


Article

Fisher's z Distribution-Based Mixture Autoregressive Model

Arifatus Solikhah ^{1,2} , Heri Kuswanto ^{1,*}, Nur Iriawan ¹ and Kartika Fithriasari ¹

¹ Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia; arifatus.solikhah@bps.go.id (A.S.); nur_i@statistika.its.ac.id (N.I.); kartika_f@statistika.its.ac.id (K.F.)

² Badan Pusat Statistik (BPS—Statistics Indonesia), Jakarta 10710, Indonesia

* Correspondence: heri_k@statistika.its.ac.id

Abstract: We generalize the Gaussian Mixture Autoregressive (GMAR) model to the Fisher's z Mixture Autoregressive (ZMAR) model for modeling nonlinear time series. The model consists of a mixture of K -component Fisher's z autoregressive models with the mixing proportions changing over time. This model can capture time series with both heteroskedasticity and multimodal conditional distribution, using Fisher's z distribution as an innovation in the MAR model. The ZMAR model is classified as nonlinearity in the level (or mode) model because the mode of the Fisher's z distribution is stable in its location parameter, whether symmetric or asymmetric. Using the Markov Chain Monte Carlo (MCMC) algorithm, e.g., the No-U-Turn Sampler (NUTS), we conducted a simulation study to investigate the model performance compared to the GMAR model and Student t Mixture Autoregressive (TMAR) model. The models are applied to the daily IBM stock prices and the monthly Brent crude oil prices. The results show that the proposed model outperforms the existing ones, as indicated by the Pareto-Smoothed Important Sampling Leave-One-Out cross-validation (PSIS-LOO) minimum criterion.



Citation: Solikhah, Arifatus, Heri Kuswanto, Nur Iriawan, and Kartika Fithriasari. 2021. Fisher's z Distribution-Based Mixture Autoregressive Model. *Econometrics* 9: 27. <https://doi.org/10.3390/econometrics9030027>

Academic Editor: In Choi

Received: 10 December 2020

Accepted: 23 June 2021

Published: 29 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Fisher's z distribution; mixture autoregressive model; the IBM stock prices; the Brent crude oil prices; Bayesian analysis; no-U-turn sampler; Stan program

1. Introduction

Many time series indicate non-Gaussian characteristics, such as outliers, flat stretches, bursts of activity, and change points (Le et al. 1996). Several methods have been proposed to deal with the presence of bursts and outliers such as applying robust or resistant estimation procedures (Martin and Yohai 1986) or omitting the outliers based on the use of diagnostics (Bruce and Martin 1989). Le et al. (1996) introduced a Mixture Transition Distribution (MTD) model to capture non-Gaussian and nonlinear patterns, using the Expectation–Maximization (EM) algorithm as its estimation method. The model was applied to two real datasets, i.e., the daily International Business Machines (IBM) common stock closing price from 17 May 1961 to 2 November 1962 and the series of consecutive hourly viscosity readings from a chemical process. The MTD model appears to capture the features of the data better than the Autoregressive Integrated Moving Average (ARIMA) models.

The Gaussian Mixture Transition Distribution (GMTD), which is a special form of MTD, was generalized to a Gaussian Mixture Autoregressive (GMAR) model by Wong and Li (2000). The model consists of a mixture of K Gaussian autoregressive components and is able to model time series with both heteroscedasticity and multimodal conditional distribution. It was applied to both the daily IBM common stock closing price from 17 May 1961 to 2 November 1962 and the Canadian lynx data for the period 1821–1934. The results indicated that the GMAR model was better than the GMTD, ARIMA, and Self-Exciting Threshold Autoregressive (SETAR) models.

The use of the Gaussian distribution in the GMAR model still leaves problems, because it is able to capture only short-tailed data patterns. Some methods developed to

overcome this problem include the use of distributions other than Gaussian, e.g., the Logistic Mixture Autoregressive with Exogenous Variables (LMARX) model (Wong and Li 2001), Student t -Mixture Autoregressive (TMAR) model (Wong et al. 2009), Laplace MAR model (Nguyen et al. 2016), and a mixture of autoregressive models based on the scale mixture of skew-normal distributions (SMSN-MAR) model (Maleki et al. 2020). Maleki et al. (2020) proposed the finite mixtures of autoregressive processes assuming that the distribution of innovations belongs to the class of Scale Mixture of Skew-Normal (SMSN) distributions. This distribution innovation can be employed in data modeling that has outliers, asymmetry, and fat tails in the distribution simultaneously. However, the SMSN distribution's mode was not stable in its location parameters (Azzalini 2014). In this paper, we propose a new MAR model called the Fisher's z Mixture Autoregressive (ZMAR) model which assumes that the distribution of innovations belongs to the Fisher's z distributions (Solikhah et al. 2021). The ZMAR model consists of a mixture of K -component Fisher's z autoregressive models, where the numbers of components are based on the number of modes in the marginal density. The Fisher's z distribution's mode is stable in its location parameters, whether it is symmetrical or skewed. Therefore, Fisher's z uses the errors in each component of the MAR model to capture the 'most likely' mode value—not the mean, median, or quantile) of the conditional distribution Y_t given the past information. The conditional mode may be a more useful summary than the conditional mean when the conditional distribution of Y_t given the past information is asymmetric. Other distributions that also have a stable mode in its location parameter are the MSNBurr distribution (Iriawan 2000; Choir et al. 2019; Pravitasari et al. 2020), the skewed Student t distribution (Fernández and Steel 1998), and the log F -distribution (Brown et al. 2002).

The Bayesian technique using Markov Chain Monte Carlo (MCMC) is proposed to estimate the model parameters. Among the algorithms in the MCMC, the Gibbs sampling (Geman and Geman 1984) and the Metropolis (Metropolis et al. 1953) algorithms are widely applied and well-known algorithms. However, these algorithms have slow convergence due to inefficiencies in the MCMC processes, especially in the case of models with many correlated parameters (Gelman et al. 2014, p. 269). Furthermore, Neal (2011) has shown that the Hamiltonian Monte Carlo (HMC) algorithm is a more efficient and robust sampler than Metropolis or Gibbs sampling for models with complex posteriors. However, the HMC suffers from a computational burden and the tuning process. The HMC can be tuned in three places (Gelman et al. 2014, p. 303), i.e., the probability distribution for the momentum variables ϕ , the step size of the leapfrog ϵ , and the number of leapfrog steps L per iteration. To overcome the challenges related to computation and tuning, the Stan program (Gelman et al. 2014, p. 307; Carpenter et al. 2015, 2017) was developed to automatically apply the HMC. Stan runs HMC using the no-U-turn sampler (NUTS) (Hoffman and Gelman 2014). Al Hakmani and Sheng (2017) used NUTS for the two-parameter mixture IRT (Mix2PL) model and discussed in more detail its performance in estimating model parameters under eight conditions, i.e., two sample sizes per class (250 and 500), two test lengths (20 and 30), and two levels of latent classes (2-class and 3-class). The results indicated that overall, NUTS performs well in retrieving model parameters. Therefore, this research applies the Bayesian method to estimate the parameters of the ZMAR model, using MCMC with the NUTS algorithm, as well as simulation studies to examine different scenarios in order to evaluate whether the proposed mixture model outperforms its counterparts. The models are applied to both the daily IBM common stock closing price from 17 May 1961 to 2 November 1962 (Box et al. 2015, p. 627) and the Brent crude oil price (World Bank 2020). For model selection, we used cross-validation Leave-One-Out (LOO) coupled with the Pareto-smoothed important sampling (PSIS), namely PSIS-LOO. This approach has very efficient computation and was stronger than the Widely Applicable Information Criterion (WAIC) (Vehtari et al. 2017).

The rest of this study is organized as follows. Section 2 describes the definition and properties of Fisher's z distribution in detail. In Section 3, we introduce the ZMAR model. Section 4 demonstrates the flexibility of the ZMAR model compared with the TMAR and

GMAR models using simulated datasets. Section 5 contains the application and comparison of the models using the daily IBM stock prices and the monthly Brent crude oil prices. The conclusion and discussion are given in Section 6.

2. Four-Parameter Fisher's z Distribution

Let Y be a random variable distributed as an F distribution with d_1 and d_2 degrees of freedom. The density of $Z = \frac{1}{2} \ln Y$ can be defined as

$$\zeta_{(d_1, d_2)}(z) = f_Z(z; d_1, d_2) = \frac{2 \left(\frac{d_2}{d_1}\right)^{\frac{1}{2}d_2} e^{-d_2 z}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) \left(1 + e^{-2z + \ln \frac{d_2}{d_1}}\right)^{(d_1+d_2)/2}}, \quad (1)$$

and the cumulative distribution function (CDF) of Z is expressed as

$$Z_{(d_1, d_2)}(z) = I_{z^*}\left(\frac{1}{2}d_2, \frac{1}{2}d_1\right) = \frac{\int_0^{z^*} t^{\frac{1}{2}d_2-1} (1-t)^{\frac{1}{2}d_1-1} dt}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)}, \quad (2)$$

where e is the exponential constant; $z^* = \frac{d_2 e^{-2z}}{d_1 + d_2 e^{-2z}}$, $I_{z^*}(\cdot)$ is the incomplete beta function ratio; and $B(\cdot)$ is the beta function, $-\infty < z < \infty$, $d_1 > 0$, $d_2 > 0$. Equations (1) and (2) are defined as a probability density function (p.d.f) and a CDF of standardized Fisher's z distribution, respectively. Let Z be a random variable distributed as a standardized Fisher's z distribution. Let μ be a location parameter, and let σ be a scale parameter. The density of $X = \sigma Z + \mu$ is (Solikhah et al. 2021)

$$f_X(x; d_1, d_2, \mu, \sigma) = \frac{2 \left(\frac{d_2}{d_1}\right)^{\frac{1}{2}d_2} e^{-d_2 \left(\frac{x-\mu}{\sigma}\right)}}{\sigma B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) \left(1 + e^{-2\left(\frac{x-\mu}{\sigma}\right) + \ln \frac{d_2}{d_1}}\right)^{(d_1+d_2)/2}}, \quad (3)$$

where $-\infty < x < \infty$, $-\infty < \mu < \infty$, and $\sigma > 0$. Equation (3) is defined as a p.d.f of Fisher's z distribution. It is denoted as $z(d_1, d_2, \mu, \sigma)$. The CDF of the Fisher's z distribution is expressed as

$$F_X(x; d_1, d_2, \mu, \sigma) = I_{x^*}\left(\frac{1}{2}d_2, \frac{1}{2}d_1\right) = \frac{1}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \int_0^{x^*} t^{\frac{1}{2}d_2-1} (1-t)^{\frac{1}{2}d_1-1} dt, \quad (4)$$

where $x^* = \frac{d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}{d_1 + d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}$. The quantile function (QF) of the Fisher's z distribution is defined as

$$x_p = \mu + \frac{\sigma}{2} \ln \frac{d_2 I_{x_p}^{-1}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)}{d_1 \left(1 - I_{x_p}^{-1}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)\right)}, \quad (5)$$

where $\left(d_2 I_{x_p}^{-1}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)\right) / \left(d_1 \left(1 - I_{x_p}^{-1}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)\right)\right)$ is the QF of the F -distribution and $I_{x_p}^{-1}(\cdot)$ is the inversion of the incomplete beta function ratio. Let $\mathcal{P}_{x_p}^{-1}(\cdot)$ be the inversion of the incomplete gamma function ratio. The QF of the Fisher's z distribution can be expressed as

$$x_p = \mu + \frac{\sigma}{2} \ln \left(\frac{d_2 \left(\mathcal{P}_{v_{1p}}^{-1}(d_1/2)\right)}{d_1 \left(\mathcal{P}_{v_{2p}}^{-1}(d_2/2)\right)} \right), \quad (6)$$

where $2 \mathcal{P}^{-1}_{v_{1p}}\left(\frac{1}{2}d_1\right)$ and $2 \mathcal{P}^{-1}_{v_{2p}}\left(\frac{1}{2}d_2\right)$ are the QF of the chi-square distribution with d_1 and d_2 degrees of freedom, respectively. The proofs of Equation (1) up to Equation (6) are postponed to Appendix A. The parameters d_1 and d_2 , known as the shape parameters, are defined for both skewness (symmetrical if $d_1 = d_2$, asymmetrical if $d_1 \neq d_2$) and fatness of the tails (large d_1 and d_2 imply thin tails). The Fisher's z distribution is also always unimodal and has the mode at $x = \mu$. Furthermore, a change in the value of the parameter μ only affects the mean of the distribution. It does not affect the variance, skewness, and kurtosis of the distribution. The detailed properties of the Fisher's z distribution are shown in Appendix B.

A useful tutorial on adding custom function to Stan is provided by [Stan Development Team \(2018\)](#) and [Annis et al. \(2017\)](#). To add a user-defined function, it is first necessary to define a block of function code. The function block must precede all other blocks of Stan code. The code for the random numbers generator function (`fisher_z_rng`) is shown in Appendix C.1, and the log probability function (`fisher_z_lpdf`) is shown in Appendix C.2. As an illustration, the p.d.f and CDF of the Fisher's z distribution with various parameter settings can be seen in Figures 1 and 2, respectively.

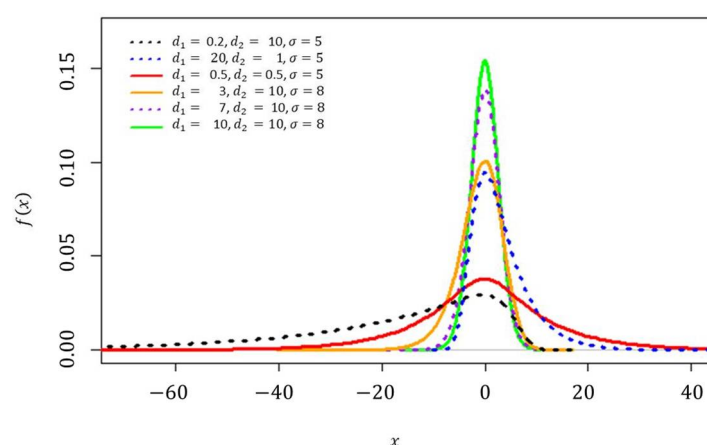


Figure 1. The p.d.f of the Fisher's z distribution when $\mu = 0$ at various choices of d_1 , d_2 , and σ .

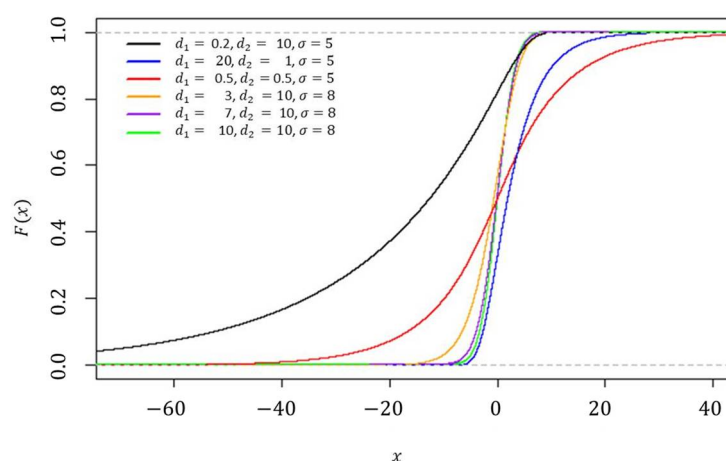


Figure 2. The CDF of the Fisher's z distribution when $\mu = 0$ at various choices of d_1 , d_2 and σ .

3. Fisher's z Mixture Autoregressive Model

3.1. Model Specification

Let y_t ; $t = 1, 2, \dots, T$ be the real-valued time series of interest; let $\mathcal{F}_{(t-1)}$ denote the information set up to time $t - 1$, let $F(y_t | \mathcal{F}_{(t-1)})$; $k = 1, 2, \dots, K$ be the conditional CDF of Y_t given the past information, evaluated at y_t ; and K is the number of components in the

ZMAR model. Let $Z_{(d_{1k}, d_{2k})}(\cdot)$ be the CDF of the standardized Fisher's z distribution with d_{1k} and d_{2k} shape parameters, given by Equation (2); let $\epsilon_{k,t}$ be a sequence of independent standardized Fisher's z random variables such that $\epsilon_{k,t}$ is independent of $\{y_{t-i}, i > 0\}$; and let σ_k is a scale parameter of the k th component. The K -component ZMAR can be defined as

$$F(y_t | \mathcal{F}_{(t-1)}) = \sum_{k=1}^K \eta_k Z_{(d_{1k}, d_{2k})} \left(\frac{y_t - \mu_{k,t}}{\sigma_k} \right), \quad (7)$$

or

$$y_t = \begin{cases} \mu_{1,t} + \sigma_1 \epsilon_{1,t}; & \text{with probability } \eta_1; \\ \mu_{2,t} + \sigma_2 \epsilon_{2,t}; & \text{with probability } \eta_2; \\ \vdots & \\ \mu_{K,t} + \sigma_K \epsilon_{K,t}; & \text{with probability } \eta_K, \end{cases} \quad (8)$$

with

$$\mu_{k,t} = \phi_{k,0} + \sum_{i=1}^{p_k} \phi_{k,i} y_{t-i}; k = 1, 2, \dots, K, \quad (9)$$

where the vector $\boldsymbol{\eta} = (\eta_1, \dots, \eta_K)$ is called the weights. $\boldsymbol{\eta}$ takes a value in the unit simplex \mathcal{E}_K , which is a subspace of $(\mathbb{R}^+)^K$, defined by the following constraint, $\eta_k > 0$ and $\sum_{k=1}^K \eta_k = 1$.

We use the abbreviation $\text{ZMAR}(K; p_1, p_2, \dots, p_K)$ for this model, with the parameter $\boldsymbol{\vartheta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_K, \boldsymbol{\eta})$; $\boldsymbol{\eta} = (\eta_k)$; $\boldsymbol{\theta}_k = (d_{1k}, d_{2k}, \sigma_k, \phi_{k,0}, \boldsymbol{\phi}_k)$; $\boldsymbol{\phi}_k = (\phi_{k,1}, \phi_{k,2}, \dots, \phi_{k,p_k})$; to each value $k \in \{1, 2, \dots, K\}$; taking values in the parameter space $\Theta_K = \Theta^K \times \mathcal{E}_K$, $\phi_{k,i}$ denotes the AR coefficient on the k th component and i th lag; $i = 1, 2, \dots, p_k$; and p_k denotes the autoregressive order of the k th component. Using the parameters $\boldsymbol{\theta}_k$, we first define the K auxiliary Fisher's z $\text{AR}(p_k)$ processes

$$f_k(y_t | \mathcal{F}_{(t-1)}) = \phi_{k,0} + \sum_{i=1}^{p_k} \phi_{k,i} y_{t-i} + \sigma_k \epsilon_{k,t}; k = 1, 2, \dots, K,$$

where the AR coefficients $\boldsymbol{\phi}_k$ are assumed to satisfy

$$1 - \sum_{i=1}^{p_k} \phi_{k,i} \mathbb{C}^i \neq 0 \text{ for } |\mathbb{C}| \leq 1; k = 1, \dots, K.$$

This condition implies that the processes $f_k(y_t | \mathcal{F}_{(t-1)})$ are stationary and that each component model in (7) or (8) satisfies the usual stationarity condition of the linear $\text{AR}(p_k)$ model.

Suppose $p = \max(p_1, p_2, \dots, p_K)$. Let a univariate time series $\mathbf{y} = (y_1, y_2, \dots, y_T)$ be influenced by a hidden discrete indicator variables $\mathbf{Q} = (Q_1, Q_2, \dots, Q_T)$, where Q_t takes values in the set $\{1, 2, \dots, K\}$. The probability of sampling from the group labeled $Q_t = k$ is equal to η_k . Suppose that the conditional density of Y_t given $Q_t = k$ is $f_k(y_t | \mathcal{F}_{(t-1)})$; ($k = 1, 2, \dots, K$). Let $\mathbf{H} = (\mathbf{H}_1', \mathbf{H}_2', \dots, \mathbf{H}_T')'$ be the unobserved random variable, where \mathbf{H}_t is a K -dimensional vector with the values $\mathbf{h} = (\mathbf{h}_1', \mathbf{h}_2', \dots, \mathbf{h}_T')'$, where $h_{k,t} = 1$ if $Q_t = k$ and $h_{k,t} = 0$, otherwise. Thus, \mathbf{H}_t is distributed according to a multinomial distribution consisting of one draw on K categories with probabilities $\eta_1, \eta_2, \dots, \eta_K$ (McLachlan and Peel 2000, p. 7); that is,

$$P(\mathbf{H}_t = \mathbf{h}_t) = \eta_1^{h_{1,t}} \eta_2^{h_{2,t}} \dots \eta_K^{h_{K,t}}. \quad (10)$$

Otherwise, this can be written as $\mathbf{H}_t \sim \text{Mult}_K(1, \boldsymbol{\eta})$; where $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_K)$. The conditional likelihood for the $\text{ZMAR}(K; p_1, p_2, \dots, p_K)$ can be formed by

$$P(y_t | \boldsymbol{\vartheta}) = \prod_{k=1}^K \prod_{t=p+1}^T [\eta_k \Delta_1 \Delta_2]^{h_{k,t}}, \quad (11)$$

where $\Delta_1 = \frac{2(d_{2k}/d_{1k})^{\frac{1}{2}} d_{2k}}{\sigma_k B(\frac{1}{2}d_{1k}, \frac{1}{2}d_{2k})}$, $\Delta_2 = \frac{\exp(-d_{2k}(a_{k,t}/\sigma_k))}{[1 + \exp(-2(a_{k,t}/\sigma_k) + \ln[d_{2k}/d_{1k}])]^{(d_{1k}+d_{2k})/2}}$, and $a_{k,t} = y_t - \mu_{k,t}$.

Let $\tau_{t,k}$ be the probability for each t -th observation, $t = (p+1), (p+2), \dots, T$, as members of the k -th component, $k = 1, 2, \dots, K$ of a mixture distribution. Suppose $\mathbf{y}^t = (y_1, y_2, \dots, y_t)$, Bayes' rule to compute the $\tau_{t,k}$ can be expressed as (Frühwirth-Schnatter 2006, p. 26)

$$P(Q_t = k | \mathbf{y}^t, \boldsymbol{\theta}) = \tau_{t,k} = \frac{\eta_k f_k(\mathbf{y}^t | \mathcal{F}_{(t-1)})}{\sum_{k=1}^K \eta_k f_k(\mathbf{y}^t | \mathcal{F}_{(t-1)})}. \quad (12)$$

Let $\zeta_{(d_{1k}, d_{2k})}(\cdot)$ be the p.d.f of the standardized Fisher's z distribution and given by Equation (1). Then, the mixing weights for the ZMAR model can be expressed as

$$\tau_{t,k} = \frac{\eta_k \sigma_k^{-1} \zeta_{(d_{1k}, d_{2k})}(a_{k,t}/\sigma_k)}{\sum_{k=1}^K \eta_k \sigma_k^{-1} \zeta_{(d_{1k}, d_{2k})}(a_{k,t}/\sigma_k)}; k = 1, 2, \dots, K.$$

Suppose $\omega_{k,t}$ and ξ_k^2 signify the conditional mean and conditional variance of the k th component, which are defined by

$$\omega_{k,t} = E[Y_t | \boldsymbol{\theta}_k] = \mu_{k,t} + \frac{\sigma_k}{2} \left(\ln \frac{d_{2k}}{d_{1k}} - \psi\left(\frac{1}{2}d_{2k}\right) + \psi\left(\frac{1}{2}d_{1k}\right) \right), \quad (13)$$

and

$$\xi_k^2 = \text{Var}[Y_t | \boldsymbol{\theta}_k] = \left(\frac{\sigma_k}{2}\right)^2 \left(\psi'\left(\frac{1}{2}d_{2k}\right) + \psi'\left(\frac{1}{2}d_{1k}\right) \right), \quad (14)$$

where $\psi(\cdot)$ is the digamma function and $\psi'(\cdot)$ is the trigamma function. The conditional mean of Y_t (Frühwirth-Schnatter 2006, p. 10) is obtained as

$$\omega_t = E[Y_t | \mathcal{F}_{(t-1)}] = \sum_{k=1}^K \eta_k \omega_{k,t}, \quad (15)$$

the conditional variance of Y_t (Frühwirth-Schnatter 2006, p. 11) is obtained as

$$\xi_t^2 = \text{Var}[Y_t | \mathcal{F}_{(t-1)}] = \sum_{k=1}^K \eta_k (\omega_{k,t}^2 + \xi_k^2) - \omega_t^2, \quad (16)$$

and the higher-order moments around the mean of Y_t (Frühwirth-Schnatter 2006, p. 11) are obtained as

$$E\left((Y - \omega_t)^3 | \mathcal{F}_{(t-1)}\right) = \sum_{k=1}^K \left((\omega_{k,t} - \omega_t)^2 + 3\xi_k^2 \right) (\omega_{k,t} - \omega_t) \eta_k, \quad (17)$$

$$E\left((Y - \omega_t)^4 | \mathcal{F}_{(t-1)}\right) = \sum_{k=1}^K \left((\omega_{k,t} - \omega_t)^4 + 6(\omega_{k,t} - \omega_t)^2 \xi_k^2 + 3\xi_k^4 \right) \eta_k. \quad (18)$$

These expressions apply to any specification of the mixing weights η_k .

3.2. Bayesian Approach for ZMAR Model

In this paper, we apply a Bayesian method to estimate the parameters $\boldsymbol{\theta}$. The Bayesian analysis requires the joint posterior density $\pi(\boldsymbol{\theta} | \mathbf{y})$, which is defined by

$$\pi(\boldsymbol{\theta} | \mathbf{y}) \propto P(y_t | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}), \quad (19)$$

where $P(y_t|\boldsymbol{\theta})$ is the conditional likelihood function given by Equation (11) and $\pi(\boldsymbol{\theta})$ is the prior of the parameter model, which is defined by

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\eta}) \prod_{k=1}^K \pi(d_{1k})\pi(d_{2k})\pi(\sigma_k)\pi(\phi_{k,0}) \prod_{i=1}^{p_k} \pi(\phi_{k,i}), \quad (20)$$

where the $\pi(\boldsymbol{\eta})$ is prior for the $\boldsymbol{\eta}$ parameter, the $\pi(d_{1k})$, $\pi(d_{2k})$, $\pi(\sigma_k)$, $\pi(\phi_{k,0})$ and $\pi(\phi_{k,i})$ are prior for the d_1 , d_2 , σ , ϕ_0 and ϕ_i parameters at the k th component; index i denotes the i th lag; $i = 1, 2, \dots, p_k$; $k = 1, 2, \dots, K$.

Various noninformative prior distributions have been suggested for the prior of AR coefficients, scale parameters, and the selection probabilities in similar models. Huerta and West (1999) analyzed and used the uniform Dirichlet distribution as the prior distribution for the latent variables related to the latent components of an autoregressive model. Gelman (2006) suggested working within the half- t family of prior distributions for variance parameters in the hierarchical modelling, which are more flexible and have better behavior near 0, compared to the inverse-gamma family. Albert and Chib (1993) used the normal distribution as the prior for the autoregressive coefficient in the Markov switching autoregressive model. Based on the findings of the previous studies, we take the singly truncated Student t distribution (positive values only) (Kim 2008) for the priors of the d_{1k} , d_{2k} , and σ_k , with the degrees of freedoms v_{1k} , v_{2k} , v_{3k} , the location parameters m_{1k} , m_{2k} , m_{3k} , and the scale parameters s_{1k}^2 , s_{2k}^2 , s_{3k}^2 , respectively. Therefore, it can be written as $d_{1k} \sim t_{v_{1k}}(m_{1k}, s_{1k}^2)I(0, \infty)$, $d_{2k} \sim t_{v_{2k}}(m_{2k}, s_{2k}^2)I(0, \infty)$, and $\sigma_k \sim t_{v_{3k}}(m_{3k}, s_{3k}^2)I(0, \infty)$. We take the Dirichlet distribution (Kotz et al. 2000, p. 485) for the prior of the $\boldsymbol{\eta}$ parameter, thus $\eta_1, \eta_2, \dots, \eta_{K-1} \sim \text{Dir}(\delta_1, \delta_2, \dots, \delta_K)$. For the priors of the $\phi_{k,0}$ and $\phi_{k,i}$, we take the normal distribution with the location parameters $u_{k,0}$, $u_{k,i}$ and the scale parameters $g_{k,0}^2$, $g_{k,i}^2$, thus $\phi_{k,0} \sim N(u_{k,0}, g_{k,0}^2)$ and $\phi_{k,i} \sim N(u_{k,i}, g_{k,i}^2)$; $i = 1, 2, \dots, p_k$. Employing the setup of prior distributions, as shown above, the natural logarithm of the joint posterior distribution of the model is given by

$$\ln \pi(\boldsymbol{\theta}|\mathbf{y}) \propto \sum_{k=1}^K \left(\left(\sum_{t=p+1}^T h_{k,t} (\ln \eta_k + \Lambda_1 + \Lambda_2) \right) + \Lambda_3 + \Lambda_4 + \Lambda_5 + \Lambda_6 + \Lambda_7 + \sum_{i=1}^{p_k} \Lambda_8 \right), \quad (21)$$

where $\Lambda_1 = d_{2k} \ln[d_{2k}/d_{1k}] + \ln[\Gamma(\frac{1}{2}(d_{1k} + d_{2k}))] - \ln[\Gamma(\frac{1}{2}d_{1k})] - \ln[\Gamma(\frac{1}{2}d_{2k})] - \ln \sigma_k$, $\Lambda_2 = -d_{2k} \left(\frac{a_{k,t}}{\sigma_k} \right) - \frac{d_{1k} + d_{2k}}{2} \ln \left[1 + \exp \left[-2 \frac{a_{k,t}}{\sigma_k} + \ln[d_{2k}/d_{1k}] \right] \right]$, $\Lambda_3 = (\delta_K - 1) \ln \left[1 - \sum_{k=1}^{K-1} \eta_k \right]$ + $\sum_{k=1}^{K-1} (\delta_k - 1) \ln \eta_k$, $\Lambda_4 = -\frac{1}{2}(v_{1k} + 1) \ln \left[1 + \frac{1}{v_{1k}} \left(\frac{d_{1k} - m_{1k}}{s_{1k}} \right)^2 \right]$, $\Lambda_5 = -\frac{1}{2}(v_{2k} + 1) \ln \left[1 + \frac{1}{v_{2k}} \left(\frac{d_{2k} - m_{2k}}{s_{2k}} \right)^2 \right]$, $\Lambda_6 = -\frac{1}{2}(v_{3k} + 1) \ln \left[1 + \frac{1}{v_{3k}} \left(\frac{\sigma_k - m_{3k}}{s_{3k}} \right)^2 \right]$, $\Lambda_7 = -\frac{1}{2} \left(\frac{\phi_{k,0} - u_{k,0}}{g_{k,0}} \right)^2$, and $\Lambda_8 = -\frac{1}{2} \left(\frac{\phi_{k,i} - u_{k,i}}{g_{k,i}} \right)^2$.

HMC requires the gradient of the ln-posterior density. In practice, the gradient must be computed analytically (Gelman et al. 2014, p. 301). The gradient of $\ell(\boldsymbol{\theta}|\mathbf{y})$ is

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\mathbf{y}) = \frac{\partial}{\partial \boldsymbol{\theta}} \ln \pi(\boldsymbol{\theta}|\mathbf{y})$$

$$\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}|\mathbf{y}) = \left(\frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial \eta_k}, \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial d_{1k}}, \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial d_{2k}}, \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial \sigma_k}, \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial \phi_{k,0}}, \frac{\partial \ell(\boldsymbol{\theta}|\mathbf{y})}{\partial \phi_{k,i}} \right) \quad (22)$$

where $i = 1, 2, \dots, p_k$; $k = 1, 2, \dots, K$. The HMC algorithm for estimating the parameters of the ZMAR model is as follows:

1. Determine the initial value of the parameter $\boldsymbol{\theta}^0$, the diagonal mass matrix \mathbf{M} , the scale factor of the leapfrog steps ϵ , the number of leapfrog steps \mathcal{N} , and the number of iterations \mathcal{R} .
2. For each iteration r ; $r = 1, 2, \dots, \mathcal{R}$ where \mathcal{R} represents the number of iterations,

- a. Generate the momentum variables $\boldsymbol{\varphi}$ with $\boldsymbol{\varphi} \sim \text{Normal}(\mathbf{0}, \mathbf{M})$;
- b. For each iteration $n = 1, 2, \dots, \mathcal{N}$,
 - (1) Use the gradient of the ln-posterior density of $\boldsymbol{\vartheta}_n$ to make a half-step of $\boldsymbol{\varphi}_n$

$$\boldsymbol{\varphi}_{n+0.5} \leftarrow \boldsymbol{\varphi}_n + \frac{1}{2} \varepsilon \frac{\partial \ln \pi(\boldsymbol{\vartheta}_n | \mathbf{y})}{\partial \boldsymbol{\vartheta}_n},$$
 - (2) Update the vector $\boldsymbol{\vartheta}_n$ using the vector $\boldsymbol{\varphi}_{n+0.5}$

$$\boldsymbol{\vartheta}_{n+0.5} \leftarrow \boldsymbol{\vartheta}_n + \varepsilon \mathbf{M}^{-1} \boldsymbol{\varphi}_{n+0.5}$$
 - (3) Update the next half-step for $\boldsymbol{\varphi}$

$$\boldsymbol{\varphi}_{n+1} \leftarrow \boldsymbol{\varphi}_{n+0.5} + \frac{1}{2} \varepsilon \frac{\partial \ln \pi(\boldsymbol{\vartheta}_{n+0.5} | \mathbf{y})}{\partial \boldsymbol{\vartheta}_{n+0.5}}.$$
- c. Labels $\boldsymbol{\vartheta}^{(r-1)}$ and $\boldsymbol{\varphi}^{(r-1)}$ as the value of the parameter and momentum vectors at the start of the leapfrog process and $\boldsymbol{\vartheta}^\#$, $\boldsymbol{\varphi}^\#$ as the value after the \mathcal{N} steps.
- d. Compute

$$b = \frac{\pi(\boldsymbol{\vartheta}^\# | \mathbf{y}) f(\boldsymbol{\varphi}^\#)}{\pi(\boldsymbol{\vartheta}^{(r-1)} | \mathbf{y}) f(\boldsymbol{\varphi}^{(r-1)})}.$$
- e. Set $\boldsymbol{\vartheta}^{(r)}$

$$\boldsymbol{\vartheta}^{(r)} = \begin{cases} \boldsymbol{\vartheta}^\# & \text{with probability } \min(b, 1) \\ \boldsymbol{\vartheta}^{(r-1)} & \text{otherwise.} \end{cases}$$
- f. Save $\boldsymbol{\vartheta}^{(r)}$; $r = 1, 2, \dots, \mathcal{R}$.

The performance of the HMC is very sensitive to two user-defined parameters, i.e., the step size of the leapfrog ε and the number of leapfrog steps L . The No-U-Turn Sampler (NUTS) could eliminate the need to set the parameter L and could adapt the step size parameter ε on the fly based on a primal-dual averaging (Hoffman and Gelman 2014). The NUTS algorithm was implemented in C++ as part of the open-source Bayesian inference package, Stan (Gelman et al. 2014, p. 304; Carpenter et al. 2017). Stan is also a platform for computing log densities and their gradients, so that the densities and gradients are easy to obtain (Carpenter et al. 2015, 2017). Stan can be called from R using the rstan package. An example of the Stan code to fit the ZMAR model can be seen in Section 4.

4. Simulation Studies

A simulation study was carried out to evaluate the performance of the ZMAR model compared to the TMAR and GMAR models. We consider the simulations to accommodate eight scenarios for the conditional density in the first component of the ZMAR model. Furthermore, the conditional densities in the second and third components are specified as the symmetric-fat-tail of the Fisher's z distributions.

We conducted a Bayesian analysis on the eight simulated datasets, whose datasets were generated by the following steps:

- Step 1: Specify the ZMAR model with three components as $\text{ZMAR}(3; 1, 1, 1)$ where

$$f(y_t | \mathcal{F}_{(t-1)}) = \sum_{k=1}^3 \eta_k f_k(y_t | \mathcal{F}_{(t-1)}) = \sum_{k=1}^3 \eta_k (\phi_{k,0} + \phi_{k,1} y_{t-1} + e_{k,t}); \eta_1 = \eta_2 = \eta_3 = \frac{1}{3}; \phi_{1,1} = -0.6; \phi_{2,1} = 0.2; \phi_{3,1} = 0.7 \text{ with } e_{k,t} \sim z(d_{1,k}, d_{2,k}, 0, \sigma_k); k = 1, 2, 3$$
 are the innovations in the first, second, and third components. The scenarios for the simulation are as follows:
 - Scenario 1: represent a mixture of a highly skewed (Bulmer 1967, p. 63) to the left and two symmetrical distributions, where $\phi_{1,0} = \phi_{2,0} = \phi_{3,0} = 0$, excess unconditional kurtosis is 5.73, $e_{1,t} \sim z(0.2, 10, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, and $e_{3,t} \sim z(30, 30, 0, 10)$;
 - Scenario 2: represent a mixture of a highly skewed to the right and two symmetrical distributions, where $\phi_{1,0} = \phi_{2,0} = \phi_{3,0} = 0$, excess unconditional kurtosis is 2.23, $e_{1,t} \sim z(20, 1, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, and $e_{3,t} \sim z(30, 30, 0, 10)$;
 - Scenario 3: represent a mixture of three symmetrical distributions, where $\phi_{1,0} = \phi_{2,0} = \phi_{3,0} = 0$, excess unconditional kurtosis is 1.67, $e_{1,t} \sim z(0.5, 0.5, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, and $e_{3,t} \sim z(30, 30, 0, 10)$;

- Scenario 4: represent a mixture of moderately skewed (Bulmer 1967, p. 63) distributions, where $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, excess unconditional kurtosis is 2.12, $e_{1.t} \sim z(3, 10, 0, 8)$, $e_{2.t} \sim z(1, 1, 0, 5)$, and $e_{3.t} \sim z(30, 30, 0, 10)$;
- Scenario 5: represent a mixture of three fairly symmetrical distributions (Bulmer 1967, p. 63), where $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, excess unconditional kurtosis is 4.25, $e_{1.t} \sim z(7, 10, 0, 8)$, $e_{2.t} \sim z(1, 1, 0, 5)$, and $e_{3.t} \sim z(30, 30, 0, 10)$;
- Scenario 6: represent a mixture of three symmetrical distributions, where $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, excess unconditional kurtosis is 4.47, $e_{1.t} \sim z(10, 10, 0, 8)$, $e_{2.t} \sim z(1, 1, 0, 5)$, and $e_{3.t} \sim z(30, 30, 0, 10)$;
- Scenario 7: represent a mixture of three symmetrical distributions, where $\phi_{1.0} = -1$, $\phi_{2.0} = \phi_{3.0} = 0$, excess unconditional kurtosis is 1.63, $e_{1.t} \sim z(0.5, 0.5, 0, 5)$, $e_{2.t} \sim z(1, 1, 0, 8)$, and $e_{3.t} \sim z(30, 30, 0, 10)$;
- Scenario 8: represent a mixture of three symmetrical distributions, where $\phi_{1.0} = -20$, $\phi_{2.0} = \phi_{3.0} = 0$, excess unconditional kurtosis is 0.67, $e_{1.t} \sim z(0.5, 0.5, 0, 5)$, $e_{2.t} \sim z(1, 1, 0, 8)$, and $e_{3.t} \sim z(30, 30, 0, 10)$;

The innovations of Scenario 7 and Scenario 8 are the same as those of Scenario 3. The comparison of graph visualizations for the innovations in the first component in Scenario 1 to Scenario 6 represented specifically for the Fisher's z distribution can be seen in Figures 1 and 2;

- Step 2: Generate $e_{1.t}$, $e_{2.t}$, and $e_{3.t}$;
- Step 3: Generate $H_t \sim \text{Mult}_3(1, \eta)$; where $\eta = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$;
- Step 4: Compute $f_k(y_t | \mathcal{F}_{(t-1)}) = \phi_{k.0} + \phi_{k.1}y_{t-1} + e_{k.t}$; $k = 1, 2, 3$;
- Step 5: Compute $y_t = f_1(y_t | \mathcal{F}_{(t-1)})^{h_{1.t}} f_2(y_t | \mathcal{F}_{(t-1)})^{h_{2.t}} f_3(y_t | \mathcal{F}_{(t-1)})^{h_{3.t}}$.

Furthermore, we generated 600 datasets and fit all of the models for each simulated dataset, respectively, to find the best performance in terms of model comparisons. Figure 3 shows the simulated datasets for Scenario 1 to Scenario 8. We implemented the models using the rstan package (Stan Development Team 2020), the R interface to Stan developed by the Stan Development Team in the R software. Suppose $p_k = p_k$; $\text{con_eta} = c(1, 1, 1)$; $\text{tau}[t, k] = \tau_{t,k}$; $\text{eta}[k] = \eta_k$; $\text{sigma}[k] = \sigma_k$; $d1[k] = d_{1k}$; $d2[k] = d_{2k}$; $\text{phik}[1] = \phi_{k.1}$; $k = 1, 2, 3$. Here is an example of the Stan code to fit the ZMAR model in Scenario 1:

```
fitMAR1 = "
functions{
  real fisher_z_lpdf(real x, real d1, real d2, real mu, real sigma){
    return (log(2)+0.5*d2*(log(d2)-log(d1))-d2*(x-mu)/sigma-log(sigma)-lbeta(0.5*d1,0.5*d2)-
      (d1+d2)/2*log1p_exp((-2*(x-mu)/sigma)+log(d2)-log(d1)));
  }
}
data {
  int<lower=0> p1;
  int<lower=0> p2;
  int<lower=0> p3;
  int T;
  vector[T] y;
  vector[3] con_eta;
}
parameters {
  simplex[3] eta; // mixing proportions
  vector<lower = 0>[3] sigma;
  vector<lower = 0>[3] d1;
  vector<lower = 0>[3] d2;
  real phi1[p1];
  real phi2[p2];
```

```

real phi3[p3];
}
model {
matrix[T,3] tau;
real lf[T];
eta ~ dirichlet(con_eta);
//priors
sigma[1] ~ student_t(3,5,0.1);
sigma[2] ~ student_t(3,8,0.1);
sigma[3] ~ student_t(3,10,0.1);
d1[1] ~ student_t(3,0.2,0.1);
d1[2] ~ student_t(3,1,0.1);
d1[3] ~ student_t(3,30,0.1);
d2[1] ~ student_t(3,10,0.1);
d2[2] ~ student_t(3,1,0.1);
d2[3] ~ student_t(3,30,0.1);
phi1[1] ~ normal(-0.6,0.1);
phi2[1] ~ normal(0.2,0.1);
phi3[1] ~ normal(0.7,0.1);
//ZMAR model
for(t in 1:T) {
if(t==1) {
tau[t,1] = log(eta[1])+fisher_z_lpdf(y[t]/sigma[1] | d1[1], d2[1],0,1)-log(sigma[1]);
tau[t,2] = log(eta[2])+fisher_z_lpdf(y[t]/sigma[2] | d1[2], d2[2],0,1)-log(sigma[2]);
tau[t,3] = log(eta[3])+fisher_z_lpdf(y[t]/sigma[3] | d1[3], d2[3],0,1)-log(sigma[3]);
} else {
real mu1 = 0;
real mu2 = 0;
real mu3 = 0;
for (i in 1:p1)
mu1 += phi1[i] * y[t-i];
for (i in 1:p2)
mu2 += phi2[i] * y[t-i];
for (i in 1:p3)
mu3 += phi3[i] * y[t-i];
tau[t,1] = log(eta[1])+fisher_z_lpdf((y[t]-mu1)/sigma[1] | d1[1],d2[1],0,1)-log(sigma[1]);
tau[t,2] = log(eta[2])+fisher_z_lpdf((y[t]-mu2)/sigma[2] | d1[2],d2[2],0,1)-log(sigma[2]);
tau[t,3] = log(eta[3])+fisher_z_lpdf((y[t]-mu3)/sigma[3] | d1[3],d2[3],0,1)-log(sigma[3]);
}
lf[t] = log_sum_exp(tau[t,]);
}
target += sum(lf);
} "

```

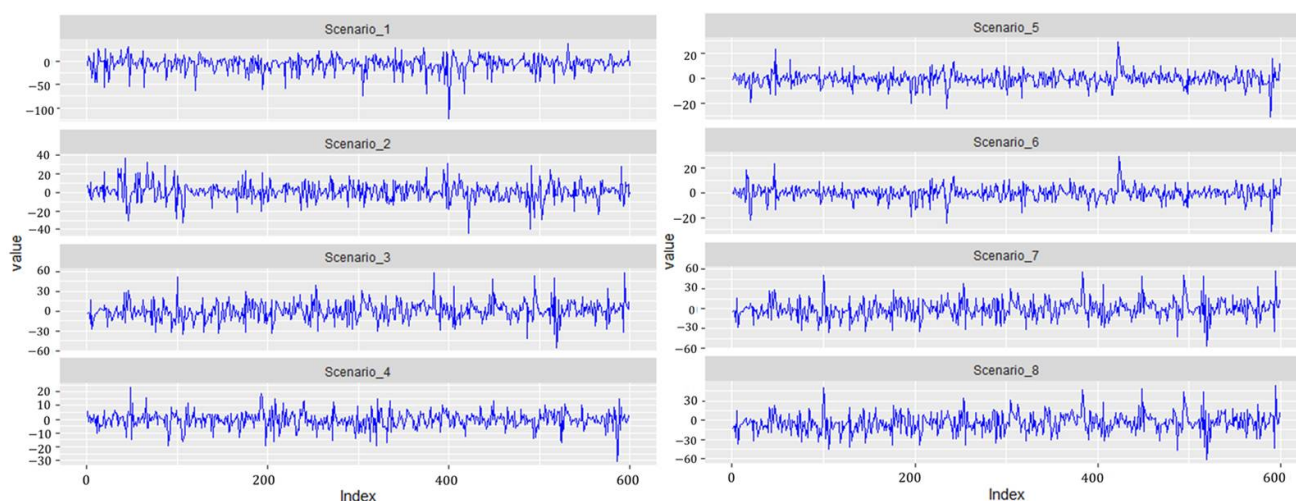


Figure 3. Simulating time-series datasets for Scenario 1 to Scenario 8.

The warm-up stage in these simulation studies was set to 1500 iterations, 3 chains with 5000 sampling iterations, and 1 thin. The `adapt_delta` parameter was set to 0.99, and the `max_treedepth` was set to 15. For all scenarios, the parameter priors of the ZMAR, TMAR, and GMAR models are shown in Appendix D, and their posterior inferences are presented in Appendix E. There are a variety of convergence diagnoses, such as the potential scale reduction factor \hat{R} (Gelman and Rubin 1992; Susanto et al. 2018; Gelman et al. 2014, p. 285; Vehtari et al. 2020) and the effective sample size n_{eff} (Gelman et al. 2014, p. 266; Vehtari et al. 2020). If the MCMC chain has reached convergence, the \hat{R} statistic is less than 1.01, and the n_{eff} statistic is greater than 400 (Vehtari et al. 2020). To compare the performance of the models, we use the PSIS-LOO.

Table 1 shows the summary simulation result for all scenarios, which indicates that the ZMAR model performs the best when the datasets are generated from the ZMAR model, in which one of the components is asymmetric. When all the components are symmetric, the ZMAR model also performs better than TMAR and GMAR, as long as the excess unconditional kurtosis is large enough or the intercept distances between the components are far apart. However, when all of the mixture components are symmetric, the excess unconditional kurtosis is small and the intercept distances between the components are close enough or the intercepts are the same, then the GMAR model plays the best. Let us now focus on the results of Scenario 3 and Scenario 6.

In the third and sixth scenarios, the datasets are generated from three symmetrical distributions. The two scenarios have the same intercepts and are generated with different unconditional kurtosis. Scenario 3 has a smaller unconditional kurtosis than Scenario 6. The best ZMAR, TMAR, and GMAR models for the two scenarios are ZMAR(3;1,1,1), TMAR(3;1,1,1), and GMAR(3;1,1,1), where the values of PSIS-LOO are 4483.30, 4483.10, and 4481.10, for the third scenario and 3490.10, 3496.20, and 3494.60, for the sixth scenario. Clearly, the PSIS-LOO value for the GMAR model is smaller than for the ZMAR and TMAR models in the third scenario, and the ZMAR model has the smallest PSIS-LOO value in the sixth scenario. When the intercepts in Scenario 3 are varied and determined as in Scenario 7, the GMAR model is also the best. However, when the intercepts in Scenario 3 are varied and determined as in Scenario 8, the ZMAR model is the best. For other scenarios, in datasets generated from asymmetric components (the first, second, fourth, and fifth scenarios), the ZMAR model is the best.

Table 1. Model comparison using pareto-smoothed important sampling leave-one-out cross-validation (PSIS-LOO) in several scenarios.

Scenarios	ZMAR		TMAR		GMAR	
	Model	PSIS-LOO	Model	PSIS-LOO	Model	PSIS-LOO
1	ZMAR(3;1,1,1)	4563.6 *	TMAR(3;1,1,1)	4639.1	GMAR(4;1,1,1,1)	4646.6
2	ZMAR(3;1,1,1)	4120.0 *	TMAR(3;1,1,1)	4139.0	GMAR(3;1,1,1)	4130.8
3	ZMAR(3;1,1,1)	4483.3	TMAR(3;1,1,1)	4483.1	GMAR(3;1,1,1)	4481.1 *
4	ZMAR(3;1,1,1)	3639.7 *	TMAR(3;1,1,1)	3652.7	GMAR(3;1,1,1)	3660.9
5	ZMAR(3;1,1,1)	3523.9 *	TMAR(3;1,1,1)	3534.1	GMAR(4;1,1,1,1)	3539.5
6	ZMAR(3;1,1,1)	3490.1 *	TMAR(3;1,1,1)	3496.2	GMAR(3;1,1,1)	3494.6
7	ZMAR(3;1,1,1)	4483.6	TMAR(3;1,1,1)	4481.5	GMAR(3;1,1,1)	4478.0 *
8	ZMAR(3;1,1,1)	4549.1 *	TMAR(3;1,1,1)	4579.4	GMAR(3;1,1,1)	4575.1

Simulation Model: $f(y_t | \mathcal{F}_{(t-1)}) = \frac{1}{3}(\phi_{1.0} + -0.6 y_{t-1} + e_{1,t}) + \frac{1}{3}(\phi_{2.0} + 0.2 y_{t-1} + e_{2,t}) + \frac{1}{3}(\phi_{3.0} + 0.7 y_{t-1} + e_{3,t})$; Scenario 1: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(0.2, 10, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 2: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(20, 1, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 3: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(0.5, 0.5, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 4: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(3, 10, 0, 8)$, $e_{2,t} \sim z(1, 1, 0, 5)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 5: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(7, 10, 0, 8)$, $e_{2,t} \sim z(1, 1, 0, 5)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 6: $\phi_{1.0} = \phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(10, 10, 0, 8)$, $e_{2,t} \sim z(1, 1, 0, 5)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 7: $\phi_{1.0} = -1$, $\phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(0.5, 0.5, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, $e_{3,t} \sim z(30, 30, 0, 10)$; Scenario 8: $\phi_{1.0} = -20$, $\phi_{2.0} = \phi_{3.0} = 0$, $e_{1,t} \sim z(0.5, 0.5, 0, 5)$, $e_{2,t} \sim z(1, 1, 0, 8)$, $e_{3,t} \sim z(30, 30, 0, 10)$; * indicates the smallest value.

5. Application for Real Data

5.1. IBM Stock Prices

To illustrate the potential of the ZMAR model, we consider the daily IBM common stock closing price from 17 May 1961 to 2 November 1962 (Box et al. 2015, p. 627). This time series has been analyzed by many researchers such as Le et al. (1996) and Wong and Li (2000). Wong and Li (2000) used the EM algorithm to estimate the parameters model and has been identified that the best GMAR model for the series was a GMAR(3;1,1,0).

Figure 4 shows that the IBM stock prices series has a trimodal marginal distribution, where the estimated locations of the modes are at 377.48, 481.78, and 547.90 points. Therefore, we choose the three-component ZMAR model for the differenced series. The orders of the autoregressive components are chosen by the minimum PSIS-LOO. The best three-component ZMAR model is ZMAR(3;0,1,1) without intercept and the value of PSIS-LOO is 2424.40. The warm-up stage for the ZMAR, TMAR, and GMAR models was set to 1500 iterations followed by 3 chains with 5000 sampling iterations and 1 thin; the adapt_delta parameter was set to 0.99, and the max_treedepth was set to 15. Table 2 shows the summary of posterior inferences for all models. The prior distributions and posterior density plots for the parameters of the ZMAR model are presented in Appendices F.1 and G.1.1, respectively. For all the parameters of the ZMAR model, the MCMC chain has reached convergence, which is shown by the \hat{R} statistic being less than 1.01 and the n_{eff} statistic being greater than 400. The three-component ZMAR model for the differenced series was then transformed into a three-component ZMAR model for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.01 Z_{(1.95, 3.90)} \left(\frac{y_t - y_{t-1}}{28.28} \right) + 0.46 Z_{(1.87, 6.41)} \left(\frac{y_t - 1.61 y_{t-1} + 0.61 y_{t-2}}{9.77} \right) + 0.53 Z_{(4.91, 1.70)} \left(\frac{y_t - 0.72 y_{t-1} - 0.28 y_{t-2}}{6.34} \right). \quad (23)$$

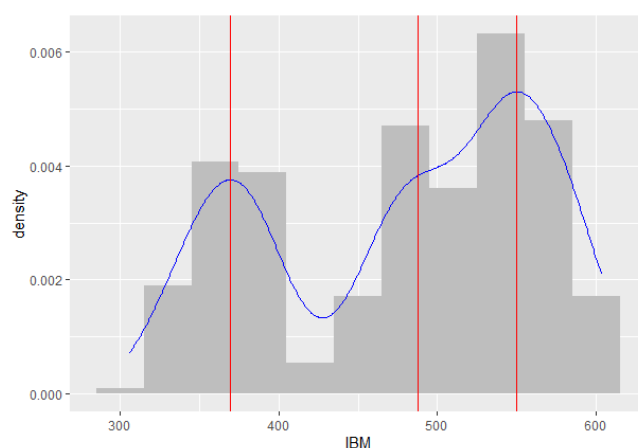


Figure 4. Histogram and marginal density plot of the IBM stock prices.

Table 2. Summary of posterior inferences for ZMAR, TMAR, and GMAR models in the IBM stock prices (first-differenced series).

Parameters	ZMAR Model					TMAR Model					GMAR Model				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
d1[1]	1.95	1.65	2.26	7185	1	-	-	-	-	-	-	-	-	-	-
d1[2]	1.87	1.65	2.21	6655	1	-	-	-	-	-	-	-	-	-	-
d1[3]	4.91	4.58	5.23	6008	1	-	-	-	-	-	-	-	-	-	-
d2[1]	3.90	3.60	4.23	5905	1	-	-	-	-	-	-	-	-	-	-
d2[2]	6.41	6.09	6.75	2798	1	-	-	-	-	-	-	-	-	-	-
d2[3]	1.70	1.50	1.95	8370	1	-	-	-	-	-	-	-	-	-	-
nu[1]	-	-	-	-	-	12.52	12.21	12.84	7072	1	-	-	-	-	-
nu[2]	-	-	-	-	-	10.77	10.46	11.09	6020	1	-	-	-	-	-
nu[3]	-	-	-	-	-	14.03	13.72	14.36	5384	1	-	-	-	-	-
eta[1]	0.01	0.00	0.05	13945	1	0.58	0.44	0.71	6826	1	0.54	0.41	0.67	5941	1
eta[2]	0.46	0.34	0.58	10248	1	0.40	0.27	0.54	6713	1	0.42	0.29	0.55	5853	1
eta[3]	0.53	0.41	0.64	10235	1	0.02	0.00	0.05	10685	1	0.04	0.01	0.09	9810	1
sigma[1]	28.28	27.97	28.58	4083	1	4.97	4.65	5.21	5145	1	4.82	4.56	5.08	7747	1
sigma[2]	9.77	9.40	10.05	2413	1	5.80	5.50	6.07	7715	1	6.01	5.72	6.29	7369	1
sigma[3]	6.34	6.05	6.63	9774	1	25.02	24.69	25.35	7061	1	19.04	18.71	19.36	4525	1
phi1[1]	-	-	-	-	-	-0.29	-0.41	-0.17	8124	1	-0.32	-0.43	-0.21	6681	1
phi2[1]	0.61	0.47	0.76	11354	1	0.68	0.52	0.84	8061	1	0.67	0.52	0.82	7592	1
phi3[1]	-0.28	-0.39	-0.17	10868	1	-	-	-	-	-	-	-	-	-	-
PSIS-LOO	2424.40					2431.30					2431.70				

Note: eta[1] = η_1 ; eta[2] = η_2 ; eta[3] = η_3 ; sigma[1] = σ_1 ; sigma[2] = σ_2 ; sigma[3] = σ_3 ; d1[1] = d_{11} ; d1[2] = d_{12} ; d1[3] = d_{13} ; d2[1] = d_{21} ; d2[2] = d_{22} ; d2[3] = d_{23} ; nu[1] = ν_1 ; nu[2] = ν_2 ; nu[3] = ν_3 ; phi1[1] = $\phi_{1.1}$; phi2[1] = $\phi_{2.1}$; phi3[1] = $\phi_{3.1}$.

We compared the ZMAR model with the TMAR and GMAR models. The best three-component TMAR and GMAR models are TMAR(3;1,1,0) and GMAR(3;1,1,0), without intercept. The PSIS-LOO values of the TMAR and GMAR models are 2431.30 and 2431.70, respectively. The prior distributions and posterior density plots for the parameters of the TMAR and GMAR models are presented in Appendices F.1 and G.1. Here is the result of the summary posterior inferences for all models

The MCMC chains for the TMAR parameter model and the GMAR parameter model have also reached convergence, which is shown by the \hat{R} statistic being less than 1.01 and the n_{eff} statistic being greater than 400. Let $\Phi(\cdot)$, and let $\mathcal{T}_{\nu_k}(\cdot)$ be the CDF of the standard normal distribution and the standardized Student t distribution with ν_k ; $k = 1, 2, 3$ degrees of freedom. The three-component TMAR model for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.58 \mathcal{T}_{12.52} \left(\frac{y_t - 0.71 y_{t-1} - 0.29 y_{t-2}}{4.97} \right) + 0.40 \mathcal{T}_{10.77} \left(\frac{y_t - 1.68 y_{t-1} + 0.68 y_{t-2}}{5.80} \right) + 0.02 \mathcal{T}_{14.03} \left(\frac{y_t - y_{t-1}}{25.02} \right). \quad (24)$$

and a three-component GMAR for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.54 \Phi\left(\frac{y_t - 0.68 y_{t-1} - 0.32 y_{t-2}}{4.82}\right) + 0.42 \Phi\left(\frac{y_t - 1.67 y_{t-1} + 0.67 y_{t-2}}{6.01}\right) + 0.04 \Phi\left(\frac{y_t - y_{t-1}}{19.04}\right). \quad (25)$$

Therefore, the ZMAR model is preferred over the TMAR and GMAR models, which are indicated by the PSIS-LOO value of the ZMAR model being the smallest.

5.2. Brent Crude Oil Prices

In December 1988, the Organization of Petroleum Exporting Countries (OPEC) decided to adopt Brent as a new benchmark, rather than the value of the Arabian light (Carollo 2012, p. 10). Since then, Brent has been one of the main benchmarks for oil purchases in the world, the other being West Texas Intermediate (WTI). Figure 5a shows the monthly Brent crude oil price from January 1989 to June 2020 (in U.S. dollars per barrel), taken from the World Bank (2020). Figure 5b shows the first-differenced series. Figure 6 shows the marginal distribution of the original series as being trimodal, where the estimated locations of the modes are at 18.57, 61.64, and 110.15 points. Therefore, we also decided to choose a three-component mixture model for the ZMAR, TMAR, and GMAR models applied to the differenced series. The best three-component mixture model estimates for each of the ZMAR, TMAR, and GMAR models are ZMAR (3;1,2,2), TMAR (3;2,1,3), and GMAR (3;3,3,3), without intercept.

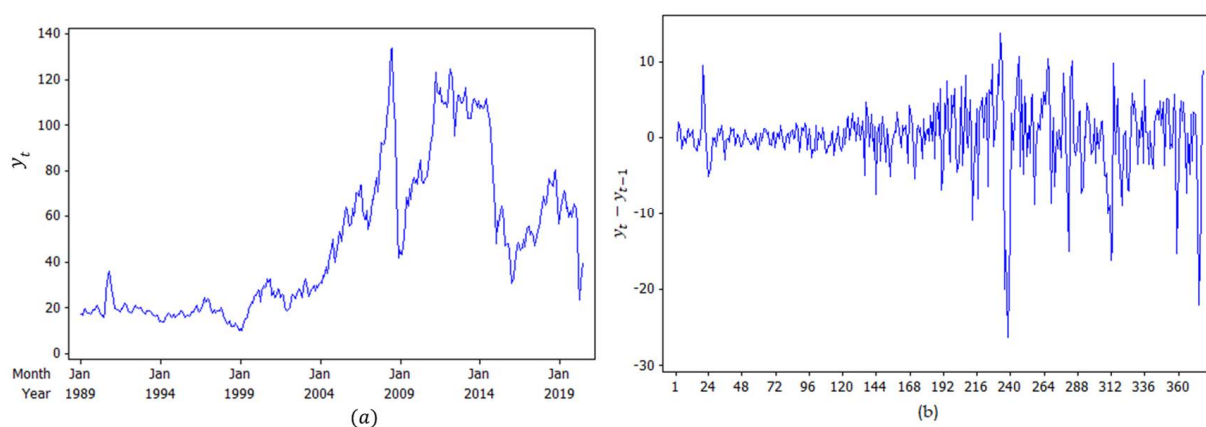


Figure 5. Monthly average Brent crude oil price series from January 1989 to June 2020: (a) original series; (b) first-differenced series.

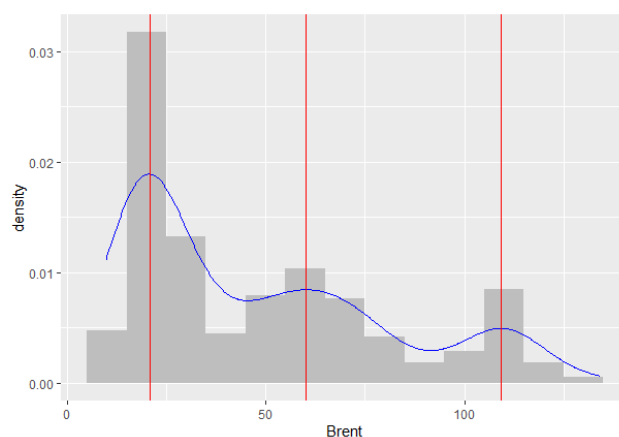


Figure 6. Histogram and density plot of the Brent monthly crude oil price series.

The warm-up steps for all the models were set to 1500 iterations with 5000 sampling iterations and 1 thin, followed by 3 chains; the adapt_delta parameters were set to 0.99, and the max_treedepts were set to 15. Table 3 shows the summary of posterior inferences for all models. The prior distributions and posterior density plots for the parameters of the ZMAR model, the TMAR model, and the GMAR model are presented, respectively, in Appendices F.2 and G.2.

Table 3. Summary of posterior inferences for ZMAR, TMAR, and GMAR models in the Brent monthly crude oil prices (first-differenced series).

Parameters	ZMAR Model					TMAR Model					GMAR Model				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
d1[1]	13.22	12.89	13.57	1136	1	-	-	-	-	-	-	-	-	-	-
d1[2]	0.99	0.78	1.16	7606	1	-	-	-	-	-	-	-	-	-	-
d1[3]	10.09	9.76	10.43	783	1	-	-	-	-	-	-	-	-	-	-
d2[1]	4.49	4.19	4.83	6163	1	-	-	-	-	-	-	-	-	-	-
d2[2]	1.91	1.64	2.21	8505	1	-	-	-	-	-	-	-	-	-	-
d2[3]	4.37	4.06	4.80	786	1	-	-	-	-	-	-	-	-	-	-
nu[1]	-	-	-	-	-	14.98	14.64	15.30	5742	1	-	-	-	-	-
nu[2]	-	-	-	-	-	12.08	11.77	12.39	6135	1	-	-	-	-	-
nu[3]	-	-	-	-	-	4.34	4.01	4.68	4090	1	-	-	-	-	-
eta[1]	0.40	0.30	0.51	7598	1	0.34	0.25	0.44	7002	1	0.45	0.35	0.55	8395	1
eta[2]	0.39	0.28	0.50	8658	1	0.40	0.29	0.49	7202	1	0.28	0.18	0.39	7317	1
eta[3]	0.21	0.12	0.31	8287	1	0.26	0.17	0.36	7343	1	0.27	0.18	0.36	9667	1
sigma[1]	5.14	4.78	5.40	5164	1	4.93	4.23	5.23	1419	1	4.18	3.96	4.52	5266	1
sigma[2]	2.82	2.57	3.08	9271	1	1.62	1.39	1.86	7816	1	1.66	1.37	1.88	6374	1
sigma[3]	7.12	6.03	7.46	674	1	1.80	1.42	2.03	4462	1	1.92	1.69	2.15	10687	1
phi1[1]	-0.37	-0.49	-0.26	8682	1	0.61	0.45	0.77	8641	1	0.14	0.01	0.28	7866	1
phi1[2]	-	-	-	-	-	-0.42	-0.56	-0.27	7932	1	-0.35	-0.48	-0.22	8284	1
phi1[3]	-	-	-	-	-	-	-	-	-	-	-0.19	-0.31	-0.07	8919	1
phi2[1]	0.68	0.55	0.81	8187	1	-0.28	-0.42	-0.15	7045	1	0.48	0.36	0.60	7799	1
phi2[2]	-0.34	-0.48	-0.21	8060	1	-	-	-	-	-	-0.17	-0.28	-0.06	9453	1
phi2[3]	-	-	-	-	-	-	-	-	-	-	0.38	0.24	0.54	7424	1
phi3[1]	0.69	0.55	0.84	9548	1	0.54	0.44	0.64	7735	1	0.57	0.47	0.67	6397	1
phi3[2]	0.72	0.57	0.86	9333	1	0.86	0.75	0.97	6974	1	0.91	0.81	1.02	7326	1
phi3[3]	-	-	-	-	-	-0.28	-0.41	-0.14	7427	1	-0.29	-0.41	-0.16	7545	1
PSIS-LOO	2024.50					2034.10					2048.70				

Note: eta[1] = η_1 ; eta[2] = η_2 ; eta[3] = η_3 ; sigma[1] = σ_1 ; sigma[2] = σ_2 ; sigma[3] = σ_3 ; d1[1] = d_{11} ; d1[2] = d_{12} ; d1[3] = d_{13} ; d2[1] = d_{21} ; d2[2] = d_{22} ; d2[3] = d_{23} ; nu[1] = ν_1 ; nu[2] = ν_2 ; nu[3] = ν_3 ; phi1[1] = $\phi_{1.1}$; phi1[2] = $\phi_{1.2}$; phi1[3] = $\phi_{1.3}$; phi2[1] = $\phi_{2.1}$; phi2[2] = $\phi_{2.2}$; phi2[3] = $\phi_{2.3}$; phi3[1] = $\phi_{3.1}$; phi3[2] = $\phi_{3.2}$; phi3[3] = $\phi_{3.3}$.

For all the parameter models, the MCMC chains reached convergence, which was shown by the \hat{R} statistic being less than 1.01 and the n_{eff} statistic being greater than 400.

The three-component ZMAR model for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.40 Z_{(13.22, 4.49)} \left(\frac{y_t - 0.63 y_{t-1} - 0.37 y_{t-2}}{5.14} \right) + 0.39 Z_{(0.99, 1.91)} \left(\frac{y_t - 1.68 y_{t-1} + 1.02 y_{t-2} - 0.34 y_{t-3}}{2.82} \right) + 0.21 Z_{(10.09, 4.37)} \left(\frac{y_t - 1.69 y_{t-1} - 0.03 y_{t-2} + 0.72 y_{t-3}}{7.12} \right). \quad (26)$$

a three-component TMAR model for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.34 \mathcal{T}_{14.98} \left(\frac{y_t - 1.61 y_{t-1} + 1.03 y_{t-2} - 0.42 y_{t-3}}{4.93} \right) + 0.40 \mathcal{T}_{12.08} \left(\frac{y_t - 0.72 y_{t-1} - 0.28 y_{t-2}}{1.62} \right) + 0.26 \mathcal{T}_{4.34} \left(\frac{y_t - 1.54 y_{t-1} - 0.32 y_{t-2} + 1.14 y_{t-3} - 0.28 y_{t-4}}{1.80} \right). \quad (27)$$

and a three-component GMAR for the original series, namely

$$F(y_t | \mathcal{F}_{(t-1)}) = 0.45 \Phi\left(\frac{y_t - 1.14 y_{t-1} + 0.49 y_{t-2} - 0.16 y_{t-3} - 0.19 y_{t-4}}{4.18}\right) + 0.28 \Phi\left(\frac{y_t - 1.48 y_{t-1} + 0.65 y_{t-2} - 0.55 y_{t-3} + 0.38 y_{t-4}}{1.66}\right) + 0.27 \Phi\left(\frac{y_t - 1.57 y_{t-1} - 0.34 y_{t-2} + 1.20 y_{t-3} - 0.29 y_{t-4}}{1.92}\right). \quad (28)$$

The PSIS-LOO values of the ZMAR, TMAR, and GMAR models are 2024.40, 2034.10, and 2048.70, respectively. Therefore, the ZMAR model is preferred over the TMAR and GMAR models, which is indicated by the PSIS-LOO value of the ZMAR model being the smallest.

6. Conclusions

We have discussed the definition and properties of the four-parameter Fisher's z distribution. The four-parameters of the Fisher's z distribution are μ , σ , d_1 , and d_2 . The μ is a location parameter, the σ is a scale parameter, and the d_1, d_2 are known as the shape parameters, defined for both skewness (symmetric if $d_1 = d_2$, asymmetric if $d_1 \neq d_2$) and fatness of the tails (large d_1 and d_2 imply thin tails). The Fisher's z distribution is always unimodal and has the mode at $x = \mu$. The value of μ only affects the mean of the distribution. It does not affect the variance, skewness, and kurtosis of the distribution. Furthermore, if $d_1 = d_2$, then the mean is equal to μ ; if $d_1 < d_2$, then the mean is less than μ ; and if $d_1 > d_2$, then the mean is greater than μ . The excess kurtosis value for this distribution is always positive.

We also discussed a new class of nonlinearity in the level (or mode) model for capturing time series with heteroskedasticity and with multimodal conditional distribution, using Fisher's z distribution as an innovation in the MAR model. The model offers great flexibility that other models, such as the TMAR and GMAR models, do not. The MCMC algorithm, using NUTS, allows for the easy estimation of the parameters in the model. The paper provides a simulation study using eight scenarios to indicate the flexibility and superiority of the ZMAR model compared with the TMAR and GMAR models. The simulation result shows that the ZMAR model is the best for representing the datasets generated from asymmetric components. When all the components are symmetrical, the ZMAR model also performs the best, as long as the excess unconditional kurtosis is large enough or the intercept distances between the components are far apart. However, when the datasets are generated from symmetrical components with small excess unconditional kurtosis and close intercept distances between the components, the GMAR model is the best. Furthermore, we compared the proposed model with the GMAR and TMAR models using two real data, namely the daily IBM stock prices and the monthly Brent crude oil prices. The results show that the proposed model outperforms the existing ones.

Fong et al. (2007) extended univariate the GMAR models to a Gaussian Mixture Vector Autoregressive (GMVAR) model. The ZMAR model can also be extended to a multivariate time-series context. Jones (2002) extended the standard multivariate F distribution to the multivariate skew t distribution and the multivariate Beta distribution. Likewise, the F distributed can also be extended to multivariate Fisher's z distribution.

Author Contributions: A.S., H.K., N.I. and K.F. analyzed and designed the research; A.S. collected, analyzed the data, and drafted the paper. All authors critically read and revised the draft and approved the final paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available from stated sources.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs of the Equations (1)–(6)

Appendix A.1. Proof of the Equation (1)

Solikhah et al. (2021) described transforming a random variable with the F distribution to the Fisher's z distribution. Let Y be a random variable distributed as an F distribution with two parameters d_1 and d_2 . The density of $Z = \frac{1}{2} \ln Y$ is (Fisher 1924; Aroian 1941)

$$f_Z(z; d_1, d_2) = \frac{2d_1^{\frac{1}{2}d_1}d_2^{\frac{1}{2}d_2}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \frac{e^{d_1z}}{(d_1e^{2z} + d_2)^{(d_1+d_2)/2}}, \quad (A1)$$

where $-\infty < z < \infty$, $d_1 > 0$, $d_2 > 0$ and $B(\cdot)$ is the beta function. Interchanging d_1 and d_2 is equivalent to replacing z with $-z$ (Fisher 1924; Aroian 1941), thus Equation (A1) can also be defined as

$$f_Z(z; d_1, d_2) = \frac{2d_1^{\frac{1}{2}d_1}d_2^{\frac{1}{2}d_2}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \frac{e^{-d_2z}}{(d_2e^{-2z} + d_1)^{(d_1+d_2)/2}}. \quad (A2)$$

If the denominator and numerator of Equation (A2) are divided by $d_1^{(d_1+d_2)/2}$ then we get

$$f_Z(z; d_1, d_2) = \frac{2\left(\frac{d_2}{d_1}\right)^{\frac{1}{2}d_2}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \frac{e^{-d_2z}}{\left(1 + \frac{d_2}{d_1}e^{-2z}\right)^{(d_1+d_2)/2}}. \quad (A3)$$

Equation (A3) can also be defined as

$$f_Z(z; d_1, d_2) = \frac{2\left(\frac{d_2}{d_1}\right)^{\frac{1}{2}d_2}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \frac{e^{-d_2z}}{\left(1 + e^{-2z + \ln \frac{d_2}{d_1}}\right)^{(d_1+d_2)/2}}$$

□

Appendix A.2. Proof of the Equation (2)

Let Y be a random variable distributed as an F distribution with d_1 and d_2 degrees of freedom, let $I_{y^*}(\cdot)$ be the incomplete beta function ratio, and let $B(\cdot)$ be the beta function. The cumulative distribution function (CDF) of the $Y = e^{2z}$ be defined as follows (Johnson et al. 1995, vol. 2, p. 327)

$$F_Y(y; d_1, d_2) = I_{y^*}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) = \frac{\int_0^{y^*} t^{\frac{1}{2}d_1-1}(1-t)^{\frac{1}{2}d_2-1}dt}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)},$$

where $y^* = \frac{d_1y}{d_2+d_1y}$. If $Z = \frac{1}{2} \ln Y$ then $Y = e^{2z}$, thus the CDF of Z is

$$F_Z(z; d_1, d_2) = I_{z^*}\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) = \frac{\int_0^{z^*} t^{\frac{1}{2}d_1-1}(1-t)^{\frac{1}{2}d_2-1}dt}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)}, \quad (A4)$$

where $z^* = \frac{d_1e^{2z}}{d_2+d_1e^{2z}}$. Equation (A4) can also be defined as

$$F_Z(z; d_1, d_2) = I_{z^*}\left(\frac{1}{2}d_2, \frac{1}{2}d_1\right) = \frac{\int_0^{z^*} t^{\frac{1}{2}d_2-1}(1-t)^{\frac{1}{2}d_1-1}dt}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)},$$

where $z^* = \frac{d_2e^{-2z}}{d_1+d_2e^{-2z}}$. □

Appendix A.3. Proof of the Equation (3)

Let Z be a random variable distributed as a standardized Fisher's z distribution with the p.d.f given by Equation (1), let μ be a location parameter, and let σ be a scale parameter. The density of $X = \sigma Z + \mu$ can be defined as

$$f_X(x; d_1, d_2, \mu, \sigma) = f_Z\left(\frac{x - \mu}{\sigma}\right) |J(x)|, \quad (\text{A5})$$

where $J(x)$ is the Jacobian of the transformation and is defined as

$$J(x) = \frac{\partial z}{\partial x} = \frac{1}{\sigma}.$$

Therefore, the p.d.f of the Fisher's z distribution can be expressed as

$$f_X(x; d_1, d_2, \mu, \sigma) = \frac{2}{\sigma} \frac{\left(\frac{d_2}{d_1}\right)^{\frac{1}{2}d_2} e^{-d_2\left(\frac{x-\mu}{\sigma}\right)}}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) \left(1 + e^{-2\left(\frac{x-\mu}{\sigma}\right) + \ln \frac{d_2}{d_1}}\right)^{(d_1+d_2)/2}}$$

□

Appendix A.4. Proof of the Equation (4)

Let Z be a random variable distributed as a standardized Fisher's z distribution with the CDF as in Equation (2), let μ be a location parameter, and let σ be a scale parameter. The CDF of $X = \sigma Z + \mu$ can be defined as

$$F_X(x; d_1, d_2, \mu, \sigma) = P(X \leq x) = P(\sigma Z + \mu \leq x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right),$$

Therefore, the CDF of the Fisher's z distribution can be expressed as

$$F_X(x; d_1, d_2, \mu, \sigma) = \frac{1}{B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right)} \int_0^{x^*} t^{\frac{1}{2}d_2-1} (1-t)^{\frac{1}{2}d_1-1} dt,$$

$$\text{where } x^* = \frac{d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}{d_1 + d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}. \quad \square$$

Appendix A.5. Proof of the Equation (5)

Let $I_{x^*}(\cdot)$ be the incomplete beta function ratio; thus, Equation (4) can be expressed as

$$F_X(x; d_1, d_2, \mu, \sigma) = I_{x^*}(d_2/2, d_1/2); \quad x^* = \frac{d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}{d_1 + d_2 e^{-2\left(\frac{x-\mu}{\sigma}\right)}}$$

The value x_p is called the p -quantile of the population, if $P(X \leq x_p) = p$ with $0 \leq p \leq 1$ (Gilchrist 2000, p. 12). Let $I_{x_p}^{-1}(\cdot)$ be the inversion of the incomplete beta function ratio, then

$$\begin{aligned} \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right) &= \frac{d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}}{d_1 + d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}}; \\ \left(d_1 + d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}\right) \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right) &= d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}; \\ (d_1) \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right) &= d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)} - d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)} \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right); \\ (d_1) \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right) &= \left(1 - \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right)\right) d_2 e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}; \\ \frac{d_1 \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right)}{d_2 \left(1 - \left(I_{x_p}^{-1}(d_2/2, d_1/2)\right)\right)} &= e^{-2\left(\frac{x_p-\mu}{\sigma}\right)}; \end{aligned}$$

$$x_p = \mu - \frac{\sigma}{2} \ln \left(\frac{d_1 \left(I^{-1}_{x_p}(d_2/2, d_1/2) \right)}{d_2 \left(1 - \left(I^{-1}_{x_p}(d_2/2, d_1/2) \right) \right)} \right);$$

Interchanging d_1 and d_2 is equivalent to replacing x with $-x$; thus, the QF can also be defined as:

$$x_p = \mu + \frac{\sigma}{2} \ln \left(\frac{d_2 \left(I^{-1}_{x_p}(d_1/2, d_2/2) \right)}{d_1 \left(1 - \left(I^{-1}_{x_p}(d_1/2, d_2/2) \right) \right)} \right)$$

□

Appendix A.6. Proof of the Equation (6)

Let us denote a chi-square random variables with d_1 and d_2 degrees of freedom by V_1^2 and V_2^2 , respectively. Let $\mathcal{P}_{v^*}(\cdot)$ be the incomplete gamma function ratio, and the CDF of V_1^2 and V_2^2 can be defined as

$$F_{V_1^2}(v_1; d_1) = \mathcal{P}_{v_1^*}(d_1/2); \quad v_1^* = \frac{v_1}{2},$$

$$F_{V_2^2}(v_2; d_2) = \mathcal{P}_{v_2^*}(d_2/2); \quad v_2^* = \frac{v_2}{2}.$$

Let $\mathcal{P}^{-1}_{v_p}(\cdot)$ be the inversion of the incomplete gamma function ratio, and then the QF of the V_1^2 and V_2^2 can be defined as

$$v_{1p} = 2\mathcal{P}^{-1}_{v_{1p}}(d_1/2),$$

$$v_{2p} = 2\mathcal{P}^{-1}_{v_{2p}}(d_2/2),$$

The Beta distribution arises naturally as the distribution of $X = V_1^2 / (V_1^2 + V_2^2)$ (Johnson et al. 1995, vol. 2, p. 212); therefore, QF of the Fisher's z distribution can be expressed as

$$x_p = \mu + \frac{\sigma}{2} \ln \left(\frac{d_2 \frac{2\mathcal{P}^{-1}_{v_{1p}}(d_1/2)}{2\mathcal{P}^{-1}_{v_{1p}}(d_1/2) + 2\mathcal{P}^{-1}_{v_{2p}}(d_2/2)}}{d_1 \left(1 - \frac{2\mathcal{P}^{-1}_{v_{1p}}(d_1/2)}{2\mathcal{P}^{-1}_{v_{1p}}(d_1/2) + 2\mathcal{P}^{-1}_{v_{2p}}(d_2/2)} \right)} \right)$$

$$x_p = \mu + \frac{\sigma}{2} \ln \left(\frac{d_2 \left(\mathcal{P}^{-1}_{v_{1p}}(d_1/2) \right)}{d_1 \left(\mathcal{P}^{-1}_{v_{2p}}(d_2/2) \right)} \right)$$

□

Appendix B. Properties of the Fisher's z Distribution and the Proofs

Appendix B.1. Properties of the Fisher's z Distribution

Let X be a random variable distributed as a Fisher's z distribution and let $M_X(\theta)$ be the Moment Generating Function (MGF) of a random variable X . The MGF of the Fisher's z distribution be expressed as

$$M_X(\theta) = E(e^{\theta x}) = e^{\theta \mu} \left(\left(\frac{d_2}{d_1} \right)^{\frac{\sigma \theta}{2}} \right) \frac{\Gamma\left(\frac{1}{2}(d_2 - \sigma \theta)\right) \Gamma\left(\frac{1}{2}(d_1 + \sigma \theta)\right)}{\Gamma\left(\frac{1}{2}d_1\right) \Gamma\left(\frac{1}{2}d_2\right)}, \quad (\text{A6})$$

where $\Gamma(\cdot)$ is a gamma function. Let $K_X(\theta)$ be the Cumulant Generating Function (CGF) of a random variable X . The CGF of the Fisher's z distribution be given by

$$K_X(\theta) = \theta\mu + \frac{\theta\sigma}{2} \ln \frac{d_2}{d_1} + \ln \Gamma\left(\frac{1}{2}(d_1 + \sigma\theta)\right) + \ln \Gamma\left(\frac{1}{2}(d_2 - \sigma\theta)\right) - \ln \Gamma\left(\frac{1}{2}d_1\right) - \ln \Gamma\left(\frac{1}{2}d_2\right). \quad (\text{A7})$$

The coefficient of $\theta!/j!$ in the Taylor expansion of the CGF is the j th cumulant of X and be denoted as κ_j . The j th cumulant, therefore, can be obtained by differentiating the expansion j times and evaluating the result at zero.

$$\kappa_j = \left. \frac{\partial^j}{\partial \theta^j} K_X(\theta) \right|_{\theta=0}; j = 1, 2, 3, \dots$$

The first cumulant of the Fisher's z distribution be defined as

$$\kappa_1 = \mu + \frac{\sigma}{2} \ln \frac{d_2}{d_1} - \frac{\sigma}{2} \psi\left(\frac{1}{2}d_2\right) + \frac{\sigma}{2} \psi\left(\frac{1}{2}d_1\right), \quad (\text{A8})$$

and the j th cumulant would be defined in Equation (A9).

$$\kappa_j = (-1)^j \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{1}{2}d_2\right) + \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{1}{2}d_1\right); j = 2, 3, 4, \dots, \quad (\text{A9})$$

where $\psi(\cdot)$ is the digamma function, $\psi'(\cdot)$ is the trigamma function, $\psi''(\cdot)$ is the tetragamma function, and $\psi^{(3)}(\cdot)$ is the pentagamma function. Generally, $\psi^{(j-1)}(\cdot)$ is the $(j+1)$ -gamma function (Johnson et al. 2005, p. 9). Let the mean and the variance of a random variable X be denoted respectively $E(X)$ and $\text{Var}(X)$. The mean of the Fisher's z distribution is given by

$$E(X) = \kappa_1 = \mu + \frac{\sigma}{2} \ln \frac{d_2}{d_1} - \frac{\sigma}{2} \psi\left(\frac{1}{2}d_2\right) + \frac{\sigma}{2} \psi\left(\frac{1}{2}d_1\right), \quad (\text{A10})$$

and the variance is defined as

$$\text{Var}(X) = \kappa_2 = \left(\frac{\sigma}{2}\right)^2 \left(\psi'\left(\frac{1}{2}d_2\right) + \psi'\left(\frac{1}{2}d_1\right) \right). \quad (\text{A11})$$

On the basis of Equation (A10), it can be concluded that

$$(d_1 = d_2) \longrightarrow (E(X) = \mu);$$

$$(d_1 < d_2) \longrightarrow (E(X) < \mu);$$

$$(d_1 > d_2) \longrightarrow (E(X) > \mu).$$

Let the skewness and the excess kurtosis of a random variable X be denoted, respectively, as γ_1 and γ_2 . The skewness of the Fisher's z distribution is given by

$$\gamma_1 = \frac{\psi''\left(\frac{1}{2}d_1\right) - \psi''\left(\frac{1}{2}d_2\right)}{\left(\psi'\left(\frac{1}{2}d_1\right) + \psi'\left(\frac{1}{2}d_2\right)\right)^{\frac{3}{2}}}, \quad (\text{A12})$$

and the excess kurtosis is

$$\gamma_2 = \frac{\left(\psi^{(3)}\left(\frac{d_2}{2}\right) + \psi^{(3)}\left(\frac{d_1}{2}\right)\right)}{\left(\psi'\left(\frac{d_2}{2}\right) + \psi'\left(\frac{d_1}{2}\right)\right)^2}. \quad (\text{A13})$$

On the basis of Equation (A13), the excess kurtosis value for this distribution is always positive, which shows that the distribution has heavier tails than the Gaussian distribution. Furthermore, based on Equation (A10) through Equation (A13), it can be seen that a change

in the value of the parameter μ only affects the mean of the distribution. It does not affect the variance, skewness, and kurtosis of the distribution.

Appendix B.2. Proof of the Properties

Appendix B.2.1. Proof of the Equation (A6)

If Z is random variables distributed as a standardized Fisher's z , then the MGF of Z is expressed as (Aroian 1941; Johnson et al. 1995)

$$M_Z(\theta) = \left(\left(\frac{d_2}{d_1} \right)^{\frac{\theta}{2}} \right) \frac{\Gamma\left(\frac{1}{2}(d_1 + \theta)\right) \Gamma\left(\frac{1}{2}(d_2 - \theta)\right)}{\Gamma\left(\frac{1}{2}d_1\right) \Gamma\left(\frac{1}{2}d_2\right)}$$

If the random variable Z is transformed to $X = \sigma Z + \mu$, then

$$M_X(\theta) = E(e^{\theta X}) = E(e^{\theta(\mu + \sigma Z)}) = e^{\theta\mu} E(e^{\theta\sigma Z}) = e^{\theta\mu} M_Z(\theta\sigma)$$

$$M_X(\theta) = e^{\theta\mu} \left(\left(\frac{d_2}{d_1} \right)^{\frac{\theta\sigma}{2}} \right) \frac{\Gamma\left(\frac{1}{2}(d_2 - \theta\sigma)\right) \Gamma\left(\frac{1}{2}(d_1 + \theta\sigma)\right)}{\Gamma\left(\frac{1}{2}d_1\right) \Gamma\left(\frac{1}{2}d_2\right)}$$

□

Appendix B.2.2. Proof of the Equation (A7)

The CGF of the random variable X is the natural logarithm of the moment generating function of X (Johnson et al. 2005, p. 54), therefore

$$K_X(\theta) = \ln M_X(\theta) = \ln \left(e^{\theta\mu} \left(\left(\frac{d_2}{d_1} \right)^{\frac{\theta\sigma}{2}} \right) \frac{\Gamma\left(\frac{1}{2}(d_1 + \theta\sigma)\right) \Gamma\left(\frac{1}{2}(d_2 - \theta\sigma)\right)}{\Gamma\left(\frac{1}{2}d_1\right) \Gamma\left(\frac{1}{2}d_2\right)} \right)$$

$$K_X(\theta) = \theta\mu + \frac{\theta\sigma}{2} \ln\left(\frac{d_2}{d_1}\right) + \ln \Gamma\left(\frac{1}{2}(d_1 + \theta\sigma)\right) + \ln \Gamma\left(\frac{1}{2}(d_2 - \theta\sigma)\right) - \ln \Gamma\left(\frac{1}{2}d_1\right) - \ln \Gamma\left(\frac{1}{2}d_2\right)$$

□

Appendix B.2.3. Proof of the Equations (A8) and (A9)

If the random variable X has the CGF in the Equation (A7), then

$$\begin{aligned} \frac{\partial}{\partial \theta} K_X(\theta) &= \mu + \frac{\sigma}{2} \ln\left(\frac{d_2}{d_1}\right) - \frac{\sigma}{2} \psi\left(\frac{d_2 - \theta\sigma}{2}\right) + \frac{\sigma}{2} \psi\left(\frac{d_1 + \theta\sigma}{2}\right) \\ \frac{\partial^2}{\partial \theta^2} K_X(\theta) &= (-1)^2 \left(\frac{\sigma}{2}\right)^2 \psi'\left(\frac{d_2 - \theta\sigma}{2}\right) + \left(\frac{\sigma}{2}\right)^2 \psi'\left(\frac{d_1 + \theta\sigma}{2}\right) \\ \frac{\partial^3}{\partial \theta^3} K_X(\theta) &= (-1)^3 \left(\frac{\sigma}{2}\right)^3 \psi''\left(\frac{d_2 - \theta\sigma}{2}\right) + \left(\frac{\sigma}{2}\right)^3 \psi''\left(\frac{d_1 + \theta\sigma}{2}\right) \\ &\vdots \\ \frac{\partial^j}{\partial \theta^j} K_X(\theta) &= (-1)^j \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{d_2 - \theta\sigma}{2}\right) + \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{d_1 + \theta\sigma}{2}\right). \end{aligned}$$

The first cumulant be defined as

$$\begin{aligned} \kappa_1 &= \frac{\partial}{\partial \theta} K_X(\theta) \Big|_{\theta=0} \\ &= \mu + \frac{\sigma}{2} \ln\left(\frac{d_2}{d_1}\right) - \frac{\sigma}{2} \psi\left(\frac{d_2 - \theta\sigma}{2}\right) + \frac{\sigma}{2} \psi\left(\frac{d_1 + \theta\sigma}{2}\right) \Big|_{\theta=0} \\ &= \mu + \frac{\sigma}{2} \ln\left(\frac{d_2}{d_1}\right) - \frac{\sigma}{2} \psi\left(\frac{d_2}{2}\right) + \frac{\sigma}{2} \psi\left(\frac{d_1}{2}\right) \end{aligned}$$

□

and the j th cumulant be defined as

$$\begin{aligned}\kappa_j &= \left. \frac{\partial^j}{\partial \theta^j} K_X(\theta) \right|_{\theta=0} \\ &= (-1)^j \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{d_2 - \theta\sigma}{2}\right) + \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{d_1 + \theta\sigma}{2}\right) \Big|_{\theta=0} \\ &= (-1)^j \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{1}{2}d_2\right) + \left(\frac{\sigma}{2}\right)^j \psi^{(j-1)}\left(\frac{1}{2}d_1\right); j = 2, 3, 4, \dots\end{aligned}$$

□

Appendix B.2.4. Proof of the Equation (A10)

The mean of the random variable X is the first cumulant (Zelen and Severo 1970), therefore

$$\begin{aligned}E(X) &= \kappa_1 \\ E(X) &= \mu + \frac{\sigma}{2} \ln\left(\frac{d_2}{d_1}\right) - \frac{\sigma}{2} \psi\left(\frac{d_2}{2}\right) + \frac{\sigma}{2} \psi\left(\frac{d_1}{2}\right)\end{aligned}$$

□

Appendix B.2.5. Proof of the Equation (A11)

The variance of the random variable X is the second cumulant (Zelen and Severo 1970), therefore

$$\begin{aligned}\text{Var}(X) &= \kappa_2 \\ \text{Var}(X) &= \left(\frac{\sigma}{2}\right)^2 \psi'\left(\frac{d_2}{2}\right) + \left(\frac{\sigma}{2}\right)^2 \psi'\left(\frac{d_1}{2}\right) = \left(\frac{\sigma}{2}\right)^2 \left(\psi'\left(\frac{d_2}{2}\right) + \psi'\left(\frac{d_1}{2}\right)\right)\end{aligned}$$

□

Appendix B.2.6. Proof of the Equation (A12)

The skewness is formed from the second and third cumulants (Zelen and Severo 1970), namely

$$\begin{aligned}\gamma_1 &= \frac{\kappa_3}{(\kappa_2)^{3/2}} = \frac{\left(\frac{\sigma}{2}\right)^3 \left(-\psi''\left(\frac{d_2}{2}\right) + \psi''\left(\frac{d_1}{2}\right)\right)}{\left(\frac{\sigma}{2}\right)^3 \left(\psi'\left(\frac{d_2}{2}\right) + \psi'\left(\frac{d_1}{2}\right)\right)^{\frac{3}{2}}}; \\ \gamma_1 &= \frac{\psi''\left(\frac{1}{2}d_1\right) - \psi''\left(\frac{1}{2}d_2\right)}{\left(\psi'\left(\frac{1}{2}d_1\right) + \psi'\left(\frac{1}{2}d_2\right)\right)^{\frac{3}{2}}}\end{aligned}$$

□

Appendix B.2.7. Proof of the Equation (A13)

The excess kurtosis can be formed from the second and fourth cumulants (Zelen and Severo 1970), namely

$$\begin{aligned}\gamma_2 &= \frac{\kappa_4}{(\kappa_2)^2} = \frac{\left(\frac{\sigma}{2}\right)^4 \left(\psi^{(3)}\left(\frac{d_2}{2}\right) + \psi^{(3)}\left(\frac{d_1}{2}\right)\right)}{\left(\frac{\sigma}{2}\right)^4 \left(\psi'\left(\frac{d_2}{2}\right) + \psi'\left(\frac{d_1}{2}\right)\right)^2}; \\ \gamma_2 &= \frac{\left(\psi^{(3)}\left(\frac{d_2}{2}\right) + \psi^{(3)}\left(\frac{d_1}{2}\right)\right)}{\left(\psi'\left(\frac{d_2}{2}\right) + \psi'\left(\frac{d_1}{2}\right)\right)^2}\end{aligned}$$

□

Appendix C. Adding the Fisher's z Distribution Functions in Stan

Appendix C.1. Random Numbers Generator Function

We can add the random numbers generator function of Fisher's z distribution (fisher_z_rng) in Stan using the following code,

```
functions{
  real fisher_z_rng(real d1, real d2, real mu, real sigma){
    return(mu+sigma*0.5*log((chi_square_rng(d1)*d2)/(chi_square_rng(d2)*d1)));
  }
}
```

where chi_square_rng(d1) and chi_square_rng(d2) are the chi-square random numbers generator with d_1 and d_2 degrees of freedoms.

Appendix C.2. Log Probability Density Function

We can also add the log probability function of the Fisher's z distribution (fisher_z_lpdf) in Stan using the following code,

```
functions{
  real fisher_z_lpdf(real x, real d1, real d2, real mu, real sigma){
    return (log(2)+0.5*d2*(log(d2)-log(d1))-d2*(x-mu)/sigma-log(sigma)-
    lbeta(0.5*d1,0.5*d2)-(d1+d2)/2*log1p_exp((-2*(x-mu)/sigma)+log(d2)-log(d1)));
  }
}
```

where lbeta(0.5*d1,0.5*d2) is the natural logarithm of the beta function applied to $\frac{1}{2}d_1$ and $\frac{1}{2}d_2$, and log1p_exp((-2*(x-mu)/sigma)+log(d2)-log(d1))) is the natural logarithm of one plus the natural exponentiation of $-2\left(\frac{x-\mu}{\sigma}\right) + \ln d_2 - \ln d_1$.

Appendix D. Priors of Parameters on the ZMAR, TMAR, and GMAR Models in the Simulation Study

Appendix D.1. Scenario 1

Table A1. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 1.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(0.2, 0.1)I(0, \infty);$	$\nu_1 \sim t_3(2.1, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(23, 0.1)I(0, \infty)$
$d_{21} \sim t_3(10, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(14.6, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.07, 0.1)$
$\sigma_1 \sim t_3(5, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.62, 0.1)$	$\sigma_2 \sim t_3(33, 0.1)I(0, \infty)$
$\phi_{1.1} \sim N(-0.6, 0.1)$	$\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(-0.47, 0.1)$
$d_{12} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_2 \sim t_3(8.1, 0.1)I(0, \infty)$	$\sigma_3 \sim t_3(11, 0.1)I(0, \infty)$
$d_{22} \sim t_3(1, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.2, 0.1)$	$\phi_{3.1} \sim N(0.6, 0.1)$
$\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$	$\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$	$\sigma_4 \sim t_3(2, 0.1)I(0, \infty)$
$\phi_{2.1} \sim N(0.2, 0.1)$	$\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$	$\phi_{4.1} \sim N(0.8, 0.1)$
$d_{13} \sim t_3(30, 0.1)I(0, \infty);$	$\phi_{3.1} \sim N(0.84, 0.1)$	$\eta \sim \text{Dir}(1, 1, 1)$
$d_{23} \sim t_3(30, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$	
$\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$		
$\phi_{3.1} \sim N(0.7, 0.1)$		
$\eta \sim \text{Dir}(1, 1, 1)$		

Appendix D.2. Scenario 2

Table A2. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 2.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(20, 0.1)I(0, \infty);$	$\nu_1 \sim t_3(2.19, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(5.01, 0.1)I(0, \infty)$
$d_{21} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(8, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.59, 0.1)$
$\sigma_1 \sim t_3(5, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.62, 0.1)$	$\sigma_2 \sim t_3(2.69, 0.1)I(0, \infty)$
$\phi_{1.1} \sim N(-0.6, 0.1)$	$\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.64, 0.1)$
$d_{12} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_2 \sim t_3(8.12, 0.1)I(0, \infty)$	$\sigma_3 \sim t_3(14.49, 0.1)I(0, \infty)$
$d_{22} \sim t_3(1, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.2, 0.1)$	$\phi_{3.1} \sim N(0.57, 0.1)$
$\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$	$\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$
$\phi_{2.1} \sim N(0.2, 0.1)$	$\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$	
$d_{13} \sim t_3(30, 0.1)I(0, \infty);$	$\phi_{3.1} \sim N(0.84, 0.1)$	
$d_{23} \sim t_3(30, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$	
$\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$		
$\phi_{3.1} \sim N(0.7, 0.1)$		
$\eta \sim \text{Dir}(1, 1, 1)$		

Appendix D.3. Scenario 3

Table A3. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 3.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(0.5, 0.1)I(0, \infty);$	$\nu_1 \sim t_3(2.1, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(10, 0.1)I(0, \infty)$
$d_{21} \sim t_3(0.5, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(14.6, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.6, 0.1)$
$\sigma_1 \sim t_3(5, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.62, 0.1)$	$\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$
$\phi_{1.1} \sim N(-0.6, 0.1)$	$\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.24, 0.1)$
$d_{12} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_2 \sim t_3(8.1, 0.1)I(0, \infty)$	$\sigma_3 \sim t_3(2, 0.1)I(0, \infty)$
$d_{22} \sim t_3(1, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.2, 0.1)$	$\phi_{3.1} \sim N(0.74, 0.1)$
$\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$	$\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$
$\phi_{2.1} \sim N(0.2, 0.1)$	$\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$	
$d_{13} \sim t_3(30, 0.1)I(0, \infty);$	$\phi_{3.1} \sim N(0.84, 0.1)$	
$d_{23} \sim t_3(30, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$	
$\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$		
$\phi_{3.1} \sim N(0.7, 0.1)$		
$\eta \sim \text{Dir}(1, 1, 1)$		

Appendix D.4. Scenario 4

Table A4. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 4.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(3, 0.1)I(0, \infty);$	$\nu_1 \sim t_3(2.19, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(10, 0.1)I(0, \infty)$
$d_{21} \sim t_3(10, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(3, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.6, 0.1)$
$\sigma_1 \sim t_3(8, 0.1)I(0, \infty)$	$\phi_{1.1} \sim N(-0.62, 0.1)$	$\sigma_2 \sim t_3(11.95, 0.1)I(0, \infty)$
$\phi_{1.1} \sim N(-0.6, 0.1)$	$\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.24, 0.1)$
$d_{12} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_2 \sim t_3(6.8, 0.1)I(0, \infty)$	$\sigma_3 \sim t_3(31, 0.1)I(0, \infty)$
$d_{22} \sim t_3(1, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.2, 0.1)$	$\phi_{3.1} \sim N(0.74, 0.1)$
$\sigma_2 \sim t_3(5, 0.1)I(0, \infty)$	$\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$
$\phi_{2.1} \sim N(0.2, 0.1)$	$\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$	
$d_{13} \sim t_3(30, 0.1)I(0, \infty);$	$\phi_{3.1} \sim N(0.84, 0.1)$	
$d_{23} \sim t_3(30, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$	
$\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$		
$\phi_{3.1} \sim N(0.7, 0.1)$		
$\eta \sim \text{Dir}(1, 1, 1)$		

Appendix D.5. Scenario 5

Table A5. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 5.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(7, 0.1)I(0, \infty);$ $d_{21} \sim t_3(10.04, 0.5)I(0, \infty);$ $\sigma_1 \sim t_3(7.99, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.6, 0.1)$ $d_{12} \sim t_3(1.01, 0.1)I(0, \infty);$ $d_{22} \sim t_3(1.26, 0.5)I(0, \infty);$ $\sigma_2 \sim t_3(4.93, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.2, 0.1)$ $d_{13} \sim t_3(30, 0.1)I(0, \infty);$ $d_{23} \sim t_3(30, 0.5)I(0, \infty);$ $\sigma_3 \sim t_3(10.07, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.7, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\nu_1 \sim t_3(2.22, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(2.38, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.62, 0.1)$ $\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(6.64, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.20, 0.1)$ $\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(1.7, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.74, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\sigma_1 \sim t_3(2.86, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.47, 0.1)$ $\sigma_2 \sim t_3(3.49, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(-0.55, 0.1)$ $\sigma_3 \sim t_3(10.94, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.6, 0.1)$ $\sigma_4 \sim t_3(2, 0.1)I(0, \infty)$ $\phi_{4.1} \sim N(0.8, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1, 1)$

Appendix D.6. Scenario 6

Table A6. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 6.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(10, 0.1)I(0, \infty);$ $d_{21} \sim t_3(10, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(8, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.6, 0.1)$ $d_{12} \sim t_3(1, 0.1)I(0, \infty);$ $d_{22} \sim t_3(1, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(5, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.2, 0.1)$ $d_{13} \sim t_3(29.96, 0.1)I(0, \infty);$ $d_{23} \sim t_3(29.83, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.7, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\nu_1 \sim t_3(2.3, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(2.38, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.62, 0.1)$ $\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(6.64, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.2, 0.1)$ $\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(1.7, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.74, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\sigma_1 \sim t_3(2.9, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.54, 0.1)$ $\sigma_2 \sim t_3(1.98, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.66, 0.1)$ $\sigma_3 \sim t_3(9.53, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.61, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$

Appendix D.7. Scenario 7

Table A7. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 7.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(0.5, 0.1)I(0, \infty);$ $d_{21} \sim t_3(0.5, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(5, 0.1)I(0, \infty)$ $\phi_{1.0} \sim N(-1, 0.1)$ $\phi_{1.1} \sim N(-0.6, 0.1)$ $d_{12} \sim t_3(1, 0.1)I(0, \infty);$ $d_{22} \sim t_3(1, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.2, 0.1)$ $d_{13} \sim t_3(30, 0.1)I(0, \infty);$ $d_{23} \sim t_3(30, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.7, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\nu_1 \sim t_3(2.1, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(14.6, 0.1)I(0, \infty)$ $\phi_{1.0} \sim N(-1, 0.1)$ $\phi_{1.1} \sim N(-0.62, 0.1)$ $\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(8.1, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.2, 0.1)$ $\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.84, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\sigma_1 \sim t_3(15.96, 0.1)I(0, \infty)$ $\phi_{1.0} \sim N(-1, 0.1)$ $\phi_{1.1} \sim N(-0.58, 0.1)$ $\sigma_2 \sim t_3(8.05, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.18, 0.1)$ $\sigma_3 \sim t_3(1.9, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.7, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$

Appendix D.8. Scenario 8

Table A8. Priors of parameters on the ZMAR, TMAR, and GMAR models in the Scenario 8.

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(0.5, 0.1)I(0, \infty);$	$\nu_1 \sim t_3(2.1, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(15.96, 0.1)I(0, \infty)$
$d_{21} \sim t_3(0.5, 0.1)I(0, \infty);$	$\sigma_1 \sim t_3(14.57, 0.1)I(0, \infty)$	$\phi_{1.0} \sim N(-20, 0.1)$
$\sigma_1 \sim t_3(5, 0.1)I(0, \infty)$	$\phi_{1.0} \sim N(-20, 0.1)$	$\phi_{1.1} \sim N(-0.58, 0.1)$
$\phi_{1.0} \sim N(-20, 0.1)$	$\phi_{1.1} \sim N(-0.62, 0.1)$	$\sigma_2 \sim t_3(8.05, 0.1)I(0, \infty)$
$\phi_{1.1} \sim N(-0.6, 0.1)$	$\nu_2 \sim t_3(3.94, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.18, 0.1)$
$d_{12} \sim t_3(1, 0.1)I(0, \infty);$	$\sigma_2 \sim t_3(8.1, 0.1)I(0, \infty)$	$\sigma_3 \sim t_3(1.9, 0.1)I(0, \infty)$
$d_{22} \sim t_3(1, 0.1)I(0, \infty);$	$\phi_{2.1} \sim N(0.2, 0.1)$	$\phi_{3.1} \sim N(0.7, 0.1)$
$\sigma_2 \sim t_3(8, 0.1)I(0, \infty)$	$\nu_3 \sim t_3(9.66, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$
$\phi_{2.1} \sim N(0.2, 0.1)$	$\sigma_3 \sim t_3(1.62, 0.1)I(0, \infty)$	
$d_{13} \sim t_3(30, 0.1)I(0, \infty);$	$\phi_{3.1} \sim N(0.84, 0.1)$	
$d_{23} \sim t_3(30, 0.1)I(0, \infty);$	$\eta \sim \text{Dir}(1, 1, 1)$	
$\sigma_3 \sim t_3(10, 0.1)I(0, \infty)$		
$\phi_{3.1} \sim N(0.7, 0.1)$		
$\eta \sim \text{Dir}(1, 1, 1)$		

Appendix E. Summary of Posterior Inferences for the Simulation Study

Appendix E.1. ZMAR Model

Table A9. Summary of posterior inferences for the ZMAR model, Scenario 1 to Scenario 8.

Parameters	Scenario 1					Scenario 2					Scenario 3				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.30	0.22	0.37	7191	1	0.30	0.23	0.37	3014	1	0.37	0.26	0.47	5026	1
eta[2]	0.35	0.28	0.44	6327	1	0.35	0.27	0.43	2580	1	0.25	0.14	0.37	5112	1
eta[3]	0.35	0.30	0.40	13068	1	0.35	0.30	0.41	4751	1	0.38	0.33	0.43	12941	1
sigma[1]	4.99	4.68	5.28	5655	1	5.01	4.73	5.34	3009	1	5.03	4.76	5.39	4640	1
sigma[2]	8.01	7.73	8.33	3849	1	8.00	7.70	8.28	3630	1	8.01	7.72	8.32	2937	1
sigma[3]	9.87	8.86	10.19	1171	1	10.11	9.81	10.86	606	1	9.98	9.65	10.28	7578	1
d1[1]	0.24	0.20	0.29	8498	1	20.00	19.70	20.33	3898	1	0.46	0.37	0.56	10341	1
d1[2]	1.00	0.75	1.27	7829	1	1.00	0.84	1.16	3979	1	0.97	0.77	1.19	9542	1
d1[3]	30.01	29.71	30.37	1180	1	30.00	29.71	30.31	3380	1	30.00	29.67	30.34	7281	1
d2[1]	10.00	9.67	10.35	2476	1	0.99	0.80	1.19	4081	1	0.47	0.38	0.58	9300	1
d2[2]	0.97	0.82	1.13	10248	1	1.04	0.87	1.25	3578	1	1.02	0.80	1.29	7985	1
d2[3]	30.01	29.70	30.35	6690	1	29.99	29.65	30.28	2935	1	30.00	29.68	30.31	4824	1
phi1[1]	-0.58	-0.69	-0.49	8111	1	-0.61	-0.67	-0.54	4283	1	-0.65	-0.79	-0.51	8638	1
phi2[1]	0.20	0.12	0.28	9659	1	0.24	0.10	0.39	3885	1	0.17	0.04	0.31	10220	1
phi3[1]	0.70	0.69	0.72	11783	1	0.69	0.65	0.73	4633	1	0.70	0.68	0.72	15160	1

Parameters	Scenario 4					Scenario 5					Scenario 6				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.29	0.19	0.38	3374	1	0.33	0.25	0.41	9017	1	0.35	0.28	0.42	10507	1
eta[2]	0.34	0.24	0.45	3208	1	0.32	0.23	0.41	7678	1	0.27	0.20	0.35	8527	1
eta[3]	0.37	0.30	0.44	6006	1	0.35	0.28	0.41	14769	1	0.38	0.31	0.44	13451	1
sigma[1]	8.00	7.71	8.31	2661	1	7.95	7.56	8.24	3108	1	7.98	7.64	8.26	4814	1
sigma[2]	4.99	4.71	5.25	5046	1	4.94	4.64	5.28	5420	1	5.03	4.76	5.37	4642	1
sigma[3]	9.99	9.69	10.28	3622	1	10.06	9.74	10.34	4557	1	10.00	9.69	10.30	7130	1
d1[1]	3.00	2.68	3.30	3017	1	7.01	6.71	7.37	2621	1	10.00	9.67	10.34	3430	1
d1[2]	0.99	0.82	1.16	4244	1	0.99	0.82	1.16	9146	1	0.94	0.75	1.12	9040	1
d1[3]	30.01	29.70	30.36	1772	1	30.00	29.68	30.34	4481	1	29.96	29.65	30.27	3118	1
d2[1]	10.00	9.69	10.31	1204	1	10.19	8.69	12.11	2936	1	10.13	8.71	11.93	5848	1
d2[2]	1.16	0.98	1.40	4328	1	1.29	0.98	1.64	9659	1	1.13	0.84	1.45	9806	1
d2[3]	30.01	29.70	30.37	2704	1	30.03	28.47	31.65	6035	1	29.81	28.13	31.41	6694	1
phi1[1]	-0.61	-0.72	-0.51	4860	1	-0.59	-0.68	-0.50	12297	1	-0.58	-0.66	-0.50	11226	1
phi2[1]	0.16	0.00	0.33	4725	1	0.19	0.04	0.34	13526	1	0.21	0.04	0.37	11801	1
phi3[1]	0.70	0.64	0.75	4788	1	0.74	0.69	0.79	13276	1	0.70	0.66	0.75	12584	1

Table A9. Cont.

Parameters	Scenario 7					Scenario 8				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.37	0.26	0.48	6313	1	0.26	0.18	0.35	5531	1
eta[2]	0.24	0.13	0.36	6050	1	0.36	0.27	0.46	5302	1
eta[3]	0.38	0.33	0.43	14608	1	0.37	0.32	0.42	15821	1
sigma[1]	5.03	4.76	5.39	5870	1	5.01	4.71	5.34	5182	1
sigma[2]	8.02	7.71	8.38	4889	1	8.04	7.75	8.44	3010	1
sigma[3]	9.98	9.66	10.25	8727	1	9.98	9.62	10.27	4889	1
d1[1]	0.46	0.38	0.56	9767	1	0.64	0.49	0.84	7428	1
d1[2]	0.97	0.75	1.19	11033	1	0.97	0.76	1.19	8864	1
d1[3]	30.00	29.69	30.32	6325	1	30.00	29.68	30.32	6680	1
d2[1]	0.47	0.38	0.58	8485	1	0.36	0.26	0.53	5370	1
d2[2]	1.02	0.79	1.30	8732	1	0.96	0.77	1.16	7925	1
d2[3]	30.00	29.70	30.31	9132	1	30.00	29.68	30.34	3818	1
phi10	−1.00	−1.20	−0.80	13066	1	−19.99	−20.18	−19.80	9972	1
phi1[1]	−0.64	−0.78	−0.50	8240	1	−0.61	−0.77	−0.45	9056	1
phi2[1]	0.17	0.04	0.31	9302	1	−0.01	−0.15	0.14	7869	1
phi3[1]	0.70	0.68	0.72	12524	1	0.70	0.68	0.71	10541	1

Note: eta[1] = η_1 ; eta[2] = η_2 ; eta[3] = η_3 ; sigma[1] = σ_1 ; sigma[2] = σ_2 ; sigma[3] = σ_3 ; d1[1] = d_{11} ; d1[2] = d_{12} ; d1[3] = d_{13} ; d2[1] = d_{21} ; d2[2] = d_{22} ; d2[3] = d_{23} ; phi10 = $\phi_{1.0}$; phi1[1] = $\phi_{1.1}$; phi2[1] = $\phi_{2.1}$; phi3[1] = $\phi_{3.1}$.

Appendix E.2. TMAR Model

Table A10. Summary of posterior inferences for the TMAR model, Scenario 1 to Scenario 8.

Parameters	Scenario 1					Scenario 2					Scenario 3				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.32	0.24	0.40	7063	1	0.37	0.27	0.45	5915	1	0.41	0.30	0.51	2372	1
eta[2]	0.34	0.25	0.43	6220	1	0.26	0.15	0.37	4889	1	0.21	0.10	0.32	2247	1
eta[3]	0.34	0.29	0.39	13537	1	0.38	0.31	0.45	8252	1	0.38	0.33	0.44	5455	1
sigma[1]	14.61	14.30	14.94	5674	1	5.29	4.90	5.57	2536	1	14.57	14.17	14.87	1277	1
sigma[2]	8.14	7.86	8.51	4161	1	8.13	7.84	8.46	5277	1	8.11	7.80	8.45	825	1
sigma[3]	1.50	1.26	1.69	7778	1	1.96	1.65	2.37	6784	1	1.66	1.51	1.85	3690	1
nu[1]	2.16	1.92	2.49	7307	1	2.24	2.00	2.57	5840	1	6.84	2.47	14.35	1137	1
nu[2]	3.91	3.53	4.22	3185	1	3.94	3.64	4.24	7714	1	3.95	3.66	4.27	3009	1
nu[3]	9.66	9.34	9.97	8297	1	9.66	9.33	9.98	7186	1	9.67	9.33	10.02	3084	1
phi1[1]	−0.52	−0.66	−0.37	8479	1	−0.60	−0.68	−0.52	7125	1	−0.61	−0.76	−0.46	3324	1
phi2[1]	0.21	0.14	0.27	10361	1	0.26	0.12	0.41	8110	1	0.18	0.05	0.30	3917	1
phi3[1]	0.71	0.69	0.72	10525	1	0.70	0.66	0.74	8983	1	0.70	0.68	0.72	4471	1
Parameters	Scenario 4					Scenario 5					Scenario 6				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.31	0.22	0.41	5102	1	0.39	0.30	0.47	4914	1	0.41	0.33	0.49	7570	1
eta[2]	0.29	0.19	0.40	3605	1	0.23	0.14	0.33	3055	1	0.18	0.11	0.26	5985	1
eta[3]	0.40	0.33	0.47	8780	1	0.39	0.32	0.46	6358	1	0.41	0.35	0.48	13283	1
sigma[1]	2.96	2.61	3.22	2891	1	2.36	2.08	2.60	5918	1	2.34	2.05	2.56	6321	1
sigma[2]	6.67	5.75	6.99	1343	1	6.49	5.18	6.85	1166	1	6.62	6.29	6.90	5457	1
sigma[3]	1.68	1.50	1.92	8383	1	1.74	1.56	1.97	6751	1	1.77	1.60	1.99	8093	1
nu[1]	2.27	2.01	2.75	1310	1	2.31	2.04	2.85	1808	1	2.36	2.09	2.80	2908	1
nu[2]	3.99	3.73	4.41	3288	1	3.97	3.69	4.34	4246	1	3.96	3.66	4.30	5643	1
nu[3]	9.66	9.34	9.98	6284	1	9.66	9.33	9.99	4339	1	9.66	9.35	9.97	11682	1
phi1[1]	−0.61	−0.70	−0.51	8823	1	−0.60	−0.67	−0.51	9092	1	−0.57	−0.65	−0.50	9497	1
phi2[1]	0.16	0.00	0.32	9719	1	0.20	0.04	0.36	9273	1	0.22	0.05	0.39	10160	1
phi3[1]	0.70	0.65	0.75	10272	1	0.73	0.68	0.78	9758	1	0.70	0.65	0.75	11448	1
Parameters	Scenario 7					Scenario 8									
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat					
eta[1]	0.42	0.32	0.52	5019	1	0.23	0.15	0.30	2952	1					
eta[2]	0.20	0.09	0.31	4956	1	0.40	0.32	0.49	3629	1					
eta[3]	0.38	0.33	0.43	12590	1	0.37	0.32	0.42	12821	1					
sigma[1]	14.57	14.20	14.86	4356	1	14.49	13.92	14.80	551	1					
sigma[2]	8.11	7.82	8.41	9509	1	9.01	8.00	11.53	1391	1					
sigma[3]	1.65	1.48	1.83	10687	1	1.64	1.47	1.82	10546	1					
nu[1]	6.99	2.80	14.63	7043	1	2.73	1.99	6.59	1412	1					
nu[2]	3.95	3.63	4.28	5590	1	3.96	3.66	4.31	1783	1					
nu[3]	9.66	9.35	10.01	8211	1	9.66	9.34	9.97	4064	1					
phi10	−1.00	−1.20	−0.81	12088	1	−19.99	−20.18	−19.79	11087	1					
phi1[1]	−0.60	−0.75	−0.45	7176	1	−0.64	−0.80	−0.47	9946	1					
phi2[1]	0.18	0.04	0.31	10476	1	−0.03	−0.16	0.10	7696	1					
phi3[1]	0.70	0.68	0.72	11455	1	0.70	0.68	0.71	12521	1					

Note: eta[1] = η_1 ; eta[2] = η_2 ; eta[3] = η_3 ; sigma[1] = σ_1 ; sigma[2] = σ_2 ; sigma[3] = σ_3 ; nu[1] = ν_1 ; nu[2] = ν_2 ; nu[3] = ν_3 ; phi10 = $\phi_{1.0}$; phi1[1] = $\phi_{1.1}$; phi2[1] = $\phi_{2.1}$; phi3[1] = $\phi_{3.1}$.

Appendix E.3. GMAR Model

Table A11. Summary of posterior inferences for the GMAR model, Scenario 1 to Scenario 8.

Parameters	Scenario 1					Scenario 2					Scenario 3				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.17	0.03	0.31	2115	1	0.32	0.26	0.39	3938	1	0.45	0.37	0.53	5106	1
eta[2]	0.10	0.02	0.20	2355	1	0.44	0.38	0.50	6055	1	0.16	0.07	0.25	4794	1
eta[3]	0.37	0.27	0.47	2998	1	0.24	0.18	0.30	3828	1	0.39	0.34	0.44	14297	1
eta[4]	0.35	0.29	0.40	4562	1	-	-	-	-	-	-	-	-	-	-
sigma[1]	22.99	22.66	23.30	2544	1	5.02	4.74	5.32	3411	1	16.27	14.82	17.95	7206	1
sigma[2]	33.00	32.71	33.32	4314	1	2.62	2.37	2.83	3843	1	8.01	7.72	8.34	8301	1
sigma[3]	10.98	10.59	11.29	2125	1	14.51	14.23	14.86	2838	1	1.90	1.69	2.09	8779	1
sigma[4]	1.65	1.32	1.97	3881	1	-	-	-	-	-	-	-	-	-	-
phi1[1]	-0.10	-0.27	0.07	5071	1	-0.59	-0.67	-0.52	5206	1	-0.58	-0.72	-0.44	7440	1
phi2[1]	-0.46	-0.66	-0.27	4691	1	0.68	0.63	0.72	5424	1	0.20	0.07	0.34	9230	1
phi3[1]	0.23	0.14	0.32	4408	1	0.42	0.26	0.58	5452	1	0.70	0.68	0.72	8926	1
phi4[1]	0.70	0.69	0.72	4702	1	-	-	-	-	-	-	-	-	-	-

Parameters	Scenario 4					Scenario 5					Scenario 6				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.47	0.37	0.56	5580	1	0.20	0.02	0.41	4372	1	0.41	0.35	0.48	4103	1
eta[2]	0.10	0.05	0.17	5157	1	0.25	0.03	0.46	4182	1	0.41	0.35	0.48	4209	1
eta[3]	0.43	0.36	0.51	7119	1	0.14	0.10	0.19	8663	1	0.17	0.13	0.23	4222	1
eta[4]	-	-	-	-	-	0.40	0.34	0.47	9784	1	-	-	-	-	-
sigma[1]	5.17	4.27	6.04	5779	1	2.87	2.58	3.20	7042	1	2.87	2.62	3.09	3711	1
sigma[2]	11.92	11.55	12.22	2196	1	3.50	3.23	3.81	5643	1	1.97	1.78	2.15	4277	1
sigma[3]	2.00	1.64	2.41	6005	1	10.92	10.58	11.19	8724	1	9.54	9.23	9.85	3810	1
sigma[4]	-	-	-	-	-	2.00	1.80	2.20	9818	1	-	-	-	-	-
phi1[1]	-0.45	-0.56	-0.33	6510	1	-0.49	-0.65	-0.28	6246	1	-0.54	-0.62	-0.46	5372	1
phi2[1]	0.25	0.06	0.43	7870	1	-0.49	-0.67	-0.30	5778	1	0.69	0.64	0.74	4489	1
phi3[1]	0.68	0.62	0.73	6904	1	0.55	0.36	0.73	9482	1	0.55	0.37	0.72	4960	1
phi4[1]	-	-	-	-	-	0.73	0.67	0.78	9638	1	-	-	-	-	-

Parameters	Scenario 7					Scenario 8				
	mean	2.5%	97.5%	n_eff	Rhat	mean	2.5%	97.5%	n_eff	Rhat
eta[1]	0.46	0.38	0.54	5739	1	0.20	0.13	0.27	9210	1
eta[2]	0.15	0.07	0.24	5976	1	0.37	0.32	0.42	13255	1
eta[3]	0.39	0.33	0.44	12235	1	0.43	0.35	0.51	8512	1
sigma[1]	15.99	15.71	16.37	4308	1	15.93	15.55	16.24	2863	1
sigma[2]	8.06	7.76	8.41	9217	1	1.81	1.54	2.12	9746	1
sigma[3]	1.84	1.64	2.00	9968	1	14.75	13.09	16.56	11288	1
phi10	-1.00	-1.20	-0.81	6321	1	-19.98	-20.18	-19.78	11821	1
phi1[1]	-0.56	-0.70	-0.43	9013	1	-0.70	-0.85	-0.54	10264	1
phi2[1]	0.18	0.05	0.31	6200	1	0.69	0.67	0.71	11262	1
phi3[1]	0.70	0.68	0.72	10471	1	0.18	0.05	0.32	10450	1

Note:eta[1] = η_1 ; eta[2] = η_2 ; eta[3] = η_3 ; eta[4] = η_4 ; sigma[1] = σ_1 ; sigma[2] = σ_2 ; sigma[3] = σ_3 ; sigma[4] = σ_4 ; phi10 = $\phi_{1.0}$; phi1[1] = $\phi_{1.1}$; phi2[1] = $\phi_{2.1}$; phi3[1] = $\phi_{3.1}$.

Appendix F. Priors of Some Parameters on the Models in the IBM Stock Prices and the Brent Crude Oil Prices

Appendix F.1. IBM Stock Prices (First-Differenced Series)

Table A12. Priors of some parameters on the models in the IBM stock prices (first-differenced series).

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(1.94, 0.1)I(0, \infty)$; $d_{21} \sim t_3(3.90, 0.1)I(0, \infty)$; $\sigma_1 \sim t_3(28.28, 0.1)I(0, \infty)$; $d_{12} \sim t_3(1.79, 0.1)I(0, \infty)$; $d_{22} \sim t_3(6.40, 0.1)I(0, \infty)$; $\sigma_2 \sim t_3(9.81, 0.1)I(0, \infty)$; $\phi_{2.1} \sim N(0.61, 0.1)$; $d_{13} \sim t_3(4.92, 0.1)I(0, \infty)$; $d_{23} \sim t_3(1.66, 0.1)I(0, \infty)$; $\sigma_3 \sim t_3(6.34, 0.1)I(0, \infty)$; $\phi_{3.1} \sim N(-0.28, 0.1)$; $\eta \sim \text{Dir}(1, 1, 1)$	$v_1 \sim t_3(12.52, 0.1)I(0, \infty)$; $\sigma_1 \sim t_3(5.01, 0.1)I(0, \infty)$; $\phi_{1.1} \sim N(-0.29, 0.1)$; $v_2 \sim t_3(10.77, 0.1)I(0, \infty)$; $\sigma_2 \sim t_3(5.82, 0.1)I(0, \infty)$; $\phi_{2.1} \sim N(0.68, 0.1)$; $v_3 \sim t_3(14.03, 0.1)I(0, \infty)$; $\sigma_3 \sim t_3(25.02, 0.1)I(0, \infty)$; $\eta \sim \text{Dir}(1, 1, 1)$	$\sigma_1 \sim t_3(4.8227, 0.1)I(0, \infty)$; $\phi_{1.1} \sim N(-0.31, 0.1)$; $\sigma_2 \sim t_3(6.0082, 0.1)I(0, \infty)$; $\phi_{2.1} \sim N(0.67, 0.1)$; $\sigma_3 \sim t_3(19.04, 0.1)I(0, \infty)$; $\eta \sim \text{Dir}(1, 1, 1)$

Appendix F.2. Brent Crude Oil Prices (First-Differenced Series)

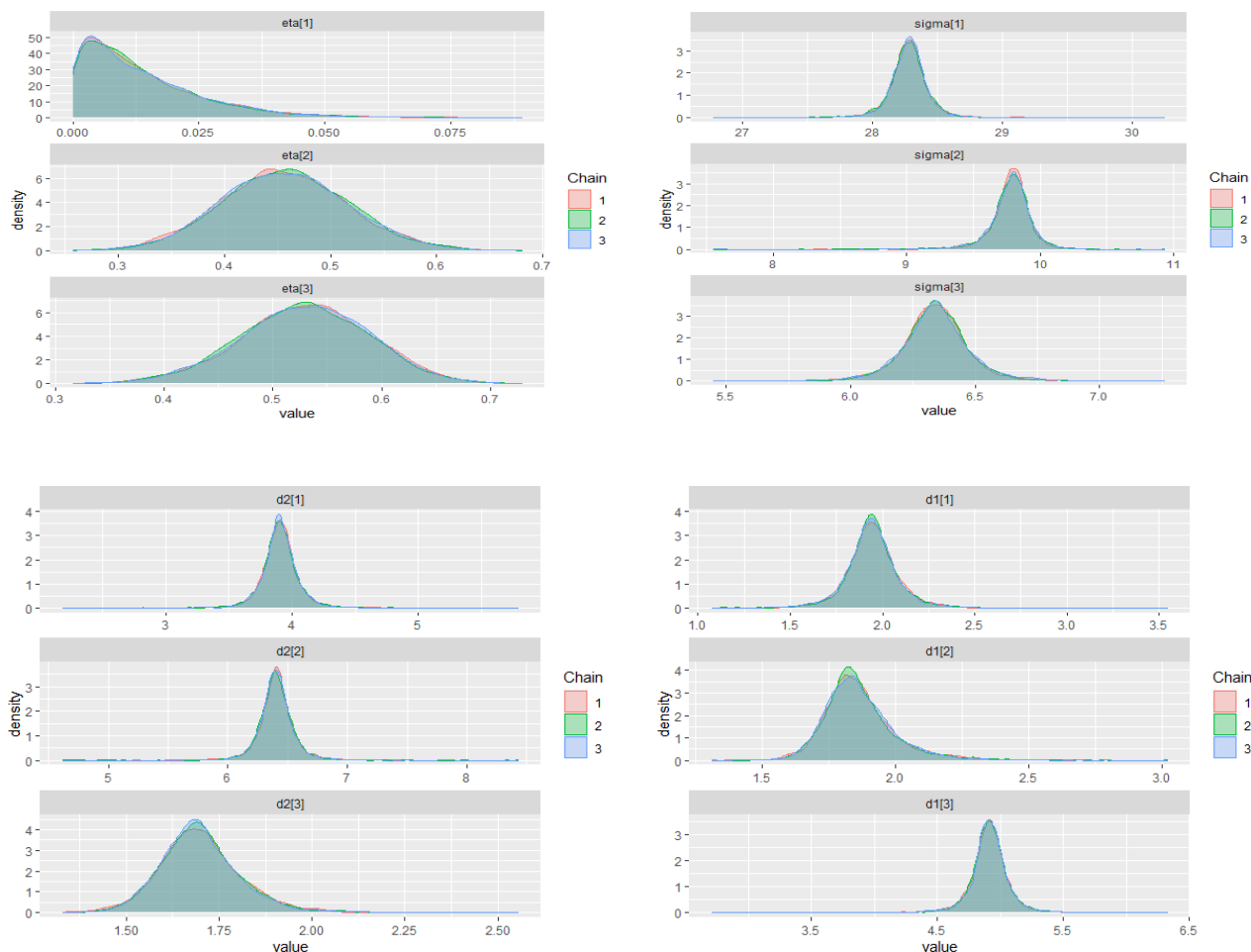
Table A13. Priors of some parameters on the models in the Brent crude oil prices (first-differenced series).

ZMAR	TMAR	GMAR
$d_{11} \sim t_3(13.21, 0.1)I(0, \infty);$ $d_{21} \sim t_3(4.47, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(5.18, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(-0.33, 0.1)$ $d_{12} \sim t_3(1.05, 0.1)I(0, \infty);$ $d_{22} \sim t_3(1.90, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(2.81, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.67, 0.1)$ $\phi_{2.2} \sim N(-0.31, 0.1)$ $d_{13} \sim t_3(10.07, 0.1)I(0, \infty);$ $d_{23} \sim t_3(4.31, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(7.24, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.69, 0.1)$ $\phi_{3.2} \sim N(0.69, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$\nu_1 \sim t_3(14.98, 0.1)I(0, \infty);$ $\sigma_1 \sim t_3(5.05, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(0.63, 0.1)$ $\phi_{1.2} \sim N(-0.39, 0.1)$ $\nu_2 \sim t_3(12.08, 0.1)I(0, \infty);$ $\sigma_2 \sim t_3(1.61, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(-0.26, 0.1)$ $\nu_3 \sim t_3(4.33, 0.1)I(0, \infty);$ $\sigma_3 \sim t_3(1.89, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.55, 0.1)$ $\phi_{3.2} \sim N(0.85, 0.1)$ $\phi_{3.3} \sim N(-0.27, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$	$t_3(4.11, 0.1)I(0, \infty)$ $\phi_{1.1} \sim N(0.18, 0.1)$ $\phi_{1.2} \sim N(-0.34, 0.1)$ $\phi_{1.3} \sim N(-0.16, 0.1)$ $\sigma_2 \sim t_3(1.70, 0.1)I(0, \infty)$ $\phi_{2.1} \sim N(0.44, 0.1)$ $\phi_{2.2} \sim N(-0.16, 0.1)$ $\phi_{2.3} \sim N(0.33, 0.1)$ $\sigma_3 \sim t_3(1.94, 0.1)I(0, \infty)$ $\phi_{3.1} \sim N(0.57, 0.1)$ $\phi_{3.2} \sim N(0.90, 0.1)$ $\phi_{3.3} \sim N(-0.28, 0.1)$ $\eta \sim \text{Dir}(1, 1, 1)$

Appendix G. Posterior Density Plots for Some Parameters on the Models in the IBM Stock Prices and the Brent Crude Oil Prices

Appendix G.1. IBM Stock Prices (First-Differenced Series)

Appendix G.1.1. ZMAR Model

**Figure A1.** Cont.

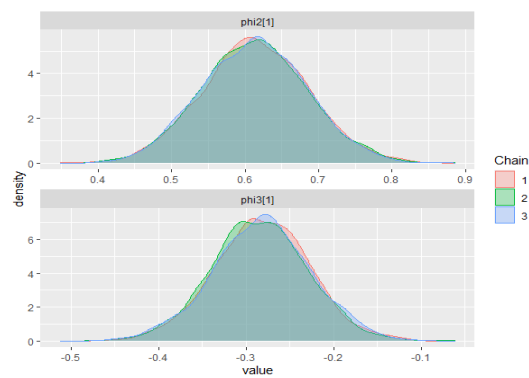


Figure A1. Posterior density plots for some parameters on the ZMAR model in the IBM stock prices (first-differenced series).

Appendix G.1.2. TMAR Model

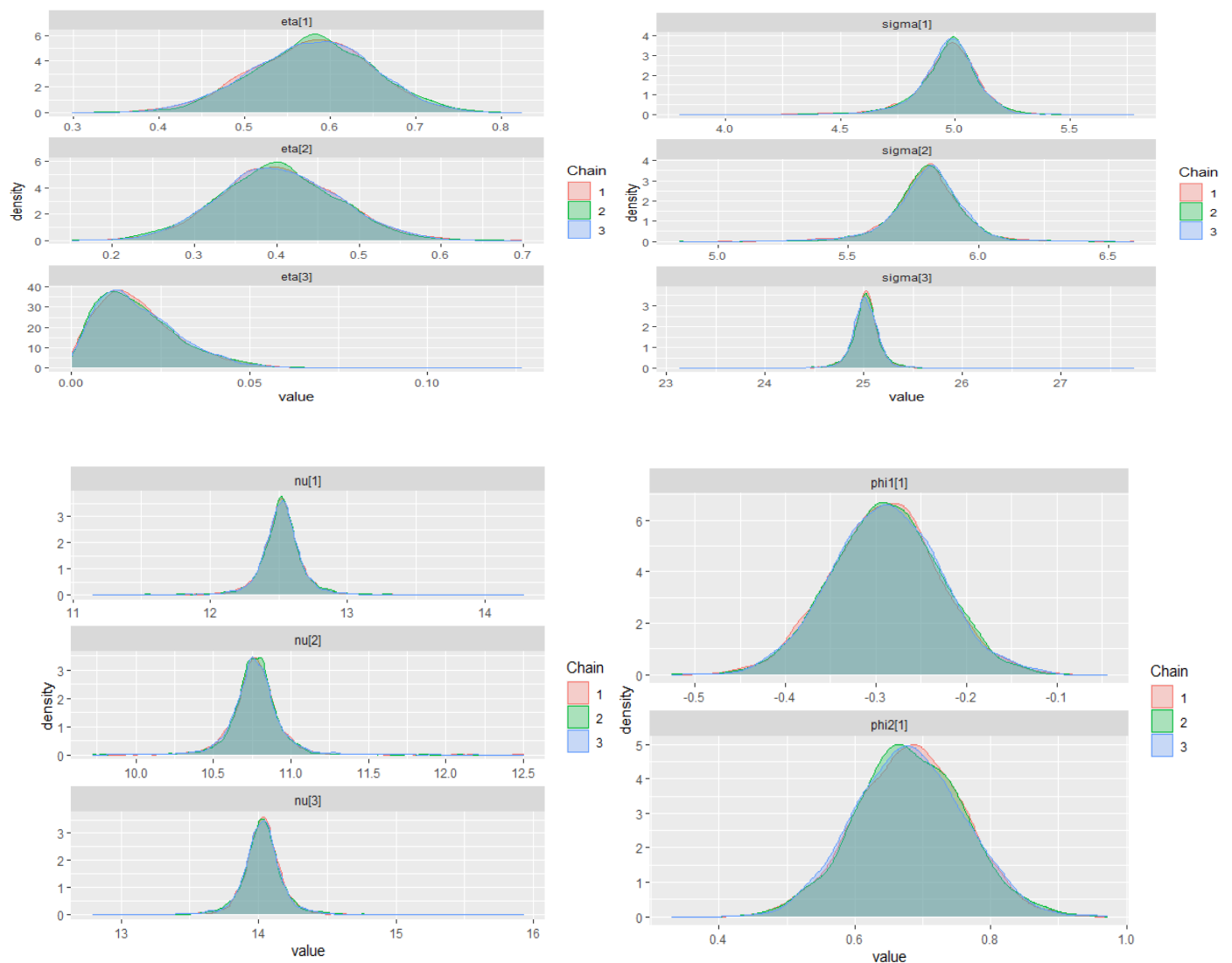


Figure A2. Posterior density plots for some parameters on the TMAR model in the IBM stock prices (first-differenced series).

Appendix G.1.3. GMAR Model

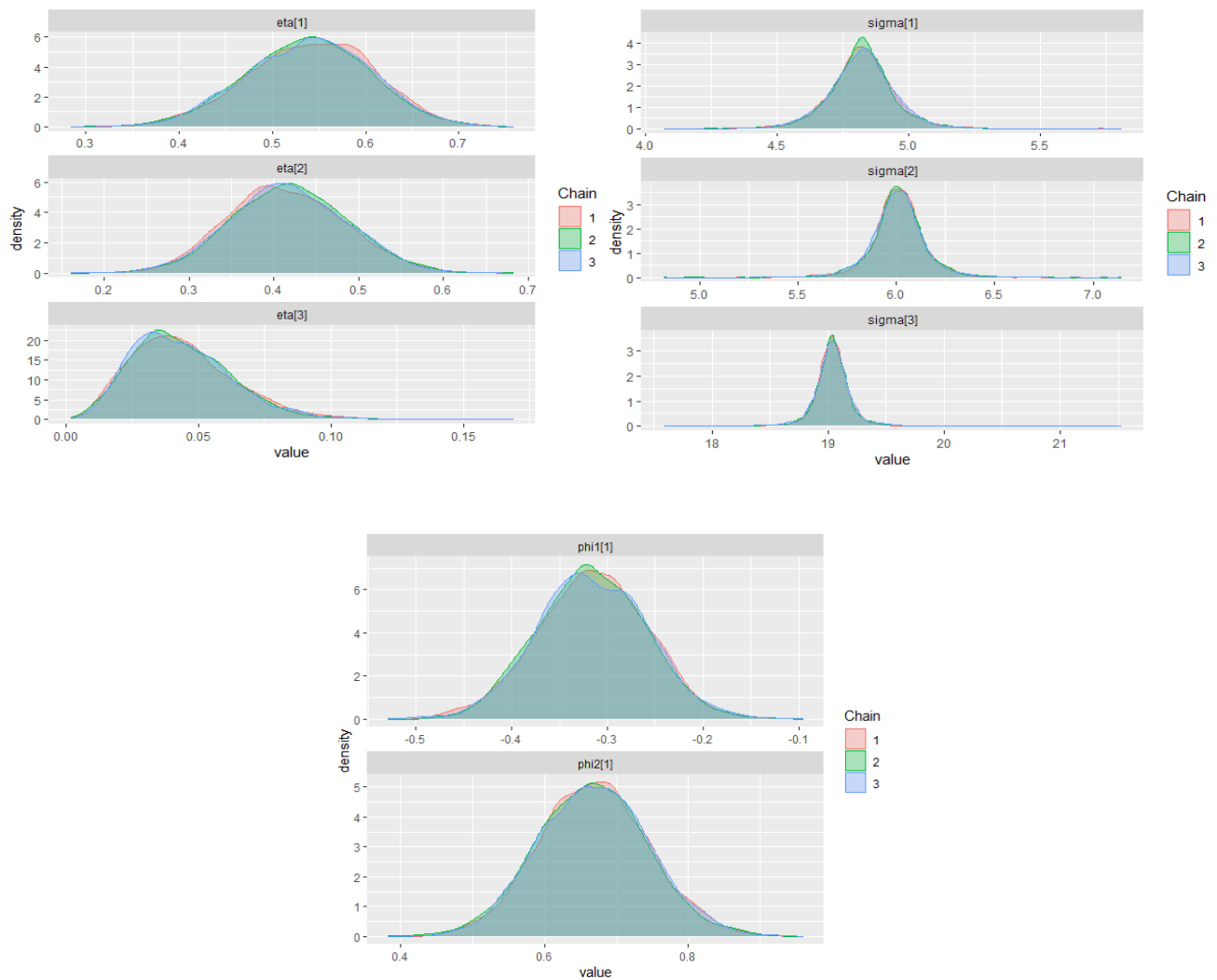


Figure A3. Posterior density plots for some parameters on the GMAR model in the IBM stock prices (first-differenced series).

Appendix G.2. Brent Crude Oil Prices (First-Differenced Series)

Appendix G.2.1. ZMAR Model

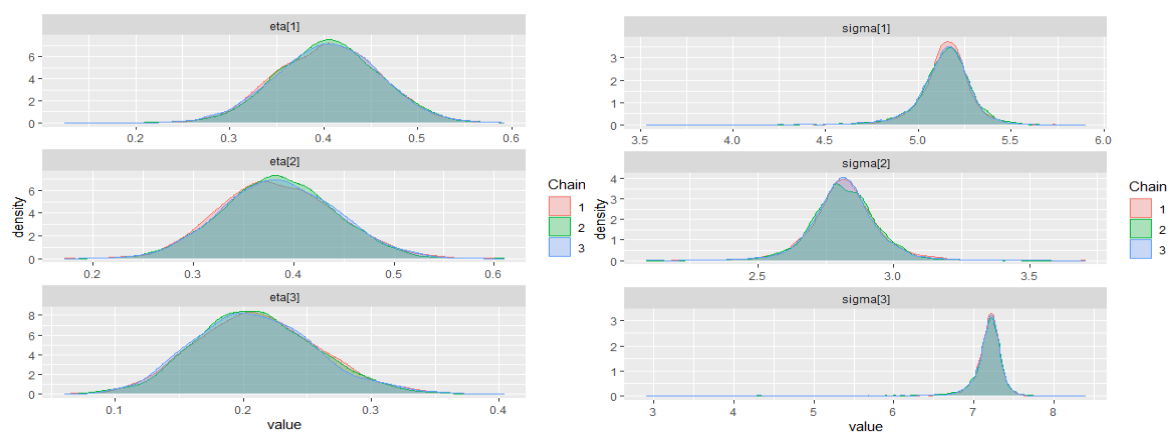


Figure A4. Cont.

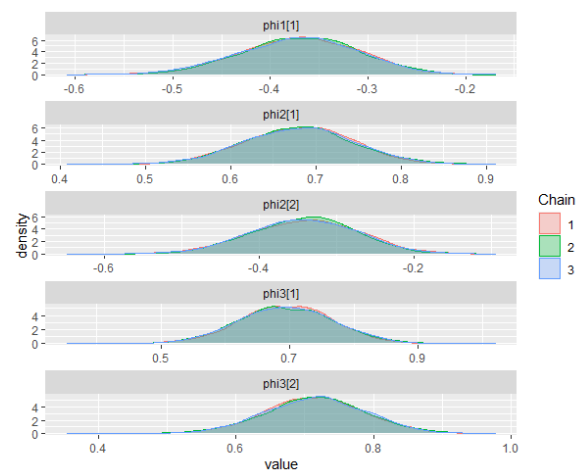


Figure A4. Posterior density plots for some parameters on the ZMAR model in the Brent crude oil prices (first-differenced series).

Appendix G.2.2. TMAR Model

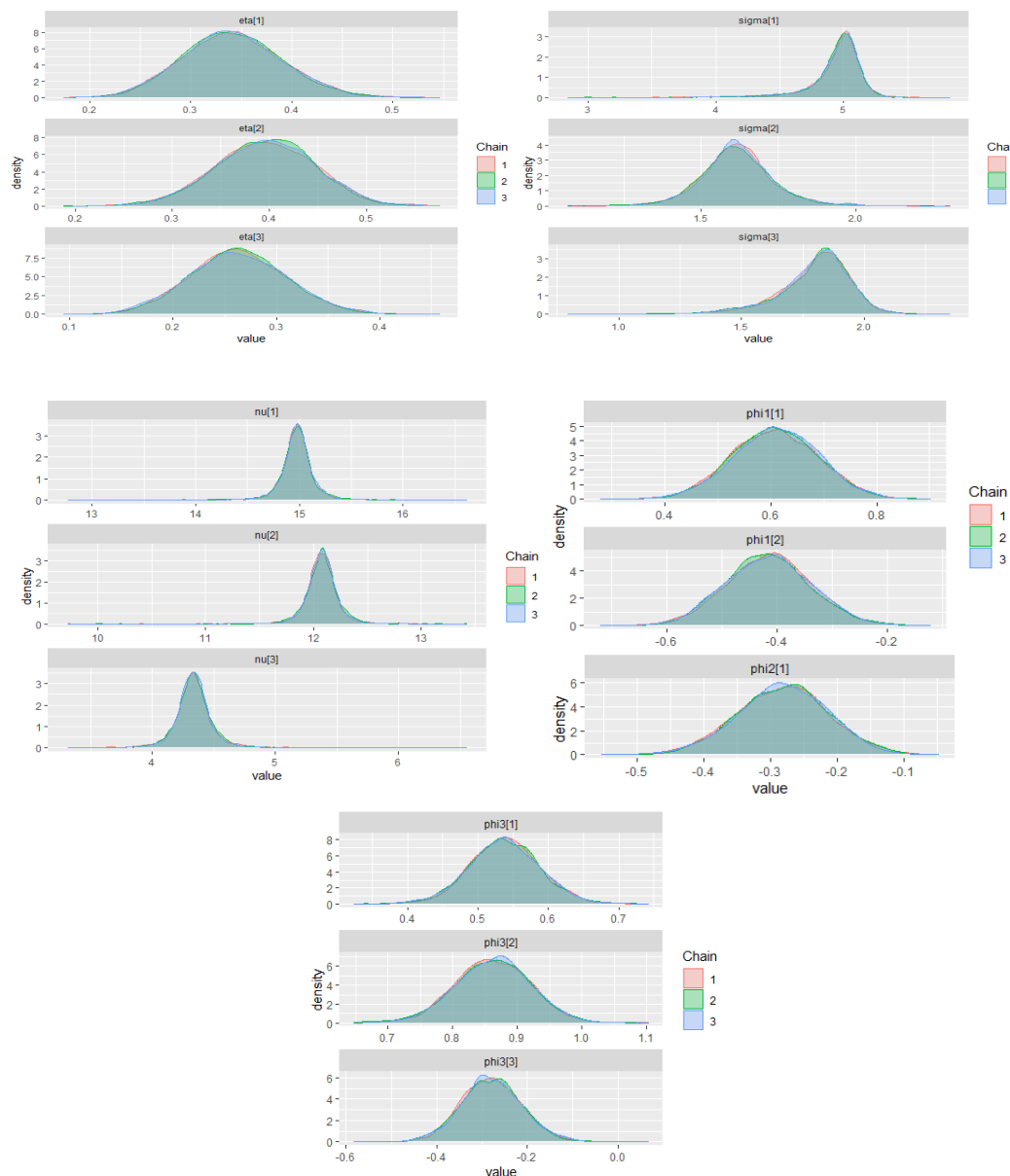


Figure A5. Posterior density plots for some parameters on the TMAR model in the Brent crude oil prices (first-differenced series).

Appendix G.2.3. GMAR Model

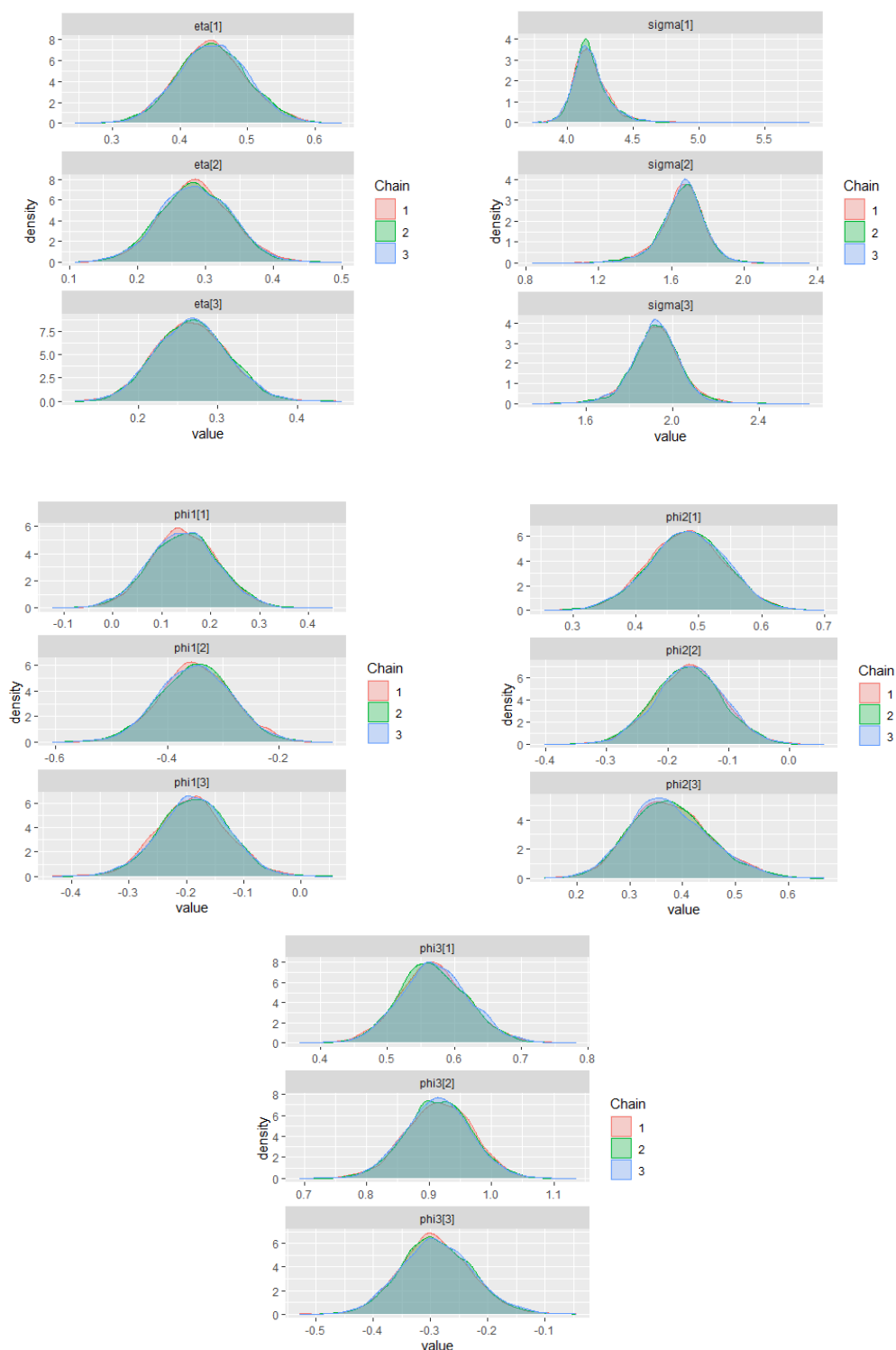


Figure A6. Posterior density plots for some parameters on the GMAR model in the Brent crude oil prices (first-differenced series).

References

- Al Hakmani, Rehab, and Yanyan Sheng. 2017. NUTS for Mixture IRT Models. Paper presented at the Annual Meeting of the Psychometric Society, Zurich, Switzerland, July 18–21; pp. 25–37.
- Albert, James H., and Siddhartha Chib. 1993. Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business & Economic Statistics* 11: 1–15.

- Annis, Jeffrey, Brent J. Miller, and Thomas J. Palmeri. 2017. Bayesian Inference with Stan: A Tutorial on Adding Custom Distributions. *Behavior Research Methods* 49: 863–86. [\[CrossRef\]](#) [\[PubMed\]](#)
- Aroian, Leo A. 1941. A Study of RA Fisher's z Distribution and the Related F Distribution. *The Annals of Mathematical Statistics* 12: 429–48. [\[CrossRef\]](#)
- Azzalini, Adelchi. 2014. *The Skew-Normal and Related Families*. Cambridge: Cambridge University Press.
- Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. Hoboken: John Wiley & Sons.
- Brown, Barry W., Floyd M. Spears, and Lawrence B. Levy. 2002. The Log F : A Distribution for All Seasons. *Computational Statistics* 17: 47–58. [\[CrossRef\]](#)
- Bruce, Andrew G., and R. Douglas Martin. 1989. Leave-K-Out Diagnostics for Time Series. *Journal of the Royal Statistical Society: Series B Methodological* 51: 363–401.
- Bulmer, Michael George. 1967. *Principles of Statistics*. Cambridge: MIT Press.
- Carollo, Salvatore. 2012. *Understanding Oil Prices: A Guide to What Drives the Price of Oil in Today's Markets*. Hoboken: John Wiley & Sons.
- Carpenter, Bob, Matthew D. Hoffman, Marcus Brubaker, Daniel Lee, Peter Li, and Michael Betancourt. 2015. The Stan Math Library: Reverse-Mode Automatic Differentiation in C++. *arXiv*, arXiv:1509.07164.
- Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 76: 1–32. [\[CrossRef\]](#)
- Choir, Achmad Syahrul, Nur Iriawan, Brodjol Sutijo Suprih Ulama, and Mohammad Dokhi. 2019. MSEPBurr Distribution: Properties and Parameter Estimation. *Pakistan Journal of Statistics and Operation Research* 15: 179–93. [\[CrossRef\]](#)
- Fernández, Carmen, and Mark F. J. Steel. 1998. On Bayesian Modeling of Fat Tails and Skewness. *Journal of the American Statistical Association* 93: 359–71.
- Fisher, Ronald Aylmer. 1924. On a Distribution Yielding the Error Functions of Several Well Known Statistics. Paper presented at the International Congress of Mathematics, Toronto, ON, Canada, April 11–16; pp. 805–13.
- Fong, Pak Wing, Wai Keung Li, C. W. Yau, and Chun Shan Wong. 2007. On a Mixture Vector Autoregressive Model. *Canadian Journal of Statistics* 35: 135–50. [\[CrossRef\]](#)
- Frühwirth-Schnatter, Sylvia. 2006. *Finite Mixture and Markov Switching Models*. Cham: Springer.
- Gelman, Andrew. 2006. Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper). *Bayesian Analysis* 1: 515–34. [\[CrossRef\]](#)
- Gelman, Andrew, and Donald B. Rubin. 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 7: 457–72. [\[CrossRef\]](#)
- Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2014. *Bayesian Data Analysis*, 3rd ed. Boca Raton: CRC Press.
- Geman, Stuart, and Donald Geman. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–41. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gilchrist, Warren. 2000. *Statistical Modelling with Quantile Functions*. Boca Raton: CRC Press.
- Hoffman, Matthew D., and Andrew Gelman. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15: 1593–623.
- Huerta, Gabriel, and Mike West. 1999. Priors and Component Structures in Autoregressive Time Series Models. *Journal of the Royal Statistical Society: Series B Statistical Methodology* 61: 881–99. [\[CrossRef\]](#)
- Iriawan, Nur. 2000. *Computationally Intensive Approaches to Inference in Neo-Normal Linear Models*. Bentley: Curtin University of Technology.
- Johnson, Norman L., Samuel Kotz, and Narayanaswamy Balakrishnan. 1995. *Continuous Univariate Distributions*, 2nd ed. Hoboken: John Wiley & Sons, vol. 2.
- Johnson, Norman L., Adrienne W. Kemp, and Samuel Kotz. 2005. *Univariate Discrete Distributions*, 3rd ed. Hoboken: John Wiley & Sons.
- Jones, Chris. 2002. Multivariate t and Beta Distributions Associated with the Multivariate F Distribution. *Metrika* 54: 215–31. [\[CrossRef\]](#)
- Kim, Hea-Jung. 2008. Moments of Truncated Student- t Distribution. *Journal of the Korean Statistical Society* 37: 81–87. [\[CrossRef\]](#)
- Kotz, Samuel, Narayanaswamy Balakrishnan, and Norman L. Johnson. 2000. *Continuous Multivariate Distributions, Volume 1: Models and Applications*. Hoboken: John Wiley & Sons.
- Le, Nhu D., R. Douglas Martin, and Adrian E. Raftery. 1996. Modeling Flat Stretches, Bursts Outliers in Time Series Using Mixture Transition Distribution Models. *Journal of the American Statistical Association* 91: 1504–15.
- Maleki, Mohsen, Arezo Hajrajab, and Reinaldo B. Arellano-Valle. 2020. Symmetrical and Asymmetrical Mixture Autoregressive Processes. *Brazilian Journal of Probability and Statistics* 34: 273–90. [\[CrossRef\]](#)
- Martin, R. Douglas, and Victor J. Yohai. 1986. Influence Functionals for Time Series. *The Annals of Statistics* 14: 781–818. [\[CrossRef\]](#)
- McLachlan, Geoffrey J., and David Peel. 2000. *Finite Mixture Models*. Hoboken: John Wiley & Sons.
- Metropolis, Nicholas, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. 1953. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* 21: 1087–92. [\[CrossRef\]](#)
- Neal, Radford M. 2011. MCMC Using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*. Edited by Steve Brooks, Andrew Gelman, Galin L. Jones and Xiao-Li Meng. Boca Raton: CRC Press, pp. 116–62.

- Nguyen, Hien D., Geoffrey J. McLachlan, Jeremy F.P. Ullmann, and Andrew L. Janke. 2016. Laplace Mixture Autoregressive Models. *Statistics & Probability Letters* 110: 18–24.
- Pravitasari, Anindya Apriliyanti, Nur Iriawan, Kartika Fithriasari, Santi Wulan Purnami, and Widiana Ferriastuti. 2020. A Bayesian Neo-Normal Mixture Model (Nenomimo) for MRI-Based Brain Tumor Segmentation. *Applied Sciences* 10: 4892. [CrossRef]
- Solikhah, Arifatus, Heri Kuswanto, Nur Iriawan, Kartika Fithriasari, and Achmad Syahrul Choir. 2021. Extending Runjags: A Tutorial on Adding Fisher's Distribution to Runjags. *AIP Conference Proceedings* 2329: 060005.
- Stan Development Team. 2018. Stan User's Guide, Version 2.18.0. Available online: https://mc-stan.org/docs/2_18/stan-users-guide/index.html (accessed on 21 October 2020).
- Stan Development Team. 2020. RStan: The R Interface to Stan. Available online: <https://cran.r-project.org/web/packages/rstan/vignettes/rstan.html> (accessed on 21 October 2020).
- Susanto, Irwan, Nur Iriawan, Heri Kuswanto, Kartika Fithriasari, Brodjol Sutija Suprih Ulama, Wahyuni Suryaningtyas, and Anindya Apriliyanti Pravitasari. 2018. On The Markov Chain Monte Carlo Convergence Diagnostic of Bayesian Finite Mixture Model for Income Distribution. *Journal of Physics: Conference Series* 1090: 012014. [CrossRef]
- Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. Practical Bayesian Model Evaluation Using Leave-One-out Cross-Validation and WAIC. *Statistics and Computing* 27: 1413–32. [CrossRef]
- Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul Christian Bürkner. 2020. Rank-Normalization, Folding, and Localization: An Improved Rhat for Assessing Convergence of MCMC. *Bayesian Analysis* 1: 1–28. [CrossRef]
- Wong, Chun Shan, and Wai Keung Li. 2000. On a Mixture Autoregressive Model. *Journal of the Royal Statistical Society: Series B Statistical Methodology* 62: 95–115. [CrossRef]
- Wong, Chun Shan, and Wai Keung Li. 2001. On a Logistic Mixture Autoregressive Model. *Biometrika* 88: 833–46. [CrossRef]
- Wong, Chun Shan, Wai-Sum Chan, and P. L. Kam. 2009. A Student T-Mixture Autoregressive Model with Applications to Heavy-Tailed Financial Data. *Biometrika* 96: 751–60. [CrossRef]
- World Bank. 2020. World Bank Commodity Price Data (The Pink Sheet). Available online: <http://pubdocs.worldbank.org/en/561011486076393416/CMO-Historical-Data-Monthly.xlsx> (accessed on 21 October 2020).
- Zelen, Marvin, and Norman C. Severo. 1970. Probability Functions. In *Handbook of Mathematical Functions*; Edited by Milton Abramowitz and Irene A. Stegun. Washington, DC: U.S. Government Printing Office, pp. 925–95.