Statistics and Its Interface Volume 16 (2023) 3–16

On dual-asymmetry linear double AR models

Songhua Tan and Qianqian Zhu*

This paper introduces a dual-asymmetry linear double autoregressive (DA-LDAR) model that can allow for asymmetric effects in both the conditional location and volatility components of time series data. The strict stationarity is discussed for the new model, for which a sufficient condition is established. A self-weighted exponential quasi-maximum likelihood estimator (EQMLE) is proposed for the DA-LDAR model, and a mixed portmanteau test for goodness-of-fit is constructed based on the self-weighted EQMLE. It is noteworthy that all the asymptotic properties for estimation and testing are established without any moment condition on the data process, which makes the new model and its inference tools applicable for heavy-tailed data. Since all inference tools need to estimate the unknown density function of innovations, we employ a random-weighting bootstrap method to facilitate accurate inference and show its asymptotic validity. Simulation studies provide support for theoretical results, and an empirical application to NASDAQ Composite Index illustrates the usefulness of the new model.

Keywords and phrases: Asymmetry effects, Bootstrap method, Double autoregressive models, Exponential QMLE, Portmanteau test, Strict stationarity.

1. INTRODUCTION

In theory and practical applications, the conditional mean and conditional variance (volatility) are two important ingredients for time series data. Many classical time series models, such as the ARMA (Box et al., 2008) and GARCH (Bollerslev, 1986) models, are proposed and widely used to model these two components separately. Many empirical findings in literatures suggest that the autoregression and volatility dynamics usually exist together in time series, for example, the return series of NASDAQ Composite Index (Kuester et al., 2006) and S&P500 index (Linton and Mammen, 2005). As a result, it is of great importance to model the conditional mean and volatility simultaneously (Li et al., 2002). Many models are introduced for this purpose, and the double autoregressive (DAR) model proposed by Ling (2004, 2007) is one of the popular specifications. The DAR model of order p has the form of

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \eta_t \sqrt{\omega + \sum_{i=1}^{p} \beta_i y_{t-i}^2}, \tag{1}$$

where $\omega > 0$, $\alpha_i \in \mathbb{R}$, $\beta_i \geq 0$ for $1 \leq i \leq p$, and $\{\eta_t\}$ are independent and identically distributed ($i.i.d.$) innovations. Model (1) is a subclass of ARMA-ARCH models in Weiss (1984), but it is different from the ARCH model in Engle (1982) if $\alpha_i \neq 0$. As the DAR model attracts growing attention, many of its variants have been widely proposed and studied, such as the threshold DAR (Li et al., 2016), the mixture DAR (Li et al., 2017), the linear DAR (Zhu et al., 2018) and the augmented DAR (Jiang et al., 2020) models. Among them, the linear DAR model extends the DAR model along the lines of linear GARCH (Taylor, 2008; Xiao and Koenker, 2009) models which assume the conditional standard deviation rather than the conditional variance of y_t in a linear structure. Particularly, the linear DAR model is defined as

$$y_t = \sum_{i=1}^{p} \alpha_i y_{t-i} + \eta_t \left(\omega + \sum_{i=1}^{p} \beta_i |y_{t-i}| \right). \tag{2}$$

Similar to the linear GARCH model which can have more robust inference than the quadratic GARCH model (Xiao and Koenker, 2009), hopefully model (2) can lead to more robust inference than model (1) as well. For model (2), Zhu et al. (2018) and Liu et al. (2020) proposed a doubly weighted quantile regression estimator and two QMLEs, and notably all these estimators are asymptotically normal under a fractional moment on y_t. Hence, the linear DAR model is also applicable for heavy-tailed data.

It is well known that financial time series especially the stock returns usually have asymmetric effects, that is, negative shocks have much larger effects on the stock price than positive shocks of the same magnitude; see Black (1976) and Francq and Zakoian (2013) for empirical evidences. Many models are introduced in literatures to capture the asymmetric effect, generally speaking, these models can be divided into threshold-type and asymmetric-type. Due to the uncertainty of the threshold and the nonlinear structure of models, the threshold-type models such as the threshold AR (Petruccelli and Woolford, 1984) and threshold DAR (Li et al., 2016) models, usually have complex theoretical properties and thus are difficult to

*Corresponding author. ORCID: 0000-0001-7791-911X.

use in practice. By contrast, asymmetric-type models such as the exponential GARCH (Nelson, 1991), GJR–GARCH (Glosten et al., 1993), threshold GARCH (Zakoian, 1994) and power GARCH (Pan et al., 2008) models, avoid the selection of threshold and make easy the statistical inference. Note that the aforementioned asymmetric-type models mainly focused on asymmetric effects in the conditional volatility. However, when analyzing the NASDAQ Composite Index and many other stock indices, we found that the asymmetry appears in both the conditional mean and volatility; see Section 6 for details. This motivates us to fill this gap and propose an asymmetric-type model along the lines of the linear DAR model (2). Moreover, to make the model more practicable, we further relax the equal order setting in model (2) and allow for different orders in the conditional mean and volatility specifications. For the proposed new model, we investigate its strict stationarity and establish a sufficient condition in Section 2. Particularly, this paper has three main contributions.

First, we propose a self-weighted exponential quasi-maximum likelihood estimator (EQMLE) for model (3) in Section 3. For the DAR model (1) and linear DAR model (2) of equal order p, under $E(y_t^\kappa) < \infty$ for $\kappa > 0$ with $E(\eta_t^4) < \infty$, their Gaussian QMLEs are shown to be asymptotically normal (Ling, 2007; Liu et al., 2020). Zhu and Ling (2013) and Liu et al. (2020) proposed the EQMLEs for these two models respectively and established the asymptotic normality under $E(y_t^\kappa) < \infty$ for $\kappa > 0$ with $E(\eta_t^2) < \infty$. It is natural to consider the EQMLE for the new model to deal with more heavy-tailed data. If the order p of the conditional location is no larger than the order q of the conditional volatility, then the asymptotic normality of EQMLE for the new model requires the same moment conditions on data process $\{y_t\}$ and innovations $\{\eta_t\}$ as for models (1) and (2). However, when $p > q$, then $E(y_t^3) < \infty$ with $E(\eta_t^2) < \infty$ is necessary. To reduce the moment condition on y_t especially for the case of $p > q$, the self-weighting method proposed by Ling (2007) is introduced to the EQMLE. As a result, a self-weighted EQMLE is investigated in framework of the new model for generality.

Secondly, we construct a mixed portmanteau test in Section 4 to check the model adequacy without any moment conditions imposed on the data process $\{y_t\}$. It is well known that diagnostic checking is one of the key steps for time series modeling (Box et al., 2008), and portmanteau tests provide a standard tool to check the overall goodness-of-fit. Specifically, the sample autocorrelation functions (ACFs) of residuals are commonly used to construct portmanteau tests for conditional mean models (Ljung and Box, 1978), while the ACFs of squared or absolute residuals are utilized for volatility models (Li, 2004; Li and Li, 2005). For the new model with both conditional mean and volatility structures, diagnostic checking should be conducted for both components, and a mixed portmanteau test can be considered accordingly (Wong and Ling, 2005). Therefore, in line with the self-weighted EQMLE, we employ the ACFs of self-weighted residuals and self-weighted absolute residuals to check the adequacy of new model, which ensures that no moment conditions will be imposed on y_t.

Finally, we employ a random-weighting bootstrap procedure for the estimation and diagnostic checking, and show that the bootstrap approach is asymptotically valid. It is noteworthy that the asymptotic covariance of the self-weighted EQMLE and the portmanteau test statistic both depend on the unknown density function of innovations. As a result, the inference for the new model necessitates some non-parametric methods such as the kernel density estimation, which makes the estimation of asymptotic distributions more involved. Moreover, the sample approximation of asymptotic distributions may not be satisfactory when the sample size is small or even moderate. Owing to the powerful computing, the bootstrap approach (Efron, 1992) has been widely used to approximate limiting distributions of statistics. Among various bootstrap methods, the random-weighting bootstrap (Zheng, 1987) recently attracted great attentions in literatures of time series settings; see also Li et al. (2014) and Zhu et al. (2020). Motivated by the above observations, this paper turns to the bootstrap approximation and considers an easy-to-use random-weighting bootstrap procedure for the new model.

The remainder of this paper is organized as follows. Section 2 proposes the new model and discusses its strict stationarity. Section 3 proposes the self-weighted EQMLE and establishes its asymptotics, and Section 4 studies the diagnostic checking for the fitted model. Moreover, a random-weighting bootstrap procedure is proposed for the estimation and diagnosis. Section 5 conducts simulation experiments to examine the finite-sample performance of the proposed statistical inference and bootstrap procedure. Section 6 presents an empirical example to illustrate the usefulness of the new model. Section 7 concludes with a short discussion. All technical details are relegated to the supplementary material. Throughout the paper, $\to_p$ and $\to_{\mathcal{L}}$ denote the convergences in probability and in distribution, respectively, and $o_p(1)$ denotes a sequence of random variables converging to zero in probability. We denote by $\|\cdot\|$ the norm of a matrix or column vector, defined as $\|A\| = \sqrt{tr(AA')} = \sqrt{\sum_{i,j} a_{ij}^2}$.

2. THE MODEL

Consider the dual-asymmetry linear DAR (DA-LDAR) model of order (p, q):

$$y_t = \sum_{i=1}^{p} \left(\alpha_{i+} y_{t-i}^{+} + \alpha_{i-} y_{t-i}^{-}\right) + \eta_t \left(\omega + \sum_{j=1}^{q} \left(\beta_{j+} y_{t-j}^{+} - \beta_{j-} y_{t-j}^{-}\right)\right), \tag{3}$$

where $\omega > 0$, $\beta_{j+}, \beta_{j-} \geq 0$ for $1 \leq j \leq q$, $\{\eta_t\}$ are $i.i.d.$ innovations, $y_t^{+} = \max\{0, y_t\}$ and $y_t^{-} = \min\{0, y_t\}$ are positive and negative parts of $\{y_t\}$, respectively, and p and q are nonnegative integers. Model (3) allows for asymmetric effects in both the conditional location and scale components, and its orders of the conditional location and volatility specifications can be different. Model (3) includes the linear DAR model (2) of Zhu et al. (2018) as a special case, that is, it will reduce to model (2) when $p = q$, $\alpha_{i+} = \alpha_{i-}$ for all i and $\beta_{j+} = \beta_{j-}$ for all j.

Since model (3) is actually nonlinear, it is difficult to derive a necessary and sufficient condition for its strict stationarity when the distribution of η_t is general. As a special case, Zhu et al. (2018) provided a sufficient condition for the linear DAR model (2). For model (3) of general settings, Theorem 2.1 below gives a sufficient condition.

Theorem 2.1. *Suppose that the density function of η_t is positive everywhere on $\mathbb{R}$, and $E(|\eta_t|^\kappa) < \infty$ for some $\kappa > 0$. Let $m = \max\{p, q\}$, $\alpha_{i-} = \alpha_{i+} = 0$ for $i > p$ and $\beta_{i-} = \beta_{i+} = 0$ for $i > q$. If the condition in (i) or (ii) holds:*
(i) for $0 < \kappa \le 1$,

$$\sum_{i=1}^{m} \max\left\{E\left(|\alpha_{i-} - \beta_{i-}\eta_t|^\kappa\right), E\left(|\alpha_{i+} + \beta_{i+}\eta_t|^\kappa\right)\right\} < 1;$$

(ii) for $\kappa \in \{2, 3, 4 \ldots\}$,

$$E\left[\left(\sum_{i=1}^{m} \max\left\{|\alpha_{i-} - \beta_{i-}\eta_t|, |\alpha_{i+} + \beta_{i+}\eta_t|\right\}\right)^\kappa\right] < 1;$$

then there exists a strictly stationary solution $\{y_t\}$ to model (3), and this solution is unique and geometrically ergodic with $E(|y_t|^\kappa) < \infty$.

Remark 2.1. *For the geometric ergodicity and existence of κth moment of $\{y_t\}$, we can alternatively use the piggyback method in Cline and Pu (2004) to obtain a more sharp sufficient condition than that in Theorem 2.1. However, an extra moment condition $\sup_x(1 + |x|)f(x) < \infty$ on the density of η_t is required. Moreover, the sufficient condition based on the piggyback method is more complicated to verify than that in Theorem 2.1. As a result, we prefer the simple sufficient conditions in Theorem 2.1.*

The stationarity region in Theorem 2.1 depends on the distribution of η_t and implies a moment condition on y_t. Because the stationarity region is at least four-dimensional, for illustration, we consider model (3) of orders $p = q = 1$ with $\alpha_{1-} = 0.8\alpha_{1+}$ and $\beta_{1-} = 0.8\beta_{1+}$. We have the following findings from Figure 1: (1) a larger value of κ in Theorem 2.1 implies a higher order moment of y_t, and hence results in a narrower stationarity region; (2) the stationarity region is different as the distribution of η_t changes; (3) model (3) of orders $p = q = 1$ can be stationary even if $|\alpha_{1+}| > 1$, which leads to a larger parameter space than AR-ARCH models.

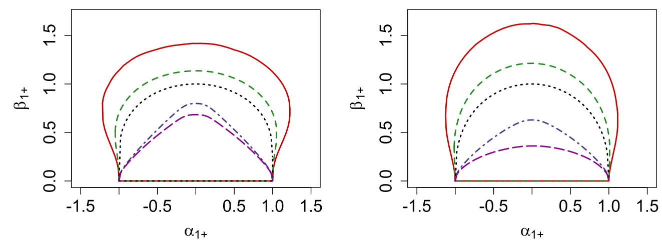

Figure 1. Stationarity regions of model (3) of order (1,1) with $\alpha_{1-} = 0.8\alpha_{1+}$ and $\beta_{1-} = 0.8\beta_{1+}$, where η_t follows the normal (left panel) or student's t_3 distribution (right panel) with $E(|\eta_t|) = 1$, and $\kappa = 0.1$ (red solid line), 0.6 (green dashed line), 1 (black dotted line), 2 (blue dotdash line) or 3 (purple longdash line).

3. SELF-WEIGHTED EQMLE

Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')' \in \mathbb{R}^d$ be the parameter vector of model (3) and $\boldsymbol{\theta}_0 = (\boldsymbol{\alpha}_0', \boldsymbol{\beta}_0')'$ be its true value, where $d = 2p + 2q + 1$, $\boldsymbol{\alpha} = (\alpha_{1+}, \ldots, \alpha_{p+}, \alpha_{1-}, \ldots, \alpha_{p-})'$ and $\boldsymbol{\beta} = (\omega, \beta_{1+}, \ldots, \beta_{q+}, \beta_{1-}, \ldots, \beta_{q-})'$. Denote the parameter space by Θ, where $\Theta \subset \mathbb{R}^{2p} \times \mathbb{R}_+^{2q+1}$ with $\mathbb{R}_+ = (0, \infty)$. Let $\boldsymbol{Y}_t = (\boldsymbol{Y}_{t+}', \boldsymbol{Y}_{t-}')'$ and $\boldsymbol{X}_t = (1, \boldsymbol{X}_{t+}', -\boldsymbol{X}_{t-}')'$, where $\boldsymbol{Y}_{t+} = (y_{t-1}^+, \ldots, y_{t-p}^+)'$, $\boldsymbol{Y}_{t-} = (y_{t-1}^-, \ldots, y_{t-p}^-)'$, $\boldsymbol{X}_{t+} = (y_{t-1}^+, \ldots, y_{t-q}^+)'$ and $\boldsymbol{X}_{t-} = (y_{t-1}^-, \ldots, y_{t-q}^-)'$. Then model (3) can be rewritten as follows

$$y_t = \boldsymbol{\alpha}'\boldsymbol{Y}_t + \boldsymbol{\beta}'\boldsymbol{X}_t\eta_t.$$

Assume that the observations $\{y_1, \ldots, y_n\}$ are generated by model (3) with the true parameter vector $\boldsymbol{\theta}_0$. Given $\{y_1, \ldots, y_n\}$, when η_t follows the standard double exponential distribution, ignoring a constant the negative weighted log-likelihood function has the form of

$$L_n(\boldsymbol{\theta}) = \sum_{t=p+q}^{n} \omega_t \ell_t(\boldsymbol{\theta}) \text{ and} \tag{4}$$
$$\ell_t(\boldsymbol{\theta}) = \ln(\boldsymbol{\beta}'\boldsymbol{X}_t) + \frac{|y_t - \boldsymbol{\alpha}'\boldsymbol{Y}_t|}{\boldsymbol{\beta}'\boldsymbol{X}_t},$$

where $\{\omega_t\}$ are positive random weights that only depend on $\{y_t\}$ itself (Ling, 2005). Let $\widehat{\boldsymbol{\theta}}_n = \text{argmin}_{\boldsymbol{\theta}\in\Theta} L_n(\boldsymbol{\theta})$. Since we do not assume η_t follows the standard double exponential distribution, $\widehat{\boldsymbol{\theta}}_n$ is called the self-weighted exponential quasi-maximum likelihood estimator (EQMLE) of $\boldsymbol{\theta}_0$; see also Zhu and Ling (2011).

Remark 3.1. *When $w_t = 1$ for all t, the self-weighted estimator becomes the common EQMLE. Using the same techniques as in Zhu and Ling (2013), we can establish the asymptotic normality of the EQMLE for model (3) under a finite fractional moment of y_t when $p \le q$. However, when $p > q$ then $E(y_t^3) < \infty$ will be required to establish the asymptotic normality, which not only leads to a much* *narrower stationarity region, but also makes the asymptotic normality inapplicable for the data without finite third or higher order moments. Since the orders are usually unknown in advance, for generality, we adopt the self-weighting approach in Ling (2005) to ensure the asymptotic normality even for heavy-tailed data with only a finite fractional moment. Note that if $p \leq q$ is known beforehand, then we can use the unweighted EQMLE with $w_t = 1$ for all t.*

Assumption 3.1. *$\{y_t\}$ is strictly stationary and ergodic with $E(|y_t|^\kappa) < \infty$ for some $\kappa > 0$.*

Assumption 3.2. *$\boldsymbol{\theta}_0$ is an interior in Θ, and Θ is compact with $\underline{\omega} \leq \omega \leq \overline{\omega}$ and $\underline{\beta} \leq \beta_{j+}, \beta_{j-} \leq \overline{\beta}$ for $j = 1, \dots, q$, where $\underline{\omega}$, $\overline{\omega}$, $\underline{\beta}$ and $\overline{\beta}$ are some positive constants.*

Assumption 3.3. *The self-weights $\{\omega_t\}$ are strictly stationary and ergodic, and ω_t is positive, bounded and measurable with respect to $\mathcal{F}_{t-1}$ with $E(\omega_t \|\boldsymbol{Y}_t\|^2 + \omega_t^2 \|\boldsymbol{Y}_t\|^3) < \infty$.*

Assumption 3.4. *(i) η_t has zero median with $E(|\eta_t|) = 1$; (ii) The density function of η_t is continuous and positive everywhere on $\mathbb{R}$ satisfying $\sup_{x\in\mathbb{R}} f(x) < \infty$; (iii) $E(\eta_t^2) < \infty$.*

Assumption 3.1 is mild for time series models, and a sufficient condition for Assumption 3.1 is provided in Theorem 2.1. Assumption 3.2 is standard in literatures of quasi-maximum likelihood estimations; see also Assumption 1 of Zhu and Ling (2013) and Assumption 3 of Guo et al. (2019). The random weight ω_t that satisfies Assumption 3.3 is introduced to reduce the moment condition on y_t in establishing the asymptotic normality for the case with $p > q$, while ω_t can be set to one for the case with $p \leq q$ and then Assumption 3.3 is not needed accordingly. In practice, we can follow Ling (2005) to choose

$$w_t = I(a_t = 0) + C^2 a_t^{-2} I(a_t \neq 0), \tag{5}$$

where $a_t = \sum_{i=1}^p |y_{t-i}| I(|y_{t-i}| \geq C)$ for some constant $C > 0$, and C is usually chosen as the 90% or 95% empirical percentile of $\{|y_t|\}_{t=1}^n$. Assumption 3.4 is a general set-up to ensure the consistency and asymptotic normality for EQMLEs; see also Assumption 2.6 of Zhu and Ling (2011) and Assumption 3 of Zhu and Ling (2013).

Theorem 3.1. *If Assumptions 3.1–3.4(i) hold, then $\widehat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ almost surely as $n \to \infty$.*

Let $\kappa_1 = E(\eta_t)$ and $\kappa_2 = E(\eta_t^2) - 1$. Define the $d \times d$ matrices as follows

$$\Sigma = \operatorname{diag}\left\{ f(0) E\left[\frac{\omega_t \boldsymbol{Y}_t \boldsymbol{Y}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right], \frac{1}{2} E\left[\frac{\omega_t \boldsymbol{X}_t \boldsymbol{X}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right] \right\} \text{ and}$$

$$\Omega = \begin{pmatrix} E\left[\dfrac{\omega_t^2 \boldsymbol{Y}_t \boldsymbol{Y}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right] & \kappa_1 E\left[\dfrac{\omega_t^2 \boldsymbol{Y}_t \boldsymbol{X}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right] \\ \kappa_1 E\left[\dfrac{\omega_t^2 \boldsymbol{X}_t \boldsymbol{Y}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right] & \kappa_2 E\left[\dfrac{\omega_t^2 \boldsymbol{X}_t \boldsymbol{X}_t'}{(\boldsymbol{\beta}_0' \boldsymbol{X}_t)^2}\right] \end{pmatrix}.$$

Theorem 3.2. *If Assumptions 3.1–3.4 hold, then $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to_{\mathcal{L}} N(0, \Xi)$ as $n \to \infty$, where $\Xi = \Sigma^{-1}\Omega\Sigma^{-1}/4$.*

Theorem 3.2 shows that the asymptotic normality of the proposed self-weighted EQMLE is established for model (3) under a fractional moment of y_t with $E(\eta_t^2) < \infty$. To estimate the asymptotic covariance of $\widehat{\boldsymbol{\theta}}_n$, matrix Ω can be approximated by sample averages with $\boldsymbol{\theta}_0$ replaced by $\widehat{\boldsymbol{\theta}}_n$, while the density $f(0)$ in Σ usually need to be estimated by non-parametric methods such as the kernel density estimation which is sensitive to the choice of bandwidths. We alternatively employ the random-weighting bootstrap method (Zheng, 1987), and define the bootstraping self-weighted EQMLE below,

$$\widehat{\boldsymbol{\theta}}_n^\star = (\widehat{\boldsymbol{\alpha}}_n^{\star\prime}, \widehat{\boldsymbol{\beta}}_n^{\star\prime})' = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} \sum_{t=p+q}^n \pi_t \omega_t \ell_t(\boldsymbol{\theta}), \tag{6}$$

where $\{\pi_t\}$ are *i.i.d.* non-negative random weights, and are independent of $\{y_t\}$; see also Jin et al. (2001), Zheng et al. (2018) and Zhu et al. (2020). Denote $\mathcal{F}_t = \sigma(y_t, y_{t-1}, \dots)$ as the σ-field generated by $\{y_s, s \leq t\}$. To show the validity of the random-weighting bootstrap method, we make the following assumption on the random weights.

Assumption 3.5. *The random weights $\{\pi_t\}$ are i.i.d. non-negative random variables with $E(\pi_t) = 1$, $\operatorname{var}(\pi_t) = \tau^2$ and $E(|\pi_t|^{2+\kappa}) < \infty$ for some $\kappa > 0$.*

Theorem 3.3. *Suppose that Assumptions 3.1–3.5 hold. Then, conditional on $\mathcal{F}_n$, $(\sqrt{n}/\tau)(\widehat{\boldsymbol{\theta}}_n^\star - \widehat{\boldsymbol{\theta}}_n) \to_{\mathcal{L}} N(0, \Xi)$ in probability as $n \to \infty$, where Ξ is defined as in Theorem 3.2.*

Theorem 3.3 shows the theoretical validity of the random-weighting bootstrap for the self-weighted EQMLE. It follows that the resulting distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n^* - \widehat{\boldsymbol{\theta}}_n)/\tau$ can provide a reasonable approximation for that of $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$. As a result, a bootstrapping estimate of Ξ, denoted by $\widehat{\Xi}$, can be obtained accordingly.

Remark 3.2. *There are many sequences of random variables satisfying Assumption 3.5 and thus can be used to generate the bootstrap weights $\{\pi_t\}$. For example, the i.i.d. sequence of standard exponential random variables, the Rademacher distributed random variables which take the value 0 or 2 with probability 0.5, and the uniform random variables following $U(0.5, 1.5)$; see also Zheng et al. (2018) and Zhu et al. (2020). Simulation results in Section 5 indicate that these choices of random weights actually have very similar performance in finite samples. Therefore, the performance of the bootstrap procedure is insensitive to the choice of random weights $\{\pi_t\}$.*

Remark 3.3. *In practice, the order (p, q) of DA-LDAR model (3) is unknown, and information criterions can be used to select it. We may consider the following Bayesian* *information criterion (BIC):*

$$BIC(p,q) = 2\sum_{t=m_{\max}+1}^{n} \ell_t(\widetilde{\boldsymbol{\theta}}_n) + d\ln(n - m_{\max}), \tag{7}$$

where $d = 2p+2q+1$, $\widetilde{\boldsymbol{\theta}}_n$ is the unweighted EQMLE with the orders set to (p,q), and $m_{\max}$ is a predetermined positive integer. Let $(\widehat{p}_n, \widehat{q}_n) = \arg\min_{1\le p,q\le m_{\max}} BIC(p,q)$. We may begin with $(\widehat{p}_n, \widehat{q}_n)$ as an initial choice of (p,q) for model fitting, and further revise the order according to the portmanteau test in Section 4 if necessary.

Remark 3.4. *In real applications, we may be interested in detecting the asymmetric effect of news on the conditional location and volatility separately and jointly. Accordingly, the Wald tests can be constructed for the following hypotheses for model (3):*

$$\begin{aligned}
&H_{01}: \forall i,\ \alpha_{i0+} = \alpha_{i0-}\ \textit{vs}\ H_{11}: \exists i,\ \alpha_{i0+} \neq \alpha_{i0-};\\
&H_{02}: \forall j,\ \beta_{j0+} = \beta_{j0-}\ \textit{vs}\ H_{12}: \exists j,\ \beta_{j0+} \neq \beta_{j0-};\\
&H_{03}: \forall i,j,\ \alpha_{i0+} = \alpha_{i0-}\ \textit{and}\ \beta_{j0+} = \beta_{j0-}\ \textit{vs}\\
&H_{13}: \exists i,j,\ \alpha_{i0+} \neq \alpha_{i0-}\ \textit{or}\ \beta_{j0+} \neq \beta_{j0-},
\end{aligned}$$

where vs is the abbreviation of versus, $i \in \{1,\ldots,p\}$ and $j \in \{1,\ldots,q\}$. Denote $R_1 = [I_p, -I_p, 0_{p\times(2q+1)}]$, $R_2 = [0_{q\times(2p+1)}, I_q, -I_q]$ and $R_3 = (R_1', R_2')'$, respectively, where $0_{m\times n}$ is the $m\times n$ zero matrix and I_m is the $m\times m$ identity matrix. Then the null hypotheses of aforementioned three tests can be represented as $H_{01}: R_1\boldsymbol{\theta}_0 = \mathbf{0}_p$, $H_{02}: R_2\boldsymbol{\theta}_0 = \mathbf{0}_q$ and $H_{03}: R_3\boldsymbol{\theta}_0 = \mathbf{0}_{p+q}$, where $\mathbf{0}_m$ is the m-dimentional zero vector. Hence, the Wald test statistics can be defined as

$$W_{in} = n\widehat{\boldsymbol{\theta}}_n' R_i'(R_i\widehat{\Xi}R_i')^{-1}R_i\widehat{\boldsymbol{\theta}}_n \text{ for } i = 1,2,3, \tag{8}$$

where $\widehat{\Xi}$ is the bootstrapping estimate of Ξ. Then under the conditions of Theorem 3.2, as $n\to\infty$, we can show that $W_{1n} \to_{\mathcal{L}} \chi^2_p$ under H_{01}, $W_{2n} \to_{\mathcal{L}} \chi^2_q$ under H_{02} and $W_{3n} \to_{\mathcal{L}} \chi^2_{p+q}$ under H_{03}, where χ^2_m is the chi-squared distribution with m degrees of freedom. As a result, if W_{in} exceeds the $(1-\tau)$th quantile of $\chi^2_{r_i}$ distribution with $r_1 = p, r_2 = q$ and $r_3 = p+q$, we can reject the null hypothesis.

4. DIAGNOSTIC CHECKING

To check whether model (3) is correctly specified, we construct a mixed portmanteau test to detect misspecifications in the conditional location and standard deviation jointly.

In accordance with the self-weighted estimation procedure in Section 3, define the self-weighted innovation $\zeta_t = w_t(\eta_t - \kappa_1)$ and the self-weighted absolute innovation $\xi_t = w_t(|\eta_t| - 1)$; see also Jiang et al. (2020). The ACFs of $\{\zeta_t\}$ and $\{\xi_t\}$ at lag k can be defined by $\rho_k = \operatorname{cov}(\zeta_t, \zeta_{t-k})/\operatorname{var}(\zeta_t)$ and $\gamma_k = \operatorname{cov}(\xi_t, \xi_{t-k})/\operatorname{var}(\xi_t)$, respectively. If the data generating process is correctly specified by model (3), then $\{\zeta_t\}$ and $\{\xi_t\}$ are uncorrelated such that $\rho_k = 0$ and $\gamma_k = 0$ hold for any $k \geq 1$. Accordingly, define the self-weighted residual $\widehat{\zeta}_t = \omega_t(\widehat{\eta}_t - \widehat{\kappa}_1)$ and the self-weighted absolute residual $\widehat{\xi}_t = \omega_t(|\widehat{\eta}_t| - 1)$, where $\widehat{\eta}_t = (y_t - \widehat{\boldsymbol{\alpha}}_n'\boldsymbol{Y}_t)/(\widehat{\boldsymbol{\beta}}_n'\boldsymbol{X}_t)$ is the residual of model (3) fitted by the self-weighted EQMLE and $\widehat{\kappa}_1$ is the sample version of κ_1, i.e. $\widehat{\kappa}_1 = (n-p-q+1)^{-1}\sum_{t=p+q}^{n}\widehat{\eta}_t$. Then the sample ACF of $\{\widehat{\zeta}_t\}$ and $\{\widehat{\xi}_t\}$ at lag k can be calculated as

$$\widehat{\rho}_k = \frac{\sum_{t=p+q+k}^{n}(\widehat{\zeta}_t - \bar{\zeta})(\widehat{\zeta}_{t-k} - \bar{\zeta})}{\sum_{t=p+q}^{n}(\widehat{\zeta}_t - \bar{\zeta})^2} \text{ and}$$

$$\widehat{\gamma}_k = \frac{\sum_{t=p+q+k}^{n}(\widehat{\xi}_t - \bar{\xi})(\widehat{\xi}_{t-k} - \bar{\xi})}{\sum_{t=p+q}^{n}(\widehat{\xi}_t - \bar{\xi})^2}$$

respectively, where $\bar{\zeta} = (n-p-q+1)^{-1}\sum_{t=p+q}^{n}\widehat{\zeta}_t$ and $\bar{\xi} = (n-p-q+1)^{-1}\sum_{t=p+q}^{n}\widehat{\xi}_t$. Note that $\widehat{\rho}_k$ (or $\widehat{\gamma}_k$) is the sample version of ρ_k (or γ_k). If the value of $\widehat{\rho}_k$ (or $\widehat{\gamma}_k$) deviates from zero significantly, it indicates possible misspecification in the conditional location (or standard deviation) of model (3).

Let $\widehat{\boldsymbol{\rho}} = (\widehat{\rho}_1,\ldots,\widehat{\rho}_M)'$, $\widehat{\boldsymbol{\gamma}} = (\widehat{\gamma}_1,\ldots,\widehat{\gamma}_M)'$ and $\widehat{\boldsymbol{\psi}} = (\widehat{\boldsymbol{\rho}}', \widehat{\boldsymbol{\gamma}}')'$, where M is a predetermined positive integer. Denote $\sigma_1^2 = \operatorname{var}(\zeta_t) = E(\omega_t^2)\left[E(\eta_t^2) - \kappa_1^2\right]$ and $\sigma_2^2 = \operatorname{var}(\xi_t) = E(\omega_t^2)\left[E(\eta_t^2) - 1\right]$. Define the $M\times d$ matrices $U_\rho = (\boldsymbol{U}_{\rho_1}', \ldots, \boldsymbol{U}_{\rho_M}')'$ and $U_\gamma = (\boldsymbol{U}_{\gamma_1}', \ldots, \boldsymbol{U}_{\gamma_M}')'$, where for $1 \le k \le M$,

$$\begin{aligned}
\boldsymbol{U}_{\rho_k} = -\Bigg(&E\left(\frac{\omega_t\omega_{t-k}(\eta_{t-k} - \kappa_1)\boldsymbol{Y}_t'}{\boldsymbol{\beta}_0'\boldsymbol{X}_t}\right),\\
&\kappa_1 E\left(\frac{\omega_t\omega_{t-k}(\eta_{t-k} - \kappa_1)\boldsymbol{X}_t'}{\boldsymbol{\beta}_0'\boldsymbol{X}_t}\right)\Bigg) \text{ and}
\end{aligned}$$

$$\boldsymbol{U}_{\gamma_k} = -\left(\mathbf{0}_{2p}', E\left(\frac{\omega_t\omega_{t-k}(|\eta_{t-k}| - 1)\boldsymbol{X}_t'}{\boldsymbol{\beta}_0'\boldsymbol{X}_t}\right)\right).$$

Denote the $2M \times (2M + d)$ matrix

$$V = \begin{pmatrix} I_M & 0_{M\times M} & U_\rho/\sigma_1^2 \\ 0_{M\times M} & I_M & U_\gamma/\sigma_2^2 \end{pmatrix}.$$

Let $\boldsymbol{G}_t = (\boldsymbol{Y}_t'[I(\eta_t < 0) - I(\eta_t > 0)], \boldsymbol{X}_t'(1 - |\eta_t|))'/(\boldsymbol{\beta}_0'\boldsymbol{X}_t)$ and $G = E(\boldsymbol{\nu}_t\boldsymbol{\nu}_t')$, where

$$\begin{aligned}
\boldsymbol{\nu}_t = (&\zeta_t\zeta_{t-1}/\sigma_1^2, \ldots, \zeta_t\zeta_{t-M}/\sigma_1^2, \xi_t\xi_{t-1}/\sigma_2^2, \ldots, \xi_t\xi_{t-M}/\sigma_2^2,\\
&-\boldsymbol{G}_t'\Sigma^{-1}/2)'.
\end{aligned}$$

Theorem 4.1. *Suppose the conditions of Theorem 3.2 hold. If model (3) is correctly specified, then $\sqrt{n}\widehat{\boldsymbol{\psi}} \to_{\mathcal{L}} N(0,\Gamma)$ as $n\to\infty$, where $\Gamma = VGV'$.*

Theorem 4.1 shows the limiting distribution of sample ACFs. To check the significance of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ individually or jointly, we need to estimate the asymptotic covariance Γ. Note that Γ depends on the unknown quantity $f(0)$, we also use the random-weighting bootstrap method to estimate Γ

as in Section 3. Specifically, define $\widehat{\zeta}_t^\star = \omega_t(\widehat{\eta}_t^\star - \widehat{\kappa}_1^\star)$ and $\widehat{\xi}_t^\star = \omega_t(|\widehat{\eta}_t^\star| - 1)$, where $\widehat{\eta}_t^\star = (y_t - \widehat{\boldsymbol{\alpha}}_n^{\star\prime}\boldsymbol{Y}_t)/(\widehat{\boldsymbol{\beta}}_n^{\star\prime}\boldsymbol{X}_t)$ with $\widehat{\boldsymbol{\alpha}}_n^\star$ and $\widehat{\boldsymbol{\beta}}_n^\star$ being the bootstrap estimator obtained by (6), and $\kappa_1^\star = (n-p-q+1)^{-1}\sum_{t=p+q}^n \eta_t^\star$. Then define the bootstrapping sample ACFs of $\widehat{\zeta}_t$ and $\widehat{\xi}_t$ at lag k below

$$
\begin{aligned}
\widehat{\rho}_k^\star &= \frac{\sum_{t=p+q+k}^n \pi_t(\widehat{\zeta}_t^\star - \bar{\zeta}^\star)(\widehat{\zeta}_{t-k}^\star - \bar{\zeta}^\star)}{\sum_{t=p+q}^n(\widehat{\zeta}_t^\star - \bar{\zeta}^\star)^2} \text{ and} \\
\widehat{\gamma}_k^\star &= \frac{\sum_{t=p+q+k}^n \pi_t(\widehat{\xi}_t^\star - \bar{\xi}^\star)(\widehat{\xi}_{t-k}^\star - \bar{\xi}^\star)}{\sum_{t=p+q}^n(\widehat{\xi}_t^\star - \bar{\xi}^\star)^2},
\end{aligned} \tag{9}
$$

where $\{\pi_t\}$ are defined as in (6), $\bar{\zeta}^\star = (n-p-q+1)^{-1}\sum_{t=p+q}^n \widehat{\zeta}_t^\star$ and $\bar{\xi}^\star = (n-p-q+1)^{-1}\sum_{t=p+q}^n \widehat{\xi}_t^\star$. Let $\widehat{\boldsymbol{\rho}}^\star = (\widehat{\rho}_1^\star, \ldots, \widehat{\rho}_M^\star)'$, $\widehat{\boldsymbol{\gamma}}^\star = (\widehat{\gamma}_1^\star, \ldots, \widehat{\gamma}_M^\star)'$ and $\widehat{\boldsymbol{\psi}}^\star = (\widehat{\boldsymbol{\rho}}^{\star\prime}, \widehat{\boldsymbol{\gamma}}^{\star\prime})'$.

Theorem 4.2. *Suppose the conditions of Theorem 3.3 hold, and the model (3) is correctly specified. Then, conditional on $\mathcal{F}_n$, $(\sqrt{n}/\tau)(\widehat{\boldsymbol{\psi}}^\star - \widehat{\boldsymbol{\psi}}) \to_{\mathcal{L}} N(0, \Gamma)$ in probability as $n \to \infty$, where Γ is defined as in Theorem 4.1.*

Theorem 4.2 ensures that the bootstrapped covariance matrix $\widehat{\Gamma}$ of $(\sqrt{n}/\tau)(\widehat{\boldsymbol{\psi}}^\star - \widehat{\boldsymbol{\psi}})$ can be used to approximate Γ. Thus we can check the significance of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ individually by constructing confidence intervals based on $\widehat{\Gamma}$. Moreover, we can construct a portmanteau test statistic to check the first M lags jointly:

$$
Q(M) = n\widehat{\boldsymbol{\psi}}'\widehat{\Gamma}^{-1}\widehat{\boldsymbol{\psi}}. \tag{10}
$$

By Theorems 4.1–4.2 and the continuous mapping theorem, we have $Q(M) \to_{\mathcal{L}} \chi^2_{2M}$ as $n \to \infty$. Therefore, if $Q(M)$ exceeds the $(1-\tau)$th quantile of χ^2_{2M} distribution, we can reject the null hypothesis that ρ_k's and γ_k's $(1 \le k \le M)$ are jointly insignificant at level τ. In summary, the random-weighting bootstrapping procedure for the estimation and portmanteau test can be summarized below:

1. Generate random weights $\{\pi_t\}$ from a non-negative distribution satisfying Assumption 3.5. Then obtain the bootstraping self-weighted EQMLE $\widehat{\boldsymbol{\theta}}_n^\star$ by (6);
2. Compute the bootstrapping sample ACFs $\widehat{\rho}_k^\star$ and $\widehat{\gamma}_k^\star$ by (9) and then $\widehat{\boldsymbol{\psi}}^\star$;
3. Calculate $E^{(1)} = (\sqrt{n}/\tau)(\widehat{\boldsymbol{\theta}}_n^\star - \widehat{\boldsymbol{\theta}}_n)$ and $T^{(1)} = (\sqrt{n}/\tau)(\widehat{\boldsymbol{\psi}}^\star - \widehat{\boldsymbol{\psi}})$. Repeat Steps 1 and 2 for $B-1$ times and obtain $\{E^{(1)}, \ldots, E^{(B)}\}$ and $\{T^{(1)}, \ldots, T^{(B)}\}$, where B is a sufficiently large number.

Then the empirical distributions of $\{E^{(b)}\}_{b=1}^B$ and $\{T^{(b)}\}_{b=1}^B$ can be used to approximate the asymptotic distributions of $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ and $\sqrt{n}\widehat{\boldsymbol{\psi}}$ respectively. As a result, the asymptotic covariance Ξ and test statistic $Q(M)$ at (10) can be calculated accordingly.

5. SIMULATION

This section presents three simulation experiments to assess the finite-sample performance of the proposed estimation, portmanteau test and their bootstrapping procedure.

The first experiment aims to examine the finite-sample performance of the self-weighted EQMLE $\widehat{\boldsymbol{\theta}}_n$. We consider the following data generating procedures (DGPs):

$$
\begin{aligned}
\text{DGP1: } y_t &= 0.2y_{t-1}^+ + 0.05y_{t-1}^- + \eta_t\sigma_t, \\
\sigma_t &= 0.5 + 0.2y_{t-1}^+ - 0.3y_{t-1}^-; \\
\text{DGP2: } y_t &= 0.2y_{t-1}^+ + 0.1y_{t-1}^- + 0.1y_{t-2}^- + \eta_t\sigma_t, \\
\sigma_t &= 0.5 + 0.3y_{t-1}^+ - 0.4y_{t-1}^-; \\
\text{DGP3: } y_t &= 0.2y_{t-1}^+ + 0.3y_{t-1}^- + \eta_t\sigma_t, \\
\sigma_t &= 0.1 + 0.3y_{t-1}^+ - 0.4y_{t-1}^- + 0.3y_{t-2}^+ - 0.4y_{t-2}^-,
\end{aligned}
$$

where $\{\eta_t\}$ are *i.i.d.* random variables following the standard Laplace(0,1), standardized Student's t_3 or standardized skewed t_3 with skew parameter -1, denoted by $st_{3,-1}$ (Fernández and Steel, 1996), and they are standarded with zero median and $E(|\eta_t|) = 1$. The sample size is $n = 500$ or 1000, with 1000 replications for each sample size. We consider two cases for the self-weights in estimation: (I) $\omega_t = 1$ for all t and (II) ω_t at (5) with C chosen as the 95% empirical percentile of $\{|y_t|\}_{t=1}^n$. Note that both self-weights (I) and (II) are applicable for DGP1 and DGP3 since they correspond to the order $p \le q$, while self-weights (II) are suggested for DGP2 due to $p > q$; see Remark 3.1 for details. The bootstrapping procedure is conducted with three types of random weights $\{\pi_t\}$: (i) *i.i.d.* Exponential(1) weights (Π_1); (ii) *i.i.d.* Rademacher weights (Π_2); (iii) *i.i.d.* U(0.5,1.5) weights (Π_3). Note that the variance $\tau^2 = 1$ for weights Π_1 and Π_2 while $\tau^2 = 1/12$ for Π_3. In addition, $B = 500$ bootstrapped samples are used to calculate the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}_n$. For comparison, the Gaussian kernel function with its rule-of-thumb bandwidth $b_n = 0.9n^{-1/5}\min\{s, \widehat{R}/1.34\}$ is employed to estimate $f(0)$ in the asymptotic covariance matrix, where s and $\widehat{R}$ are the sample standard deviation and interquartile of the residuals $\{\widehat{\eta}_t\}$, respectively.

Table 4 lists the biases, empirical standard deviations (ESDs) and asymptotic standard deviations (ASDs) of $\widehat{\boldsymbol{\theta}}_n$ for different self-weights $\{\omega_t\}$ and random weights $\{\pi_t\}$ when η_t follows the standard Laplace$(0,1)$ distribution and the data generating process is DGP1. We have the following findings: (1) as the sample size increases, the biases, ESDs and ASDs decrease, and ESDs and ASDs get closer to each other; (2) ESDs and ASDs corresponding to $\omega_t = 1$ are slightly smaller than that of ω_t at (5), which suggests that introducing self-weights in estimation could result in efficiency loss and thus unweighted EQMLE is preferred if $p \le q$ is known; (3) the ASDs calculated by three sets of random weights are quite similar when the sample size is as small as 500, which indicates that the bootstrap method is insensitive to the choice of random weights; (4) the ASDs calculated by the bootstrap method are closer to ESDs than that of the kernel-based method for the conditional location, and the distinction weakens for the conditional standard deviation. This implies that the bootstrap method provides more accurate approximation than the kernel-based method in finite samples. These findings still hold for other distributions of η_t. Moreover, the biases, ESDs and ASDs of DGP2 with η_t following Laplace(0,1) are also reported in Table 5 for comparison. Note that self-weights (I) are applicable for DGP2 with η_t following Laplace(0,1), since $E(y_t^3) < \infty$ can be verified for this case by Theorem 1. The simulation findings are unchanged for Table 5 in comparison with that of Table 4.

Based on aforementioned findings (3) and (4), we prefer to use the bootstrap method to calculate ASDs and only report the ASDs using exponential random weights for DGP2 and DGP3. Table 6 reports the biases, ESDs and ASDs of $\widehat{\boldsymbol{\theta}}_n$ for DGP2 and DGP3 with different innovation distributions, where the self-weights (II) are used for estimation. It can be seen that the biases, ESDs and ASDs become smaller and the bootstrapping approximation gets better for all DGPs as the sample size increases. In view of the ESDs and ASDs for different innovation distributions, the heavy-tailedness of innovations can worsen the estimation efficiency, but skewness cannot.

In the second experiment, we evaluate the performance of the bootstrap approximation for the residual ACF $\widehat{\rho}_k$ and the absolute residual ACF $\widehat{\gamma}_k$ in Section 4. The DGPs and other settings are preserved as in the first experiment. The biases, ESDs and ASDs of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ at lags $2, 4$ and 6 are reported in Table 7 for different self-weights and random weights, while that of different innovation distributions and DGPs are presented in Table 8. We have the following findings: (a) the biases are close to zero, and the sample size increases, ESDs and ASDs become smaller and the bootstrapping approximation gets better; (b) different random weights lead to similar results in approximating the asymptotic covariance matrix of $\widehat{\boldsymbol{\psi}}$; (c) ESDs and ASDs of different innovation distributions are very similar, which demonstrates the robustness of the proposed method.

The third experiment is to evaluate the mixed portmanteau test $Q(M)$ at (10) with the bootstrapping procedure. The data are generated from the following two DGPs:

$$\begin{aligned}
\text{DGP4: } y_t =& 0.2y_{t-1}^+ + 0.1y_{t-1}^- + 0.1y_{t-2}^- + d_1 y_{t-3} + \eta_t\sigma_t,\\
\sigma_t =& 0.5 + 0.3y_{t-1}^+ - 0.4y_{t-1}^- + d_2|y_{t-2}|;\\
\text{DGP5: } y_t =& 0.2y_{t-1}^+ + 0.3y_{t-1}^- + d_1 y_{t-2} + \eta_t\sigma_t,\\
\sigma_t =& 0.1 + 0.3y_{t-1}^+ - 0.4y_{t-1}^- + 0.3y_{t-2}^+ - 0.4y_{t-2}^-\\
&+ d_2|y_{t-3}|,
\end{aligned}$$

where the innovations $\{\eta_t\}$ are defined as in the first experiment and the departure $d_1, d_2 \in \{0, 0.2, 0.4\}$. We fit a DA-LDAR model with $(p, q) = (2, 1)$ for DGP4 and $(p, q) = (1, 2)$ for DGP5, so the case of $d_1 = d_2 = 0$ corresponds to the size of the test, the case of $d_1 \neq 0$ corresponds to misspecifications in the conditional location, and the case of $d_2 \neq 0$ corresponds to misspecifications in the conditional standard deviation. Table 9 reports the rejection rates of $Q(6)$ at 5% significance level based on 1000 replications, for sample size $n = 500$ or 1000. The bootstrapping portmanteau test performs well in terms of size and power: the size are closer to the nominal level 5%, and the power gets larger as either the sample size n or the departure d_1 or d_2 increases.

6. EMPIRICAL ANALYSIS

This section analyzes the weekly closing prices of NASDAQ Composite Index, denoted as p_t, from February 1971 to June 2020. The dataset is downloaded from the website of Yahoo Finance (https://hk.finance.yahoo.com), with 2579 observations in total. Let $r_t = 100(\ln p_t - \ln p_{t-1})$ be the log return in percentage, and denote y_t as the centered series. The time plot of $\{y_t\}$ is shown in Figure 2, and some summary statistics are listed in Table 1. Figure 2 illustrates volatility clustering of $\{y_t\}$, and the sample skewness -1.08 and kurtosis 12.95 indicate that the series $\{y_t\}$ are skewed and heavy-tailed. Moreover, as shown by Figure 3, the ACFs and PACFs of $\{y_t\}$ are significant at the first few lags, this together with the slowly decaying ACF of $\{|y_t|\}$ suggests that both the autocorrelation and conditional heteroscedasticity appear in $\{y_t\}$. The above findings motivate us to investigate $\{y_t\}$ by our proposed model and inference tools.

In view of the ACF and PACF plots of $\{y_t\}$ and $\{|y_t|\}$ in Figure 3, we use the BIC at (7) to select an initial order for (p, q) searched over $1 \leq p, q \leq m_{\max} = 15$. The resulting initial order is $(\widehat{p}_n, \widehat{q}_n) = (3, 6)$. Since $\widehat{p}_n < \widehat{q}_n$, we consider to fit a DA-LDAR(3, 6) model for $\{y_t\}$ using the unweighted

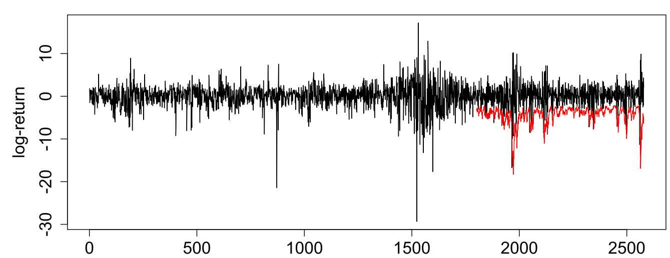

Figure 2. Time plot for weekly log returns in percentage (black line) of NASDAQ Composite Index from February 1971 to June 2020, with one-week negative VaR forecasts at the level of 5% (red line) from August 2005 to June 2020.

Table 1. Summary statistics for NASDAQ returns

Mean	Median	Std.Dev.	Skewness	Kurtosis	Min	Max
0.00	0.16	2.75	-1.08	12.53	-29.35	17.19

Table 2. The fitted coefficients of the DA-LDAR$(3,6)$ model (upper panel) and DA-LDAR$(3,8)$ model (lower panel), where the subscripts are the standard errors of the corresponding estimates

lag	1	2	3	4	5	6	7	8
DA-LDAR$(3,6)$ $\omega:$ $0.78_{0.05}$								
α_{i+}	$0.15_{0.04}$	$-0.01_{0.04}$	$0.15_{0.03}$					
α_{i-}	$0.04_{0.04}$	$0.12_{0.05}$	$-0.02_{0.03}$					
β_{j+}	$0.04_{0.02}$	$0.03_{0.02}$	$0.05_{0.02}$	$0.09_{0.02}$	$0.04_{0.02}$	$0.11_{0.02}$		
β_{j-}	$0.19_{0.03}$	$0.17_{0.02}$	$0.14_{0.03}$	$0.14_{0.02}$	$0.07_{0.02}$	$0.05_{0.02}$		
DA-LDAR$(3,8)$ $\omega:$ $0.71_{0.05}$								
α_{i+}	$0.15_{0.04}$	$-0.01_{0.04}$	$0.15_{0.03}$					
α_{i-}	$0.04_{0.03}$	$0.12_{0.05}$	$-0.01_{0.03}$					
β_{j+}	$0.02_{0.02}$	$0.01_{0.01}$	$0.04_{0.02}$	$0.07_{0.02}$	$0.02_{0.02}$	$0.09_{0.02}$	$0.02_{0.02}$	$0.09_{0.02}$
β_{j-}	$0.18_{0.03}$	$0.16_{0.02}$	$0.13_{0.02}$	$0.13_{0.02}$	$0.06_{0.02}$	$0.03_{0.02}$	$0.06_{0.02}$	$0.06_{0.02}$

Table 3. Empirical coverage rate (%) and p-values of two VaR backtests of three models at the 5%, 10%, 90% and 95% conditional quantiles. M1, M2 and M3 represent the DA-LDAR, LDAR and TDAR models, respectively

	$\tau = 5\%$			$\tau = 10\%$			$\tau = 90\%$			$\tau = 95\%$		
	ECR	CC	DQ	ECR	CC	DQ	ECR	CC	DQ	ECR	CC	DQ
M1	4.87	0.98	0.27	10.52	0.86	0.15	88.60	0.06	0.25	95.63	0.64	0.95
M2	6.54	0.02	0.02	11.68	0.27	0.39	86.64	0.01	0.01	93.32	0.01	0.01
M3	2.95	0.01	0.01	5.51	0.01	0.01	94.35	0.01	0.01	96.91	0.03	0.01

Table 4. Biases $(\times 10)$, ESDs, ASD$_k$ and ASD$_i$ $(i = 1, 2, 3)$ of $\widehat{\boldsymbol{\theta}}_n$ for DGP1 when the innovations follow the Laplace$(0, 1)$ with different self-weights ω_t and random weights π_t, where ASD$_k$ and ASD$_i$'s are the ASDs calculated by the kernel method and bootstrap method with random weights $\{\Pi_i\}$

	n	Bias	ESD	ASD$_k$	ASD$_1$	ASD$_2$	ASD$_3$
$\omega_t = 1$							
α_+	500	-0.014	0.056	0.072	0.060	0.060	0.063
	1000	-0.003	0.038	0.048	0.040	0.039	0.041
α_-	500	-0.016	0.063	0.090	0.069	0.068	0.074
	1000	-0.001	0.044	0.062	0.048	0.048	0.053
β_0	500	0.160	0.034	0.038	0.036	0.037	0.040
	1000	0.021	0.023	0.026	0.025	0.025	0.027
β_+	500	0.044	0.061	0.064	0.060	0.062	0.065
	1000	-0.003	0.042	0.045	0.042	0.043	0.046
β_-	500	-0.111	0.068	0.077	0.070	0.073	0.079
	1000	-0.025	0.047	0.055	0.050	0.051	0.056
ω_t at (5)							
α_+	500	-0.025	0.058	0.074	0.061	0.061	0.064
	1000	-0.001	0.039	0.050	0.041	0.041	0.043
α_-	500	-0.015	0.063	0.093	0.071	0.070	0.077
	1000	-0.010	0.045	0.064	0.049	0.049	0.054
β_0	500	0.139	0.033	0.039	0.036	0.037	0.039
	1000	-0.011	0.023	0.027	0.025	0.025	0.027
β_+	500	0.059	0.062	0.067	0.062	0.063	0.068
	1000	0.007	0.044	0.047	0.044	0.045	0.048
β_-	500	-0.080	0.069	0.080	0.072	0.074	0.081
	1000	0.012	0.048	0.057	0.051	0.052	0.058

EQMLE in Section 3 with the self-weights $\omega_t = 1$ for all t. The fitted coefficients are reported in Table 2. To check the goodness-of-fit of the fitted DA-LDAR$(3,6)$ model, the mixed portmanteau tests $Q(M)$ in Section 4 are employed for $M = 6, 12$ and 18, and their p-values are $0.34, 0.10$ and 0.01, respectively. This together with the residual ACF plots of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ in Figure 4, indicates that the lack of fit is due to the conditional volatility component. Therefore, we fur-

Table 5. Biases ($\times 10$), ESDs, ASD$_k$ and ASD$_i$ ($i = 1,2,3$) of $\widehat{\boldsymbol{\theta}}_n$ for DGP2 when the innovations follow the Laplace$(0,1)$ with different self-weights ω_t and random weights π_t, where ASD$_k$ and ASD$_i$'s are the ASDs calculated by the kernel method and bootstrap method with random weights $\{\Pi_i\}$

	n	Bias	ESD	ASD$_k$	ASD$_1$	ASD$_2$	ASD$_3$
$\omega_t = 1$							
α_{1+}	500	-0.016	0.058	0.090	0.065	0.063	0.078
	1000	0.009	0.044	0.062	0.045	0.044	0.052
α_{2+}	500	-0.009	0.049	0.063	0.052	0.052	0.058
	1000	-0.020	0.032	0.043	0.035	0.035	0.039
α_{1-}	500	-0.019	0.059	0.102	0.068	0.066	0.083
	1000	0.000	0.044	0.070	0.047	0.046	0.057
α_{2-}	500	-0.022	0.049	0.066	0.054	0.054	0.061
	1000	-0.002	0.033	0.045	0.036	0.036	0.040
β_0	500	0.014	0.034	0.040	0.037	0.037	0.043
	1000	-0.001	0.024	0.029	0.026	0.026	0.029
β_+	500	-0.058	0.060	0.070	0.062	0.063	0.075
	1000	-0.026	0.043	0.050	0.044	0.044	0.051
β_-	500	-0.065	0.064	0.079	0.067	0.068	0.086
	1000	-0.014	0.044	0.057	0.047	0.047	0.057
ω_t at (5)							
α_{1+}	500	-0.010	0.066	0.098	0.071	0.070	0.083
	1000	0.012	0.047	0.067	0.049	0.048	0.055
α_{2+}	500	-0.042	0.053	0.071	0.057	0.057	0.065
	1000	-0.023	0.036	0.048	0.039	0.039	0.042
α_{1-}	500	-0.007	0.069	0.108	0.074	0.073	0.090
	1000	-0.003	0.049	0.074	0.052	0.051	0.060
α_{2-}	500	-0.005	0.053	0.074	0.059	0.058	0.066
	1000	-0.007	0.037	0.051	0.040	0.040	0.044
β_0	500	-0.002	0.036	0.041	0.038	0.038	0.044
	1000	-0.006	0.025	0.030	0.027	0.027	0.030
β_+	500	-0.026	0.068	0.077	0.068	0.069	0.081
	1000	-0.013	0.048	0.055	0.048	0.048	0.055
β_-	500	-0.047	0.070	0.085	0.073	0.074	0.091
	1000	-0.001	0.049	0.061	0.051	0.052	0.060

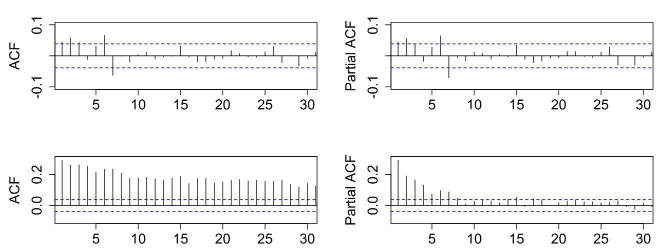

Figure 3. The ACF and PACF plots of $\{y_t\}$ (upper panel) and $\{|y_t|\}$ (lower panel), where the blue dashed lines are the corresponding 95% confidence bounds.

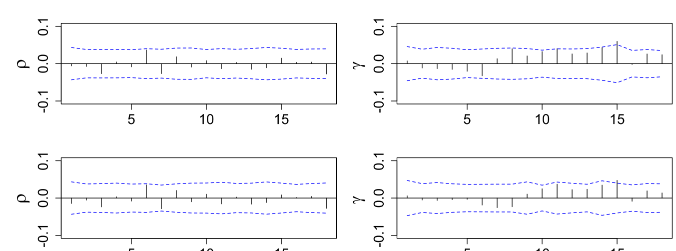

Figure 4. Residual ACF plots of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ for the fitted DA-LDAR$(3,6)$ model (upper panel) and DA-LDAR$(3,8)$ model (lower panel), where the blue dashed lines are the corresponding 95% confidence bounds.

ther revise the order by increasing the value of q. Finally, we fit a DA-LDAR$(3,8)$ model for $\{y_t\}$ owing to its satisfactory goodness-of-fit with p-values of $Q(M)$ being $0.77, 0.62$ and 0.41 for $M = 6, 12$ and 18, respectively. Moreover, the parameter estimates are summarized in Table 2, and the residual ACF plots of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ are displayed in Figure 4. It can be seen that α_{i+} and α_{i-} are clearly different for all i, while β_{j+} and β_{j-} are distinct for the first four lags, which suggests the asymmetric effects in the conditional location and volatility of y_t. The Wald tests in Remark 3.4 are conducted for the fitted model and all their p-values are less than 0.001, which corroborates the asymmetric effects

Table 6. Biases (×10), ESDs, ASDs of $\widehat{\boldsymbol{\theta}}_n$ for DGP2 and DGP3 when the innovations follow the Laplace$(0,1)$, standardized t_3 or standardized $st_{3,-1}$, where ASDs are calculated by bootstrap method with $i.i.d.$ exponential random weights $\{\Pi_1\}$

		Laplace$(0,1)$			t_3			$st_{3,-1}$		
	n	Bias	ESD	ASD	Bias	ESD	ASD	Bias	ESD	ASD
DGP2										
α_{1+}	500	0.004	0.065	0.069	-0.071	0.074	0.079	0.022	0.071	0.078
	1000	0.001	0.043	0.047	-0.017	0.055	0.056	0.006	0.052	0.056
α_{2+}	500	-0.023	0.051	0.055	-0.015	0.058	0.062	-0.038	0.060	0.062
	1000	-0.007	0.034	0.037	-0.019	0.041	0.043	-0.029	0.040	0.043
α_{1-}	500	0.028	0.062	0.072	0.037	0.084	0.084	0.007	0.081	0.084
	1000	0.003	0.045	0.050	0.001	0.055	0.059	0.001	0.056	0.059
α_{2-}	500	-0.047	0.051	0.056	-0.056	0.060	0.065	-0.036	0.062	0.065
	1000	0.029	0.035	0.038	-0.005	0.042	0.045	-0.012	0.042	0.045
β_0	500	0.004	0.035	0.037	-0.018	0.042	0.042	0.010	0.044	0.043
	1000	0.001	0.025	0.026	-0.010	0.032	0.031	-0.001	0.031	0.031
β_+	500	-0.034	0.063	0.066	-0.007	0.079	0.077	-0.048	0.079	0.078
	1000	-0.031	0.045	0.046	0.001	0.062	0.057	-0.040	0.055	0.056
β_-	500	-0.062	0.068	0.070	-0.057	0.083	0.083	-0.063	0.088	0.084
	1000	-0.033	0.046	0.050	-0.046	0.062	0.060	-0.040	0.062	0.060
DGP3										
α_+	500	0.022	0.040	0.054	-0.013	0.037	0.055	0.001	0.034	0.055
	1000	-0.005	0.019	0.027	0.002	0.020	0.029	0.001	0.019	0.029
α_-	500	0.002	0.036	0.050	0.035	0.034	0.053	0.033	0.037	0.055
	1000	0.001	0.017	0.025	0.008	0.020	0.027	-0.003	0.017	0.027
β_0	500	0.007	0.008	0.011	0.003	0.010	0.012	0.006	0.010	0.013
	1000	0.004	0.006	0.006	0.002	0.007	0.008	0.004	0.007	0.008
β_{1+}	500	0.033	0.035	0.048	0.053	0.037	0.053	0.044	0.035	0.055
	1000	0.021	0.020	0.024	0.039	0.025	0.031	0.019	0.019	0.028
β_{2+}	500	0.032	0.034	0.049	0.047	0.041	0.054	0.049	0.039	0.056
	1000	0.025	0.019	0.026	0.039	0.027	0.031	0.025	0.021	0.031
β_{1-}	500	0.017	0.027	0.046	0.010	0.035	0.052	0.015	0.033	0.057
	1000	0.002	0.013	0.022	0.001	0.020	0.028	0.013	0.019	0.027
β_{2-}	500	0.027	0.028	0.046	0.023	0.033	0.052	0.034	0.041	0.053
	1000	0.005	0.017	0.023	0.004	0.021	0.030	0.009	0.021	0.029

in both parts. In addition, the residual ACFs $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ almost fall within their corresponding 95% confidence bounds up to lag 18, which further indicates that the fitted DA-LDAR$(3,8)$ model is adequate.

We next evaluate the performance of the fitted model in forecasting the Value-at-Risk (VaR) using a rolling forecasting procedure with a fixed moving window. VaR is a commonly used risk measure for financial assets, and its negative value is actually the conditional quantile of return series $\{y_t\}$. Hence, we use the fitted model to forecast the conditional quantile of y_t for evaluation. Specifically, we fit a DA-LDAR$(3,8)$ model using the EQMLE for each moving window of size 1800, and compute the one-step-ahead forecast of the τth conditional quantile of y_{t+1}, given by $Q_{y_{t+1}}(\tau \mid \mathcal{F}_t) = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\widehat{b}_\tau$, where $\widehat{\mu}_{t+1}$ and $\widehat{\sigma}_{t+1}$ are the predicted conditional location and standard deviation, respectively, and $\widehat{b}_\tau$ is the τth sample quantile of the residuals. Then we move the window forward by one and repeat the above procedure until all data are used. Finally, we obtain 779 one-week-ahead negative VaRs for each τ. For illustration, the rolling forecasts at $\tau = 5\%$ are displayed in Figure 2.

In addition, we compare the forecasting performance of the proposed DA-LDAR model with two comparable models in the literature: the linear double AR (LDAR) model of order (p,q) fitted by EQMLE (Liu et al., 2020) and the one-regime threshold double AR (TDAR) model of order $(p_1,p_2;q_1,q_2)$ fitted by the Gaussian QMLE (Li et al., 2016). Note that the LDAR model ignores the asymmetric effects while the TDAR model accounts for threshold effects. For comparison, the order of LDAR is the same as DA-LDAR i.e. $(p,q) = (3,8)$, while the order of TDAR is choose by BIC, and $(p_1,p_2;q_1,q_2) = (1,4;1,4)$ is selected based on $p_{\max} = q_{\max} = 10$. For the TDAR model, the delay parameter d is searched among $\{1,2,3\}$, and the threshold parameter r is searched among a compact grid with the empirical percentiles of y_t from the 10th quantile to the 90th quantile. Their VaR forecasts are computed in the same way as for the DA-LDAR model. To evaluate the forecasting performance of three models, we calculate the empirical coverage rate (ECR), and conduct VaR backtests for the

Table 7. Biases ($\times 10$), ESDs and ASD$_i$ ($i = 1, 2, 3$) of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ for DGP1 at $k = 2, 4$ and 6, when the innovations follow the Laplace$(0, 1)$ with different self-weights ω_t and random weights π_t, where ASD$_i$'s are the ASDs calculate by bootstrap method with random weights $\{\Pi_i\}$

	n	Bias	ESD	ASD$_1$	ASD$_2$	ASD$_3$
$\omega_t = 1$						
$\widehat{\rho}_2$	500	-0.050	0.042	0.042	0.043	0.043
	1000	-0.008	0.031	0.030	0.031	0.031
$\widehat{\rho}_4$	500	-0.027	0.045	0.043	0.044	0.044
	1000	-0.035	0.031	0.031	0.031	0.031
$\widehat{\rho}_6$	500	-0.029	0.044	0.043	0.043	0.044
	1000	-0.010	0.031	0.031	0.031	0.031
$\widehat{\gamma}_2$	500	-0.031	0.043	0.042	0.042	0.043
	1000	-0.011	0.031	0.030	0.030	0.030
$\widehat{\gamma}_4$	500	-0.028	0.044	0.043	0.044	0.044
	1000	-0.006	0.032	0.030	0.031	0.031
$\widehat{\gamma}_6$	500	-0.020	0.042	0.043	0.043	0.044
	1000	-0.016	0.030	0.030	0.031	0.031
ω_t at (5)						
$\widehat{\rho}_2$	500	-0.040	0.041	0.041	0.042	0.042
	1000	-0.009	0.030	0.030	0.030	0.030
$\widehat{\rho}_4$	500	-0.019	0.044	0.043	0.044	0.044
	1000	-0.039	0.031	0.031	0.031	0.031
$\widehat{\rho}_6$	500	-0.029	0.044	0.043	0.044	0.044
	1000	-0.012	0.031	0.031	0.031	0.031
$\widehat{\gamma}_2$	500	-0.027	0.042	0.041	0.042	0.042
	1000	-0.009	0.030	0.029	0.030	0.030
$\widehat{\gamma}_4$	500	-0.021	0.044	0.043	0.044	0.044
	1000	-0.007	0.032	0.030	0.031	0.031
$\widehat{\gamma}_6$	500	-0.017	0.041	0.043	0.043	0.044
	1000	-0.015	0.030	0.030	0.031	0.031

VaR forecasts at $\tau = 5\%$, 10%, 90% and 95%. Specifically, ECR is calculated as the proportion of observations that fall below the corresponding conditional quantile forecast for the last 779 data points. Two VaR backtests, i.e. the likelihood ratio test for correct conditional coverage (CC) in Christoffersen (1998) and the dynamic quantile (DQ) test in Engle and Manganelli (2004) are employed. Denote the hit by $H_t = I(y_t < Q_{y_t}(\tau \mid \mathcal{F}_{t-1}))$. The null hypothesis of CC test is that, conditional on $\mathcal{F}_{t-1}$, $\{H_t\}$ are *i.i.d.* Bernoulli random variables with success probability being τ. For the DQ test, following Engle and Manganelli (2004), we regress H_t on regressors including a constant, four lagged hits $H_{t-i}, i = 1, 2, 3, 4$, and the contemporaneous VaR forecast. The null hypothesis of DQ test is that all regression coefficients are zero and the intercept equals to the quantile level τ. The forecasting performance will be satisfactory if the null hypotheses of both VaR backtests cannot be rejected and the ECR is close to τ.

Table 3 reports ECRs and p-values of two VaR backtests for the one-step-ahead forecasts by the fitted DA-LDAR, LDAR and TDAR models at the lower and upper 5% and 10% conditional quantiles, i.e. 5% and 10% VaRs for long and short positions. In terms of backtests, the proposed DA-LDAR model performs satisfactorily at three quantile levels with p-values no less than 0.15, while the LDAR model only performs acceptably at the level of 10%, and the TDAR model fails in the backtests at all levels. With respect to ECRs, it can be seen that those of the DA-LDAR model are closest to the nominal quantile levels. Therefore, we conclude that the proposed DA-LDAR model dominates the other two competitors in forecasting VaRs for NASDAQ Composite Index.

7. CONCLUSION AND DISCUSSION

This paper introduces a dual-asymmetry linear double AR (DA-LDAR) model to capture the possible asymmetric effects in both the conditional location and volatility for time series data. A sufficient condition for the strict stationarity of the new model is established, and robust estimation and diagnostic checking tools are proposed without any moment conditions on the data process. An asymptotically valid random-weighting bootstrap procedure is employed to approximate covariance matrices involving unknown density. The necessity of the new model and its robust inference is corroborated by an empirical analysis on stock returns.

Our research can be extended in two directions. First, accurate VaR forecasts of financial assets is of vital importance in practice, and it is useful to study the conditional

Table 8. Biases ($\times 10$), ESDs and ASDs of $\widehat{\rho}_k$ and $\widehat{\gamma}_k$ for DGP2 and DGP3 at $k = 2, 4$ and 6 when the innovations follow the Laplace$(0,1)$, standardized t_3 or standardized $st_{3,-1}$ distribution, where ASDs are calculated by bootstrap method with i.i.d. exponential random weights $\{\Pi_1\}$

		Laplace(0, 1)			t_3			$st_{3,-1}$		
	n	Bias	ESD	ASD	Bias	ESD	ASD	Bias	ESD	ASD
DGP2										
$\widehat{\rho}_2$	500	-0.019	0.030	0.039	-0.004	0.028	0.035	-0.017	0.028	0.034
	1000	-0.007	0.021	0.027	0.002	0.020	0.024	-0.010	0.020	0.024
$\widehat{\rho}_4$	500	-0.046	0.042	0.043	-0.033	0.043	0.041	-0.044	0.041	0.041
	1000	-0.007	0.030	0.031	-0.007	0.029	0.029	-0.008	0.031	0.029
$\widehat{\rho}_6$	500	-0.007	0.043	0.044	-0.020	0.045	0.042	-0.001	0.042	0.042
	1000	0.026	0.031	0.031	-0.004	0.031	0.030	-0.019	0.029	0.030
$\widehat{\gamma}_2$	500	-0.021	0.038	0.038	-0.011	0.032	0.031	-0.020	0.032	0.031
	1000	-0.003	0.025	0.026	0.007	0.021	0.021	0.010	0.021	0.022
$\widehat{\gamma}_4$	500	-0.035	0.044	0.043	-0.005	0.040	0.040	-0.040	0.041	0.039
	1000	-0.008	0.031	0.030	-0.007	0.029	0.028	0.001	0.031	0.028
$\widehat{\gamma}_6$	500	-0.009	0.043	0.043	-0.010	0.042	0.041	-0.010	0.042	0.040
	1000	0.012	0.031	0.031	-0.014	0.029	0.029	-0.021	0.030	0.029
DGP3										
$\widehat{\rho}_2$	500	-0.026	0.045	0.043	-0.041	0.042	0.041	-0.036	0.042	0.041
	1000	-0.020	0.031	0.031	0.009	0.030	0.029	-0.006	0.028	0.029
$\widehat{\rho}_4$	500	-0.034	0.042	0.043	-0.021	0.041	0.041	-0.029	0.041	0.041
	1000	-0.011	0.029	0.030	-0.005	0.029	0.029	-0.002	0.031	0.029
$\widehat{\rho}_6$	500	-0.004	0.042	0.043	-0.026	0.041	0.042	-0.039	0.042	0.042
	1000	-0.016	0.029	0.031	-0.007	0.030	0.029	-0.018	0.028	0.029
$\widehat{\gamma}_2$	500	-0.018	0.041	0.046	-0.031	0.039	0.042	-0.046	0.037	0.042
	1000	-0.017	0.028	0.033	0.008	0.026	0.030	0.006	0.027	0.030
$\widehat{\gamma}_4$	500	-0.030	0.043	0.043	-0.020	0.041	0.040	-0.014	0.041	0.040
	1000	-0.013	0.030	0.031	-0.009	0.030	0.028	0.037	0.030	0.028
$\widehat{\gamma}_6$	500	-0.008	0.042	0.043	-0.025	0.040	0.040	-0.039	0.039	0.041
	1000	-0.014	0.030	0.031	-0.014	0.028	0.028	-0.024	0.029	0.028

Table 9. Rejection rates of the tests $Q(6)$ for DGP4 and DGP5 at the 5% significance level, where the innovations follow the Laplace$(0,1)$, standardized t_3 or standardized $st_{3,-1}$ distribution

		Laplace(0, 1)		t_3		$st_{3,-1}$	
d_1	d_2	500	1000	500	1000	500	1000
DGP4							
0.0	0.0	0.031	0.040	0.044	0.050	0.033	0.050
0.2	0.0	0.774	0.987	0.731	0.980	0.765	0.975
0.4	0.0	1.000	1.000	1.000	1.000	1.000	1.000
0.0	0.2	0.265	0.774	0.255	0.654	0.234	0.645
0.0	0.4	0.919	1.000	0.860	0.998	0.864	0.996
DGP5							
0.0	0.0	0.034	0.049	0.036	0.046	0.040	0.042
0.2	0.0	0.397	0.787	0.391	0.771	0.409	0.755
0.4	0.0	0.989	1.000	0.987	1.000	0.985	1.000
0.0	0.2	0.285	0.746	0.239	0.615	0.242	0.603
0.0	0.4	0.885	0.999	0.816	0.998	0.843	0.994

quantile estimation in the framework of DA-LDAR models. Second, the DA-LDAR model can be generalized to a vector form to jointly model multivariate time series, and it is interesting to investigate whether the robust inference tools are still applicable for the multivariate DA-LDAR model. We leave these topics for future research.

SUPPLEMENTARY MATERIAL

To view the supplementary material for this article, please visit: http://intlpress.com/site/pub/files/_supp/sii/2023/0016/0001/SII-2023-0016-0001-s001.pdf.

ACKNOWLEDGEMENTS

We are grateful to the editor, guest editor and two anonymous referees for their valuable comments and suggestions, which led to substantial improvement of this article. Zhu's research was supported by an NSFC grant 12001355, Shanghai Pujiang Program 2019PJC051 and Shanghai Chenguang Program 19CG44.

Received 1 March 2021

REFERENCES

Black, F. (1976). Studies of stock market volatility changes. In *Proceedings of the 1967 meetings of the business and economic statistics section*. American Statistical Association.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* **31** 307–327. MR0853051

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2008). *Time Series Analysis, Forecasting and Control*. Wiley, New York. MR2419724

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review* **39** 841–862. MR1661906

Cline, D. B. and Pu, H. (2004). Stability and the lyapounov exponent of threshold AR-ARCH models. *The Annals of Applied Probability* **14** 1920–1949. MR2099657

Efron, B. (1992). *Bootstrap methods: another look at the jackknife*. Springer, New York. MR0659849

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121

Engle, R. F. and Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics* **22** 367–381. MR2091566

Fernández, C. and Steel, M. F. (1996). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* **93** 359–371. MR1614601

Francq, C. and Zakoian, J. M. (2013). Inference in nonstationary asymmetric GARCH models. *The Annals of Statistics* **41** 1970–1998. MR3127855

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* **48** 1779–1801.

Guo, S., Li, D., and Li, M. (2019). Strict stationarity testing and GLAD estimation of double autoregressive models. *Journal of Econometrics* **211** 319–337. MR3980126

Jiang, F., Li, D., and Zhu, K. (2020). Non-standard inference for augmented double autoregressive models with null volatility coefficients. *Journal of Econometrics* **215** 165–183. MR4066067

Jin, Z., Ying, Z., and Wei, L. J. (2001). A simple resampling method by perturbing the minimand. *Biometrika* **88** 381–390. MR1844838

Kuester, K., Mittnik, S., and Paolella, M. (2006). Value-at-risk prediction: a comparison of alternative strategies. *Journal of Financial Econometrics* **4** 53–89.

Li, D., Ling, S., and Zhang, R. (2016). On a threshold double autoregressive model. *Journal of Business and Economic Statistics* **34** 68–80. MR3450051

Li, G., Leng, C., and Tsai, C.-L. (2014). A hybrid bootstrap approach to unit root tests. *Journal of Time Series Analysis* **35** 299–321. MR3226447

Li, G. and Li, W. K. (2005). Diagnostic checking for time series models with conditional heteroscedasticity estimated by the least absolute deviation approach. *Biometrika* **92** 691–701. MR2202655

Li, G., Zhu, Q., Liu, Z., and Li, W. K. (2017). On mixture double autoregressive time series models. *Journal of Business and Economic Statistics* **35** 306–317. MR3622839

Li, W. K. (2004). *Diagnostic checks in time series*. Chapman and Hall, Boca Raton.

Li, W. K., Ling, S., and McAleer, M. (2002). Recent theoretical results for time series models with GARCH errors. *Journal of Economic Surveys* **16** 245–269. MR2042568

Ling, S. (2004). Estimation and testing stationarity for double-autoregressive models. *Journal of the Royal Statistical Society: Series B* **66** 63–78. MR2035759

Ling, S. (2005). Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society: Series B* **67** 381–393. MR2155344

Ling, S. (2007). A double AR(p) model: structure and estimation. *Statistica Sinica* **17** 161–175. MR2352507

Linton, O. and Mammen, E. (2005). Estimating semiparametric ARCH (∞) models by kernel smoothing methods. *Econometrica* **73** 771–836. MR2135143

Liu, H., Tan, S., and Zhu, Q. (2020). Quasi-maximum likelihood inference for linear double autoregressive models. arXiv:2010.06103.

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

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59** 347–370. MR1097532

Pan, J., Wang, H., and Tong, H. (2008). Estimation and tests for power-transformed and threshold GARCH models. *Journal of Econometrics* **142** 352–378. MR2408740

Petruccelli, J. D. and Woolford, S. W. (1984). A threshold AR(1) model. *Journal of Applied Probability* **21** 270–286. MR0741130

Taylor, S. J. (2008). *Modelling financial time series*. World Scientific, New York.

Weiss, A. A. (1984). ARMA models with ARCH errors. *Journal of Time Series Analysis* **5** 129–143.

Wong, H. and Ling, S. (2005). Mixed portmanteau tests for time-series models. *Journal of Time Series Analysis* **26** 569–579. MR2188833

Xiao, Z. and Koenker, R. (2009). Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models. *Journal of the American Statistical Association* **104** 1696–1712. MR2750586

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control* **18** 931–955. MR1462740

Zheng, Y., Zhu, Q., Li, G., and Xiao, Z. (2018). Hybrid quantile regression estimation for time series models with conditional heteroscedasticity. *Journal of the Royal Statistical Society: Series B* **80** 975–993. MR3874306

Zheng, Z. (1987). Random weighting method. *ACTA Mathematicae Applicatae Sinica* **10** 247–253. MR0901365

Zhu, K. and Ling, S. (2011). Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models. *The Annals of Statistics* **39** 2131–2163. MR2893864

Zhu, K. and Ling, S. (2013). Quasi-maximum exponential likelihood estimators for a double AR(p) model. *Statistica Sinica* **23** 251–270. MR3076167

Zhu, Q., Zeng, R., and Li, G. (2020). Bootstrap inference for GARCH models by the least absolute deviation estimation. *Journal of Time Series Analysis* **41** 21–40. MR4048680

Zhu, Q., Zheng, Y., and Li, G. (2018). Linear double autoregression. *Journal of Econometrics* **207** 162–174. MR3856765

Songhua Tan
School of Statistics and Management
Shanghai University of Finance and Economics
Shanghai 200433
China
E-mail address: tansonghua@163.sufe.edu.cn

Qianqian Zhu
School of Statistics and Management
Shanghai University of Finance and Economics
Shanghai 200433
China
E-mail address: zhu.qianqian@mail.shufe.edu.cn