1} := \theta'\mathbf{Y}_{i-1}$. Recall $Z_{i-1}, i \ge 1$ is a stationary process and G denotes the d.f. of Z_0.

Let $\zeta, \zeta_i, i = 1, \cdots$, P_ζ and E_ζ be as in the proof of Lemma 2.1. Then by the assumed symmetry of $\{\eta_i\}$, $\eta_i = \zeta_i|\eta_i|, i \ge 1$, in distribution. Then, by the Hoeffding (1964) inequality, $\forall\, x > 0, \forall\, n \ge 1$,

$$\begin{aligned} P_\zeta\Big\{|\mathcal{V}_n(z,\delta_0) - \mathcal{V}_n(y,\delta_0)| > x\Big\} &= P_\zeta\Big\{\Big|\sum_{i=1}^n \zeta_i \frac{|\eta_i|}{\sqrt{n}}(I(y < Z_{i-1} \le z)\Big| > x\Big\} \\ &\le 2\exp\big(-x^2/2d_*^2(y,z)\big), \end{aligned}$$

where

$$d_*^2(y,z) = \frac{1}{n}\sum_{i=1}^n \eta_i^2 I(y < Z_{i-1} \le z).$$

Note that because of the independence of η_i and $\boldsymbol{Y}_{i-1}$, $Ed_*^2(y,z) = \sigma^2\rho(y,z)$, and by (3.7), $d_*^2(y,z) \to d^2(y,z)$, a.s., from which the lemma follows.

Proof of Lemma 2.2 contd. Use Corollary 2.2.8 of VW and its proof together with (3.8) to obtain that for every $r > 0$ and all sufficiently large n,

$$E \sup_{d(y,z)\le r} |\mathcal{V}_n(z,\delta_0) - \mathcal{V}_n(y,\delta_0)| \le K\sqrt{2\sigma^2}\int_0^r \sqrt{\ln D(\tau,d)}d\tau,$$

where $D(\tau,d) = 2\sigma^2/\tau^2$ is the maximal packing number of $(\mathbb{R}^+, d)$, which is the maximal number of τ-separated points; K is an absolute constant. Straightforward algebra shows that $\sqrt{2\sigma^2}\int_0^r \sqrt{\ln D(\tau,d)}d\tau = \sigma^2\int_{\ln(2\sigma^2/r^2)}^\infty \sqrt{v}e^{-v}dv$. Then, by the Markov inequality, for any $\gamma > 0$,

$$P\{\sup_{d(y,z)\le r} |\mathcal{V}_n(z,\delta_0) - \mathcal{V}_n(y,\delta_0)| > \gamma\} \le K\sigma^2\gamma^{-1}\int_{\ln(2\sigma^2/r^2)}^\infty \sqrt{v}e^{-v}dv.$$

Since the integral is bounded by $\Gamma(3/2)$ then for any $\gamma > 0$ and $\gamma_1 > 0$ with $\gamma\gamma_1 \ge K\sigma^2\Gamma(3/2)$ and for every $r > 0$ and all sufficiently large n,

$$P\{\sup_{d(y,z)\le r} |\mathcal{V}_n(z,\delta_0) - \mathcal{V}_n(y,\delta_0)| > \gamma\} \le \gamma_1,$$

thereby completing the proof of the tightness of the process $\mathcal{V}_n(\cdot,\delta_0)$ when η_i are symmetric around zero. This fact together with Lemma 3.2 in turn proves the tightness of this process for general η_i's. This in turn completes the proof of Lemma 2.2.

Proof of (2.13). Recall $\delta_0 = (\theta_0, \theta')'$ and G is the d.f. of $\theta'\mathbf{Y}_0$. Fix $0 \le y_1 < y_2 < \infty$. Note that $EJ_n(y,\delta_0) \equiv 0$ and by (2.3),

$$\begin{aligned}\mathrm{Cov}\Big(\mathcal{W}_n(y_1,\delta_0), \mathcal{W}_n(y_2,\delta_0)\Big) &= \sigma^2 G(y_1\wedge y_2) - E\mathcal{V}_n(y_1,\delta_0)J_n(y_2,\delta_0)\\ &\quad - E\mathcal{V}_n(y_2,\delta_0)J_n(y_2,\delta_0) + EJ_n(y_1,\delta_0)J_n(y_2,\delta_0).\end{aligned}$$

Observe that

$$J_n(y,\delta_0) = n^{-1/2}\sum_{i=1}^n \eta_i \int_0^y \nu(z,\delta_0)'A_{\delta_0}^{-1}(z)\nu(\theta'\mathbf{Y}_{i-1},\delta_0)I(\theta'\mathbf{Y}_{i-1} \ge z)dG(z).$$

Fix $0 \le y_1 \le y_2 < \infty$. Then, by the Fubini Theorem,

$$\begin{aligned}&EJ_n(y_1,\delta_0)J_n(y_2,\delta_0)\\ &= \sigma^2 E\Big\{\int_0^{y_1} \nu(s,\delta_0)'A_{\delta_0}^{-1}(s)\nu(\theta'\mathbf{Y}_0,\delta_0)I(\theta'\mathbf{Y}_0 \ge s)dG(s)\\ &\qquad \times \int_0^{y_2} \nu(\theta'\mathbf{Y}_0,\delta_0)'I(\theta'\mathbf{Y}_0 \ge z)A_{\delta_0}^{-1}(z)\nu(z,\delta_0)dG(z)\Big\}\\ &= \sigma^2\int_0^{y_1} \nu(s,\delta_0)'A_{\delta_0}^{-1}(s)\int_0^{y_2} A(z\vee s,\delta_0)A_{\delta_0}^{-1}(z)\nu(z,\delta_0)dG(z)dG(s)\end{aligned}$$

$$
\begin{aligned}
&= 2\sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z,\delta_0)' A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(z) dG(s) \\
&\quad +\sigma^2 \int_0^{y_1} \nu(s,\delta_0)' A_{\delta_0}^{-1}(s) \int_{y_1}^{y_2} \nu(z,\delta_0) dG(z)\, dG(s).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&E\mathcal{V}_n(y_1,\delta_0) J_n(y_2,\delta_0) \\
&= \sigma^2 \int_0^{y_2} EI(\theta'\mathbf{Y}_0 \le y_1)\nu(\theta'\mathbf{Y}_0,\delta_0)' I(\theta'\mathbf{Y}_0 \ge s) A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(s) \\
&= \sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z,\delta_0)' A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(z) dG(s),
\end{aligned}
$$

$$
\begin{aligned}
&E\mathcal{V}_n(y_2,\delta_0) J_n(y_1,\delta_0) \\
&= \sigma^2 \int_0^{y_1} E\nu(\theta'\mathbf{Y}_0,\delta_0)' I(\theta'\mathbf{Y}_0 \ge s, \theta'\mathbf{Y}_0 \le y_2) A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(s) \\
&= \sigma^2 \int_0^{y_1} \int_s^{y_1} \nu(z,\delta_0)' A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(z) dG(s) \\
&\quad +\sigma^2 \int_0^{y_1} \int_{y_1}^{y_2} \nu(z,\delta_0)' A_{\delta_0}^{-1}(s)\nu(s,\delta_0) dG(z) dG(s).
\end{aligned}
$$

From the above derivations one readily sees that

$$
E\mathcal{V}_n(y_1,\theta) J_n(y_2,\theta) + E\mathcal{V}_n(y_2,\theta) J_n(y_1,\theta) = EJ_n(y_1,\theta) J_n(y_2,\theta),
$$

thereby completing the proof of (2.13).

References

[1] Engle, R. F. (2002). New frontiers for ARCH models. *J. Applied Econometrics*, **17**, 425–446.

[2] Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, **66**(5), 1127–1162.

[3] Hall, P. and Heyde, C. C. (1980). *Martingale limit theory and its applications.* Academic Press, New York.

[4] Hautsch, N. (2012). *Econometrics of Financial High-Frequency Data.* Springer, Heidelberg, Germany.

[5] Khmaladze, Estate V. and Koul, Hira L. (2004). Martingale transforms goodness-of-fit tests in regression models. *Ann. Statist.*, **32**, 995-1034.

[6] Khmaladze, Estate V. and Koul, Hira L. (2009). Goodness-of-fit problem for errors in nonparametric regression: distribution free approach. *Ann. Statist.*, **37**, 3165–3185.

[7] Koul, H. L. and Song, Weixing. (2012). Large sample results for varying kernel regression estimates. *J. Nonparametr. Stat.*, **25(4)**, 829–853.

[8] Koul, H. L. and Stute, W. (1999). Nonparametric model checks for time series, *Ann. Statist.* **27**, 204-236.

[9] Koul, H. L., Perera, I. and Silvapulle, M. J. (2012). Lack-of-fit testing of the conditional mean function in a class of Markov multiplicative error models. *Econometric Theory*, **28(6)**, 1283–1312.

[10] Mnatsakanov, R., and Sarkisian, K. (2012). Varying kernel density estimation on $\mathbb{R}_+$. *Statistics & Probability Letters*, **82**, 1337–1345.

[11] Pacurar, M. (2008). Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *J. Economic Surveys*, **22**, 711–751.

[12] Stute, W. (1997), Model checks in statistics: an innovation process approach, *IMS Lecture Notes* **31**, 373-383.

[13] Stute, W. and Zhu, Li-xing. (2002). Model checks for generalized linear models, *Scand. J. Statist.* **29**, 535–545.

[14] Stute, W,, Thies, S. and Zhu, L. (1998). Model checks for regression: an innovation process approach. *Annal. Statist.*, **26** (1998), no. 5, 19161934.

[15] Stute, W., Presedo Q.M., Gonzlez Manteiga, W. and Koul, H.L. (2006). Model checks of higher order time series. *Statist. Probab. Lett.* **76**, 1385–1396.

[16] van der Vaart, A.W. and Wellner, J. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer-Verlag, New York.