Asymptotic inference for a sign-double autoregressive (SDAR) model of order one

Emma M. Iglesias*

University of A Coruña

September 2024

Abstract

We propose an extension of the double autoregressive (DAR) model: the sign-double autoregressive (SDAR) model, in the spirit of the GJR-GARCH model (also named the sign-ARCH model). Our model shares the important property of DAR models where a unit root does not imply nonstationarity and it allows for asymmetry, as other alternatives in the literature such as the GJR-GARCH or asymmetric linear DAR and dual-asymmetry linear DAR models. We establish consistency and asymptotic normality of the quasi-maximum likelihood estimator in the context of the SDAR model. Furthermore, it is shown by simulations that the asymptotic properties also apply in finite samples. Finally, an empirical application shows the usefulness of our model specially in periods of supply/demand crises of oil disruptions, where spikes of volatility are very likely to be predominant.

Keywords: Sign-double autoregressive model; Asymptotic normality; Asymptotic Theory; Consistency; Stationarity; Quasi maximum likelihood estimation.

MSC Classification: 62F12; 62M10; 62P20. **JEL classification:** C12; C13; C22.

---

*Address: Department of Economics, Universidade da Coruña, ECOBAS, Campus de A Coruña. 15071, A Coruña. Spain. Corresponding author email address: emma.iglesias@udc.es. I wish to thank the Editor and two referees for very helpful and constructive comments. I also thank participants at the 24th International Conference on Computational Statistics (COMPSTAT) held in Bologna, August 23-26, 2022 and COMPSTAT 2023 held in London, August 22-25, 2023, for very helpful comments. The author is also very grateful for the financial support from the Spanish Ministry of Science and Innovation -project PID2022-137923NB-I00- and from Xunta de Galicia -project ED431C 2024/16-.

1 Introduction

Double autoregressive models (i.e. DAR models; see the pioneering work of Ling (2004, 2007), Ling and Li (2008) and Zhu and Ling (2013)) have attracted recently an important attention, mainly because they are simple models where a unit root alone does not imply nonstationarity (see e.g. Borkovec (2000), Borkovec and Klüppelberg (2001), Liu et al (2018) and Li et al (2019)). Recently, Gouriéroux and Zakoïan (2017) have also shown that volatility induced mean reverting martingales can be constructed from DAR models and therefore they provide an alternative to non-causal models used for the analysis of bubble formations. Moreover, testing in DAR models can be based on standard large theory (see e.g. Ling (2004)), while testing in non-causal models is much more complicated in practice since it is based on large sample distributions with heavy tails (see Gouriéroux and Zakoïan (2017) and Cavaliere, Nielsen and Rahbek (2020)). Following these important properties of DAR models, we can find in the literature extensions of DAR models such as factor DAR models (Guo, Ling and Zhu (2014)), threshold DAR models (Li, Ling and Zakoïan (2015) and Li, Ling and Zhang (2016)), mixture DAR models (Li, Zhu et al (2017)), vector DAR models (Zhu et al (2017)), a DAR in Mean model (Christensen, Dahl and Iglesias (2012)), a linear DAR model (Zhu, Zheng, and Li (2018)) and a quantile DAR model (Zhu and Li (2022)).

There are many alternative ways we could parameterize an asymmetric type of DAR model, but we may learn from the literature on Generalized Autoregressive Conditional Heteroskedastic (GARCH) Models (see Engle (1982) and Bollerslev (1986)). In this paper, we propose to follow the spirit of DAR models as in Ling (2004) and the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993) jointly. Then, we can write the *sign-double autoregressive* (i.e. the *sign-DAR* -also known as SDAR-) model of order s (i.e. SDAR(s)) given by

$$y_t = \phi y_{t-s} + \sqrt{w + \alpha\left(|y_{t-s}| - \gamma y_{t-s}\right)^2} z_t, \tag{1}$$

where a unit root alone does not imply nonstationarity, as we will see in the following section. For $s = 1$, the parameter conditions are given in Assumption A below. The SDAR model given by (1) is the novel proposal in this paper, and it has some parallelisms with a special case of the linear DAR model of order s of Zhu, Zheng, and Li (2018) given by

$$y_t = \phi^* y_{t-s} + \left(w^* + \alpha^* |y_{t-s}|\right) z_t, \tag{2}$$

as we will see in the following sections.

In order to understand the motivation of our proposed model, as stated in Bollerslev (2010, page 140), we can find a family of asymmetric GARCH type models parameterized in the form of the Asymmetric Power ARCH (APARCH) model of Ding, Granger and Engle (1993) where, the simple

APARCH(1) is given by

$$y_t = \sigma_t z_t = \varepsilon_t, \tag{3}$$

$$\sigma_t^{\delta^*} = w^* + \alpha^* \left(|\varepsilon_{t-1}| - \gamma^* \varepsilon_{t-1}\right)^{\delta^*}, \tag{4}$$

where if $\delta^* = 2$ and $0 \leq \gamma^* \leq 1$ we obtain the GJR-ARCH model, while if $\delta^* = 1$ and $0 \leq \gamma^* \leq 1$ we have the threshold ARCH (TARCH) of Zakoïan (1994); and z_t is a sequence of independent and identically distributed (i.i.d.) random variables. Inspired by the TARCH and double TARCH (i.e., DTARCH, see Li and Li (1996) and Jiang et al (2014)) models, Tan and Zhu (2022, 2023) proposed the asymmetric linear DAR and the dual-asymmetry linear DAR models, respectively. Instead of focusing on the TARCH, our novel proposal will be based on the spirit of the GJR-ARCH. If in the mean equation of such a model we introduce an AR(1) process, the traditional GJR-ARCH (also named the *sign-ARCH model*, see Bollerslev (2010, pages 150-151)) would be

$$y_t = \phi^* y_{t-1} + \sigma_t z_t = \phi^* y_{t-1} + \varepsilon_t, \tag{5}$$

$$\sigma_t^2 = w^* + \alpha^* \left(|\varepsilon_{t-1}| - \gamma^* \varepsilon_{t-1}\right)^2. \tag{6}$$

Sampid et al (2018), Batten et al (2019), Bedoui et al (2019), Caporin and Costola (2019) and Jiang et al (2019) are recent examples, where the GJR-GARCH model has shown its usefulness in empirical work in economics and finance.

The structure of the paper is as follows. In Section 2 we present the asymmetric sign-DAR (SDAR) model and provide a simple proof of asymptotic normality and consistency of the quasi-maximum likelihood estimator (QMLE) for this model. Section 3 provides a simulation exercise showing the properties of our estimator in finite samples and the usefulness of the asymptotic approximation. Section 4 shows the usefulness of our model in an empirical application. Finally Section 5 concludes.

2 The sign-double autoregressive model of order 1 (SDAR(1))

The asymmetric sign-DAR (SDAR) model of order $s = 1$ is given by (1) for $t = 1, ..., T$. Let us denote the parameter vector of interest by $\theta = (\phi, w, \alpha, \gamma)'$ and let the true parameter values be given by $\theta_0 = (\phi_0, w_0, \alpha_0, \gamma_0)'$. The quasi log likelihood function (conditional on past values of y_t) associated with (1) is given as

$$l_T(\theta) = \sum_{t=1}^{T} l_t(\theta) = -\frac{1}{2}\sum_{t=1}^{T} \ln\left[w + \alpha\left(|y_{t-1}| - \gamma y_{t-1}\right)^2\right] - \frac{1}{2}\sum_{t=1}^{T} \frac{\left(y_t - \phi y_{t-1}\right)^2}{w + \alpha\left(|y_{t-1}| - \gamma y_{t-1}\right)^2}, \tag{7}$$

where we define $\sigma_t^2(\theta) = w + \alpha\left(|y_{t-1}| - \gamma y_{t-1}\right)^2$ and we are conditioning on the initial value y_0 as in Jensen and Rahbek (2004a). We proceed under the following set of maintained assumptions, general enough to allow the results to hold in heavy-tailed case that is important for a number of key economic and financial variables, including financial returns

Assumption A

A1 *(i) The density function of* z_t *is continuous and positive everywhere on* $\mathbb{R}$, *(ii) and* $E\left(|z_t|^{\kappa}\right) < \infty$ *for some* $\kappa > 0$.

A2 **(i)** $z_t \sim i.i.d.\,(0,1)\,,\ E\left((1-z_t^2)^2\right) = \zeta \in (0,\infty)$ *(ii) and* z_t *has a symmetric density on* $\mathbb{R}$,

A3 **(i)** $\underline{w} \leq w_0 \leq \widetilde{w}$, $\underline{\alpha} \leq \alpha_0 \leq \widetilde{\alpha}$, *(ii)* $|\phi_0| < \widetilde{\phi}$, *(iii)* $-1 \leq \gamma_0 < 1$, *with* $\underline{w}, \widetilde{w}, \underline{\alpha}, \widetilde{\alpha}$ *and* $\widetilde{\phi}$ *being some finite positive constants, where the true parameter vector* θ_0 *is an interior point in the compact parameter space* Θ**.**

A4 y_t *is strictly stationary with* $E\left(|y_t|^{\kappa}\right) < \infty$ *for some* $\kappa > 0$.

Borkovec and Klüppelberg (2001), Ling (2004) and Alsmeyer (2016) required symmetry of the innovations to characterize the existence and uniqueness of a stationary distribution for the DAR model. Our Assumption A1 is the same assumption as required in Zhu, Zheng, and Li (2018, Theorem 2) and Tan and Zhu (2022, Theorem 1), where the symmetry requirement is relaxed, in order to allow for more flexibility in empirical applications. A1(i) is imposed for identifying the unique maximizer of $E\left[l_t\left(\theta\right)\right]$ at θ_0. Assumptions A2(i) and A3(i) are very common in the traditional ARCH literature (see e.g. Jensen and Rahbek (2004a, 2004b)). With Assumption A2(ii) we follow the seminal paper of Ling (2004, Assumption 1) where the information matrix has a block diagonal form. Assumption A2(ii) can be relaxed at the cost of not having the block diagonality as in Tan and Zhu (2022, Theorem 2). A3(ii) is a special condition that also holds for the traditional DAR(1) model where we see that we do not need to restrict $|\phi_0| < 1$ as in the traditional AR(1) model and therefore a unit root does not imply nonstationarity. In our A3(iii), we need to restrict $-1 \leq \gamma_0 < 1$ while in the traditional GJR-GARCH model (see Bollerslev (2010, page 140) with $\delta^* = 2$ in (4), we need $0 \leq \gamma_0^* \leq 1$. We cannot allow for $\gamma_0 = 1$ in the SDAR model and the reasons are shown in the proof of our Lemma A in Proposition 1 in the Appendix. We cannot allow either for $\gamma_0 < -1$ and the justification is given in the proof of Lemma 1. Assumption A3 is required to ensure the log-likelihood function, score function and information matrix to be bounded due to the variance without moment restrictions on y_t as in Ling (2004, 2007), and this justifies intuitively the rationale behind the asymptotic normality of the QMLE in our SDAR model. Finally, a sufficient condition for the strict stationarity of y_t as generated by (1) in A4 is provided in the following Lemma 1:

Lemma 1 Let Assumption A1 hold. If either of the following conditions hold

$$
\begin{aligned}
(i)\ \max\left\{E\left(\left|\phi_0-\alpha_0^{1/2}\left(1+\gamma_0\right)z_t\right|^{\kappa}\right), E\left(\left|\phi_0+\alpha_0^{1/2}\left(1-\gamma_0\right)z_t\right|^{\kappa}\right)\right\} &< 1, \quad \text{for } 0<\kappa\leq 1; \\
(ii)\ E\left[\left(\max\left\{\left|\phi_0-\alpha_0^{1/2}\left(1+\gamma_0\right)z_t\right|, \left|\phi_0+\alpha_0^{1/2}\left(1-\gamma_0\right)z_t\right|\right\}\right)^{\kappa}\right] &< 1, \quad \text{for } \kappa\in\{2,3,4...\}.
\end{aligned}
$$

then, there exists a strictly stationarity of y_t as generated by (1), and this solution is unique and geometrically ergodic with $E\left(|y_t|^{\kappa}\right)<\infty$.

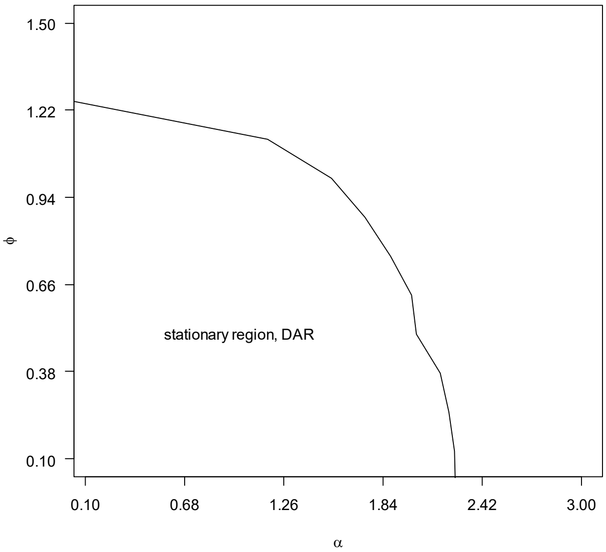
Proof of Lemma 1 Given in the Appendix.

Remark 1 *For simplicity reasons, if z_t is also symmetric in addition to A1, then the condition in Lemma 1 simplifies to* $E\left(\left|\phi_0+\alpha_0^{1/2}\left(1-\gamma_0\right)z_t\right|^{\kappa}\right)<1$ *for* $0<\kappa\leq 1$ *and* $E\left[\left(|\phi_0|+\alpha_0^{1/2}\left(1-\gamma_0\right)|z_t|\right)^{\kappa}\right]<1$ *for* $\kappa\in\{2,3,4...\}$.

Remark 2 *If z_t is also symmetric in addition to A1 and $\gamma_0=0$, then in Lemma 1 we obtain the result of Borkovec and Klüppelberg (2001, Proposition 2) for the traditional DAR(1).*

Remark 3 *Note that Klüppelberg and Pergamenchtchivkov (2004, see Lemmas 3.1 and 3.2) showed the assumptions for the strict stationarity condition of the double autoregressive model of order q. Using the same methodology as the one used in Lemma 1 we may generalize the results of Klüppelberg and Pergamenchtchivkov (2004) to find the conditions of strict stationarity of the SDAR model of order q including q lags.*

Following the pioneering work of Ling (2004), we proceed to show, using numerical methods, the set which ensures strict stationarity for the SDAR. Following Figure 1 in Guo et al (2019), we simulate first the case of a SDAR(1) model with $\gamma=0$ and $z_t\sim N(0,\pi/2)$, where the (strictly) stationary region is determined by $\left\{(\phi,\alpha): E\ln\left|\phi+\alpha^{1/2}z_t\right|<0\right\}$, that corresponds to the DAR(1). Due to the symmetry of Figure 1 of Guo et al (2019), we only focus in our simulations on the positive parameter space of ϕ and for $\alpha>0$ in order to satisfy our assumption A, and this corresponds to our Figure 1. Our Figure 1 replicates Figure 1 in Guo et al (2019), showing that for $\phi=0$, the α parameter can take a value of approximately 2.3 and the ϕ parameter can take values a bit higher than 1.22 and still being in the stationary region. Figure 2 (with $\gamma=-0.1$) and Figure 3 (with $\gamma=0.1$) show the case of a SDAR(1) model with $z_t\sim N(0,\pi/2)$ and the (strictly) stationary regions determined by $\left\{(\phi,\alpha,\gamma): E\ln\left|\phi+\alpha^{1/2}\left(1-\gamma\right)z_t\right|<0\right\}$. Figures 2 and 3 show that the SDAR model (the same as the DAR model) does not need to restrict $|\phi|<1$ versus the GJR-(G)ARCH. The difference of the SDAR with the DAR model is in relation to the α values: negative (positive) values of γ allow for smaller (larger) values of α in the stationarity region versus the DAR model of Figure 1.



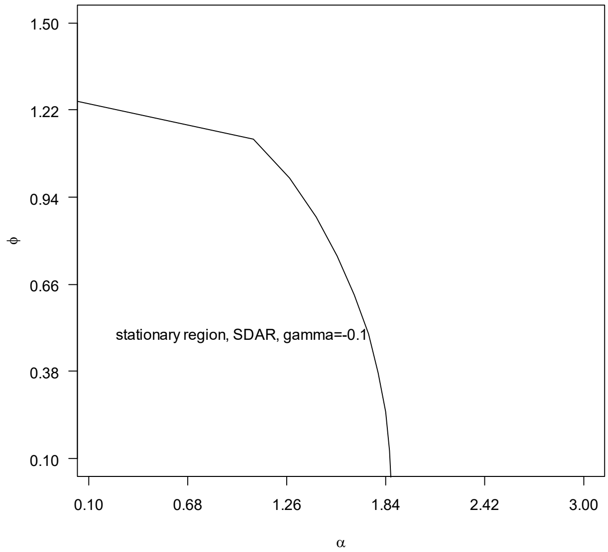


Figure 1: stationary region of DAR model. Figure 2: stationary region of SDAR model with $\gamma = -0.1$.

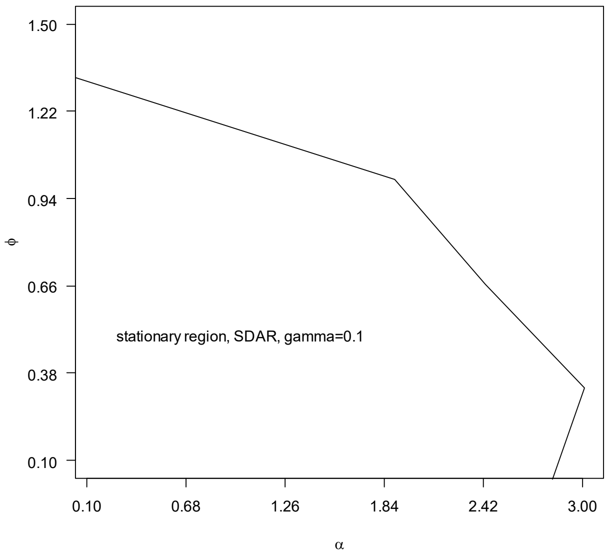


Figure 3: stationary region of SDAR model with $\gamma = 0.1$.

Then, the main result of the paper regarding the limiting distribution of the QML estimator in the SDAR model can be established in the following Theorem.

Theorem 1 Define $u_{1t}(\theta_0) = \left(\frac{y_{t-1}}{\sigma_t(\theta_0)}\right)$, $u_{2t}(\theta_0) = \left(\frac{1}{\sigma_t^2(\theta_0)}\right)$, $u_{3t}(\theta_0) = \left(\frac{w_0}{\alpha_0}\right)\left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)}\right)$ and $u_{4t}(\theta_0) = \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0\right)}\right)\left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)}\right)$, and let Assumptions A1(i) and A2-A4 hold. Consider the quasi log likelihood function given by (7). Then, there exists a fixed open neighborhood $U = U(\theta_0)$

of θ_0 such that with probability tending to one as $T \longrightarrow \infty$, $l_T(\theta)$ has a unique maximum point $\widehat{\theta}$ in U. In addition, the QML estimator $\widehat{\theta}$ is consistent and asymptotically normal

$$\sqrt{T}\left[\widehat{\theta} - \theta_0\right]' \xrightarrow{d} N\left(0, \left(\zeta^2/4\right)\Lambda^{-1}\right),$$

where

$$\Lambda = \zeta \begin{bmatrix} \zeta^{-1}\overline{m}_{11} & 0 & 0 & 0 \\ 0 & \frac{1}{2}\overline{m}_{22} & \frac{1}{2}\overline{m}_{23} & -\overline{m}_{24} \\ 0 & \frac{1}{2}\overline{m}_{23} & \frac{1}{2}\overline{m}_{33} & -\overline{m}_{34} \\ 0 & -\overline{m}_{24} & -\overline{m}_{34} & 2\overline{m}_{44} \end{bmatrix} > 0,$$

and $\overline{m}_{ij} = E\left(u_{it}(\theta_0)\, u_{jt}(\theta_0)\right)$ for $i = 1, 2, 3, 4$ and $j = 1, 2, 3, 4$.

Proof of Theorem 1 The proof of Theorem 1 is given in the Appendix.

Note that Ling (2004) showed the consistency and asymptotic normality of the QML estimator in the DAR(1) model only under the strict stationarity condition. Ling (2007) generalized this result to the DAR(p). Here we show the asymptotic normality of the QML in the SDAR model also under minimal assumptions given in Assumption A.

From Lemma 1, we can also characterize, following the literature on extreme values (see e.g. Basrak et al (2002) and Dahl and Iglesias (2022)), the tail behaviour of the *sign-double autoregressive model* in the following Lemma 2.

Lemma 2 *If Assumption A holds, then y_t given by (1) have a regularly varying tail with tail index $\kappa_1(\phi, \alpha, \gamma) > 0$ given as the unique positive solution to*

$$(i)\ \max\left\{E\left(\left|\phi - \alpha^{1/2}(1+\gamma) z_t\right|^{\kappa}\right), E\left(\left|\phi + \alpha^{1/2}(1-\gamma) z_t\right|^{\kappa_1/2}\right)\right\} = 1, \quad \text{for } 0 < \kappa_1(\phi, \alpha, \gamma) \leq 1;$$

$$(ii)\ E\left[\left(\max\left\{\left|\phi_0 - \alpha_0^{1/2}(1+\gamma_0) z_t\right|, \left|\phi_0 + \alpha_0^{1/2}(1-\gamma_0) z_t\right|\right\}\right)^{\kappa_1/2}\right] = 1, \quad \text{for } \kappa_1(\phi, \alpha, \gamma) \in \{2, 3, 4...\}.$$

Proof of Lemma 2 Given in the Appendix.

3 Simulations

In this section we evaluate and discuss, based on simulations, how well the asymptotic results of $\widehat{\theta}$ given by our Theorem 1 can approximate the finite sample properties. $1\{.\}$ is an indicator function.

We start by simulating the SDAR of order 1 given in (1) with $s = 1$, that may have different representations given by

$$y_t = \phi y_{t-1} + \sigma_t z_t = \phi y_{t-1} + \varepsilon_t, \tag{8}$$

$$\sigma_t^2 = w + \alpha \left(|y_{t-1}| - \gamma y_{t-1} \right)^2 \tag{9}$$

$$= w + \alpha_1 y_{t-1}^2 + \alpha_2 y_{t-1}^2 1 \{ y_{t-1} < 0 \} \tag{10}$$

$$= w + \alpha (1 - \gamma)^2 y_{t-1}^2 + 4\alpha\gamma y_{t-1}^2 1 \{ y_{t-1} < 0 \} \tag{11}$$

$$= w + \alpha (1 - \gamma)^2 y_{t-1}^2 1 \{ y_{t-1} \geq 0 \} + \alpha (1 + \gamma)^2 y_{t-1}^2 1 \{ y_{t-1} < 0 \} . \tag{12}$$

where $\alpha_1 = \alpha (1 - \gamma)^2 = \alpha (1 + \gamma^2) - 2\alpha\gamma$ and $\alpha_1 + \alpha_2 = \alpha (1 + \gamma^2) + 2\alpha\gamma \Longrightarrow \alpha (1 + \gamma^2) - 2\alpha\gamma + \alpha_2 = \alpha (1 + \gamma^2) + 2\alpha\gamma \Longrightarrow \alpha_2 = 4\alpha\gamma$. Therefore, $\frac{(1-\gamma)^2}{4\gamma} = \frac{\alpha_1}{\alpha_2}$ and $\alpha = \frac{\alpha_2}{4\gamma}$. We use the specification of estimating α_1 and α_2 in (8) and (10), since this is the form that it is widely used in practice to estimate the traditional GJR-GARCH model (see e.g. Zavadska et al (2020)). Moreover, for each value of α_1 and α_2 in (8) and (10), we can also find the corresponding values for γ and α in (8) and (9) that we also compute in Table 1, and we show that they justify our Assumption A3. We carry out 10000 simulations. Table 1 shows the results for the SDAR(1) and different sample sizes with z_t being drawn from a normal distribution with zero mean and unit variance (i.e. $z_t \sim N(0, 1)$). Tables 1 and 2 show that our Theorem 1 provides a reasonable approximation in finite samples and the robustness to distributional assumptions of z_t by drawing from a t distribution with 4 degrees of freedom (t_4). The finite sample properties when drawing from innovations that follow a t distribution are moderately worse than when drawing from the normal distribution, however both Tables 1 and 2 provide support that our Theorem 1 provides a reasonable approximation in finite samples. Tables 1 and 2 confirm that the asymptotic results of Theorem 1 are a sensible approximation of the finite sample results in practice. In Table 1, the asymptotic standard deviations (AD) from Theorem 1 and the empirical standard deviations (SD) are quite different; while in Table 2, they are more similar when the sample size increases.

Table 1: Bias and standard deviations of MLEs for the SDAR(1) model in (10). $w = 1$ and $z_t \sim N(0,1)$.

ϕ	α_1	α_2		$T=200$				$T=400$			
				$\widehat{\phi}$	$\widehat{w}$	$\widehat{\alpha_1}$	$\widehat{\alpha_2}$	$\widehat{\phi}$	$\widehat{w}$	$\widehat{\alpha_1}$	$\widehat{\alpha_2}$
0.5	0.5	0.2	*BIAS*	−0.003	−0.202	−0.174	0.030	−0.002	−0.121	−0.102	0.011
	γ	α	*SD*	0.109	0.476	0.232	0.220	0.091	0.409	0.209	0.168
	0.084	0.596	*AD*	0.056	0.134	0.139	0.222	0.044	0.101	0.109	0.168
0.5	0.5	0.5	*BIAS*	−0.005	0.017	−0.045	−0.011	−0.004	0.016	−0.002	−0.003
	γ	α	*SD*	0.144	0.352	0.473	0.283	0.126	0.207	0.459	0.221
	0.172	0.729	*AD*	0.065	0.173	0.170	0.308	0.047	0.123	0.124	0.220
0.5	0.5	0.7	*BIAS*	−0.007	0.083	−0.009	−0.051	−0.003	0.055	−0.002	−0.018
	γ	α	*SD*	0.161	0.336	0.561	0.281	0.141	0.260	0.543	0.223
	0.215	0.812	*AD*	0.063	0.185	0.174	0.337	0.045	0.127	0.122	0.239
1	0.5	0.2	*BIAS*	−0.303	0.303	−0.321	0.009	−0.287	0.238	−0.176	0.001
	γ	α	*SD*	0.250	1.126	0.284	0.260	0.230	1.026	0.280	0.216
	0.084	0.596	*AD*	0.038	0.166	0.091	0.143	0.030	0.129	0.079	0.107
1	0.5	0.5	*BIAS*	−0.380	0.360	−0.220	−0.117	−0.243	0.245	−0.157	−0.090
	γ	α	*SD*	0.224	1.249	0.412	0.341	0.169	0.943	0.402	0.297
	0.172	0.729	*AD*	0.042	0.225	0.120	0.199	0.032	0.170	0.092	0.167
1	0.5	0.7	*BIAS*	−0.409	0.477	−0.143	−0.213	−0.294	0.235	−0.082	−0.163
	γ	α	*SD*	0.208	1.336	0.542	0.365	0.154	1.033	0.523	0.323
	0.215	0.812	*AD*	0.042	0.258	0.136	0.229	0.029	0.184	0.097	0.163
0.5	1	0.2	*BIAS*	−0.005	−0.125	−0.301	0.073	−0.004	−0.105	−0.147	0.051
	γ	α	*SD*	0.164	0.894	1.049	0.297	0.161	0.690	1.034	0.249
	0.046	1.098	*AD*	0.037	0.151	0.172	0.266	0.026	0.110	0.135	0.203
0.5	1	0.5	*BIAS*	−0.003	0.243	−0.062	−0.012	−0.009	0.180	−0.031	−0.001
	γ	α	*SD*	0.201	1.131	1.128	0.345	0.193	0.878	1.014	0.300
	0.101	1.237	*AD*	0.041	0.203	0.221	0.363	0.036	0.138	0.162	0.263
0.5	1	0.7	*BIAS*	−0.008	0.288	0.056	−0.083	−0.005	0.211	0.049	−0.055
	γ	α	*SD*	0.217	1.291	1.265	0.349	0.210	1.126	1.219	0.307
	0.132	1.327	*AD*	0.039	0.226	0.242	0.399	0.033	0.148	0.170	0.284

BIAS denotes the empirical bias. AD denotes asymptotic standard deviation from Theorem 1 and SD denotes the empirical standard deviations.

Table 2: Bias and standard deviations of MLEs for the SDAR(1) model in (10). $z_t \sim N(0,1)$, $w = 1$ and $z_t \sim t_4$.

ϕ	α_1	α_2		$T = 1000$ *and* $z_t \sim N(0,1)$				$T = 1000$ *and* $z_t \sim t_4$			
				$\widehat{\phi}$	$\widehat{w}$	$\widehat{\alpha_1}$	$\widehat{\alpha_2}$	$\widehat{\phi}$	$\widehat{w}$	$\widehat{\alpha_1}$	$\widehat{\alpha_2}$
0.5	0.5	0.2	*BIAS*	−0.001	−0.025	−0.009	0.007	−0.002	−0.031	−0.011	−0.009
	γ	α	*SD*	0.069	0.131	0.129	0.115	0.093	0.183	0.136	0.143
	0.084	0.596	*AD*	0.030	0.049	0.103	0.114	0.023	0.058	0.094	0.098
0.5	0.5	0.5	*BIAS*	−0.005	0.015	−0.001	−0.002	−0.007	−0.024	−0.009	−0.011
	γ	α	*SD*	0.101	0.128	0.116	0.140	0.120	0.131	0.131	0.161
	0.172	0.729	*AD*	0.028	0.076	0.096	0.137	0.025	0.066	0.083	0.094
0.5	0.5	0.7	*BIAS*	−0.006	0.038	−0.001	−0.006	−0.004	−0.020	−0.009	−0.012
	γ	α	*SD*	0.123	0.190	0.201	0.155	0.135	0.221	0.224	0.201
	0.215	0.812	*AD*	0.026	0.077	0.094	0.147	0.024	0.068	0.085	0.134
1	0.5	0.2	*BIAS*	−0.004	0.026	−0.011	0.001	−0.002	−0.035	−0.015	−0.010
	γ	α	*SD*	0.070	0.131	0.129	0.106	0.093	0.203	0.214	0.196
	0.084	0.596	*AD*	0.029	0.070	0.069	0.106	0.023	0.068	0.054	0.098
1	0.5	0.5	*BIAS*	−0.007	0.027	−0.009	−0.010	−0.002	−0.039	−0.011	−0.012
	γ	α	*SD*	0.101	0.128	0.116	0.140	0.120	0.197	0.192	0.201
	0.172	0.729	*AD*	0.029	0.076	0.077	0.137	0.025	0.066	0.063	0.114
1	0.5	0.7	*BIAS*	−0.010	0.039	−0.014	−0.006	−0.004	−0.020	−0.019	−0.009
	γ	α	*SD*	0.123	0.242	0.240	0.155	0.135	0.271	0.275	0.220
	0.215	0.812	*AD*	0.026	0.078	0.074	0.148	0.024	0.068	0.055	0.104
0.5	1	0.2	*BIAS*	−0.002	−0.031	−0.014	0.006	−0.003	−0.037	−0.021	0.009
	γ	α	*SD*	0.075	0.190	0.210	0.196	0.138	0.227	0.234	0.246
	0.046	1.098	*AD*	0.016	0.073	0.093	0.136	0.013	0.050	0.076	0.092
0.5	1	0.5	*BIAS*	−0.007	0.033	−0.011	−0.001	−0.005	−0.034	−0.019	−0.011
	γ	α	*SD*	0.129	0.136	0.217	0.215	0.134	0.169	0.238	0.236
	0.101	1.237	*AD*	0.034	0.093	0.104	0.167	0.044	0.079	0.092	0.144
0.5	1	0.7	*BIAS*	−0.001	0.032	0.022	−0.020	−0.007	−0.044	0.034	−0.031
	γ	α	*SD*	0.098	0.203	0.203	0.174	0.103	0.221	0.234	0.197
	0.132	1.327	*AD*	0.031	0.132	0.153	0.165	0.032	0.094	0.137	0.153

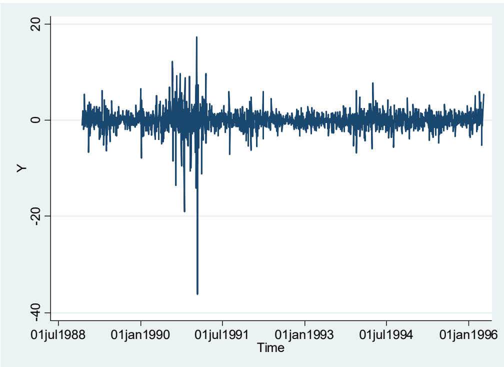
BIAS denotes the empirical bias. AD denotes asymptotic standard deviation from Theorem 1 and SD denotes the empirical standard deviations.

4 Empirical application

Iglesias and Rivera-Alonso (2022) showed that the results in Zavadska et (2020) were robust to estimating GARCH and GJR-GARCH-type models subject to positivity restrictions of the conditional variance, when applied to oil BRENT returns (given by y_t), and they are defined as the daily closing spot oil BRENT returns calculated as

$$100ln(BRENT\_S_t/BRENT\_S_{t-1}),$$

with $BRENT\_S_t$ being the BRENT daily spot oil prices at time t. Their main results supported the theory that spikes of volatility (which are highly erratic) are produced during periods of supply/demand crises of oil disruptions while periods where economic/financial/stock market crises are the predominant trigger are associated to higher volatility persistence. However, Zavadska et (2020) and Iglesias and Rivera-Alonso (2022) did not consider models with dynamics in the mean equation. We carried out two extensions of the results of Zavadska et (2020) and Iglesias and Rivera-Alonso (2022) as follows.



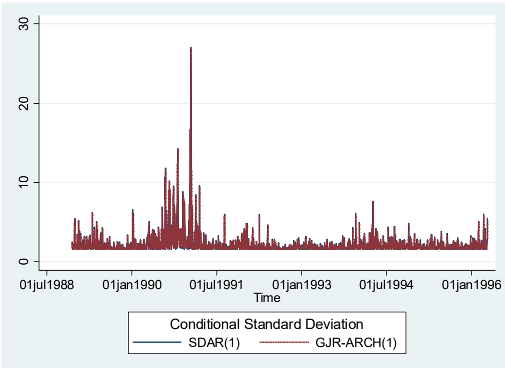


Figure 4: y_t during Gulf-war period. Figure 5: Conditional Standard Deviations.

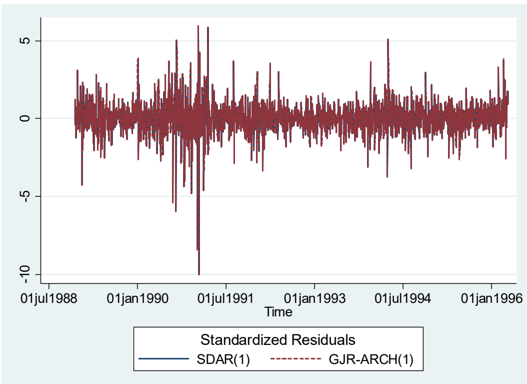


Figure 6: Standardized Residuals.

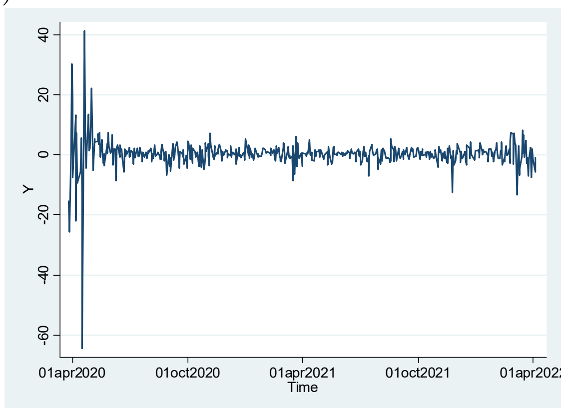
First, we analyzed all periods of supply/demand crises of oil disruptions provided in Iglesias and Rivera-Alonso (2022), since specially in those cases is where spikes of volatility are very likely to be predominant (and therefore the ARCH-type term is larger and more relevant than the GARCH-type term); what is the best situation where our SDAR model may be useful in practice (since the SDAR model does not contain past values of the volatility term). We find that indeed for the periods of the 2001 US-terrorist attack and the 2014/2016 oil conflict of US and Saudi Arabia, our SDAR model did not provide statistical significant relationships. However, our SDAR model was useful to analyze the Gulf-war period (dated from 1988-12-07 to 1996-04-12 in Zavadska et (2020)) and the results are provided in Table 3. Figure 4 shows the evolution of the BRENT returns from 1988-12-07 to 1996-04-12. For that period, we started by estimating different GARCH-type models such as GARCH, Exponential-GARCH (EGARCH) of Nelson (1991), GJR-GARCH and the Log-GARCH of Geweke (1986), Pantula (1986) and Milhøj (1987), all of them with different lags in the mean equation. As expected from the results of Zavadska et al (2020) and Iglesias and Rivera-Alonso (2022), the GARCH term is never statistically significant, and the model that provides statistically significant relationships for all the parameters and estimated standardized residuals with supporting evidence of white noise with the portmanteau (or Q) test for white noise of Box and Pierce (1970) and Ljung and Box (1978) is the AR(1)-GJR-ARCH(1) model. Therefore, we focus on two competitor models: the AR(1)-GJR-ARCH(1) model given as

$$y_t = \phi^* y_{t-1} + \sigma_t z_t = \phi^* y_{t-1} + \varepsilon_t, \tag{13}$$

$$\sigma_t^2 = w^* + \alpha_1^* \varepsilon_{t-1}^2 + \alpha_2^* \varepsilon_{t-1}^2 1\left\{\varepsilon_{t-1} < 0\right\} \tag{14}$$

and the SDAR(1) model given at (8) and (10) (SDAR models with lags higher than 1 did not provide statistically significant relationships). Table 3 shows the estimation results when estimating by QML both models, and where all the parameters estimates are statistically significant at the 1% level. Table 4 shows as well the estimated values of γ and α. We notice that the estimated γ value for our SDAR model is negative (see Table 4), what is allowed from our assumption A3; and from our Figure 2, estimated values of ϕ, γ and α for the Gulf-war period are inside the stationarity region, satisfying our assumption A4. In the traditional AR-GJR-ARCH model we also get a negative estimated γ^* value (see Table 4), and from the best of our knowledge (see Bollerslev (2010, page 140) with $\delta^* = 2$ in (4)), we need $0 \leq \gamma_0^* \leq 1$, what implies that the theoretical properties of such a case are unknown in relation to the asymptotic normality of the QMLE, and this provides evidence of the usefulness of our model in practice. Table 5 shows diagnostic test statistics for the standardized residuals including descriptive statistics (mean, standard deviation, minimum and maximum values) and the portmanteau (or Q) test for white noise of Box and Pierce (1970) and Ljung and Box (1978) for the standardized residuals ($\widehat{z}_t$), with p-values in brackets. We show that we cannot reject the null hypothesis that the residuals are white noise neither for the SDAR(1) nor for the AR(1)-GJR-ARCH(1) model (we use lags 20 and 40; other intermediate lags do not provide evidence against the null hypothesis of white noise). However, there is evidence that the residuals are slightly closer to a mean of zero and a standard deviation of 1 for the SDAR(1) than for the AR(1)-GJR-ARCH(1) (although in both cases they are not statistically different from 0 and 1 respectively). Following Ling (2004) -where he compared the log-likelihood values of the AR(1) and the DAR(1)-, the value of the Log-likelihood of the estimated SDAR(1) model is -3945.2 while if we estimate an AR(1) model is -4183.0, showing that the SDAR(1) model is much better fit than the AR(1) model. The Log-Likelihood value of the estimated AR(1)-GJR-ARCH(1) is -3924,3, however, we cannot compare this value with the one of the SDAR(1) model since they are non-nested. The Vuong (1984) test for non-nested models, where the null hypothesis states that the two competing models (the SDAR(1) versus the AR(1)-GJR-ARCH(1) models) are equally close to the true data generating process, shows a p-value that allows not to reject the null hypothesis, what means that both models are complementary; although in the case of the SDAR model, we can provide the asymptotic theory covering the negative estimated γ value[1].

Figures 5 and 6 show that the estimated standard deviations and standardized residuals of both models have a similar behaviour, although the standardized residuals of the AR(1)-GJR-ARCH(1) present larger extreme values (see also the comparison of minimum and maximum values in Table 5).



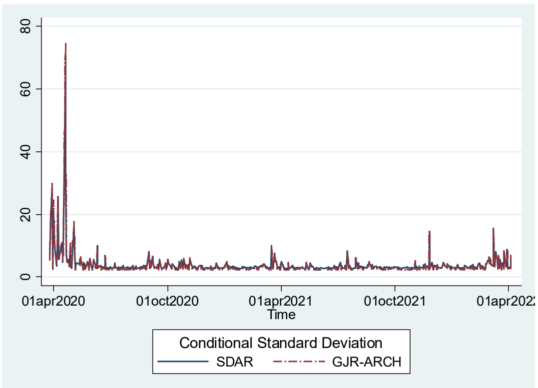


Figure 7: y_t from 2020-03-26 to 2022-04-08. Figure 8: Conditional Standard Deviations.

[1] The Vuong (1984) test when applied to compare a GARCH(1,1) model versus an AR(1)-GJR-ARCH(1), or also versus a GJR-GARCH(1,1) or an EGARCH or Log-GARCH does not allow either to find a superior model.

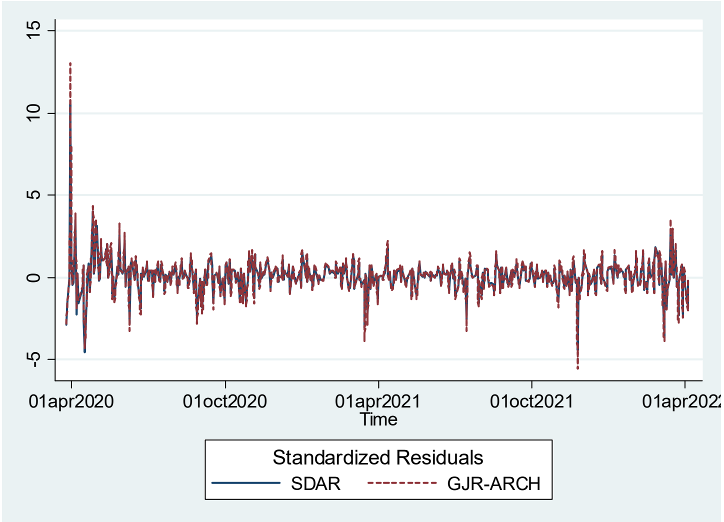


Figure 9: Standardized residuals.

Second, the last sub-sample analyzed in Iglesias and Rivera-Alonso (2022) covered from 2020-03-26 to 2021-03-10, and it corresponded to a period of supply/demand disruptions during the 2020 Russia-Saudi Arabia war and Covid-19. We now extend the end of this period from 2021-03-10 to 2022-04-08[2]. Figure 7 shows the evolution of the BRENT returns. We again started by estimating different GARCH-type models such as GARCH, Exponential-GARCH (EGARCH) of Nelson (1991), GJR-GARCH and the Log-GARCH of Geweke (1986), Pantula (1986) and Milhøj (1987), all of them with dynamics in the mean equation. As expected from the results of Zavadska et al (2020) and Iglesias and Rivera-Alonso (2022), the GARCH term is never statistically significant in all cases, and the model that provides statistically significant relationships for all the parameters and estimated standardized residuals with evidence of white noise is the AR(1)-GJR-ARCH(1) model. Therefore, we focus again in our two competing models: the AR(1)-GJR-ARCH(1) and the SDAR(1). The theoretical argument in Zavadska et al (2020) and Iglesias and Rivera-Alonso (2022) states that spikes of volatility (which are highly erratic) are produced during periods of supply/demand crises of oil disruptions (where the estimated values of α_1^* and $\alpha_1^* + \alpha_2^*$ in (13) and (14) and α_1 and $\alpha_1 + \alpha_2$ in (8) and (10) tend to be quite large) while periods where economic/financial/stock market crises are the predominant trigger are associated to higher volatility persistence (where the estimated α_1^* and $\alpha_1^* + \alpha_2^*$ in (13) and (14) and α_1 and $\alpha_1 + \alpha_2$ in (8) and (10) tend to be quite low). Table 4 shows the estimated values of α_1^* and $\alpha_1^* + \alpha_2^*$ for the GJR-ARCH (and α_1 and $\alpha_1 + \alpha_2$ for the SDAR(1)). Table 5 shows diagnostic test statistics for the standardized residuals including descriptive statistics (mean, standard deviation, minimum and maximum values) and the portmanteau (or Q) test for white noise of Box and Pierce (1970) and Ljung and Box (1978) for the standardized residuals ($\widehat{z}_t$),

[2] We checked that the results are robust both when we include and exclude the Ukraine-Russia war period, that started on February 24, 2022.

with p-values in brackets. We show that we cannot reject the null hypothesis that the residuals are white noise neither for the SDAR(1) nor for the AR(1)-GJR-ARCH(1) model (we use lags 20 and 40; other intermediate lags do not provide evidence against the null hypothesis of white noise). However, there is evidence that the residuals are slightly closer to a mean of zero and a standard deviation of 1 for the SDAR(1) than for the AR(1)-GJR-ARCH(1) (although in both cases they are not statistically different from 0 and 1 respectively). The value of the Log-likelihood of the estimated SDAR(1) model is -1340,6 while if we estimate an AR(1) model is -1533,7, showing that the SDAR(1) model is much better fit than the AR(1) model. The Log-Likelihood value of the estimated AR(1)-GJR-ARCH(1) is -1285,7, however, we cannot compare this value with the one of the SDAR(1) model since they are non-nested. The Vuong (1984) test for non-nested models, where the null hypothesis states that the two competing models (the SDAR(1) versus the AR(1)-GJR-ARCH(1) models) are equally close to the true data generating process, shows a p-value of 0.242 that allows not to reject the null hypothesis, what means that both models are complementary; although in the case of the SDAR model, we can provide the asymptotic theory covering the negative estimated γ value[3]. Therefore, in this application, both models are complementary and there is no support to find a superior model. However, there are differences when using the two models, since the SDAR(1) model does not restrict $|\phi_0| < 1$ as stated in assumptions A3 and A4 and Figures 2 and 3; and also it is straightforward to check the stationarity condition even when γ is negative. Figures 8 and 9 show the estimated standard deviations and standardized residuals of both models, showing a similar behaviour, although the standardized residuals of the AR(1)-GJR-ARCH(1) model tend to be larger in absolute value (see also the comparison of minimum and maximum values in Table 5).

Table 3: Estimated results for different crises periods. Dependent variable: Brent oil returns.

Period	$SDAR(1)$ model				$AR(1) - GJR - ARCH(1)$ model			
	ϕ	w	α_1	α_2	ϕ^*	w^*	α_1^*	α_2^*
1988-12-07	0.071***	2.474***	0.918***	-0.359***	0.198***	2.379***	0.896***	-0.343***
to 1996-04-12	(0.010)	(0.068)	(0.024)	(0.048)	(0.013)	(0.066)	(0.024)	(0.048)
2020-03-26	-0.076***	7.667***	0.476***	0.883***	-0.241***	4.793***	0.658***	0.683***
to 2022-04-08	(0.006)	(0.369)	(0.079)	(0.158)	(0.012)	(0.259)	(0.090)	(0.132)

*** Denotes statistical significance at the 1% level. Standard errors are given in parenthesis.

[3]The Vuong (1984) test when applied to compare a GARCH(1,1) model versus an AR(1)-GJR-ARCH(1), or also versus a GJR-GARCH(1,1) or an EGARCH does not allow either to find a superior model.

Table 4: Estimated results for different crises periods. Dependent variable: Brent oil returns.

Period	$SDAR(1)$ model				$AR(1)-GJR-ARCH(1)$ model			
	α_1	$\alpha_1+\alpha_2$	γ	α	α_1^*	$\alpha_1^*+\alpha_2^*$	γ^*	α^*
1988-12-07 to 1996-04-12	0.918	0.559	-0.123	0.727	0.896	0.553	-0.120	0.714
2020-03-26 to 2022-04-08	0.476	1.358	0.256	0.861	0.658	1.342	0.176	0.969

Table 5: Diagnostic tests of standardized residuals

Model/Period	1988-12-07 to 1996-04-12	2020-03-26 to 2022-04-08
$SDAR(1)$ model		
mean	0.015	0.100
sd	1.014	0.990
minimum	-9.912	-4.507
maximum	5.856	10.536
Q(20)	[0.229]	[0.272]
Q(40)	[0.082]	[0.239]
$AR(1)-GJR-ARCH(1)$ model		
mean	0.016	0.116
sd	1.031	1.150
minimum	-10.042	-5.553
maximum	5.928	13.055
Q(20)	[0.235]	[0.450]
Q(40)	[0.084]	[0.451]

Q denotes the portmanteau test for white noise, with p-values in brackets.

5 Conclusion

In the traditional volatility GARCH literature, TARCH and GJR-ARCH models are very important tools and they are both needed in empirical applications. Inspired in the TARCH, Tan and Zhu (2022, 2023) have recently proposed the asymmetric linear DAR and the dual-asymmetry linear DAR models. Inspired in the GJR-ARCH model, in this paper we propose the *sign-double autoregressive (SDAR)* model and we establish consistency and asymptotic normality of the QML estimator in such a model with one lag. We also show in simulations that the asymptotic theory provides a good approximation in finite samples. An empirical application shows the usefulness of our new model. A generalization for the SDAR model with lags higher than one is left for future research. Klüppelberg et (2002), Ling (2004) and Guo et al (2019) provided testing procedures for the stationarity of AR-(G)ARCH and DAR processes. Along the same lines, we could also develop new testing procedures for the stationarity of SDAR processes and this is also a subject of future research since it is not straightforward.

6 Data Availability Statement

The data required to reproduce the above findings cannot be publicly shared due to confidentiality reasons.

Appendix

We start with the proof of Lemmas 1 and 2.

Proof of Lemma 1 Before proving Lemma 1, we provide a short review of the current literature of conditions for stationarity in models related to our new SDAR model. A second and alternative representation for the GJR-GARCH model in model (3)-(4) with $\delta^* = 2$ and $0 \leq \gamma^* \leq 1$ is where we replace $\sigma_t^2 = w^* + \alpha^* \left(|\varepsilon_{t-1}| - \gamma^* \varepsilon_{t-1}\right)^2$ by $\sigma_t^2 = w^* + \alpha_1^* \varepsilon_{t-1}^2 + \alpha_2^* \varepsilon_{t-1}^2 1\{\varepsilon_{t-1} < 0\}$ and $1\{.\}$ is an indicator function (see Glosten, Jagannathan and Runkle (1993)) where $1\{\varepsilon_{t-1} < 0\}$ is equal to 1 when $\varepsilon_{t-1} < 0$ and 0 otherwise. Ling and McAleer (2002), McAleer, Chan and Marinova (2007) and McAleer, Hoti and Chan (2009) showed that the strict stationarity condition for the GJR-ARCH model is given by $E\left[\ln\left(\left(\alpha_1^* + \alpha_2^* 1\{z_t < 0\}\right) z_t^2\right)\right] < 0$. We can also find a relationship between the parameters α^*, γ^* and α_1^*, α_2^* and therefore we can find a third representation for the traditional GJR-GARCH model for the volatility given by $\sigma_t^2 = w^* + \alpha^* (1 - \gamma^*)^2 \varepsilon_{t-1}^2 1\{\varepsilon_{t-1} \geq 0\} + \alpha^* (1 + \gamma^*)^2 \varepsilon_{t-1}^2 1\{\varepsilon_{t-1} < 0\}$ where now the condition for strict stationarity is

$$E\left[\ln\left(\left(\alpha^* (1 - \gamma^*)^2 1\{z_t \geq 0\} + \alpha^* (1 + \gamma^*)^2 1\{z_t < 0\}\right) z_t^2\right)\right] < 0.$$

In relation to the DAR model, given by $y_t = \phi y_{t-1} + \sqrt{w + \alpha y_{t-1}^2} z_t$, the strict stationarity condition was first provided by Borkovec and Klüppelberg (2001), where under the assumption that the innovations z_t are symmetric with a continuous Lebesgue measure, z_t has a full support $\mathbb{R}$ and the second moment of z_t exists, they required that $E \ln\left|\phi + \alpha^{1/2} z_t\right| < 0$. More recently, Alsmeyer (2016, pages 221-224) provided an alternative and stronger condition given by $E \ln\left(\phi + \alpha^{1/2} |z_t|\right) < 0$ that only requires symmetry of z_t and it does not require that z_t has a continuous Lebesgue density, a finite second moment and that its support is the whole real line.

If we focus now on the case of the linear DAR model of order 1, given by $y_t = \phi y_{t-1} + (w + \alpha |y_{t-1}|) z_t$ and under our A1, Zhu, Zheng, and Li (2018, Theorem 2) and Tan and Zhu (2022, Theorem 1) showed that the strictly stationary solution required $\max\{E(|\phi - \alpha z_t|^\kappa), E(|\phi + \alpha z_t|^\kappa)\} < 1$ for $0 < \kappa \leq 1$ and $E[(\max\{|\phi - \alpha z_t|, |\phi + \alpha z_t|\})^\kappa] < 1$ for $\kappa \in \{2, 3, 4...\}$. If z_t is also symmetric in addition to A1, then the condition simplifies to $E(|\phi + \alpha z_t|^\kappa) < 1$ for $0 < \kappa \leq 1$ and $E[(|\phi| + \alpha |z_t|)^\kappa] < 1$ for $\kappa \in \{2, 3, 4...\}$. There is an important similarity between models (1) and (2) since both include the absolute value function. Following the proofs of Zhu, Zheng, and Li (2018, Theorem 2), and Borkovec and Klüppelberg (2001), we can show that the process is a homogeneous Markov chain with state space $\mathbb{R}$ with the Borel $\sigma-$algebra. We find that the transition kernel density for the SDAR is given by

$$P(y_1 \in dx | y_0 = y) = \frac{1}{\sqrt{w + \alpha(|y| - \gamma y)^2}} p\left(\frac{x - \phi y}{\sqrt{w + \alpha(|y| - \gamma y)^2}}\right) dx, \quad y \in \mathbb{R}$$

and under A1, because of the strict positivity and continuity of the transition density, y_t is a $\upsilon-$irreducible Feller chain. Also, we can show that for $0 < \kappa \leq 1$, and assuming $w = 1$ as Zhu, Zheng, and Li (2018, Theorem 2) without loss of generality

$$\begin{aligned}
& E(|y_{t+1}|^\kappa | y_t = x) \\
= \; & E\left(\left|\phi x + \sqrt{1 + \alpha(|x| - \gamma x)^2} z_{t+1}\right|^\kappa\right) \\
\leq \; & E\left(\left|\phi x + \left(1 + \sqrt{\alpha(|x| - \gamma x)^2}\right) z_{t+1}\right|^\kappa\right) \\
= \; & E\left(\left|\phi x + \sqrt{\alpha(|x| - \gamma x)^2} z_{t+1} + z_{t+1}\right|^\kappa\right) \\
= \; & E\left(\left|\phi x + \sqrt{\alpha}(|x| - \gamma x) z_{t+1} + z_{t+1}\right|^\kappa\right) \\
\leq \; & E\left(\left|\phi sign(x) + \sqrt{\alpha}(1 - \gamma sign(x)) z_{t+1}\right|^\kappa\right) |x|^\kappa + E(|z_{t+1}|^\kappa) \\
\leq \; & a|x|^\kappa + E(|z_{t+1}|^\kappa)
\end{aligned}$$

where

$$a = \max\left\{E\left(\left|\phi_0 - \alpha_0^{1/2}(1 + \gamma_0) z_t\right|^\kappa\right), E\left(\left|\phi_0 + \alpha_0^{1/2}(1 - \gamma_0) z_t\right|^\kappa\right)\right\}$$

and that Tweedie's drift criterion (Tweedie (1983), Theorem 4) holds. Note that we cannot allow above for $\gamma < -1$ when $x < 0$ since in that case $(|x| - \gamma x)$ is negative. Moreover, from Theorem 4(ii) in Tweedie (1983) and Theorems 1 and 2 in Feigin and Tweedie (1985), y_t is geometrically ergodic with a unique stationary distribution that implies that $E(|y_t|^\kappa) < \infty$.

Then, the condition of the existence of a unique stationary distribution for the SDAR model is given by

$$(i) \;\; \max\left\{E\left(\left|\phi_0 - \alpha_0^{1/2}(1 + \gamma_0) z_t\right|^\kappa\right), E\left(\left|\phi_0 + \alpha_0^{1/2}(1 - \gamma_0) z_t\right|^\kappa\right)\right\} < 1, \quad \text{for } 0 < \kappa \leq 1,$$

and following Tan and Zhu (2022, Theorem 1), we can show that for $\kappa \in \{2, 3, 4...\}$,

$$(ii)\ E\left[\left(\max\left\{\left|\phi_0 - \alpha_0^{1/2}(1+\gamma_0) z_t\right|, \left|\phi_0 + \alpha_0^{1/2}(1-\gamma_0) z_t\right|\right\}\right)^{\kappa}\right] < 1, \quad \text{for } \kappa \in \{2, 3, 4...\}.$$

■

Proof of Lemma 2 We analyze the stochastic difference equation

$$y_t = \phi y_{t-1} + \sqrt{w + \alpha\left(|y_{t-1}| - \gamma y_{t-1}\right)^2} z_t,$$

that satisfies the fixpoint equation (along the lines of equation (2.4) of Borkovec and Klüppelberg (2001))

$$y \stackrel{d}{=} \phi y + \sqrt{w + \alpha\left(|y| - \gamma y\right)^2} z,$$

where y is independent from z. Without loss of generality in the proof, under symmetry of the innovations and for $0 < \kappa_1(\phi, \alpha, \gamma) \leq 1$, we can define the function

$$h_{\phi,\alpha,\gamma}(u) := E\left(\left|\phi + \alpha^{1/2}(1-\gamma) z\right|^{u}\right) < 1, \qquad u \geq 0,$$

and if the parameters ϕ, α, γ are chosen such that

$$h'_{\phi,\alpha,\gamma}(0) := E\left(\ln\left|\phi + \alpha^{1/2}(1-\gamma) z\right|\right) < 0,$$

following Proposition 2 of Borkovec and Klüppelberg (2001), there exists a unique solution to the equation $h_{\phi,\alpha,\gamma}(u) = 1$. The same as in Proposition 2 of Borkovec and Klüppelberg (2001), the result follows from the properties of moment generating functions. The case of $\rho = \gamma = 0$ is given in Lemma 8.4.6 of Embrechts, Klüppelberg and Mikosch (1997).

■

The analytical expressions for the first, second and third order derivatives of the quasi log likelihood function are given in a Supplementary Appendix. We provide now three important propositions that we need in order to prove Theorem 1. The proof technique for the QMLE utilizes the classic Cramér type conditions for consistency and asymptotic normality (central limit theorem for the score, convergence of the Hessian and uniformly bounded third-order derivatives); see e.g. Lehmann (1999). The result holds by establishing the regularity conditions (A.1), (A.2) and (A.3) in Jensen and Rahbek (2004b, Lemma 1), which are classical Cramer type conditions addressing first, second and third order differentials of the log-likelihood function.

Proposition 1 Let $u_{jt}(\theta_0)$ be defined as in Theorem 1. Under Assumption A, the joint distribution of the score functions evaluated at $\theta = \theta_0$ are asymptotically Gaussian,

$$\frac{1}{\sqrt{T}}\frac{\partial}{\partial\theta}l_T(\theta_0) \xrightarrow{d} N(0,\Lambda),$$

where

$$\Lambda = \zeta \begin{bmatrix} \zeta^{-1}\overline{m}_{11} & 0 & 0 & 0 \\ 0 & \frac{1}{2}\overline{m}_{22} & \frac{1}{2}\overline{m}_{23} & -\overline{m}_{24} \\ 0 & \frac{1}{2}\overline{m}_{23} & \frac{1}{2}\overline{m}_{33} & -\overline{m}_{34} \\ 0 & -\overline{m}_{24} & -\overline{m}_{34} & 2\overline{m}_{44} \end{bmatrix} > 0,$$

and $\overline{m}_{ij} = E(u_{it}(\theta_0)u_{jt}(\theta_0))$ for $i = 1,2,3,4$ and $j = 1,2,3,4$.

Proof of Proposition 1 For the proof of Proposition 1, we need first the following 2 Lemmas.

Lemma A Let Assumption A hold and define $u_{1t}(\theta_0) = \left(\frac{y_{t-1}}{\sigma_t(\theta_0)}\right)$, $u_{2t}(\theta_0) = \left(\frac{1}{\sigma_t^2(\theta_0)}\right)$, $u_{3t}(\theta_0) = \left(\frac{w_0}{\alpha_0}\right)\left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)}\right)$ and $u_{4t}(\theta_0) = \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0\right)}\right)\left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)}\right)$. Then $u_{it}(\theta_0)$ is a stationary and ergodic sequence. In addition $\frac{1}{T}\sum_{t=1}^{T} u_{it}(\theta_0) \xrightarrow{p} E(u_{it}(\theta_0)) \equiv \overline{u}_i$ and $\frac{1}{T}\sum_{t=1}^{T} u_{it}^2(\theta_0) \xrightarrow{p} E(u_{it}^2(\theta_0)) \equiv \overline{m}_{ii}$ for $i = 1,2,3,4$.

Proof of Lemma A Define $I_t = \{y_t, z_t, y_{t-1}, z_{t-1}, y_{t-2}, z_{t-2}, \ldots\}$. Note first that

$$|u_{1t}(\theta_0)| \leq \left|\frac{y_{t-1}}{\sqrt{w_0 + \alpha_0\left(|y_{t-1}| - \gamma_0 y_{t-1}\right)^2}}\right| = \left|\frac{1}{\sqrt{\frac{w_0}{y_{t-1}^2} + \alpha_0\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0\frac{y_{t-1}}{y_{t-1}}\right)^2}}\right|,$$

hence

$$E|u_{1t}(\theta_0)| \leq \frac{1}{(1-\gamma_0)\sqrt{\alpha_0}} < \infty,$$

where we have used that under A3, $-1 \leq \gamma_0 < 1$. Hence we can write

$$u_{1t}(\theta_0) \equiv g_1\left(y_{t-1}, \sigma_t^2(\theta_0)\right),$$

where g_1 is a I_t-measurable function and where all arguments y_{t-1} and $\sigma_t^2(\theta_0)$ are stationary and ergodic as a consequence of Lemma 1. This implies that $u_{1t}(\theta_0)$ is stationary and ergodic by Theorem 3.35 in White (1984). Consequently $\frac{1}{T}\sum_{t=1}^{T} u_{1t}(\theta_0) \xrightarrow{p} E(u_{1t}(\theta_0))$ follows by the Ergodic Theorem.

Similarly, it follows straightforwardly that $E\left|u_{2t}(\theta_0)\right| \leq \left(\frac{1}{w_0}\right)$, $E\left|u_{3t}(\theta_0)\right| \leq \left(\frac{2}{\alpha_0}\right)$ and

$$\begin{aligned} E\left|u_{4t}(\theta_0)\right| &= E\left|\left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}}-\gamma_0\right)}\right)\left(\frac{1}{w_0}-\frac{1}{\sigma_t^2(\theta_0)}\right)\right| \leq E\left|\left(\frac{w_0}{(1-\gamma_0)}\right)\left(\frac{1}{w_0}+\frac{1}{\sigma_t^2(\theta_0)}\right)\right| \\ &\leq E\left|\left(\frac{w_0}{(1-\gamma_0)}\right)\left(\frac{2}{w_0}\right)\right| = \frac{2}{(1-\gamma_0)}. \end{aligned}$$

We can write $u_{2t}(\theta_0) \equiv g_2(\sigma_t^2(\theta_0))$, $u_{3t}(\theta_0) \equiv g_3(\sigma_t^2(\theta_0))$ and $u_{4t}(\theta_0) \equiv g_4(\sigma_t^2(\theta_0))$ and as above conclude that $(u_{2t}(\theta_0), u_{3t}(\theta_0), u_{4t}(\theta_0))$ is stationary and ergodic, and hence $\frac{1}{T}\sum_{t=1}^{T} u_{it}(\theta_0) \xrightarrow{p} E(u_{it}(\theta_0))$ for $i = 2, 3, 4$. Second, notice that

$$\left|u_{1t}^2(\theta_0)\right| = \left|\frac{1}{\frac{w_0}{y_{t-1}^2}+\alpha_0\left(\frac{|y_{t-1}|}{y_{t-1}}-\gamma_0\frac{y_{t-1}}{y_{t-1}}\right)^2}\right|,$$

such that

$$E\left|u_{1t}^2(\theta_0)\right| \leq \frac{1}{(1-\gamma_0)^2\alpha_0} < \infty.$$

In addition, $E\left|u_{2t}^2(\theta_0)\right| \leq \left(\frac{1}{w_0^2}\right)$, $E\left|u_{3t}^2(\theta_0)\right| \leq \left(\frac{4}{\alpha_0^2}\right)$ and $E\left|u_{4t}^2(\theta_0)\right| \leq \left(\frac{4}{(1-\gamma_0)^2}\right)$. We can therefore conclude, by Theorem 3.35 in White (1984), that since $u_{it}(\theta_0)$ is stationary and ergodic then so is $u_{it}^2(\theta_0)$ for $i = 1, 2, 3, 4$. Furthermore as $E|u_{it}^2(\theta_0)|$ is bounded then $\frac{1}{T}\sum_{t=1}^{T} u_{1t}^2(\theta_0) \xrightarrow{p} E(u_{1t}^2(\theta_0))$ for $i = 1, 2, 3, 4$ follows from the ergodicity theorem. This completes the proof of Lemma A.■

Lemma B Under Assumption A, the marginal distributions of the score functions (see the supplementary appendix) evaluated at $\theta = \theta_0$ are asymptotically Gaussian,

$$\frac{1}{\sqrt{T}}\frac{\partial}{\partial\phi}l_T(\theta_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t u_{1t}(\theta_0) \xrightarrow{d} N(0, \overline{m}_{11}), \tag{17}$$

$$\frac{1}{\sqrt{T}}\frac{\partial}{\partial w}l_T(\theta_0) = \frac{-1}{\sqrt{T}}\sum_{t=1}^{T} \frac{1}{2}\left(1-z_t^2\right)u_{2t}(\theta_0) \xrightarrow{d} N\left(0, \frac{\zeta}{4}\overline{m}_{22}\right), \tag{18}$$

$$\frac{1}{\sqrt{T}}\frac{\partial}{\partial\alpha}l_T(\theta_0) = \frac{-1}{\sqrt{T}}\sum_{t=1}^{T} \frac{1}{2}\left(1-z_t^2\right)u_{3t}(\theta_0) \xrightarrow{d} N\left(0, \frac{\zeta}{4}\overline{m}_{33}\right), \tag{19}$$

$$\frac{1}{\sqrt{T}}\frac{\partial}{\partial\gamma}l_T(\theta_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \left(1-z_t^2\right)u_{4t}(\theta_0) \xrightarrow{d} N(0, \zeta\overline{m}_{44}), \tag{20}$$

where $\overline{m}_{ii}$, $i = 1, 2, 3, 4$ and ζ are defined by Lemma A and A2 respectively.

Proof of Lemma B We will prove (17) in detail. The results in (18), (19) and (19) hold by identical arguments. Define again $I_t = \{y_t, z_{t,} y_{t-1}, z_{t-1,} y_{t-2}, z_{t-2,} ...\}$ and recall from Result 1 that

$$s_{1t}(\theta_0) = z_t u_{1t}(\theta_0).$$

Consequently

$$
\begin{aligned}
E\left(s_{1t}|I_{t-1}\right) &= E\left(z_t u_{1t}\left(\theta_0\right)|I_{t-1}\right) = E\left(z_t\right) u_{1t}\left(\theta_0\right) \\
&= 0. \qquad (21)
\end{aligned}
$$

Since $\{s_{1t}, I_t\}$ is an adapted stochastic sequence the result in (21) implies that $\{s_{1t}, I_t\}$ is a martingale difference sequence according to Definition 3.75 in White (1984). Further, notice that

$$
V_{1T}^2\left(\theta_0\right) = \sum_{t=1}^{T} E\left(s_{1t}^2\left(\theta_0\right)|I_{t-1}\right) = \sum_{t=1}^{T} E\left(z_t^2\right) u_{1t}^2\left(\theta_0\right) = \sum_{t=1}^{T} u_{1t}^2\left(\theta_0\right).
$$

Hence,

$$
E(V_{1T}^2\left(\theta_0\right)) = \sum_{t=1}^{T} E\left(u_{1t}^2\left(\theta_0\right)\right) = T\overline{m}_{11}.
$$

Furthermore, according to Lemma A we have that

$$
\frac{1}{T}\sum_{t=1}^{T} u_{1t}^2\left(\theta_0\right) \xrightarrow{p} \overline{m}_{11},
$$

implying that

$$
\frac{1}{T}V_{1T}^2\left(\theta_0\right) \xrightarrow{p} \overline{m}_{11}.
$$

From this we see that

$$
\left(V_{1T}^2\left(\theta_0\right)\right)\left(E(V_{1T}^2\left(\theta_0\right))\right)^{-1} \xrightarrow{p} 1. \qquad (22)
$$

Importantly, the result given by equation (22) corresponds to Condition (1), page 60 in Brown (1971)[4].

Finally, we need to prove that the Lindeberg type condition, which is Condition (2) in Brown (1971). In particular, we need to show that

$$
\left(E(V_{1T}^2\left(\theta_0\right))\right)^{-1}\sum_{t=1}^{T} E\left(s_{1t}^2\left(\theta_0\right) 1\left\{\left|s_{1t}\left(\theta_0\right)\right| > \epsilon\sqrt{E(V_{1T}^2\left(\theta_0\right))}\right\}\right) \xrightarrow{p} 0,
$$

for all $\epsilon > 0$. By inserting the expression for s_{1t}^2 and $E(V_{1T}^2\left(\theta_0\right))$ we get

$$
\begin{aligned}
& \lim_{T\to\infty}\frac{1}{T\overline{m}_{11}}\sum_{t=1}^{T} E\left(s_{1t}^2\left(\theta_0\right) 1\left\{\left|s_{1t}\left(\theta_0\right)\right| > \epsilon\sqrt{T\overline{m}_{11}}\right\}\right) \\
= & \lim_{T\to\infty}\frac{1}{\overline{m}_1} E\left(\left(z_t^2 u_{1t}^2\left(\theta_0\right)\right) 1\left\{\left|z_t^2 u_{1t}^2\left(\theta_0\right)\right| > \sqrt{T\overline{m}_{11}}\right\}\right) \to 0,
\end{aligned}
$$

[4]Note that since $s_{1t}\left(\theta_0\right)$ is a MDS, we may use Billingsley (1961)´s Central Limit Theorem (CLT) instead of Brown (1971)´s CLT.

for all $\overline{m}_{11}$ because, from Lemma A and A2, $u_{1t}^2(\theta_0)$ and z_t^2 have finite moments and are stationary and ergodic. Consequently, the Lindeberg condition holds.

According to Theorem 2, page 60, in Brown (1971) we can therefore conclude that

$$\frac{1}{\sqrt{T\overline{m}_{11}}}\sum_{t=1}^{T} s_{1t}(\theta_0) \xrightarrow{d} N(0,1),$$

which completes the proof.

Along the same lines

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{2t}^2 \mid I_{t-1}\right) &= \frac{1}{T}\sum_{t=1}^{T}\frac{\zeta}{4}\left(\frac{1}{w_0+\alpha_0\left(|y_{t-1}|-\gamma_0 y_{t-1}\right)^2}\right)^2 \xrightarrow{p} \frac{\zeta}{4w_0^2} > 0,\\
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{3t}^2 \mid I_{t-1}\right) &= \frac{1}{T}\sum_{t=1}^{T}\frac{\zeta}{4}\left(\left(\frac{w_0}{\alpha_0}\right)\left(\frac{1}{w_0}-\frac{1}{\sigma_t^2(\theta_0)}\right)\right)^2 \xrightarrow{p} \frac{\zeta}{4\alpha_0^2} > 0,\\
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{4t}^2 \mid I_{t-1}\right) &= \frac{1}{T}\sum_{t=1}^{T}\zeta\left(\left(\left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}}-\gamma_0\right)}\right)\right)\left(\frac{1}{w_0}-\frac{1}{\sigma_t^2(\theta_0)}\right)\right)^2 \xrightarrow{p} \frac{\zeta}{(1-\gamma_0)^2} > 0
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{2t}^2 1\left\{|s_{2t}| > \sqrt{T}\delta\right\}\right) &\le E\left(\left(\frac{(1-z_t^2)^2}{4w_0^2}\right)1\left\{\left|\frac{(1-z_t^2)}{2w_0}\right| > \sqrt{T}\delta\right\}\right) \to 0,\\
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{3t}^2 1\left\{|s_{3t}| > \sqrt{T}\delta\right\}\right) &\le E\left(\left(\frac{(1-z_t^2)^2}{4\alpha_0^2}\right)1\left\{\left|\frac{(1-z_t^2)}{2\alpha_0}\right| > \sqrt{T}\delta\right\}\right) \to 0,\\
\frac{1}{T}\sum_{t=1}^{T} E\left(s_{4t}^2 1\left\{|s_{4t}| > \sqrt{T}\delta\right\}\right) &\le E\left(\left(\frac{(1-z_t^2)^2}{(1-\gamma_0)^2}\right)1\left\{\left|\frac{(1-z_t^2)}{(1-\gamma_0)}\right| > \sqrt{T}\delta\right\}\right) \to 0
\end{aligned}$$

for some $\delta > 0$ and as T tends to ∞.■

Proof of Proposition 1 In order to fully characterize the asymptotic distribution we need to determine the off-diagonal elements of the variance covariance matrix of the score vectors given by Λ. In particular, because $u_{1t}(\theta_0)$, $u_{2t}(\theta_0)$, $u_{3t}(\theta_0)$ and $u_{4t}(\theta_0)$ are all stationary and ergodic with finite first moments (from Lemma A) it follows straightforwardly that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} s_{1t}(\theta_0)\, s_{2t}(\theta_0) &= -\frac{1}{T}\sum_{t=1}^{T} z_t \frac{1}{2}\left(1-z_t^2\right) u_{1t}(\theta_0)\, u_{2t}(\theta_0) \xrightarrow{p} 0, \\
\frac{1}{T}\sum_{t=1}^{T} s_{1t}(\theta_0)\, s_{3t}(\theta_0) &= -\frac{1}{T}\sum_{t=1}^{T} \frac{1}{2} z_t \left(1-z_t^2\right) u_{1t}(\theta_0)\, u_{3t}(\theta_0) \xrightarrow{p} 0, \\
\frac{1}{T}\sum_{t=1}^{T} s_{1t}(\theta_0)\, s_{4t}(\theta_0) &= \frac{1}{T}\sum_{t=1}^{T} z_t \left(1-z_t^2\right) u_{1t}(\theta_0)\, u_{4t}(\theta_0) \xrightarrow{p} 0, \\
\frac{1}{T}\sum_{t=1}^{T} s_{2t}(\theta_0)\, s_{3t}(\theta_0) &= \frac{1}{T}\sum_{t=1}^{T} \frac{1}{4}\left(1-z_t^2\right)^2 u_{2t}(\theta_0)\, u_{3t}(\theta_0) \xrightarrow{p} \frac{1}{4}\zeta\overline{m}_{23}, \\
\frac{1}{T}\sum_{t=1}^{T} s_{2t}(\theta_0)\, s_{4t}(\theta_0) &= -\frac{1}{T}\sum_{t=1}^{T} \frac{1}{2}\left(1-z_t^2\right)^2 u_{2t}(\theta_0)\, u_{4t}(\theta_0) \xrightarrow{p} -\frac{1}{2}\zeta\overline{m}_{24}, \\
\frac{1}{T}\sum_{t=1}^{T} s_{3t}(\theta_0)\, s_{4t}(\theta_0) &= -\frac{1}{T}\sum_{t=1}^{T} \frac{1}{2}\left(1-z_t^2\right)^2 u_{3t}(\theta_0)\, u_{4t}(\theta_0) \xrightarrow{p} -\frac{1}{2}\zeta\overline{m}_{34}.
\end{aligned}$$

Since all the elements in the score vector are asymptotically normal (see Lemma B), we also follow Lemma A.3 in Pedersen and Rahbek (2019) to show that all linear combinations of the components have Gaussian asymptotic distributions by defining the sequence $c'\frac{\partial}{\partial\theta}l_T(\theta_0)\, c$ for any vector c with the same dimension as θ_0, and the result follows directly from application of the Cramer-Wold device, see for example Proposition 5.1 in White (1984), which completes the proof.■

Proposition 2 Let $u_{jt}(\theta_0)$ be defined as in Theorem 1. Under Assumptions A and B, the observed information evaluated at $\theta = \theta_0$ converges in probability, i.e.,

$$-\frac{1}{T}\frac{\partial^2}{\partial\theta\partial\theta'}l_T(\theta_0) \xrightarrow{p} \Omega,$$

where

$$\Omega = \begin{bmatrix} \overline{m}_{11} & 0 & 0 & 0 \\ 0 & \frac{1}{2}\overline{m}_{22} & \frac{1}{2}\overline{m}_{23} & -\overline{m}_{24} \\ 0 & \frac{1}{2}\overline{m}_{23} & \frac{1}{2}\overline{m}_{33} & -\overline{m}_{34} \\ 0 & -\overline{m}_{24} & -\overline{m}_{34} & 2\overline{m}_{44} \end{bmatrix} > 0,$$

and $\overline{m}_{ij} = E\left(u_{it}(\theta_0)\, u_{jt}(\theta_0)\right)$ for $i = 1,2,3,4$ and $j = 1,2,3,4$.

Proof of Proposition 2 We start with the second order derivative of γ since it is one of the most complicated ones. Recall from Result 2 (see the supplementary appendix) that

$$-\frac{1}{T}\frac{\partial^2}{\partial\gamma^2}l_T(\theta_0) = -\frac{2}{T}\sum_{t=1}^{T}\left(1-2z_t^2\right)u_{4t}^2(\theta_0) + \frac{1}{T}\sum_{t=1}^{T}\left(1-z_t^2\right)\frac{\alpha_0 y_{t-1}^2}{\sigma_t^2(\theta_0)}.$$

Since z_t^2 and $u_{1t}^2(\theta_0)$ are independent, the first term on the right hand side converges to $2\overline{m}_{44}$ by Lemma A. Furthermore, since $\frac{\alpha_0 y_{t-1}^2}{\sigma_t^2(\theta_0)}$ has bounded moments, it is ergodic and stationary and since $E(1-z_t^2)=0$, it follows from the ergodic theorem that the last term on the right hand side converges in probability to zero. Therefore, the result follows. Also, $-\frac{1}{T}\frac{\partial^2}{(\partial\phi\partial w)}l_T(\theta_0) \xrightarrow{p} 0$, $-\frac{1}{T}\frac{\partial^2}{(\partial\phi\partial\alpha)}l_T(\theta_0) \xrightarrow{p} 0$, $-\frac{1}{T}\frac{\partial^2}{(\partial\phi\partial\gamma)}l_T(\theta_0) \xrightarrow{p} 0$. Using identical arguments we find

$$-\frac{1}{T}\frac{\partial^2}{\partial\phi^2}l_T(\theta_0) = \frac{1}{T}\sum_{t=1}^{T} u_{1t}^2(\theta_0) \xrightarrow{p} \overline{m}_{11},$$

$$-\frac{1}{T}\frac{\partial^2}{\partial w^2}l_T(\theta_0) = -\frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\left(1-2z_t^2\right)u_{2t}^2(\theta_0) \xrightarrow{p} \frac{1}{2}\overline{m}_{22},$$

$$-\frac{1}{T}\frac{\partial^2}{\partial\alpha^2}l_T(\theta_0) = -\frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\left(1-2z_t^2\right)u_{3t}^2(\theta_0) \xrightarrow{p} \frac{1}{2}\overline{m}_{33},$$

$$-\frac{1}{T}\frac{\partial^2 l_t(\theta_0)}{\partial w\partial\alpha} = -\frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\left(1-2z_t^2\right)u_{2t}(\theta_0)\,u_{3t}(\theta_0) \xrightarrow{p} \frac{1}{2}\overline{m}_{23},$$

$$-\frac{1}{T}\frac{\partial^2}{(\partial\gamma\partial w)}l_T(\theta_0) = -\frac{1}{T}\sum_{t=1}^{T}\left(-1+2z_t^2\right)u_{2t}(\theta_0)\,u_{3t}(\theta_0) \xrightarrow{p} -\overline{m}_{24},,$$

$$-\frac{1}{T}\frac{\partial^2}{(\partial\gamma\partial\alpha)}l_T(\theta_0) = -\frac{1}{T}\sum_{t=1}^{T}\left(-1+2z_t^2\right)u_{3t}(\theta_0)\,u_{4t}(\theta_0) - \frac{1}{T\alpha_0}\sum_{t=1}^{T}\left(1-z_t^2\right)u_{4t}(\theta_0) \xrightarrow{p} -\overline{m}_{34},$$

.

We proceed now to show that Λ is positive definite. Λ will be positive definite if for any non-zero column vector z with entries a, b, c and d, we show that $z^T\Lambda z > 0$. In our case

$$\begin{aligned}
z^T\Lambda z &= \zeta\begin{pmatrix} a\zeta^{-1}E\left(u_{1t}^2\right) \\ \frac{b}{2}E\left(u_{2t}^2\right)+\frac{c}{2}E\left(u_{2t}u_{3t}\right)-dE\left(u_{2t}u_{4t}\right) \\ \frac{b}{2}E\left(u_{2t}u_{3t}\right)+\frac{c}{2}E\left(u_{3t}^2\right)-dE\left(u_{3t}u_{4t}\right) \\ -bE\left(u_{2t}u_{4t}\right)-cE\left(u_{3t}u_{4t}\right)+2dE\left(u_{4t}^2\right)\end{pmatrix}^T\begin{pmatrix} a \\ b \\ c \\ d\end{pmatrix} \\
&= \zeta[a^2\zeta^{-1}E\left(u_{1t}^2\right)+\frac{b^2}{2}E\left(u_{2t}^2\right)+\frac{bc}{2}E\left(u_{2t}u_{3t}\right)-bdE\left(u_{2t}u_{4t}\right)+\frac{bc}{2}E\left(u_{2t}u_{3t}\right) \\
&\quad +\frac{c^2}{2}E\left(u_{3t}^2\right)-dcE\left(u_{3t}u_{4t}\right)-bdE\left(u_{2t}u_{4t}\right)-cdE\left(u_{3t}u_{4t}\right)+2d^2E\left(u_{4t}^2\right)] \\
&= \zeta[a^2\zeta^{-1}E\left(u_{1t}^2\right)+\frac{b^2}{2}E\left(u_{2t}^2\right)+\frac{c^2}{2}E\left(u_{3t}^2\right)+2d^2E\left(u_{4t}^2\right) \\
&\quad +bcE\left(u_{2t}u_{3t}\right)-2bdE\left(u_{2t}u_{4t}\right)-2cdE\left(u_{3t}u_{4t}\right)]
\end{aligned}$$

where we have written $u_{it}(\theta_0) = u_{it}$ for simplicity reasons. Since ζ, by Assumption A2, is always positive and larger than zero, and also $a^2\zeta^{-1}E\left(u_{1t}^2\right) > 0$, then, we need to show if the following term is strictly positive

$$
\begin{aligned}
& \frac{b^2}{2}E\left(u_{2t}^2\right)+\frac{c^2}{2}E\left(u_{3t}^2\right)+2d^2E\left(u_{4t}^2\right)+bcE\left(u_{2t}u_{3t}\right)-2bdE\left(u_{2t}u_{4t}\right)-2cdE\left(u_{3t}u_{4t}\right) \\
= \; & E\left(\sqrt{2}du_{4t}-\frac{b}{\sqrt{2}}u_{2t}-\frac{c}{\sqrt{2}}u_{3t}\right)^2>0.
\end{aligned}
$$

Finally notice that since $\Omega=2\Lambda\zeta^{-1}$, then $\Omega>0$. This completes the proof of Proposition 2.■

Proposition 3 Define the lower and upper values for each parameter in θ_0 as $\gamma_L<\gamma_0<\gamma_U, w_L<w_0<w_U$, $\phi_L<\phi_0<\phi_U$ and $\alpha_L<\alpha_0<\alpha_U$, respectively and the neighborhood $N\left(\theta_0\right)$ around θ_0 as

$$
N\left(\theta_0\right)=\left\{\theta\backslash\gamma_L\leq\gamma\leq\gamma_U, w_L\leq w\leq w_U,\ \phi_L<\phi_0<\phi_U \text{ and } \alpha_L\leq\alpha\leq\alpha_U\right\}.
$$

Under Assumption A, there exists a neighborhood $N\left(\theta_0\right)$ for which for $i,j,k=1,2,3,4$

$$
\sup_{\theta\in N(\theta_0)}\left|\frac{1}{T}\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_k}l_T\left(\theta\right)\right|\leq\frac{1}{T}\sum_{t=1}^{T}w_{ijkt},
$$

where w_{ijkt} is stationary. Furthermore $\frac{1}{T}\sum_{t=1}^{T}w_{ijkt}\xrightarrow{a.s.}E\left(w_{ijkt}\right)<\infty$ for $\forall ijk$.

Proof of Proposition 3 Let us start from the components of $\left|\frac{1}{T}\frac{\partial^3}{\partial\gamma^3}l_T\left(\theta\right)\right|$ defined in Result 3 (see the supplementary appendix) since this is one of the most complicated third order derivatives. Part I (which is also defined in Result 3) can be written as

$$
\begin{aligned}
& \left|\frac{6}{T}\sum_{t=1}^{T}\left(1-2\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha^2y_{t-1}^3\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)}{\left(w+\alpha\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)^2\right)^2}\right| \\
\leq & \left|\frac{6}{T}\sum_{t=1}^{T}\left(1+2\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha^2y_{t-1}^3\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)}{\alpha^2\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)^4}\right| \\
\leq & \left|\frac{6}{T}\sum_{t=1}^{T}\left(1+2\frac{w+\alpha\left(\left|y_{t-1}\right|-\gamma y_{t-1}\right)^2}{w_0+\alpha_0\left(\left|y_{t-1}\right|-\gamma_0 y_{t-1}\right)^2}z_t^2\right)\frac{1}{\left(1-\gamma_L\right)^3}\right| \\
\leq & \left|\frac{6}{T}\sum_{t=1}^{T}\left(1+2\left(\frac{w_U}{w_L}+\frac{\alpha_U\left(1-\gamma_L\right)^2}{\alpha_L\left(1-\gamma_U\right)^2}\right)z_t^2\right)\frac{1}{\left(1-\gamma_L\right)^3}\right|,
\end{aligned}
$$

where the result follows by the law of large numbers (see Jensen and Rahbek (2004a), Lemma 5). We have made use of the definition of z_t in (1) for $s=1$ following similar arguments as in Jensen and Rahbek (2004a, 2004b). Part II follows the same argument.

Along the same lines for $\left|\frac{1}{T}\frac{\partial^3}{\partial\alpha^3}l_T(\theta)\right|$

$$\begin{aligned}
\left|\frac{1}{T}\frac{\partial^3}{\partial\alpha^3}l_T(\theta)\right| &= \left|\frac{1}{T}\sum_{t=1}^{T}\left(3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}-1\right)\frac{(|y_{t-1}|-\gamma y_{t-1})^6}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3}\right| \\
&\leq \left|\frac{1}{T}\sum_{t=1}^{T}\left(3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}+1\right)\right|\frac{(1-\gamma_L)^6}{\alpha_L^3(1-\gamma_U)^6} \\
&\leq \frac{1}{T}\sum_{t=1}^{T}\left(3\left(\frac{w_U}{w_L}+\frac{\alpha_U(1-\gamma_L)^2}{\alpha_L(1-\gamma_U)^2}\right)z_t^2+1\right)\frac{(1-\gamma_L)^6}{\alpha_L^3(1-\gamma_U)^6}.
\end{aligned}$$

The rest of the cases follow directly using the same argument. This completes the proof of Proposition 3.■

Proof of Theorem 1 Given the conditions provided by Propositions 1 - 3, Theorem 1 follows from Lumsdaine (1996, pages 593-595, Theorem 3), the ergodic theorem and Lemma 1 in Jensen and Rahbek (2004b, pages 1206-1207, based in Lehmann (1999)).■

References

Alsmeyer , G. (2016), On the Stationarity Tail Index of Iterated Random Lipschitz Functions, *Stochastic Processes and their Applications* 126, 209-233.

Basrak, B., R. David and T. Mikosch (2002), Regular Variation of GARCH Processes, *Stochastic Processes and Their Applications* 99, 95-115.

Batten, J. A., H. Kinateder, P. G. Szilagyi and N. F. Wagner (2019), Liquidity, surprise volume and return premia in the oil market, *Energy Economics* 77, 93-104.

Bedoui, R., S. Braiek, K. Guesmi and J. Chevallier (2019), On the conditional dependence structure between oil, gold and USD exchange rates: Nested copula based GJR-GARCH model, *Energy Economics* 80, 876-889.

Billingsley, P. (1961), The Lindeberg-Lévy Theorem for Martingales, *Proc. Amer. Math. Soc.* 12, 788-792.

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics* 31, 307-327.

Bollerslev, T. (2010), Glossary to ARCH (GARCH), in Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle (eds. T. Bollerslev, J. R. Russell and M. W. Watson), Chapter 8, 137-163. Oxford, UK.: Oxford University Press.

Borkovec, M. (2000), Extremal behavior of the autoregressive process with ARCH(1) errors, *Stochastic Processes and Their Applications* 85, 189-207.

Borkovec, M. and C. Klüppelberg (2001), The Tail of the Stationary Distribution of An Autoregressive Process with ARCH(1) Errors, *The Annals of Applied Probability* 11, 4, 1220-1241.

Box, G. E. P. and D. A. Pierce (1970), Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American Statistical Association* 65, 1509–1526.

Brown, B. M. (1971), Martingale Central Limit Theorems, *Annals of Mathematical Statistics* 42, 59-66.

Cavaliere, G., H. B. Nielsen and A. Rahbek (2020), Bootstrapping Noncausal Autoregressions: With Applications to Explosive Bubble Modeling, *Journal of Business and Economic Statistics* 38, 1, 55-67.

Caporin, M. and M. Costola (2019), Asymmetry and leverage in GARCH models: a News Impact Curve perspective, *Applied Economics* 51, 31, 3345-3364.

Christensen, B.-J., C. M. Dahl and E. M. Iglesias (2012), Semiparametric Inference in a GARCH-in-Mean Model, *Journal of Econometrics* 167, 2, 458-472.

Dahl, C. M. and E. M. Iglesias (2022), The Tail Behaviour due to the Presence of the Risk Premium in AR-GARCH-in-Mean, GARCH-AR and Double-Autoregressive-in-Mean Models, *Journal of Financial Econometrics* 20, 1, 139-159.

Ding, Z., C. W. J. Granger and R. F. Engle (1993), A Long Memory Property of Stock Market Returns and a New Models, *Journal of Empirical Finance* 1, 1, 83-106.

Embrechts, P., C. Klüppelberg and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

Engle, R. F. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987-1007.

Feigin, P. D. and R. L. Tweedie (1985), Random coefficient autoregressive processes: a markov chain analysis of stationarity and finiteness of moments, *Journal of Time Series Analysis* 6, 1–14.

Geweke, J. (1986), Modelling the Persistence of Conditional Variance: A Comment, *Econometric Reviews* 5, 57–61.

Glosten, L. R., R. Jagannathan and D. E. Runkle (1993), On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks, *Journal of Finance* 48, 1779-1801.

Gouriéroux, C. and J.-M. Zakoïan (2017), Local Explosion Modelling by Non-Causal Process, *Journal of the Royal Statistical Society, Series B* 79, 3, 737-756.

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

Guo, S., S. Ling and K. Zhu (2014), Factor Double Autoregressive Models with Application to Simultaneous Causality Testing, *Journal of Statistical Planning and Inference* 148, 82-94.

Iglesias, E.M. and D. Rivera-Alonso (2022), Brent and WTI oil prices volatility during major crises and Covid-19, *Journal of Petroleum Science and Engineering* 211, article 110182.

Jensen, S. T. and A. Rahbek (2004a), Asymptotic Normality of the QML Estimator of ARCH in the Nonstationary Case, *Econometrica* 72, 2, 641-646.

Jensen, S. T. and A. Rahbek (2004b), Asymptotic Inference for Nonstationary GARCH, *Econometric Theory* 20, 6, 1203-1226.

Jiang, J., X. Jiang and X. Song (2014), Weighted composite quantile regression estimation of DTARCH models, *The Econometrics Journal* 17, 1-23.

Jiang Y., C. Jiang, H. Nie and B. Mo (2019), The Time-Varying Linkages between Global Oil Market and China's Commodity Sectors: Evidence from DCC-GJR-GARCH Analyses, *Energy* 166, 577-586.

Klüppelberg, C., R. A. Maller, M. Van De Vyver and D. Wee (2002), Testing for Reduction to Random Walk in Autoregressive Conditional Heteroskedasticity Models, *The Econometrics Journal* 5, 2, 387-416.

Klüppelberg, C. and S. Pergamenchtchivkov (2004), The Tail of the Stationary Distribution of A Random Coefficient AR(q) Model, *The Annals of Applied Probability* 14, 2, 971-1005.

Lehmann, E. L. (1999), *Elements of Large Sample Theory*, New York: Springer Verlag.

Li, D., S. Guo and K. Zhu (2019), Double AR Model Without Intercept: An Alternative to Modeling Nonstationarity and Heteroscedasticity, *Econometric Reviews* 38, 3, 319-331.

Li, C. W. and W. K. Li (1996). On a double-threshold autoregressive heteroscedastic time series model, *Journal of Applied Econometrics* 11, 253–74.

Li, D., S. Ling and J.-M. Zakoïan (2015), Asymptotic inference in multiple-threshold double autoregressive models, *Journal of Econometrics* 189, 2, 415-427.

Li, D., S. Ling and R. Zhang (2016), On a Threshold Double Autoregressive Model, *Journal of Business and Economic Statistics* 34, 1, 68-80.

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

Ling, S. (2004), Estimation and Testing Stationarity for Double-Autoregressive Models, *Journal of the Royal Statistical Society Series B* 66, 1, 63-78.

Ling, S. (2007), A Double AR(p) Model: Structure and Estimation, *Statistica Sinica* 17, 161-175.

Ling, S. and D. Li (2008): Asymptotic Inference For a Nonstationary Double AR(1) Model, *Biometrika*, 95, 257–263.

Ling, S. and M. McAleer (2002), Stationarity and the Existence of Moments of a Family of GARCH Processes, *Journal of Econometrics* 106, 109-117.

Liu, F., D. Li and X. Kang (2018), Sample path properties of an explosive double autoregressive model, *Econometric Reviews* 37, 5, 484-490.

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

Lumsdaine, R. L. (1996), Asymptotic Properties of the Maximum Likelihood Estimator in GARCH(1,1) and IGARCH(1,1) Models, *Econometrica* 64, 3, 575-596.

McAleer, M., F. Chan and D. Marinova (2007), An Econometric Analysis of Asymmetric Volatility: Theory and Application to Patents, *Journal of Econometrics* 139, 259-284.

McAleer, M., S. Hoti and F. Chan (2009), Structure and Asymptotic Theory for Multivariate Asymmetric Conditional Volatility, *Econometric Reviews* 28, 5, 422-440.

Milhøj, A. (1987), A Multiplicative Parameterization of ARCH Models. *Research Report* 101, University of Copenhagen: Institute of Statistics.

Nelson, D. B. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.

Sampid, M.G., H. M. Hasim and H. Dai (2018), Refining value-at-risk estimates using a Bayesian Markov-switching GJR-GARCH copula-EVT model, *PLoS ONE* 13, 6: e0198753.

Pantula, S. (1986). Modelling the Persistence of Conditional Variance: A Comment, *Econometric Reviews* 5, 71–73.

Pedersen, R. and A. Rahbek (2019), Testing GARCH-X type models, *Econometric Theory*, 35, 5, 1012-1047.

Tan, S. and Q. Zhu (2022), Asymmetric linear double autoregression, *Journal of Time Series Analysis* 43, 3, 371-388.

Tan, S. and Q. Zhu (2023), On dual-asymmetry linear double AR models, *Statistics and Its Interface* 16, 3-16.

Tweedie, R. L. (1983), Criteria for rates of convergence of Markov chains, with application to queueing and storage theory. In J. F. C. Kingman and G. E. H. Reuter (Eds.), *Probability, Statistics and Analysis*, pp. 260–276. Cambridge: Cambridge University Press.

Vuong, Q. H. (1989), Likelihood Ratio Test for Model Selection and Non-Nested Hypotheses, *Econometrica* 57, 307-333.

White H. (1984), *Asymptotic Theory For Econometricians*, New York: Academic Press.

Zavadska, M., L. Morales and J. Coughlan (2020), Brent crude oil prices volatility during major crisis, *Finance Research Letters* 32, 101078.

Zakoïan, J.-M. (1994), Threshold Heteroskedastic Models, *Journal of Economics Dynamics and Control* 18, 5, 931-955.

Zhu, Q. and G. Li (2022), Quantile Double Autoregression, *Econometric Theory* 38, 793-839.

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

Zhu, H., X. Zhang, X. Liang and Y. Li (2017), On a Vector Double Autoregressive Model, *Statistics and Probability Letters* 129, 86-95.

Zhu, Q., Y. Zheng, and G. Li (2018), Linear double autoregression, *Journal of Econometrics* 207, 1, 162–174.

1 Supplementary Technical Appendix for "Asymptotic inference for a sign-double autoregressive (SDAR) model of order one", by Emma M. Iglesias.

We collect now the first, second and third order derivatives of the loglikelihood function.

Result 1

The first order derivatives are given by

$$
\begin{aligned}
(1)\ \frac{\partial}{\partial \phi} l_T(\theta) &= \sum_{t=1}^{T} s_{1t}(\theta) = \sum_{t=1}^{T} \left(\frac{(y_t - \phi y_{t-1})\, y_{t-1}}{w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2} \right), \\
(2)\ \frac{\partial}{\partial w} l_T(\theta) &= \sum_{t=1}^{T} s_{2t}(\theta) = -\sum_{t=1}^{T} \frac{1}{2} \left(1 - \frac{(y_t - \phi y_{t-1})^2}{w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2} \right) \frac{1}{\left(w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2\right)}, \\
(3)\ \frac{\partial}{\partial \alpha} l_T(\theta) &= \sum_{t=1}^{T} s_{3t}(\theta) = -\sum_{t=1}^{T} \frac{1}{2} \left(1 - \frac{(y_t - \phi y_{t-1})^2}{w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2} \right) \frac{\left(|y_{t-1}| - \gamma y_{t-1}\right)^2}{\left(w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2\right)}, \\
(4)\ \frac{\partial}{\partial \gamma} l_T(\theta) &= \sum_{t=1}^{T} s_{4t}(\theta) = \sum_{t=1}^{T} \left(1 - \frac{(y_t - \phi y_{t-1})^2}{w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2} \right) \frac{\alpha \left(|y_{t-1}| - \gamma y_{t-1}\right) y_{t-1}}{\left(w + \alpha \left(|y_{t-1}| - \gamma y_{t-1}\right)^2\right)}.
\end{aligned}
$$

In particular, $\sigma_t^2(\theta_0) = w_0 + \alpha_0 \left(|y_{t-1}| - \gamma_0 y_{t-1}\right)^2$ and

$$
\begin{aligned}
s_{1t}(\theta_0) &= z_t \frac{y_{t-1}}{\sigma_t(\theta_0)}, \\
s_{2t}(\theta_0) &= -\frac{1}{2}\left(1 - z_t^2\right) \frac{1}{\sigma_t^2(\theta_0)}, \\
s_{3t}(\theta_0) &= -\frac{1}{2}\left(1 - z_t^2\right) \left(\frac{w_0}{\alpha_0}\right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right), \\
s_{4t}(\theta_0) &= \left(1 - z_t^2\right) \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right)} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right).
\end{aligned}
$$

Result 2

The second order derivatives evaluated at evaluated at $\theta = \theta_0$ are given by

$$\frac{\partial^2}{\partial \phi^2} l_T(\theta_0) = \sum_{t=1}^{T} \left(\frac{-y_{t-1}^2}{w_0 + \alpha_0 \left(|y_{t-1}| - \gamma_0 y_{t-1} \right)^2} \right),$$

$$\frac{\partial^2}{\partial w^2} l_T(\theta_0) = \frac{1}{2} \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \left(\frac{1}{\sigma_t^2(\theta_0)} \right)^2,$$

$$\frac{\partial^2}{\partial \alpha^2} l_T(\theta_0) = \frac{1}{2} \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \left(\frac{w_0}{\alpha_0} \right)^2 \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right)^2,$$

$$\frac{\partial^2}{\partial \gamma^2} l_T(\theta_0) = \sum_{t=1}^{T} \left(1 - z_t^2\right) \frac{-\alpha y_{t-1}^2}{\sigma_t^2(\theta_0)} + 2 \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right)} \right)^2 \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right)^2,$$

$$\frac{\partial^2}{\partial \phi \partial w} l_T(\theta_0) = - \sum_{t=1}^{T} z_t \frac{y_{t-1}}{\sigma_t^3(\theta_0)},$$

$$\frac{\partial^2}{\partial \phi \partial \alpha} l_T(\theta_0) = - \sum_{t=1}^{T} z_t \frac{y_{t-1}}{\sigma_t(\theta_0)} \left(\frac{w_0}{\alpha_0} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right),$$

$$\frac{\partial^2}{\partial \phi \partial \gamma} l_T(\theta_0) = -2 \sum_{t=1}^{T} z_t \frac{y_{t-1}}{\sigma_t(\theta_0)} \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right)} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right),$$

$$\frac{\partial^2}{\partial w \partial \alpha} l_T(\theta_0) = \frac{1}{2} \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \frac{1}{\sigma_t^2(\theta_0)} \left(\frac{w_0}{\alpha_0} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right),$$

$$\frac{\partial^2}{\partial w \partial \gamma} l_T(\theta_0) = - \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \frac{1}{\sigma_t^2(\theta_0)} \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right)} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right),$$

$$\begin{aligned} \frac{\partial^2}{\partial \alpha \partial \gamma} l_T(\theta_0) = & \sum_{t=1}^{T} \left(1 - z_t^2\right) \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right) \alpha_0} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right) \\ & - \sum_{t=1}^{T} \left(1 - 2z_t^2\right) \left(\frac{w_0}{\alpha_0} \right) \left(\frac{w_0}{\left(\frac{|y_{t-1}|}{y_{t-1}} - \gamma_0 \right) \alpha_0} \right) \left(\frac{1}{w_0} - \frac{1}{\sigma_t^2(\theta_0)} \right)^2. \end{aligned}$$

Result 3

The third order derivatives are given by

$$\begin{aligned} \frac{\partial^3}{\partial \phi^3} l_T(\theta) &= 0, \\ \frac{\partial^3}{\partial w^3} l_T(\theta) &= - \sum_{t=1}^{T} \left(1 - 3 \frac{\left(y_t - \phi y_{t-1}\right)^2}{w + \alpha \left(|y_{t-1}| - \gamma y_{t-1} \right)^2} \right) \frac{1}{\left(w + \alpha \left(|y_{t-1}| - \gamma y_{t-1} \right)^2 \right)^3}, \end{aligned}$$

$$
\begin{aligned}
\frac{\partial^3}{\partial\alpha^3} l_T(\theta) &= -\sum_{t=1}^{T}\left(1-3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{(|y_{t-1}|-\gamma y_{t-1})^6}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},\\
\frac{\partial^3}{\partial\gamma^3} l_T(\theta) &= -6\sum_{t=1}^{T}\left(1-2\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{\alpha^2 y_{t-1}^3(|y_{t-1}|-\gamma y_{t-1})}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2}\\
&\quad +8\sum_{t=1}^{T}\left(1-3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{\alpha^3 y_{t-1}^3(|y_{t-1}|-\gamma y_{t-1})^3}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3}\\
&= I+II,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial\phi^2\partial w} l_T(\theta) &= \frac{1}{2}\sum_{t=1}^{T}\frac{y_{t-1}^2}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2},\\
\frac{\partial^3}{\partial\phi^2\partial\alpha} l_T(\theta) &= \sum_{t=1}^{T}\frac{y_{t-1}^2(|y_{t-1}|-\gamma y_{t-1})^2}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2},\\
\frac{\partial^3}{\partial\phi^2\partial\gamma} l_T(\theta) &= -\sum_{t=1}^{T}\frac{2\alpha y_{t-1}^3(|y_{t-1}|-\gamma y_{t-1})}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial\alpha\partial w^2} l_T(\theta) &= -\sum_{t=1}^{T}\left(1-3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{(|y_{t-1}|-\gamma y_{t-1})^2}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},\\
\frac{\partial^3}{\partial\gamma\partial w^2} l_T(\theta) &= -2\sum_{t=1}^{T}\left(1-3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{(|y_{t-1}|-\gamma y_{t-1})\,y_{t-1}}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},\\
\frac{\partial^3}{\partial\gamma\partial\alpha^2} l_T(\theta) &= -2\sum_{t=1}^{T}\left(1-2\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{(|y_{t-1}|-\gamma y_{t-1})^3\,y_{t-1}}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2}+\\
&\quad +2\sum_{t=1}^{T}\left(1-3\frac{(y_t-\phi y_{t-1})^2}{w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2}\right)\frac{\alpha y_{t-1}(|y_{t-1}|-\gamma y_{t-1})^5}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial\phi\partial w^2} l_T(\theta) &= \sum_{t=1}^{T}\frac{2(y_t-\phi y_{t-1})\,y_{t-1}}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},\\
\frac{\partial^3}{\partial\phi\partial\alpha^2} l_T(\theta) &= \sum_{t=1}^{T}\frac{2(y_t-\phi y_{t-1})\,y_{t-1}(|y_{t-1}|-\gamma y_{t-1})^4}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3},\\
\frac{\partial^3}{\partial\phi\partial\gamma^2} l_T(\theta) &= -\sum_{t=1}^{T}\Big(\frac{2\alpha(y_t-\phi y_{t-1})\,y_{t-1}^3}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^2}+\frac{8\alpha^2(y_t-\phi y_{t-1})\,y_{t-1}^3(|y_{t-1}|-\gamma y_{t-1})^2}{\left(w+\alpha(|y_{t-1}|-\gamma y_{t-1})^2\right)^3}\Big),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial\gamma\partial w\partial\phi} l_T(\theta) &= \sum_{t=1}^{T} \frac{4\alpha\left(y_t-\phi y_{t-1}\right) y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}, \\
\frac{\partial^3}{\partial\alpha\partial w\partial\phi} l_T(\theta) &= \sum_{t=1}^{T} \frac{2\left(y_t-\phi y_{t-1}\right) y_{t-1}\left(|y_{t-1}|-\gamma y_{t-1}\right)}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}, \\
\frac{\partial^3}{\partial\gamma\partial\alpha\partial\phi} l_T(\theta) &= \sum_{t=1}^{T} \left(\frac{2\left(y_t-\phi y_{t-1}\right) y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^2} - \frac{4\alpha\left(y_t-\phi y_{t-1}\right) y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)^3}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}\right), \\
\frac{\partial^3}{\partial\gamma\partial\alpha\partial w} l_T(\theta) &= -\sum_{t=1}^{T}\left(1-2\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\left(|y_{t-1}|-\gamma y_{t-1}\right) y_{t-1}}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^2} \\
&\quad +2\sum_{t=1}^{T}\left(1-3\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\left(|y_{t-1}|-\gamma y_{t-1}\right)^3 \alpha y_{t-1}}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^3}{\partial w\partial\alpha^2} l_T(\theta) &= -\sum_{t=1}^{T}\left(1-3\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\left(|y_{t-1}|-\gamma y_{t-1}\right)^4}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}, \\
\frac{\partial^3}{\partial\alpha\partial\gamma^2} l_T(\theta) &= -\sum_{t=1}^{T}\left(1-\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{y_{t-1}^2}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)} \\
&\quad +5\sum_{t=1}^{T}\left(1-2\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^2} \\
&\quad -4\sum_{t=1}^{T}\left(1-3\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha^2 y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)^4}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}, \\
\frac{\partial^3}{\partial w\partial\gamma^2} l_T(\theta) &= \sum_{t=1}^{T}\left(1-2\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha y_{t-1}^2}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^2} \\
&\quad -4\sum_{t=1}^{T}\left(1-3\frac{\left(y_t-\phi y_{t-1}\right)^2}{w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}\right)\frac{\alpha^2 y_{t-1}^2\left(|y_{t-1}|-\gamma y_{t-1}\right)^2}{\left(w+\alpha\left(|y_{t-1}|-\gamma y_{t-1}\right)^2\right)^3}.
\end{aligned}
$$