Article:

Koul, H.L. and Perera, I. (2021) A minimum distance lack-of-fit test in a Markovian multiplicative error model. Journal of Statistical Theory and Practice, 15 (2). 34. ISSN 1559-8608

https://doi.org/10.1007/s42519-021-00168-1

A Minimum Distance Lack-of-fit Test in a Markovian Multiplicative Error Model[1]

Hira L. Koul and Indeewara Perera

Michigan State University and University of Sheffield

Abstract

This paper proposes a lack-of-fit test for a parametric specification of the conditional mean function in a Markovian multiplicative error time series model. The proposed test is based on a minimized distance obtained using an integral of the square of a certain marked residual process. The asymptotic null distribution of the proposed test is model dependent and is not free from the underlying nuisance parameters. We propose a bootstrap method to implement the test, and establish that the proposed bootstrap method is asymptotically valid. A finite sample simulation study that evaluates the empirical level and power is included. It compares the finite sample performance of the proposed test with several competing tests from the literature.

1 Introduction

Multiplicative error models (MEMs) of Engle [6] provide a general framework for modeling non-negative dynamic processes. This family of models generalizes the autoregressive conditional duration models introduced by Engle and Russell [8] to more general non–negative time series. These models show good performance in capturing the stylized facts of various non-negative valued time series, including, autoregressive financial duration processes in Engle and Russell [8], trading volume of orders in Manganelli [22], high-low range of asset prices in Chou [5], and realized volatility in Engle and Gallo [7]. See Engle [6], Pacurar [24] and Hautsch [14] for several other applications and some key methods of statistical inference in these models.

Proceeding a bit more precisely, let Y_i, $i \in \mathbb{Z} := \{0, \pm 1, \cdots\}$ be a time series of non-negative random variables (r.v.'s) having finite expectation and $\mathcal{H}_{i-1}$ denote the information available through time $i-1$ for forecasting Y_i. In the MEM of interest here the non-negative stationary time series $Y_i, i \in \mathbb{Z}$ obeys the model

$$Y_i = E(Y_i|\mathcal{H}_{i-1})\varepsilon_i, \quad i \in \mathbb{Z}, \tag{1.1}$$

where $\varepsilon_i, i \in \mathbb{Z}$ are independent and identically distributed (i.i.d.) non-negative r.v.'s, $E(\varepsilon_0) = 1$, $0 < \mathrm{Var}(\varepsilon_0) = \sigma^2 < \infty$ and for each $i \in Z$, ε_i is independent of $\mathcal{H}_{i-1}$.

In this paper we develop a minimum distance (m.d.) lack-of-fit test for fitting a Markovian parametric model to the conditional mean function $E(Y_i|\mathcal{H}_{i-1})$ in the model (1.1). More specifically, let Z_{i-1} be a known positive function of $Y_{i-1}, \cdots, Y_{i-p}$, $i \in \mathbb{Z}$, where $p \geq 1$ is a given positive integer. Let q be a given positive integer and $\psi(z, \vartheta)$, $z \geq 0, \vartheta \in \Theta \subset \mathbb{R}^q$ be a family of positive functions.

[1]Corresponding author: Hira L. Koul: koul@msu.edu

AMS Subject Classification: Primary 62M02, Secondary 62F40

Keywords & phrases: Marked empirical processes, bootstrap distribution

Our goal here is to test the lack of fit of the parametric specification $\psi(Z_{i-1}, \theta)$ for the conditional mean $E(Y_i|\mathcal{H}_{i-1})$ in (1.1), $i \in \mathbb{Z}$; i.e, to test the hypothesis

$$(1.2) \qquad H_0: \ E[Y_i \mid \mathcal{H}_{i-1}] = \psi(Z_{i-1}, \theta), \quad \forall\, i \in \mathbb{Z}, \text{ for some } \theta \in \Theta, \text{ a.s.,}$$

against the alternative $H_1: \ H_0$ *is not true*, based on $Y_{-p}, Y_{1-p}, \cdots, Y_n$ obtained from an observable stationary process $\{Y_i\}$ obeying (1.1).

It is implicitly assumed that the parameter space Θ is such that when $\theta \in \Theta$ is true under H_0, the above time series $Y_i, i \in \mathbb{Z}$ obeying (1.1)–(1.2) is stationary. An example of this model, with Z_{i-1}'s different from Y_{i-1}'s, is provided in Example 2 of Perera and Koul [25], where the time series $\{Y_i\}$ consists of daily annualized realized volatility measures, constructed from a series of intraday spot price data for the S&P500 index. More precisely, they modelled the series $\{Y_i\}$ in the form (1.1)–(1.2) by taking $Z_{i-1} = \sum_{t=1}^{p} w_t Y_{i-t}$, a weighted sum of past realized volatilities with known $\{w_i\}$ and p. Other examples of this model, with $Z_{i-1} = Y_{i-1}$, are discussed in Koul, Perera and Silvaphulle [19] and Guo and Li [12], amongst others.

Koul, Perera and Balakrishna [18] (KPB) proposed m.d. estimators for θ in the model (1.1)–(1.2) based on the integrated square of the marked residual empirical process

$$U_n(z, \vartheta) := n^{-1/2} \sum_{i=1}^{n} \Big(\frac{Y_i}{\psi(Z_{i-1}, \vartheta)} - 1\Big) I(Z_{i-1} \le z), \quad z \ge 0, \ \vartheta \in \Theta,$$

where I is the indicator function. A motivation for basing inference in the above MEM on an analog of the process $U_n(z, \vartheta), z \ge 0, \vartheta \in \Theta$ appears in Koul *et al.* [19], where the authors proposed tests for testing H_0 against H_1, for the case $Z_{i-1} = Y_{i-1}, i \in \mathbb{Z}$, based on $\sup_{z\ge 0} |U_n(z, \hat\theta)|$ with $\hat\theta$ being a $n^{1/2}$-consistent estimator of θ, under H_0. In a simulation study, the performance of this test in terms of the empirical level and power was found to be significantly superior to that of Box–Pierce–Ljung's Portmanteau test (Ljung and Box [20]), Lagrange Multiplier (LM) tests (Meitz and Teräsvirta [23]), and a generalized moment M test (Chen and Hsieh [4]). In this simulation study, $\hat\theta$ was taken to be the quasi maximum likelihood estimator based on the standard exponential error distribution. The approach of Koul *et al.* [19] has its roots in Stute [27], Stute, Thies and Zhu [28] and Koul and Stute [17].

In the current paper, we propose an alternative method for testing H_0 against H_1, based on a particular minimized distance with its large values being significant. To introduce this test statistic, let $G_n(z) := n^{-1} \sum_{i=1}^{n} I(Z_{i-1} \le z)$, denote the empirical d.f. of $Z_{i-1}, 1 \le i \le n$. Define

$$(1.3) \qquad M_n(\vartheta) := \int U_n^2(z, \vartheta) dG_n(z), \qquad \widehat\theta_n := \text{argmin}_\vartheta M_n(\vartheta), \qquad \vartheta \in \Theta.$$

The m.d. test statistic of our interest is given by $M_n(\widehat\theta_n)$. The m.d. estimator $\widehat\theta_n$ in (1.3) was proposed in KPB. This estimator was shown to have desirable finite sample properties, in particular, in comparison with the quasi maximum likelihood estimator for the standard exponential errors. Furthermore, $\widehat\theta_n$ is also root-n consistent and asymptotically normal under fairly general conditions, provided that the innovation has unit mean and a finite second moment, even if the true innovation distribution is unknown.

In this paper we derive the limiting null distribution of $M_n(\widehat{\theta}_n)$, which turns out to be model dependent and not free from the underlying nuisance parameters. Hence critical values to implement the test even for the large samples cannot be tabulated for general use. As an alternative, we propose a bootstrap method to implement the test. The paper contains the proof of the asymptotic validity of the proposed bootstrap testing procedure. In a finite sample comparison, the new test dominates some of the competing tests, both in terms of empirical level and power, and complements the test proposed by Koul *et al.* [19] for model validation in Markovian MEMs.

The rest of this paper is structured as follows. Section 2 presents the regularity conditions needed for deriving the asymptotic null distribution of the test statistic and outlines the bootstrap algorithm for implementing the lack-of-fit test. A proof of the asymptotic validity of the bootstrap test and the needed additional regularity conditions are presented in Section 4. Section 3 contains the findings of a simulation study, where we evaluate the empirical level and power of the proposed test in comparison with several competing tests from the literature. Section 5 concludes the paper.

2 Asymptotic null distribution of the test statistic

In this section we describe some sufficient conditions under which the asymptotic null distribution of the m.d. test statistic $M_n(\widehat{\theta}_n)$ is obtained. Let $\|\cdot\|$ denote the Euclidean norm and $\mathcal{N}_n(b) := \{\vartheta \in \Theta, n^{1/2}\|\vartheta - \theta\| \le b\}$, $0 < b < \infty$. The following assumptions are needed for deriving the large sample null distribution of $M_n(\widehat{\theta}_n)$. They are the same as used in KPB for deriving the asymptotic distribution of $\widehat{\theta}_n$.

C.1. *There exists a positive constant* $C < \infty$ *such that* $\inf_{z\ge 0, \vartheta\in\Theta} \psi(z,\vartheta) \ge C > 0$.

C.2. *There exists a* q*-vector* $\dot{\psi}(z,\vartheta)$ *such that* $E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty$ *and for every* $0 < b < \infty$,

$$\sup_{1\le i\le n,\, \|t\|\le b} \sqrt{n}|\psi(Z_{i-1}, \theta + n^{-1/2}t) - \psi(Z_{i-1},\theta) - n^{-1/2}t'\dot{\psi}(Z_{i-1},\theta)| = o_p(1).$$

C.3. $\forall\, \epsilon > 0$, $0 < \eta < \infty$, $\exists\, N_{\epsilon,\eta}$, $0 < b \equiv b_{\epsilon,\eta} < \infty$, *such that* $\forall\, n > N_{\epsilon,\eta}$,

$$P\Big(\inf_{\|t\|>b} M_n(\theta + n^{-1/2}t) \ge \eta\Big) \ge 1 - \epsilon.$$

The assumptions (C.1) and (C.2) are used to show that the defining dispersion $M_n(\vartheta)$ is AULQ (asymptotically uniformly locally quadratic) in $n^{1/2}(\vartheta - \theta)$ for $\vartheta \in \mathcal{N}_n(b)$, for every $0 < b < \infty$, while assumption (C.3) is used to show that $\|n^{1/2}(\widehat{\theta}_n - \theta)\| = O_p(1)$.

Recall that for given $0 < b < \infty$, any $\|\vartheta\| > b$ can be written as $\vartheta = re$, for some unit vector $e \in \mathbb{R}^q, \|e\| = 1$ and a real number r such that $|r| > b$. Lemma 2.1 of KPB shows that (C.1), (C.2) and the following condition (2.1) imply (C.3).

(2.1) $\psi(z, \theta + n^{-1/2}re)$ is monotonic in r, $\forall\, z \ge 0$ and $\forall\, e \in \mathbb{R}^q$, $\|e\| = 1$.

Let $\varphi(z) := \dot{\psi}(z,\theta)/\psi(z,\theta)$. Assumptions (C.1) and (C.2) imply that

(2.2)
$$E\Big(\frac{\|\dot{\psi}(Z_0,\theta)\|^2}{\psi^j(Z_0,\vartheta)}\Big) \le C^{-j}E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty, \quad \forall\, j \ge 1,$$
$$E\|\varphi(Z_0)\|^2 < C^{-2}E\|\dot{\psi}(Z_0,\theta)\|^2 < \infty.$$

To proceed further, we need some more notation as in KPB. We write $U_n(z)$ for $U_n(z,\theta)$, where θ is as in H_0. Let G denote the d.f. of Z_0 and let

$$\begin{aligned}
&\Psi_n(z) := n^{-1}\sum_{i=1}^n \varphi(Z_{i-1})\, I(Z_{i-1}\le z), \qquad \Psi(z) := E\big(\varphi(Z_0)I(Z_0\le z)\big),\\
&\tilde{M}_n(\theta) := \int U_n^2 dG, \quad V_n := \int U_n(z)\Psi(z)dG(z), \quad \mathcal{G} := \int \Psi(z)\Psi(z)' dG(z),\\
&Q_n(t) := \int \big(U_n(z) - t'\Psi(z)\big)^2 dG(z) = \tilde{M}_n(\theta) - 2t'V_n + t'\mathcal{G}t,\\
&\widetilde{t}_n := \text{argmin}_t Q_n(t), \qquad \xi(x) := \int_{z\ge x} \Psi(z)dG(z), \quad x\ge 0.
\end{aligned}$$

From KPB we obtain the following. Let Z_{01}, Z_{02} denote independent copies of Z_0. Then, by the Fubini Theorem,

$$\mathcal{G} = \int E\big(\varphi(Z_0)I(Z_0\le z)\big)E\big(\varphi(Z_{01})'I(Z_{01}\le z)\big)dG(z) = E\big(\varphi(Z_0)\varphi(Z_{01})'\big[1 - G\big((Z_0\vee Z_{01})-\big)\big]\big).$$

The Fubini Theorem, G being a d.f. and (2.2) imply

$$\begin{aligned}
\xi(x) &= \int_{z\ge x} E\big(\varphi(Z_0)I(Z_0\le z)\big)dG(z) = E\Big(\varphi(Z_0)\big[1 - G\big((Z_0\vee x)-\big)\big]\Big),\\
\sup_{x\ge 0}\|\xi(x)\|^2 &\le E\big(\|\varphi(Z_0)\|^2\big) < \infty.
\end{aligned}$$

Hence, $\Sigma := E(\xi(Z_0)\xi(Z_0)')$ is well defined and one can rewrite

$$\Sigma = E\big\{E\big(\varphi(Z_0)\varphi(Z_{01})'[1 - G((Z_0\vee Z_{02})-)][1 - G((Z_{01}\vee Z_{02})-)]\big|Z_{02}\big)\big\}.$$

Moreover,

$$\begin{aligned}
&V_n := n^{-1/2}\sum_{i=1}^n(\varepsilon_i - 1)\int I(Z_{i-1}\le z)\Psi(z)dG(z) = n^{-1/2}\sum_{i=1}^n(\varepsilon_i-1)\xi(Z_{i-1}),\\
&EV_n \equiv 0, \qquad E(V_nV_n') \equiv \sigma^2 E(\xi(Z_0)\xi(Z_0)') = \sigma^2\Sigma.
\end{aligned}$$

The following lemma describes the AULQ property of $M_n(\theta + n^{-1/2}t)$ in $\|t\|\le b$ and obtains that $n^{1/2}(\widehat{\theta}_n - \theta)$ is bounded in probability under H_0, which is used to derive the limiting null distribution of $M_n(\widehat{\theta}_n)$. It appears as Theorem 2.1 in KPB.

Lemma 2.1. *Suppose that H_0 holds, and the assumptions (C.1) and (C.2) are satisfied. Then, the following AULQ result holds.*

$$\sup_{\|t\|\le b}\big|M_n(\theta + n^{-1/2}t) - Q_n(t)\big| = o_p(1), \quad \forall\, 0<b<\infty. \tag{2.3}$$

If, in addition, the condition (C.3) holds, then $\|n^{1/2}(\widehat{\theta}_n - \theta)\| = O_p(1)$. If, further $\mathcal{G}$ is positive definite, then $\widetilde{t}_n = \mathcal{G}^{-1}V_n$ and $\|n^{1/2}(\widehat{\theta}_n - \theta) - \widetilde{t}_n\| = o_p(1)$.

We also need the following lemma, where B denotes the standard Brownian motion on $[0,\infty)$.

Lemma 2.2. *Suppose that H_0 holds. Then, under the above set up, U_n converges weakly to $\sigma B \circ G$, in Skorokhod space $D[0,\infty]$ and uniform metric.*

The proof of this lemma is similar to that of Lemma 2.2 in Balakrishna, Koul, Sakhanenko and Ossiander [1], and hence is omitted. In particular it implies $M_n(\theta) \le \sup_{z\ge 0} U_n^2(z) = O_p(1)$.

The following theorem gives the asymptotic null distribution of the test statistic $M_n(\widehat{\theta}_n)$.

Theorem 2.1. *Suppose that H_0 holds, and assumptions (C.1), (C.2), and (C.3) are satisfied. Further assume that $\mathcal{G}$ is positive definite. Then, $M_n(\widehat{\theta}_n) \to_D \sigma^2 \mathcal{Z}$, where the r.v.*

$$\mathcal{Z} := \int_0^1 B^2(u)du - \int_0^1 B(u)\Psi(G^{-1}(u))'du\ \mathcal{G}_2^{-1} \int_0^1 B(u)\Psi(G^{-1}(u))du.$$

Proof. Under H_0 and the assumptions (C.1) and (C.2), by using (2.3) of Lemma 2.1, we obtain $|M_n(\widehat{\theta}_n) - Q_n(\widetilde{t}_n)| = o_p(1)$. But $\widetilde{t}_n = \mathcal{G}^{-1}V_n$ implies that

$$\begin{aligned} Q_n(\widetilde{t}_n) &= \tilde{M}_n(\theta) - 2V_n'\mathcal{G}^{-1}V_n + V_n'\mathcal{G}^{-1}\mathcal{G}\mathcal{G}^{-1}V_n \\ &= \int U_n^2 dG - \int U_n \Psi' dG\ \mathcal{G}^{-1} \int U_n \Psi dG. \end{aligned}$$

Hence, by Lemma 2.2, the continuous mapping theorem and G being continuous, we obtain

$$M_n(\widehat{\theta}_n) \to_D \sigma^2 \Big[\int B^2(G(z))dG(z) - \int B(G(z))\Psi(z)'dG(z)\ \mathcal{G}^{-1} \int B(G(z))\Psi(z)dG(z) \Big] = \sigma^2 \mathcal{Z}. \ \square$$

A consistent estimator of σ^2, under H_0, is given by

$$s_n^2 := n^{-1}\sum_{i=1}^n \Big(\Big\{ n^{-1}\sum_{t=1}^n \widehat{\varepsilon}_t \Big\}^{-1} \widehat{\varepsilon}_i - 1 \Big)^2, \quad \widehat{\varepsilon}_i = Y_i/\psi(Z_{i-1}, \widehat{\theta}_n), \quad i = 1, \cdots, n.$$

Let $T_n := s_n^{-2} M_n(\widehat{\theta}_n)$. For a given $0 < \alpha < 1$, let $z_\alpha := \inf\{z \ge 0 : P(\mathcal{Z} \le z) \ge 1 - \alpha\}$. Then the asymptotic level of the test that rejects H_0, in favor of H_1, whenever $T_n > z_\alpha$, is α. However, the distribution of $\mathcal{Z}$ is model dependent and is not free from the underlying nuisance parameters (G, θ). Therefore, we cannot compute z_α for general use $(0 < \alpha < 1)$. To implement the test for the large samples, we propose the following bootstrap testing procedure.

2.1 The bootstrap algorithm for implementing the $M_n(\widehat{\theta}_n)$-test

To carry out the resampling under the null hypothesis we need to first scale the residuals as

$$\widetilde{\varepsilon}_i = \Big\{ n^{-1}\sum_{t=1}^n \widehat{\varepsilon}_t \Big\}^{-1} \widehat{\varepsilon}_i, \qquad \widehat{\varepsilon}_i = Y_i/\psi(Z_{i-1}, \widehat{\theta}_n), \qquad i = 1, \cdots, n, \tag{2.4}$$

so that the empirical distribution of $\{\widetilde{\varepsilon}_i\}_{i=1}^n$ has mean 1.

The bootstrap algorithm for implementing the lack-of-fit test based on $M_n(\widehat{\theta}_n)$ is as follows:

Step 1: Compute $\{\widehat{\theta}_n, T_n\}$ for the realized sample $Y_{-p}, Y_{1-p}, \cdots, Y_n$.

Step 2: Compute $\widetilde{\varepsilon}_i$, $i = 1, \cdots, n$ as in (2.4). Then draw a random sample (with replacement) of

size $n+p+1$, say $\{\varepsilon^*_{-p}, \varepsilon^*_{1-p}, \cdots, \varepsilon^*_n\}$, from $\{\widetilde{\varepsilon}_i; 1 \le i \le n\}$. This ensures that $n+p+1$ independent observations are generated from the empirical distribution function $\widetilde{F}_n(\cdot) := n^{-1}\sum_{i=1}^n I(\widetilde{\varepsilon}_i \le \cdot)$.
Step 3: Generate the bootstrap sample $\{Y^*_{-p}, Y^*_{1-p}, \cdots, Y^*_n\}$ at $(\widehat{\theta}_n, \widetilde{F}_n)$ recursively, by using the model equation

$$Y^*_i = \psi^*_i(Z^*_{i-1}, \widehat{\theta}_n)\varepsilon^*_i, \quad i = -p, 1-p, \cdots, n,$$

conditional on starting values, e.g., $Y^*_{-p-t} = \bar{Y}$, $t = 1, 2, \cdots, p$, where $\bar{Y} := n^{-1}\sum_{i=1}^n Y_i$. Recall that Z^*_{i-1} can be computed as the known positive function of $Y^*_{i-1}, \cdots, Y^*_{i-p}$, $i = -p, 1-p, \cdots, n$.
Step 4: Based on $\{Y^*_{-p}, Y^*_{1-p}, \cdots, Y^*_n\}$, compute $\widehat{\theta}^*_n$ the bootstrap analog of $\widehat{\theta}_n$.
Step 5: Compute the bootstrap test statistic $T^*_n := (s^*_n)^{-2} M^*_n(\widehat{\theta}^*_n)$, where s^*_n and $M^*_n(\vartheta)$ are the bootstrap analogs of s_n and $M_n(\vartheta)$, respectively.

Let P^*_n denote the probability measure induced by the above bootstrap, conditional on $\mathcal{Y} = (Y_{-p}, Y_{1-p}, \cdots, Y_n)'$ and z^*_α denote the α-level critical value of T^*_n under P^*_n, i.e.,

$$z^*_\alpha := \inf\{z \ge 0 : P^*_n(T^*_n \le z) \ge 1-\alpha\}, \quad 0 < \alpha < 1. \tag{2.5}$$

Since the distribution of T^*_n is not available, to approximate z^*_α, repeat steps 2–5 a large number of times, say for $k = 1, \cdots, K$, and compute the bootstrap statistic, say $T^{*(k)}_n$, for each k. Then, approximate z^*_α by the value z^{*K}_α that satisfies

$$\sum_{k=1}^K I\big(T^{*(k)}_n > z^{*K}_\alpha\big) = [\alpha K], \tag{2.6}$$

where $[\alpha K]$ denotes the integer part of αK. The bootstrap test rejects H_0 at level α if $T_n > z^{*K}_\alpha$. By the Glivenko–Cantelli theorem $\sup_z \big|K^{-1}\sum_{k=1}^K I\big(T^{*(k)}_n > z\big) - P^*_n(T^*_n > z)\big| \to 0$, P^*_n-a.s., and hence z^*_α can be approximated by z^{*K}_α, as accurately as desired, by selecting K large enough.

3 Simulation study

In this section we present the findings of a Monte Carlo simulation study that evaluates the finite sample performance of the above proposed bootstrap lack-of-fit test and compare it with several competing tests, namely, the test proposed by Koul *et al.* [19], which we denote by KPS, the Ljung-Box Q test (Ljung and Box [20]), a LM test (Meitz and Teräsvirta [23]), and the generalized moment M test of Chen and Hsieh [4]. The Ljung-Box Q test, although not originally designed for MEMs, is routinely applied for evaluating MEMs; see, Pacurar [24] and Hautsch [14]. This test evaluates the lack-of-fit of a given MEM by testing the significance of the serial dependence of the residuals estimated from the fitted model. The other three aforementioned tests are designed specifically for MEMs.

Computational formula. Here we provide a computational formula for $M_n(\widehat{\theta}_n)$ as given in KPB. We use this in the computation of the test statistic $T_n := s_n^{-2} M_n(\widehat{\theta}_n)$. Order $Z_{i-1}, 1 \le i \le n$ as $Z_{(0)} \le Z_{(1)} \le \cdots \le Z_{(n-1)}$. Let $Y^\dagger_i$ denote the Y_i corresponding to $Z_{(i-1)}$, for $1 \le i \le n$. Then, using the fact $G_n\big(Z_{(i-1)}-\big) = (i-1)/n$, for all $1 \le i \le n$, we obtain

$$M_n(\vartheta) = n^{-2}\sum_{i=1}^{n}(n-i+1)\Big(\frac{Y_i^{\dagger}}{\psi(Z_{(i-1)},\vartheta)}-1\Big)^2 + 2n^{-2}\sum_{i=1}^{n}\sum_{j=i+1}^{n}(n-j+1)\Big(\frac{Y_i^{\dagger}}{\psi(Z_{(i-1)},\vartheta)}-1\Big)\Big(\frac{Y_j^{\dagger}}{\psi(Z_{(j-1)},\vartheta)}-1\Big).$$

Design of the simulation study. In this simulation study, we use $q = 2$ and $Z_{i-1} = Y_{i-1}$, and test the null hupothesis

$$(3.1)\qquad H_0: \quad \psi(Z_{i-1};\theta) = \theta_1 + \theta_2 Z_{i-1}, \quad \text{for some} \ \ \theta_1 > 0,\ 0 \le \theta_2 < 1,\ \theta = (\theta_1, \theta_2)',$$

against the alternative H_1 : Not H_0. Since $\theta \in \Theta := (0,\infty)\times[0,1)$ in (3.1), the process under H_0 is stationary. Similar specifications for the conditional mean have been previously considered in Koul *et al.* [19], Guo and Li [12] and KPB in the MEM setting. Also note that the model $\psi(z,\vartheta) = \vartheta_1 + \vartheta_2 z$ satisfies the assumptions (C.1), (C.2) and (2.1) trivially.

For the data generating processes to evaluate the level performances of the tests, we consider

$$(3.2)\qquad \text{MEM}(1,0): \quad Y_i = \tau_i\varepsilon_i, \quad \tau_i = \psi(Y_{i-1},\theta) = 0.2 + 0.1Y_{i-1},$$

For the error distribution, we consider the following families of densities.

1. **Exponential** [E]: $f(x) := e^{-x}, x > 0$.
2. **Gamma** [G]: $f(x) := a^a\Gamma(a)^{-1}x^{a-1}e^{-ax}$, $x > 0, a > 0$.
3. **Weibull** [W]: $f(x) = (\kappa/c)(x/c)^{\kappa-1}\exp\{-(x/c)^{\kappa}\}$, $\kappa > 0$, $c = [\Gamma(1+\kappa^{-1})]^{-1}$.
4. **Generalized gamma** [GG]:
   $f(x) = c\{\sigma\Gamma(a)\}^{-1}(x/\sigma)^{ac-1}\exp\{-(x/\sigma)^c\}$, $a, c > 0$, $\sigma = \{\Gamma(a+c^{-1})\}^{-1}\Gamma(a)$.
5. **Burr** [B]: $f(x) = (a/\sigma)(x/\sigma)^{a-1}[1+b(x/\sigma)^a]^{-(1+b^{-1})}$, $a > b > 0$, and
   $\sigma = \{\Gamma(1+a^{-1})\Gamma(b^{-1}-a^{-1})\}^{-1}b^{(1+a^{-1})}\Gamma(1+b^{-1})$.

Note that the above exponential and gamma distributions a priori satisfy the model assumption that $E(\varepsilon) = 1$. The restriction imposed on the parameters in the other three families of the above distributions ensure the satisfaction of this requirement.

Some important theoretical properties and applications of these families of error distributions, in the context of MEMs,are discussed in Engle and Russell [8], Grammig and Maurer [11], Lunde [21] and Engle and Gallo [7].

To evaluate the empirical powers of the tests, we consider data generating processes based on the following Markov MEM as considered in Koul *et al.* [19]:

$$(3.3)\qquad \mathcal{M}(\omega,\beta,\gamma): \quad Y_i = \tau_i\varepsilon_i, \quad \tau_i = \omega + \beta Y_{i-1} + \gamma\sqrt{Y_{i-1}}.$$

The data are generated from $\mathcal{M}(0.1, 0.2, 0.3)$, $\mathcal{M}(0.1, 0.2, 0.5)$ and $\mathcal{M}(0.1, 0.2, 0.7)$ models while considering standard exponential [E] and generalized gamma [GG] as the error distribution.

For the sample sizes, $n = 200, 500$ and 1000 are considered. To ensure that the effect of initialization is negligible, in each Monte Carlo replication, we generate $(n+\ell+1)$ observations with $\ell = 300$, discard the first ℓ observations and use the remaining $n+1$ observations as $Y_0, Y_1, \cdots, Y_n$. For each replication and data generating process, we first compute the m.d. estimator $\widehat{\theta}_n$ and $M_n(\widehat{\theta}_n)$. The optimization procedure to implement the constrained minimization problem in (1.3) is carried out by using the `fmincon` function in Matlab. To implement the proposed m.d. test we use the bootstrap algorithm outlined in Section 2.1, while adopting the 'Warp-Speed' Monte Carlo method of Giacomini, Politis and White [10] to reduce the computational burden. More precisely, as formally proved by Giacomini *et al.* [10], since the test statistic satisfies (4.1) below, it is sufficient to generate only one bootstrap sample in each Monte Carlo replication, say $r = 1, \ldots, R$, and then evaluate the test against the empirical distribution of the R bootstrap statistics $T_n^{*(1)}, \ldots, T_n^{*(R)}$. To this end, at any given level α, $0 < \alpha < 1$, the empirical rejection rate of the null hypothesis, say $\alpha_{n,R}$, is computed as

$$\alpha_{n,R} \equiv R^{-1} \sum_{r=1}^{R} I\big\{T_n^{(r)} > z_\alpha^{(n,R)}\big\}, \qquad z_\alpha^{(n,R)} \equiv \inf\big\{z \geq 0 : R^{-1} \sum_{r=1}^{R} I(T_n^{(*r)} \leq z) \geq 1-\alpha\big\},$$

where $T_n^{(r)}$ denotes the test statistic at the rth Monte Carlo replication, $r = 1, \ldots, R$.

Note that, in the standard Monte Carlo bootstrap, at each Monte Carlo replication, one generates K bootstrap samples and compute the α-level bootstrap critical value satisfying (2.6), say $z_\alpha^{*(K,r)}$, $r = 1, \ldots, R$, and then compute the empirical rejection rate of the null hypothesis as

$$\alpha_{n,R}^{(K)} \equiv R^{-1} \sum_{r=1}^{R} I\big\{T_n^{(r)} > z_\alpha^{*(K,r)}\big\}.$$

Clearly one can compute $\alpha_{n,R}$ much more efficiently than computing $\alpha_{n,R}^{(K)}$, because in the computation of $\alpha_{n,R}$ the bootstrap statistic is only computed R times.

It is also of interest to give an explanation as to why it is reasonable to use $\alpha_{n,R}$ in place of $\alpha_{n,R}^{(K)}$. To this end, for the rth Monte Carlo replication, $r = 1, \ldots, R$, let $z_\alpha^{*(r)}$ denote the α-level critical value (2.5) from the bootstrap distribution of $T_n^{(*r)}$. Then, as outlined in Section 2.1, the critical value $z_\alpha^{*(r)}$ can be approximated by $z_\alpha^{*(K,r)}$, as accurately as desired, by selecting K large enough, $r = 1, \ldots, R$. By this fact and Theorem 1 and Corollary 2 in Giacomini *et al.* [10], $|\alpha_{n,R} - \alpha_{n,R}^{(K)}|$ can be made arbitrarily small with large probability for all large enough n, R and K.

The Ljung-Box Q statistic of a lag length λ is $\text{LBQ}(\lambda) = n(n+2) \sum_{k=1}^{\lambda} (n-k)^{-1} \rho_k^2$, where ρ_k^2 is the squared sample autocorrelation of the residuals at lag k. As in Engle and Russell [8] we use the χ_λ^2 distribution to obtain the critical values for $\text{LBQ}(\lambda)$. The other three aforementioned tests are implemented as in Koul *et al.* [19] by using the critical values given by their respective asymptotic null distributions. To evaluate the empirical level and power of the tests we compute the frequency of times the null hypothesis is rejected by each test over the $R = 2000$ Monte Carlo replications, when the data generating process (DGP) is given by H_0 and when DGP is given by the chosen three alternatives, respectively.

Summary of the results. The simulation findings are given in Tables 1 and 2. In these tables, KPS, MDT, $\text{LBQ}(\lambda)$, LM and M denote, respectively, the Koul *et al.* [19] test, the proposed m.d. test of this paper, the Ljung-Box Q test with lag length λ, the Lagrange Multiplier test and the generalized moment test. Also, F denotes the d.f. of ε. Note that, the DGP based on MEM(1,0) in (3.2) is under the null hypothesis, whereas the DGP's based on $\mathcal{M}(\omega, \beta, \gamma)$ in (3.3) are under the alternative hypothesis.

From Table 1 we see that the proposed MDT test entails the best overall performance in terms of the empirical level among all the competing tests considered in the simulation study. Although the empirical level of the KPS test is significantly better than that of LBQ(λ), LM and M tests for all the sample sizes, its performance is not as good as that of the MDT test. In particular, the empirical level of the MDT test is closer to the nominal level than that of the KPS test for a majority of the chosen error distributions and for both chosen sample sizes.

From Table 2 we see that for the standard exponential error distribution, the empirical power of the MDT test is significantly higher than that of the other six tests, for each DGP and sample sizes, while for the generalized-gamma error distribution, the KPS test performs significantly better than the MDT test, for all sample sizes considered. But, the MDT test outperform each of the LBQ(λ), LM and M tests for both exponential and generalized-gamma error distributions.

Table 1: Empirical level: Percentage of times H_0 is rejected when the DGP is MEM(1,0) of (3.2), the model under H_0.

α(%)	F	KPS	MDT	LBQ(5)	LBQ(10)	LBQ(15)	LM	M
		DGP: MEM(1,0), n: 200						
5	E	2.5	4.4	1.4	2.8	4.1	2.4	21.8
	W	3.0	4.3	2.8	3.4	3.7	22.0	22.8
	G	3.6	3.1	2.0	2.2	3.2	0.4	22.1
	GG	2.9	4.4	2.8	3.6	3.3	7.7	22.1
	B	2.6	3.9	2.5	3.3	3.8	6.1	19.9
10	E	6.5	9.2	4.0	5.5	7.5	7.5	27.2
	W	6.4	8.7	4.9	6.2	5.8	30.7	26.6
	G	6.4	6.6	4.8	6.3	7.6	1.4	28.5
	GG	5.9	9.0	4.7	6.4	7.5	14.3	26.6
	B	5.6	8.7	4.6	5.8	6.4	11.5	24.6
		DGP: MEM(1,0), n: 500						
5	E	4.9	4.4	2.9	2.9	2.9	5.1	8.2
	W	4.4	5.3	4.1	4.2	4.3	21.9	9.6
	G	4.5	3.8	2.3	4.2	3.7	0.3	10.9
	GG	3.7	3.4	2.6	3.5	3.7	8.6	9.9
	B	3.9	4.3	2.6	3.5	3.5	7.4	9.4
10	E	8.4	9.4	5.0	7.8	7.0	9.2	13.2
	W	8.1	10.8	7.0	7.8	7.5	29.6	13.8
	G	10.2	8.1	5.7	7.0	8.0	1.6	17.2
	GG	7.3	8.4	5.0	6.0	6.7	15.5	14.7
	B	6.6	9.2	4.5	5.5	6.9	14.0	13.7

Table 2: Empirical power: Percentage of times H_0 is rejected when the DGP's are under the alternative $\mathcal{M}(\omega, \beta, \gamma)$ in (3.3).

n	α(%)	F	KPS	MDT	LBQ(5)	LBQ(10)	LBQ(15)	LM	M
			DGP: $\mathcal{M}$(0.1, 0.2, 0.3)						
500	5	E	15.5	18.1	3.4	4.1	5.0	0.0	4.4
	10	E	26.3	28.1	6.5	7.9	8.6	0.3	10.7
1000	5	E	28.2	37.0	3.1	4.4	4.4	0.1	8.8
	10	E	42.2	47.1	5.0	7.6	9.2	0.3	14.7
500	5	GG	33.0	14.4	0.2	2.8	3.4	3.4	9.8
	10	GG	45.0	20.8	1.0	6.6	5.0	6.0	16.6
1000	5	GG	54.4	24.2	0.6	4.8	4.8	5.2	12.4
	10	GG	68.0	33.8	3.6	10.8	8.4	9.6	19.4
			DGP: $\mathcal{M}$(0.1, 0.2, 0.5)						
500	5	E	26.7	39.2	3.3	3.7	4.8	0.1	7.6
	10	E	43.8	50.2	6.5	7.5	9.0	0.4	13.2
1000	5	E	55.1	72.7	5.4	6.4	6.2	0.1	8.2
	10	E	70.1	79.3	9.8	11.4	11.8	1.2	14.8
500	5	GG	45.6	22.5	1.0	2.8	3.4	3.6	10.8
	10	GG	59.6	31.2	2.2	6.8	5.0	6.6	16.8
1000	5	GG	75.2	48.6	3.6	6.0	5.4	6.2	12.6
	10	GG	85.2	58.5	7.2	12.0	10.6	10.8	19.8
			DGP: $\mathcal{M}$(0.1, 0.2, 0.7)						
500	5	E	40.4	48.9	4.6	5.8	5.7	0.2	7.5
	10	E	56.4	60.4	10.3	11.3	10.9	0.5	13.1
1000	5	E	76.5	85.6	7.6	7.3	6.3	0.3	8.9
	10	E	86.8	90.2	14.1	13.2	14.0	2.1	16.0
500	5	GG	53.0	24.6	1.6	2.4	3.4	3.6	11.0
	10	GG	69.8	39.2	3.6	6.8	5.2	6.6	16.2
1000	5	GG	84.0	60.1	6.0	7.6	6.6	6.6	11.6
	10	GG	91.0	72.1	15.4	14.0	12.6	11.8	18.4

4 Asymptotic validity of the bootstrap test

In this section we establish the asymptotic validity of the above proposed bootstrap test. To this end, let $O_{p^*}(1)$, in probability, $o_{p^*}(1)$, in probability, and E^* denote the usual stochastic orders of magnitude and expectation, respectively, with respect to the bootstrap probability measure P_n^*.

To establish the asymptotic validity of bootstrap we need to show that the conditional distribution, given $\mathcal{Y} = (Y_{-p}, Y_{1-p}, \cdots, Y_n)'$, of the bootstrap statistic T_n^*, consistently estimates the distribution of $\mathcal{Z}$ under H_0. That is, $T_n^* \to_{D^*} \mathcal{Z}$, in probability, under H_0, where "$\to_{D^*}$" denotes

$$P_n^*[T_n^* \le z \mid \mathcal{Y}] \to_p \mathcal{T}(z) \quad \text{as} \quad n \to \infty, \tag{4.1}$$

at each continuity point z of $\mathcal{T}(z) := \mathsf{Pr}(\mathcal{Z} \le z)$. Moreover, under the alternative H_1, T_n^* should be of order $O_{p^*}(1)$, in probability, to have good asymptotic power properties.

First, we introduce some additional notation and assumptions. Let $M_n^*(\vartheta)$, $\widehat{\theta}_n^*$, $G_n^*(z)$, $U_n^*(z,\vartheta)$ denote the bootstrap analogs of $M_n(\vartheta)$, $\widehat{\theta}_n$, $G_n(z)$, $U_n(z,\vartheta)$, $z \ge 0$, $\vartheta \in \Theta$ defined by

$$\begin{aligned} M_n^*(\vartheta) &:= \int \{U_n^*(z,\vartheta)\}^2 dG_n^*(z), \qquad \widehat{\theta}_n^* := \text{argmin}_\vartheta M_n^*(\vartheta), \qquad \vartheta \in \Theta, \\ G_n^*(z) &:= n^{-1} \sum_{i=1}^n I(Z_{i-1}^* \le z), \quad z \ge 0, \\ U_n^*(z,\vartheta) &:= n^{-1/2} \sum_{i=1}^n \Big(\frac{Y_i^*}{\psi(Z_{i-1}^*, \vartheta)} - 1 \Big) I(Z_{i-1}^* \le z), \quad z \ge 0,\ \vartheta \in \Theta. \end{aligned} \tag{4.2}$$

Consider the following additional assumptions.

B.1. *$\exists\, N \in \mathcal{N}$ such that* $\sup_{n \ge N} E^* \|\dot\psi(Z_0^*, \widehat{\theta}_n)\|^2 < \infty$*, a.s. and*

$$\sup_{1 \le i \le n,\, \|t\| \le b} \sqrt{n} |\psi(Z_{i-1}^*, \widehat{\theta}_n + n^{-1/2}t) - \psi(Z_{i-1}^*, \widehat{\theta}_n) - n^{-1/2} t' \dot\psi(Z_{i-1}^*, \widehat{\theta}_n)| = o_{p^*}(1),$$

in probability, for every $0 < b < \infty$.

B.2. *Under the null hypothesis* H_0, $E^*[G_n^*(z)] = G_n(z) + o_p(1)$, $z \ge 0$*, and for every relatively compact set* $\mathcal{U} \subset D[0,\infty]$, $\sup_{y \ge 0, \alpha \in \mathcal{U}} \big| \int_0^y \alpha(x)[dG_n^*(x) - dG_n(x)] \big| = o_{p^*}(1)$*, in probability.*

B.3. $\forall\, \epsilon > 0$, $0 < \eta < \infty$, $\exists\, N_{\epsilon,\eta}$, $0 < b \equiv b_{\epsilon,\eta} < \infty$, *such that* $\forall\, n > N_{\epsilon,\eta}$,

$$\mathsf{Pr}\Big[P_n^*\Big(\inf_{\|t\| > b} M_n^*(\widehat{\theta}_n + n^{-1/2}t) \ge \eta \Big) \ge 1 - \epsilon \Big] \ge 1 - \epsilon,$$

The assumptions (B.1) and (B.3) are the bootstrap analogs of (C.2) and (C.3), respectively.

Let $T_n^*(t) := M_n^*(\widehat{\theta}_n + n^{-1/2}t)$, $t \in \mathbb{R}^q$. We shall write $U_n^*(z)$ for $U_n^*(z, \widehat{\theta}_n)$. Note that

$$U_n^*(z) = n^{-1/2} \sum_{i=1}^n (\varepsilon_i^* - 1) I(Z_{i-1}^* \le z), \quad E^* U_n^*(z) \equiv 0, \quad E^*\big(U_n^*(z)\big)^2 \equiv s_n^2 E^*\big(G_n^*(z)\big).$$

Because $s_n^2 = \sigma^2 + o_p(1)$, $E^*[G_n^*(z)] = G_n(z) + o_p(1)$, and $\sup_{z \ge o} |G_n(z) - G(z)| = o_p(1)$, the continuity of B implies that analogous to Lemma 2.2, that conditionally, given the original sample,

$$U_n^* \text{ converges weakly to } \sigma B(G), \quad \text{in probability,} \tag{4.3}$$

in Skorokhod space and uniform metric. Hence

$$T_n^*(0) = M_n^*(\widehat{\theta}_n) \le \sup_{z\ge 0} \big(U_n^*(z)\big)^2 = O_{p^*}(1), \quad \text{in probability.} \tag{4.4}$$

This fact is useful in showing that $\Delta_n^* := n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n) = O_{p^*}(1)$, in probability, which in turn is needed for deriving the asymptotic distribution of T_n^*.

Let $\varphi(z,\vartheta) := \dot{\psi}(z,\vartheta)/\psi(z,\vartheta)$, $\vartheta \in \Theta$. Then as in Section 2, by (C.1) and (B.1),

$$\sup_{n\ge N} E^*\Big(\frac{\|\dot{\psi}(Z_0^*,\widehat{\theta}_n)\|^2}{\psi^j(Z_0^*,\vartheta)}\Big) < \infty, \quad j\ge 1, \qquad \sup_{n\ge N} E^*\|\varphi(Z_0^*,\widehat{\theta}_n)\|^2 < \infty, \quad \text{a.s.} \tag{4.5}$$

To proceed further, let

$$\begin{aligned}
&\Psi_n^*(z) := n^{-1}\sum_{i=1}^n \varphi(Z_{i-1}^*,\widehat{\theta}_n)\, I(Z_{i-1}^*\le z), \qquad \Psi^*(z) := E^*\big(\varphi(Z_0^*,\widehat{\theta}_n) I(Z_0^*\le z)\big),\\
&\tilde{M}_n^*(\widehat{\theta}_n) := \int \big(U_n^*\big)^2 dG, \quad V_n^* := \int U_n^*(z)\Psi^*(z)dG(z), \quad \mathcal{G}^* := \int \Psi^*(z)\Psi^*(z)'dG(z),\\
&Q_n^*(t) := \int \big(U_n^*(z) - t'\Psi^*(z)\big)^2 dG(z) = \tilde{M}_n^*(\widehat{\theta}_n) - 2t'V_n^* + t'\mathcal{G}^* t,\\
&\widetilde{t}_n^* := \text{argmin}_t Q_n^*(t), \qquad \xi^*(x) := \int_{z\ge x}\Psi^*(z)dG(z), \quad x\ge 0.
\end{aligned}$$

The Fubini Theorem, G being a d.f. and (4.5) imply

$$\begin{aligned}
&\xi^*(x) = \int_{z\ge x} E^*\big(\varphi(Z_0^*,\widehat{\theta}_n) I(Z_0^*\le z)\big)dG(z) = E^*\Big(\varphi(Z_0^*,\widehat{\theta}_n)\big[1 - G\big((Z_0^*\vee x)-\big)\big]\Big),\\
&\sup_{x\ge 0,\, n\ge N} \|\xi^*(x)\|^2 \le \sup_{n\ge N} E^*\big(\|\varphi(Z_0^*,\widehat{\theta}_n)\|^2\big) < \infty, \quad \text{a.s.}
\end{aligned}$$

Hence $\Sigma^* := E^*(\xi^*(Z_0^*)\xi_2(Z_0^*)')$ is well defined. Further, for two independent copies of Z_0^*, say Z_{01}^* and Z_{02}^*, one can rewrite

$$\Sigma^* = E^*\Big\{E^*\Big(\varphi(Z_0^*,\widehat{\theta}_n)\varphi(Z_{01}^*,\widehat{\theta}_n)'\big[1 - G\big((Z_0^*\vee Z_{02}^*)-\big)\big]\big[1 - G\big((Z_{01}^*\vee Z_{02}^*)-\big)\big]\Big| Z_{02}^*\Big)\Big\}.$$

Moreover,

$$\begin{aligned}
V_n^* &:= n^{-1/2}\textstyle\sum_{i=1}^n(\varepsilon_i^*-1)\int I(Z_{i-1}^*\le z)\Psi^*(z)dG(z) = n^{-1/2}\sum_{i=1}^n(\varepsilon_i^*-1)\xi^*(Z_{i-1}^*),\\
&E^*V_n^*\equiv 0, \qquad E[V_n^*(V_n^*)'] \equiv s_n^2 E^*(\xi^*(Z_0^*)\xi^*(Z_0^*)') = s_n^2\Sigma^*.
\end{aligned} \tag{4.6}$$

Proposition 4.1. *Suppose that H_0 holds, and (C.1) and (B.1) are satisfied. Then,*

$$\sup_{\|t\|\le b}\big|M_n^*(\widehat{\theta}_n + n^{-1/2}t) - Q_n^*(t)\big| = o_{p^*}(1), \ \textit{in probability}, \quad \forall\, 0<b<\infty. \tag{4.7}$$

If, in addition, (B.3) also holds, then $\|n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n)\| = O_{p^*}(1)$, in probability. If, further $\mathcal{G}^*$ is positive definite, then $\widetilde{t}_n^* = (\mathcal{G}^*)^{-1}V_n^*$ and $\|n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n) - \widetilde{t}_n^*\| = o_{p^*}(1)$, in probability.

In the proof of Proposition 4.1 we make use of the following preliminary result given in KPB.

Lemma 4.1. *Let $\mathcal{U}$ be a relatively compact subset of $D[0,\infty]$. Let μ_n, μ be a sequence of random nondecreasing right continuous functions on $[0,\infty)$ having left limits and such that $\mu_n(\infty)+\mu(\infty) < \infty$, a.s., and $\sup_{x\geq 0}\big|\mu_n(x)-\mu(x)\big| \to 0$, a.s.. Then*

$$\sup_{y\geq 0,\alpha\in\mathcal{U}}\Big|\int_0^y \alpha(x)[d\mu_n(x)-d\mu(x)]\Big| \to_p 0.$$

Proof of Proposition 4.1. The proof of (4.7) is given below. Since V_n^* is a vector of the sums of martingale difference arrays satisfying (4.6), by extending the arguments in the proof of Theorem 5.4.1 of Koul [15] to a triangular array setup, one verifies that (4.4), (B.3) and (4.7) imply $\|n^{1/2}(\widehat{\theta}_n^*-\widehat{\theta}_n)\| = O_{p^*}(1)$, and that $\|n^{1/2}(\widehat{\theta}_n^*-\widehat{\theta}_n)-\widetilde{t}_n^*\| = o_{p^*}(1)$, in probability. By the positive definiteness of $\mathcal{G}^*$ we clearly have $\widetilde{t}_n^* = (\mathcal{G}^*)^{-1}V_n^*$. We prove (4.7) by making use of some arguments from the proofs of Theorems 2.1 and 2.2 in KPB adapted to the current bootstrap setup. The details are given below.

Proof of (4.7). Let $\theta_{nt} := \widehat{\theta}_n + n^{-1/2}t$, $t\in\mathbb{R}^q$. Define, for $z\geq 0$, $t\in\mathbb{R}^q$,

$$W_n^*(z,t) := n^{-1/2}\sum_{i=1}^n\Big[\frac{\psi(Z_{i-1}^*,\widehat{\theta}_n)}{\psi(Z_{i-1}^*,\theta_{nt})}-1\Big]\,\varepsilon_i^*\,I(Z_{i-1}^*\leq z), \tag{4.8}$$

$$S_n^*(z) := n^{-1}\sum_{i=1}^n\varphi(Z_{i-1}^*,\widehat{\theta}_n)\,(\varepsilon_i^*-1)\,I(Z_{i-1}^*\leq z),$$

$$\tilde{V}_n^* := \int U_n^*(z)\Psi^*(z)dG_n^*(z), \quad \mathcal{G}_n^* := \int \Psi^*(z)\Psi^*(z)'dG_n^*(z),$$

$$\tilde{Q}_n^*(t) := \int \big(U_n^*(z)-t'\Psi^*(z)\big)^2 dG_n^*(z) = M_n^*(\widehat{\theta}_n) - 2t'\tilde{V}_n^* + t'\mathcal{G}_n^* t,$$

$$T_n^*(t) := M_n^*(\theta_{nt}), \qquad T_{n1}^*(t) := \int (W_n^*(z,t)+t'\Psi^*(z))^2 dG_n^*(z),$$

$$T_{n2}^*(t) := \int (W_n^*(z,t)+t'\Psi^*(z))(U_n^*(z)-t'\Psi^*(z))dG_n^*(z).$$

Use the model assumption $\varepsilon_i^* = Y_i^*/\psi(Z_{i-1}^*,\widehat{\theta}_n)$ to obtain

$$\begin{aligned} U_n^*(z,\theta_{nt}) &= n^{-1/2}\sum_{i=1}^n\Big[\frac{Y_i^*}{\psi(Z_{i-1}^*,\theta_{nt})}-1\Big]I(Z_{i-1}^*\leq z) \\ &= n^{-1/2}\sum_{i=1}^n\Big[\frac{Y_i^*}{\psi(Z_{i-1}^*,\theta_{nt})}-\frac{Y_i^*}{\psi(Z_{i-1}^*,\widehat{\theta}_n)}+\frac{Y_i^*}{\psi(Z_{i-1}^*,\widehat{\theta}_n)}-1\Big]I(Z_{i-1}^*\leq z) \\ &= n^{-1/2}\sum_{i=1}^n\Big[\frac{\psi(Z_{i-1}^*,\widehat{\theta}_n)}{\psi(Z_{i-1}^*,\theta_{nt})}-1\Big]\,\varepsilon_i^*\,I(Z_{i-1}^*\leq z) + n^{-1/2}\sum_{i=1}^n(\varepsilon_i^*-1)I(Z_{i-1}^*\leq z) \\ &= W_n^*(z,t)+U_n^*(z), \qquad \forall\, z\geq 0,\ t\in\mathbb{R}. \end{aligned} \tag{4.9}$$

Hence, one can write

$$\begin{aligned} T_n^*(t) &= \int \{U_n^*(z,\theta_{nt})\}^2 dG_n^*(z) = \int \Big(W_n^*(z,t)+t'\Psi^*(z)+U_n^*(z)-t'\Psi^*(z)\Big)^2 dG_n^*(z) \\ &= T_{n1}^*(t)+2T_{n2}^*(t)+\tilde{Q}_n^*(t). \end{aligned} \tag{4.10}$$

We shall shortly prove the following facts. For every $0 < b < \infty$,

$$\text{(4.11)} \quad \text{(a)} \quad \sup_{\|t\|\le b} T^*_{n1}(t) = o_{p^*}(1), \ \text{in probability}, \quad \text{(b)} \quad \sup_{\|t\|\le b} \tilde{Q}^*_n(t) = O_{p^*}(1), \ \text{in probability.}$$

By (4.11) and the Cauchy-Schwarz inequality,

$$\sup_{\|t\|\le b} \left|T^*_{n2}(t)\right|^2 \le \sup_{\|t\|\le b} T^*_{n1}(t) \sup_{\|t\|\le b} \tilde{Q}^*_n(t) = o_{p^*}(1),$$

in probability, and hence the claim (4.7) follows.

Proof of (4.11)(a). Rewrite

$$\begin{aligned}
W^*_n(z,t) &:= n^{-1/2} \sum_{i=1}^n \Big[\frac{\psi(Z^*_{i-1},\widehat{\theta}_n)}{\psi(Z^*_{i-1},\theta_{nt})} - 1\Big]\, \varepsilon^*_i\, I(Z^*_{i-1} \le z) \\
&= -n^{-1/2} \sum_{i=1}^n \frac{1}{\psi(Z^*_{i-1},\theta_{nt})} \Big[\psi(Z^*_{i-1},\theta_{nt}) - \psi(Z^*_{i-1},\widehat{\theta}_n) - n^{-1/2} t' \dot{\psi}(Z^*_{i-1},\widehat{\theta}_n)\Big]\, \varepsilon^*_i\, I(Z^*_{i-1} \le z) \\
&\quad - t' n^{-1} \sum_{i=1}^n \Big[\frac{1}{\psi(Z^*_{i-1},\theta_{nt})} - \frac{1}{\psi(Z^*_{i-1},\widehat{\theta}_n)}\Big] \dot{\psi}(Z^*_{i-1},\widehat{\theta}_n)\, \varepsilon^*_i\, I(Z^*_{i-1} \le z) \\
&\quad - t' n^{-1} \sum_{i=1}^n \varphi(Z^*_{i-1},\widehat{\theta}_n)(\varepsilon^*_i - 1)\, I(Z^*_{i-1} \le z) \\
&\quad - t' n^{-1} \sum_{i=1}^n \Big[\varphi(Z^*_{i-1},\widehat{\theta}_n)\, I(Z^*_{i-1} \le z) - E^*\big(\varphi(Z^*_0,\widehat{\theta}_n) I(Z^*_0 \le z)\big)\Big] - t'\Psi^*(z).
\end{aligned}$$

Let $d^*_{it} := \psi(Z^*_{i-1},\theta_{nt}) - \psi(Z^*_{i-1},\widehat{\theta}_n)$ and $\delta^*_{it} := d^*_{it} - n^{-1/2} t' \dot{\psi}(Z^*_{i-1},\widehat{\theta}_n)$. Then the above identity is equivalent to

$$\begin{aligned}
\text{(4.12)} \quad & W^*_n(z,t) + t'\Psi^*(z) \\
&= -n^{-1/2} \sum_{i=1}^n \frac{\delta^*_{it}}{\psi(Z^*_{i-1},\theta_{nt})}\, \varepsilon^*_i\, I(Z^*_{i-1} \le z) \\
&\quad + t' n^{-1} \sum_{i=1}^n \frac{d^*_{it}}{\psi(Z^*_{i-1},\theta_{nt})} \varphi(Z^*_{i-1},\widehat{\theta}_n)\, \varepsilon^*_i\, I(Z^*_{i-1} \le z) - t' S^*_n(z) - t'\big(\Psi^*_n(z) - \Psi^*(z)\big) \\
&= A^*_1(z,t) + A^*_2(z,t) - t' S^*_n(z) - t'\big(\Psi^*_n(z) - \Psi^*(z)\big), \qquad \text{say.}
\end{aligned}$$

By (C.1), $\inf_{z\ge 0, \vartheta\in\Theta} \psi(z,\vartheta) \ge C > 0$ and by (B.1), for every $0 < b < \infty$,

$$\sup_{1\le i\le n, \|t\|\le b} n^{1/2} \big|\delta^*_{it}\big| = o_{p^*}(1),$$

in probability. Moreover, because $E^*(\varepsilon^*_0) = 1$, we have $n^{-1}\sum_{i=1}^n \varepsilon^*_i = O_{p^*}(1)$ in probability. Hence

$$\text{(4.13)} \quad \sup_{z\ge 0, \|t\|\le b} |A^*_1(z,t)| \le C^{-1} \sup_{1\le i\le n, \|t\|\le b} n^{1/2} \big|\delta^*_{it}\big| n^{-1} \sum_{i=1}^n \varepsilon^*_i = o_{p^*}(1), \ \text{in probability.}$$

Since $\sup_{n\ge N} E^*\|\dot\psi(Z_0^*,\widehat\theta_n)\|^2<\infty$, a.s., we obtain that $n^{-1/2}\max_{1\le i\le n}\|\dot\psi(Z_{i-1}^*,\widehat\theta_n)\|=o_{p^*}(1)$, in probability. Hence (C.1) and (B.1) imply that

$$D_n^* \;:=\; \sup_{1\le i\le n,\|t\|\le b}|d_{it}^*|\le \sup_{1\le i\le n,\|t\|\le b}\left|\delta_{it}^*\right|+bn^{-1/2}\max_{1\le i\le n}\|\dot\psi(Z_{i-1}^*,\widehat\theta_n)\|=o_{p^*}(1),$$

in probability. Moreover, $\sup_{n\ge N}E^*\big(n^{-1}\sum_{i=1}^n\|\varphi(Z_{i-1}^*,\widehat\theta_n)\|\,\varepsilon_i^*\big)=\sup_{n\ge N}E^*\|\varphi(Z_0^*,\widehat\theta_n)\|<\infty$ a.s. and hence by the Markov inequality, $n^{-1}\sum_{i=1}^n\|\varphi(Z_{i-1}^*,\widehat\theta_n)\|\,\varepsilon_i^*=O_{p^*}(1)$, and

$$\sup_{z\ge0,\|t\|\le b}|A_2^*(z,t)|\le bC^{-1}D_n^*n^{-1}\sum_{i=1}^n\|\varphi(Z_{i-1}^*,\widehat\theta_n)\|\,\varepsilon_i^*=o_{p^*}(1),\ \text{in probability.}\tag{4.14}$$

Next, consider $S_n^*(z)$. Observe that $S_n^*(z)$ is a vector of marked empirical processes of $Z_0^*,\cdots,Z_{n-1}^*$ with marks $n^{-1/2}\varphi(Z_0^*,\widehat\theta_n)\,(\varepsilon_1^*-1),\cdots,n^{-1/2}\varphi(Z_{n-1}^*,\widehat\theta_n)\,(\varepsilon_n^*-1)$. Further, the summands of each component are stationary and ergodic, and $E^*S_n^*(z)\equiv0$. Hence, a Glivenko-Cantelli Lemma type argument yields $\sup_{z\ge0}\|S_n^*(z)\|=o_{p^*}(1)$, in probability. Similarly, one also obtains $\sup_{z\ge0}\big\|\Psi_n^*(z)-\Psi^*(z)\big\|^2=o_{p^*}(1)$, in probability. For details see Koul [16]. Upon combining these two facts with (4.12), (4.13) and (4.14) we obtain that for every $0<b<\infty$,

$$\sup_{z\ge0,\|t\|\le b}\big|W_n^*(z,t)+t'\Psi^*(z)\big|=o_{p^*}(1),\quad\text{in probability.}$$

This fact combined with the definition of T_{n1}^* and G_n^* being a d.f. readily yields (4.11)(a).

Proof of (4.11)(b). Recall that $\tilde M_n^*(\widehat\theta_n):=\int\{U_n^*(z,\widehat\theta_n)\}^2dG(z)$, $\mathcal{G}^*:=\int\Psi^*(z)\Psi^*(z)'dG(z)$. Note that $\|\mathcal{G}^*\|\le\int\|\Psi^*(z)\|^2dG(z)\le E^*\|\varphi(Z_0^*,\widehat\theta_n)\|^2=O_{p^*}(1)$, in probability, and we have $\tilde M_n^*(\widehat\theta_n)\le\sup_{z\ge0}\{U_n^*(z,\widehat\theta_n)\}^2=O_{p^*}(1)$, in probability, by (4.4). Hence, for every $0<b<\infty$,

$$\begin{aligned}\sup_{\|t\|\le b}Q_n^*(t)\;&:=\;\sup_{\|t\|\le b}\int\Big(U_n^*(z,\widehat\theta_n)-t'\Psi^*(z)\Big)^2dG(z)\;\le\;2\tilde M_n^*(\widehat\theta_n)+2b^2\|\mathcal{G}^*\|\\&=\;O_{p^*}(1),\quad\text{in probability.}\end{aligned}$$

Hence, it suffices to show that for every $0<b<\infty$,

$$\sup_{\|t\|\le b}\big|\tilde Q_n^*(t)-Q_n^*(t)\big|=o_{p^*}(1),\quad\text{in probability.}\tag{4.15}$$

To prove (4.15), let

$$\begin{aligned}B_{n1}^*&:=\Big|\int(U_n^*)^2\big[dG_n-dG\big]\Big|, &\quad B_{n2}^*&:=\Big\|\int U_n^*\Psi^*\big[dG_n-dG\big]\Big\|,\\ \widetilde B_{n1}^*&:=\Big|\int(U_n^*)^2\big[dG_n^*-dG_n\big]\Big|, &\quad \widetilde B_{n2}^*&:=\Big\|\int U_n^*\Psi^*\big[dG_n^*-dG_n\big]\Big\|.\end{aligned}$$

Note that the left hand side of (4.15) is bounded from the above by

$$\begin{aligned}&\Big|\int(U_n^*)^2\big[dG_n^*-dG\big]\Big|+2b\Big\|\int U_n^*\Psi^*\big[dG_n^*-dG\big]\Big\|+b^2\Big\|\int\Psi^*(\Psi^*)'\Big[dG_n^*-dG\Big]\Big\|\\&\le(B_{n1}^*+\widetilde B_{n1}^*)+2b(B_{n2}^*+\widetilde B_{n2}^*)+b^2\Big\|\int\Psi^*(\Psi^*)'\big[dG_n^*-dG\big]\Big\|.\end{aligned}$$

Since $E^*[G_n^*(z)] = G_n(z) + o_p(1)$ and $\sup_{x\geq 0}\left|G_n(x) - G(x)\right| \to 0$, a.s., by the Ergodic Theorem and Assumption B.2, $\left\|\int \Psi^*(\Psi^*)'\left[dG_n^* - dG\right]\right\| = o_{p^*}(1)$, in probability. Since $\sup_{x\geq 0}\left|G_n(x) - G(x)\right| \to 0$, a.s., by Lemmas 2.2 and 4.1, B_{n1}^* and B_{n2}^* tend to zero in probability. Similarly, by Lemma 2.2 and Assumption B.2, $\widetilde{B}_{n1}^*$ and $\widetilde{B}_{n2}^*$ are of order $o_{p^*}(1)$ in probability. This completes the proof of (4.15), and hence (4.11)(b) follows. □

Next we establish the asymptotic validity of the bootstrap testing procedure. To this end we make use of some ideas from Cavaliere, Perera and Rahbek [3].

First, we generalize the DGP specified by model (1.1)–(1.2) for an arbitrary true value $\vartheta \in \Theta$ and a given error d.f. L, with unit mean and finite variance, as follows:

$$Y_i^{(\vartheta,L)} = \psi(Z_{i-1}^{(\vartheta,L)}, \vartheta)\varepsilon_i^{(L)}, \tag{4.16}$$

where $\varepsilon_i^{(L)} = L^{-1}(U_i) := \inf\{y \in \mathbb{R} : L(y) \geq U_i\}$ and $\{U_i, i \in \mathbb{Z}\}$ are i.i.d. uniform(0,1) r.v.'s.

Let F denote the d.f. of ε_i in the model (1.1)–(1.2). Let F_n, $n \in \mathbb{N}$, be a sequence of cumulative distribution functions, with unit mean and finite variance, and $\theta_n \in \Theta$, such that $\|\theta_n - \theta\| \to 0$ and $d_2(F_n, F) \to 0$ as $n \to \infty$, where $d_2(F_X, F_Y)$ is the Mallows metric for the distance between two probability distributions F_X and F_Y, with $d_2(F_X, F_Y) = \inf\{E|X - Y|^2\}^{1/2}$, where the infimum is over all square integrable random variables X and Y with marginal d.f.'s F_X and F_Y; see Freedman [9] and Bickel and Freedman [2], Section 8, for a detailed account of this metric.

Note that, for $(\vartheta, L) = (\theta, F)$ the model (4.16) is equivalent to the null model (1.1)–(1.2). Hence, the d.f. G of Z_0 in (1.1)–(1.2) is the same as the stationary distribution of the Z-process induced by (θ, F) in (4.16). We need to make the following additional assumption.

B.4. *Under H_0, for every nonrandom sequence $(\theta_n, F_n) \to (\theta, F)$ where $\|\theta_n - \theta\| \to 0$ and $d_2(F_n, F) \to 0$, $\sup_z |\tilde{G}_n(z) - G(z)| = o(1)$ and $\sup_x \|\varphi(x, \theta_n) - \varphi(x, \theta)\| = o(1)$, where $\tilde{G}_n$ is the stationary distribution of the Z-process induced by (θ_n, F_n) in* (4.16).

The next theorem shows that the conditional distribution, given $\mathcal{Y} = (Y_{-p}, Y_{1-p}, \cdots, Y_n)'$, of the bootstrap statistic T_n^*, consistently estimates the limiting null distribution of T, in probability.

Theorem 4.1. *Suppose that H_0 holds, and assumptions (C.1)–(C.3) and (B.1)–(B.4) are satisfied. Further assume that $\mathcal{G}$, $\mathcal{G}^*$ are positive definite. Then, $T_n^* := (s_n^*)^{-2} M_n^*(\widehat{\theta}_n^*) \to_{D^*} \mathcal{Z}$, in probability, where $\mathcal{Z}$ is as in Theorem 2.1.*

Proof of Theorem 4.1. Note that

$$|M_n^*(\widehat{\theta}_n^*) - Q_n^*(\widetilde{t}_n^*)| \leq |M_n^*(\widehat{\theta}_n^*) - M_n^*(\widehat{\theta}_n + n^{-1/2}\widetilde{t}_n^*)| + |M_n^*(\widehat{\theta}_n + n^{-1/2}\widetilde{t}_n^*) - Q_n^*(\widetilde{t}_n^*)|. \tag{4.17}$$

Let $\tilde{G}_n^*$ denote the stationary d.f. of Z_i^*-process, conditional on $\mathcal{Y} = (Y_{-p}, Y_{1-p}, \cdots, Y_n)'$. Then

$$\begin{aligned} &\Big\| \int_0^z \varphi(x, \widehat{\theta}_n) d\tilde{G}_n^*(x) - \int_0^z \varphi(x, \theta) dG(x) \Big\| \\ &\leq \sup_x \|\varphi(x, \widehat{\theta}_n) - \varphi(x, \theta)\| + \Big\| \int_0^z \varphi(x, \theta)[d\tilde{G}_n^*(x) - dG(x)] \Big\|. \end{aligned} \tag{4.18}$$

Lemma 2.1 implies $\|\widehat{\theta}_n - \theta\| \to_p 0$. Further, by arguing as in the proof of Lemma 6 in Perera and Silvapulle [26] one obtains $d_2(\widetilde{F}_n, F) \to_p 0$. These two facts together with condition (B.4) yield

$$(4.19) \qquad \text{(a)} \quad \sup_z |\tilde{G}_n^*(z) - G(z)| = o_p(1), \qquad \text{(b)} \quad \sup_x \|\varphi(x, \widehat{\theta}_n) - \varphi(x, \theta)\| = o_p(1).$$

By (4.19)(b) the first term of the upper bound (4.18) is of order $o_p(1)$. Since $\tilde{G}_n^*$ and G are uniformly bounded, the second term can be argued to tend to zero in probability as in Lemma 4.1 by using (4.19)(a). Therefore, $E^*\big[\varphi(Z_0^*, \widehat{\theta}_n)I(Z_0^* \le z)\big] \to_p E\big[\varphi(Z_0, \theta)I(Z_0 \le z)\big]$, $z \ge 0$, and hence $\Sigma^* := E^*(\xi^*(Z_0^*)\xi_2(Z_0^*)') \to_p \Sigma$ and $\mathcal{G}^* := \int \Psi^*(z)\Psi^*(z)'dG(z) \to_p \mathcal{G}$.

Since V_n^* is a vector of the sums of martingale difference arrays satisfying (4.6) and $s_n^2 \to_p \sigma^2$, then by the martingale central limit theorem, see Hall and Heyde [13], $V_n^* \to_{D^*} N(0, \sigma^2\Sigma)$, in probability. Hence $\widetilde{t}_n^* = \{\mathcal{G}^*\}^{-1}V_n^* = O_{p^*}(1)$, in probability. Therefore, by (4.7) of Proposition 4.1,

$$|M_n^*(\widehat{\theta}_n + n^{-1/2}\widetilde{t}_n^*) - Q_n^*(\widetilde{t}_n^*)| = o_{p^*}(1), \quad \text{in probability.}$$

Since Proposition 4.1 yields $\|n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n) - \widetilde{t}_n^*\| = o_{p^*}(1)$, in probability, by assumptions (C.1), (B.1) and (B.3) one also obtains that $|M_n^*(\widehat{\theta}_n^*) - M_n^*(\widehat{\theta}_n + n^{-1/2}\widetilde{t}_n^*)| = o_{p^*}(1)$, in probability. These two facts in combination with (4.17) imply that $|M_n^*(\widehat{\theta}_n^*) - Q_n^*(\widetilde{t}_n^*)| = o_{p^*}(1)$, in probability.

Since $\widetilde{t}_n^* = \{\mathcal{G}^*\}^{-1}V_n^*$ one can rewrite

$$(4.20) \qquad \begin{aligned} Q_n^*(\widetilde{t}_n^*) &= \int \big(U_n^*(z) - (\widetilde{t}_n^*)'\Psi^*(z)\big)^2 dG(z) = \tilde{M}_n^*(\widehat{\theta}_n) - 2(\widetilde{t}_n^*)'V_n^* + (\widetilde{t}_n^*)'\mathcal{G}^*(\widetilde{t}_n^*) \\ &= \tilde{M}_n^*(\widehat{\theta}_n) - 2(V_n^*)'\{\mathcal{G}^*\}^{-1}V_n^* + (V_n^*)'\{\mathcal{G}^*\}^{-1}\mathcal{G}^*\{\mathcal{G}^*\}^{-1}V_n^* \\ &= \tilde{M}_n^*(\widehat{\theta}_n) - (V_n^*)'\{\mathcal{G}^*\}^{-1}V_n^* \\ &= \int \{U_n^*\}^2 dG - \int U_n^*(\Psi^*)'dG\ \{\mathcal{G}^*\}^{-1}\int U_n^*\Psi^* dG. \end{aligned}$$

Since $(s_n^*)^2 = s_n^2 + o_{p^*}(1)$, in probability, and $s_n^2 \to_p \sigma^2$, by (4.3) and the Slutsky's Theorem, $(s_n^*)^{-1}U_n^*$ converges weakly to $B \circ G$, in probability, in the Skorokhod space and uniform metric, where B is the standard Brownian motion on $[0, \infty)$, and hence $(s_n^*)^{-1}U_n^*$ is of order $O_{p^*}(1)$, in probability. Moreover, as above $\Psi^*(z) := E^*\big(\varphi(Z_0^*, \widehat{\theta}_n)I(Z_0^* \le z)\big) \to_p \Psi(z)$, $z \ge 0$, and $\mathcal{G}^* \to_p \mathcal{G}$. By these facts and (4.20),

$$(s_n^*)^{-2}Q_n^*(\widetilde{t}_n^*) = \int \{(s_n^*)^{-1}U_n^*\}^2 dG - \int (s_n^*)^{-1}U_n^*\Psi' dG\ \mathcal{G}^{-1}\int (s_n^*)^{-1}U_n^*\Psi dG + o_{p^*}(1),$$

in probability. Hence, one obtains

$$\begin{aligned} (s_n^*)^{-2}M_n^*(\widehat{\theta}_n^*) &\to_{D^*} \int B^2(G(z))dG(z) - \int B(G(z))\Psi(z)'dG(z)\ \mathcal{G}^{-1}\int B(G(z))\Psi(z)dG(z) \\ &= \int_0^1 B^2(u)du - \int_0^1 B(u)\Psi(G^{-1}(u))'du\ \mathcal{G}^{-1}\int_0^1 B(u)\Psi(G^{-1}(u))du \\ &= \mathcal{Z}, \quad \text{in probability.} \qquad \square \end{aligned}$$

5 Conclusion

This paper advances the current state of econometric methodology in MEMs for nonnegative time series. In particular, we propose a new minimum distance lack-of-fit test for a parametric specification of the conditional mean in a Markovian MEM. We derive the asymptotic null distribution of the proposed test statistic under fairly general and easily verifiable conditions. Since the limiting distribution of the test statistic is model dependent and is not free from nuisance parameters, we propose a bootstrap method to implement the test and establish that the proposed bootstrap method is asymptotically valid. Our bootstrap test is easy to implement and is flexible enough to be applied to a wide range of parametric specifications of the conditional mean in the Markovian MEM setup. The simulation findings about empirical level and power of the test demonstrate that the new testing procedure performs better than some of the competing ones.

6 No Conflict

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

[1] Balakrishna, N., Koul, H.L., Sakhanenko, L. and Ossiander, M. (2019). Fitting a pth order parametric generalized linear autoregressive multiplicative error model. *Sankhya: The Indian Journal of Statistics*, **81-B**, Supplement 1, pp. S103–S122.

[2] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.*, **9**, 1196–21217.

[3] Cavaliere, G., Perera, I. and Rahbek, A. (2021). Specification tests for GARCH processes. *Working paper.*

[4] Chen, Y.-T. and C.-S. Hsieh (2010). Generalized moment tests for autoregressive conditional duration models. *J. Financ. Econom.*, **8**, 345–391.

[5] Chou, R.Y. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. *J. Money, Credit Banking*, **37(3)**, 561–582.

[6] Engle, R. (2002) New frontiers for ARCH models. *J. of App. Econometrics*, **17**, 425–446.

[7] Engle, R.F. and Gallo, G.M. (2006). A multiple indicators model for volatility using intra-daily data. *J. Econometrics*, **131**, 3–27.

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

[9] Freedman D.A. (1981). Bootstrapping regression models. *Ann. Statist.*, **9**, 1218–1228.

[10] Giacomini, R., Politis, D. N. and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, **29**, 567–589.

[11] Grammig, J. and Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal*, **3**(1), 16–38.

[12] Guo, B. and Li, S. (2018). Diagnostic checking of markov multiplicative error models. *Economics Letters*, **170**, 139 – 142.

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

[14] Hautsch, N. (2012). *Econometrics of financial high-frequency data.* Springer, Heidelberg, Germany.

[15] Koul, Hira L. (2002). *Weighted empirical processes in dynamic nonlinear models. 2nd Edition.* **Lecture Notes Series in Statistics, 166,** Springer, New York, N.Y., USA.

[16] Koul, Hira L. (2019). A uniform convergence result for weighted residual empirical processes. RM#719, Department of Statistics and Probability, MSU.

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

[18] Koul, H.L., Perera, I. and Balakrishna, N. (2021). A class of minimum distance estimators in markovian multiplicative error models. *Working paper.*

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

[20] Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

[21] Lunde, A. (1999). A generalized gamma autoregressive conditional duration model. *Discussion paper, Aalborg University.*

[22] Manganelli, S. (2005). Duration, volume and volatility impact of trades. *J. Financ. Mark.*, **8(4)**, 377–399.

[23] Meitz, M. and Teräsvirta, T. (2006). Evaluating models of autoregressive conditional duration. *J. Bus. Econom. Statist.*, **24**, 104–124.

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

[25] Perera, I. and Koul, H.L. (2017). Fitting a two phase threshold multiplicative error model. *J. Econometrics*, **197**, 348–367.

[26] Perera, I. and Silvapulle, M. J. (2021). Bootstrap based probability forecasting in multiplicative error models. *J. Econometrics*, **221**, 1–24.

[27] Stute, W. (1997). Nonparametric model checks for regression. *Ann. Statist.*, **25(2)**, 613–641.

[28] Stute, W., S. Thies, & Zhu, L.X. (1998). Model checks for regression: an innovation process approach. *Ann. Statist.*, **26**, 1916–1934.

Hira L. Koul
Department of Statistics and Probability
Michigan State University
East Lansing, MI 48824-1027
U.S.A.
E-mail: koul@msu.edu

Indeewara Perera
Department of Economics
The University of Sheffield
Sheffield S1 4DT
U.K.
e-mail: i.perera@sheffield.ac.uk