

Research Article

Bayesian Semiparametric Double Autoregressive Modeling

Min Wang ^{1,2} and Ji-Chun Liu¹

¹*School of Mathematical Sciences, Xiamen University, Xiamen 361005, China*

²*School of Mathematical Sciences, Guizhou Normal University, Guiyang 550001, China*

Correspondence should be addressed to Min Wang; wampum@gznu.edu.cn

Received 19 January 2019; Revised 27 March 2019; Accepted 28 April 2019; Published 15 July 2019

Academic Editor: Frederico Cruz

Copyright © 2019 Min Wang and Ji-Chun Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a Bayesian semiparametric modeling approach for the return distribution in double autoregressive models. Monte Carlo investigation of finite sample properties and an empirical application are presented. The results indicate that the semiparametric model developed in this paper is valuable and competitive.

1. Introduction

The original Bayesian theory is a parametric method. The parametric model has been long applied in classical statistical and Bayesian statistical inference studies, and its estimation is based on the unknown parameters of the overall distribution. In a review of Bayesian statistics, Lindley [1] identified nonparametric methods as a notable area in that it lacks Bayesian progress. Our purpose here is to present work with mixture priors that substantially contributes to both parametric and nonparametric density estimation. With the development of statistical theory, statisticians have found the need to reduce the overall distribution assumptions in many cases. We find that estimates of volatility using nonparametric models can differ dramatically from those using a normal return distribution if there is evidence of a heavy-tailed return distribution. When using a normal distribution or Weibull distribution in general, we often need to assume that the population distribution is unimodal (see [2]). The support set must be sufficiently large to capture the fat tail. We find that the Dirichlet process mixture (DPM) model is able to adapt to several frequently used distribution models and also accurately estimates the posterior distribution of the volatilities of the returns, without assuming any underlying distribution.

Mixture models are noted for their flexibility and are widely used in the statistical literature. Nonparametric methods are usually combined Bayesian approaches with the

development of more mature technologies in density estimation, regression, survival analysis, hierarchical models, and model validation [3]. The Dirichlet process (DP) is one type of nonparametric method often used in density estimation. Ferguson [4] took Dirichlet as a random probability measure and noted that if using the mixture distribution to approximate an unknown distribution, the posterior inference must be “controlled metaphysically” and the support set must be sufficiently large. Escobar and West [5] were the first to propose a modeling style involving the DP to introduce uncertainty in the number of components. The wide application of DPM priors is one of the major success stories of modern Bayesian statistics.

A central property of economic time series that is common to many financial time series is that their volatility varies over time. Describing the volatility of an asset is a key issue in financial economics. The most popular class of models for time-varying volatility is represented by GARCH-type models [6]. GARCH models are commonly used to describe, estimate, and predict the dynamics of financial returns. Recent surveys of the existing GARCH model literature can be found in Davidson [7] and Rombouts et al. [8]. In contrast to Engle’s [9] ARCH model, a double autoregressive (DAR) model, which is a special case of the ARMA-ARCH models in Weiss [10] and an example of weak ARMA models in Francq and Zakoian [11, 12], is also drawing attention from researchers. Under the assumption of the disturbance following a normal distribution, Ling

[13] considered the structure and the maximum likelihood estimation. In addition, stochastic volatility (SV) models have enjoyed great popularity in analyzing financial data in the last couple of decades (see [14], etc.). Delatola and Griffin [15] presented a method for Bayesian nonparametric analysis of the return distribution in a stochastic volatility model, which allows noninformative prior distributions for the parameters and assumes that the mixing component variances are constant. In addition, they [16] extended the work and developed a Bayesian semiparametric model with a leverage effect allowing for stylized facts such as heavy tails of the distribution of returns and correlation between the returns and changes in volatility. Jensen [17] solved a semiparametric, Bayesian estimator of the long-memory stochastic volatility model using a Markov chain Monte Carlo (MCMC) algorithm. Jensen and Maheu [18] extended the existing fully parametric Bayesian literature on stochastic volatility to allow for more general return distributions.

Jensen and Maheu [19, 20] proposed a Bayesian nonparametric modeling approach for the return distribution multivariate GARCH models (MGARCH). Virbickaitė [21] considered an asymmetric dynamic conditional correlation (ADCC) model for time-varying correlations with GJR-GARCH models for individual volatilities, denoted ADCC-GJR-GARCH, and provided a substantially more realistic evaluation of the comovements of the assets' returns than standard symmetric MGARCH models. Ausín et al. [22] developed a Bayesian semiparametric approach based on modeling the innovations using the class of scale mixtures of Gaussian distributions with a DP prior on the mixing distribution. Jensen and Maheu [19, 20] extended the asymmetric, stochastic, and volatility models by modeling the return volatility distribution nonparametrically. The novelty of their paper was in modeling the distribution with an infinite mixture of normals, where the mixture of unknowns have a DPM prior. This motivated us to establish a DAR with a DPM model (DAR-DPM).

The remainder of this paper is organized as follows. In Section 2, we describe the DP mixed process and introduce the DAR model with a nonparametric DPM prior for the unknown distribution. An MCMC scheme is used to sample from the posteriors, and the parametric portions of the model are presented in Section 3. Section 4 provides some simulation studies regarding the performance of the proposed method. An empirical example is reported in Section 5, and Section 6 concludes the paper.

2. DAR Models with Dirichlet Process Mixture Prior

The usual structure of a DAR model [13] assumes that a return series, denoted by r_t , can be written as follows:

$$r_t = \sum_{i=1}^p \phi_i r_{t-i} + h_t^{1/2} \varepsilon_t, \quad (1)$$

$$h_t = \omega + \sum_{i=1}^q \beta_i r_{t-i}^2,$$

where h_t are the conditional variances of r_t given the past history, $t \in \{-p, \dots, 0, 1, \dots\}$, $\omega, \beta_i > 0$; ε_t is a sequence of independent and identically distributed (IID) random variables with zero mean and unit variance; and r_s is independent of $\{\varepsilon_t : t \geq 1\}$ for $s \leq 0$. Ling [13] gave the stationary ergodic conditions of the general model and obtained the necessary and sufficient conditions for the strict stationarity and weak stationarity of the DAR model. Let A_t be a random matrix of $p \times p$:

$$\begin{pmatrix} \phi_1 + \sqrt{\beta_1} \varepsilon_{1t} \cdots \phi_{p-1} + \sqrt{\beta_{p-1}} \varepsilon_{p-1,t} & \phi_p + \sqrt{\beta_p} \varepsilon_{pt} \\ I_{p-1} & O_{(p-1) \times 1} \end{pmatrix} \quad (2)$$

where I_r is the unit matrix of $r \times r$ and $O_{r \times s}$ is the zero matrix of $r \times s$. To facilitate the presentation of results, we need to give the concept of the Lyapunov exponent here, which is defined as

$$r = \inf_n \frac{1}{n} E \ln \|A_1 \cdots A_n\| \quad (3)$$

From $\varepsilon_{it} \sim N(0, 1)$, we know that there is $E \ln^+ \|A_1\| < \infty$, where $\ln^+ x = \max\{x, 0\}$, subordinate to the Theorem of Subadditive Ergodicity [23], and

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \|A_1 \cdots A_n\|. \quad (4)$$

Ling [13] gave a theorem in which the necessary and sufficient condition for the existence of a strictly stationary solution of the DAR model is $\gamma < 0$. When $\beta = 1$, $\gamma = E \ln |\phi_1 + \sqrt{\beta_1} \varepsilon_t|$, the condition is reduced to $\phi_1^2 + \beta_1 < 1$. From the study of Lu and Jiang [24], the sufficient condition for the weakest stationarity of the model is $\sum_{i=1}^p |\phi_i| + E|\xi_t| \sum_{i=1}^p \beta_i < 1$.

As noted by Ausín et al. [22], typical models for the innovation distribution include the Gaussian, Student, and Gaussian mixture models. The purpose of this paper is to construct a robust alternative to these distributional assumptions. Generally, the unit variance restriction of ε_t in (1) makes it difficult to undertake semiparametric Bayesian inference. To avoid this, we rewrite (1) as follows:

$$r_t = \sum_{i=1}^p \phi_i r_{t-i} + \tilde{h}_t^{1/2} \xi_t, \quad (5)$$

$$\tilde{h}_t = 1 + \sum_{i=1}^q \tilde{\beta}_i r_{t-i}^2$$

where $\tilde{h}_t = h_t/\omega$ is a rescaled volatility and $\tilde{\beta}_i = \beta_i/\omega$, and $\xi_t = \omega^{1/2} \varepsilon_t$ are sequences of IID random variables with mean 0 and variance ω .

We relax all assumptions concerning the distribution of ξ_t in (5), allowing its distribution to be completely unknown and random as if the distribution were an unknown in addition to the parameters and latent volatilities of the DAR model. This strategy suggests using a Gaussian scale mixture model with a density function given by

$$f_\xi(\xi_t | G) = \int \phi(\xi_t | 0, \Lambda_t) dG(\Lambda_t), \quad (6)$$

where $\phi(\xi_t | 0, \Lambda_t)$ denotes the density function of the Gaussian distribution with mean zero and variance Λ_t and G is a random probability measure of a DP; G_0 is a measure function in probability space, and α is a strength parameter. The distribution of G is based on the unknown prior of the DP(α, G_0). Employing a Lo [25] type DPM prior, assume that

$$\xi_t | \Lambda_t \sim N(0, \Lambda_t) \quad (7)$$

$$\Lambda_t | G \sim G \quad (8)$$

$$G | \alpha, G_0 \sim DP(\alpha, G_0) \quad (9)$$

to model the unknown distribution, where $G_0 \sim \text{Inv-Gamma}(v/2, c/2)$ and $\alpha \sim \text{Gamma}(a, b)$. In (7)-(9), ξ_t are normally distributed with a mean zero but a random variance Λ_t , distributed as G . The nonnegative value α is the strength parameter, and G_0 is the base distribution. By the definition of DP(α, G_0) in Sethuraman [26], G is almost necessarily equal to a discrete distribution $G(d\Lambda) = \sum_{j=1}^{\infty} \pi_j \delta_{\sigma_j^2}(d\Lambda)$, where $\delta_{\sigma_j^2}(\cdot)$ is a degenerative distribution on σ_j^2 with the random variance extracted from the prior distribution G_0 of DP. To ensure the conjugation, G_0 is a gamma distribution with the parameters $v/2$ and $c/2$, i.e., the probability distribution of $G_0 \sim \text{Inv-Gamma}(v/2, c/2)$. The probability of Λ_t being equal to a particular σ_j^2 is π_j , which satisfies

$$\begin{aligned} \pi_1 &= \pi'_1, \\ \pi_j &= \pi'_j \prod_{l < j} (1 - \pi'_l) \end{aligned} \quad (10)$$

under $\alpha > 0$, $\pi'_j \sim \text{Beta}(1, \alpha)$.

The DAR-DPM model above can also be introduced into the potential distribution of variables $S = (S_1, S_2, \dots, S_T)$, while $S_t = j$, Λ_t equals a particular σ_j^2 . Under the prior of the DP, the distribution of S_t is $S_t \sim \sum_{j=1}^{\infty} \pi_j \delta_j$, where the probabilistic weight π_j , $j = 1, 2, \dots$, defined by (10). Incorporating S_t into the definition of the DAR-DPM model, we propose rescaling the mode defined in (5), (7), (8), and (9) as follows:

$$\begin{aligned} r_t &= \sum_{i=1}^p \phi_i r_{t-i} + \xi_t \sqrt{\left(1 + \sum_{i=1}^q \beta_i r_{t-i}^2\right)} \\ \xi | S_t, \Lambda_t &\sim N(0, \Lambda_{S_t}) \\ S_t &\sim \sum_{j=1}^{\infty} \pi_j \delta_j \\ \sigma_{S_t}^2 &\sim G_0. \end{aligned} \quad (11)$$

Finally, to complete the Bayesian inference of the model, we also need to deduce distributions for the parameters of the DAR-DPM-type models.

3. Bayesian Inference for DAR-DPM-Type Models

This section describes how to perform Bayesian inference for DAR-DPM-type models. Given an observed time series $r = \{r_1, r_2, \dots, r_T\}$ and the priors defined in the previous section, we consider the joint posterior distribution of the model parameters. Unfortunately, the joint posterior distribution does not have a closed analytical expression. To this end, the MCMC algorithms of the DPM model in Escobar and West [5], which produce samples from a Markov chain whose stationary distribution is the joint posterior distribution of the model parameters, are employed.

Let $\Theta = (\phi, \beta, \{\Lambda_t\}, \alpha)$, where $\phi = (\phi_1, \phi_2, \dots, \phi_p)$ and $\beta = (\beta_1, \beta_2, \dots, \beta_q)$ are the full model parameters and $\Theta \setminus \theta$ represent a set of parameters that do not contain the parameter θ . We block the unknown parameters such that the conditional posterior distribution of the controllable block is either known or given by a manageable form.

$$\begin{aligned} \pi(\phi, \beta, \{\Lambda_t\}, \alpha | r) &\propto f(r | \phi, \beta, \{\Lambda_t\}, \alpha) \\ &\cdot \pi(\phi | \beta, \alpha) \pi(\beta | \alpha) \pi(\{\Lambda_t\} | \alpha) \pi(\alpha). \end{aligned} \quad (12)$$

Assuming that other parameters and latent variables have been obtained, we can construct a Markov chain on each piece of the posterior distributions for iterative sampling.

3.1. *Sample from $\{\Lambda_t\} | r, \phi, \beta, \alpha$.* $\{\Lambda_t\} | r, \phi, \beta, \alpha$ can be sampled using the two-step method based on the Polya method [27]. Set $\xi = (\xi_1, \xi_2, \dots, \xi_T)'$, where $\xi_t = (r_t - \sum_{i=1}^p \phi_i r_{t-i}) / \sqrt{1 + \sum_{i=1}^q \beta_i r_{t-i}^2}$. The posterior distribution of Λ_t can be written as

$$\begin{aligned} \Lambda_t | \{\Lambda_{t'}, t' \neq t\}, \xi_t, \alpha &\sim \frac{\alpha}{\alpha + T + 1} g(\xi_t) G(d\Lambda | \xi_t) \\ &+ \frac{1}{\alpha + T - 1} \sum_{t' \neq t} f_N(\xi_t | \Lambda_{t'}) \delta_{\Lambda_{t'}}(d\Lambda_t) \end{aligned} \quad (13)$$

where $g(\xi_t) \equiv \int f_N(\xi_t | 0, \Lambda) G_0(\Lambda) d\Lambda$. Under the conditional probability rule,

$$G(d\Lambda | \xi_t) \propto f_N(\xi_t | 0, \Lambda) G_0(d\Lambda) \quad (14)$$

whose density function is

$$\begin{aligned} g(\Lambda | \xi_t) &\propto \Lambda^{-1/2} \exp\left\{-\frac{\xi_t^2}{2\Lambda}\right\} \\ &\cdot \Lambda^{-(1+v/2)} \exp\left\{-\frac{c}{2\Lambda}\right\} \\ &= \Lambda^{-1+(v+1)/2} \exp\left\{-\frac{\xi_t^2 + c}{2\Lambda}\right\}. \end{aligned} \quad (15)$$

Assuming known ϕ, β, α , the posterior distribution is

$$\pi(\{\sigma_j\}, S | \xi, \alpha) \equiv \pi(\{\sigma_j\}, S | r, \phi, \beta, \alpha), \quad (16)$$

where σ_j , $j = 1, 2, \dots, k$, $k \leq T$, are the distinct Λ_t , $t = 1, 2, \dots, T$.

The two steps are sampled as follows.

Step 1.

$$S_t \mid \{\Lambda_{t'} : t' \neq t\}, \xi_t, \alpha \sim \begin{cases} \frac{\alpha}{\alpha + T - 1} g(\xi_t) \delta_0(dS_t) \\ \frac{1}{\alpha + T - 1} \sum_{j=1}^{k^{(t)}} f_N(\xi_t \mid 0, \sigma_j) \delta_j(dS_t) \end{cases} \quad (17)$$

$k^{(t)}$ is the distinct number of $\{\Lambda_{t'}, t' \neq t\}$ presenting different σ_j . If S_t is drawn to be 0, k increases by 1, S_t is set for the new value k , and the new σ_k is extracted from (15); otherwise, S_t is the random sample value j .

Step 2. Discarding σ in the extraction of Step 1, new σ_j , $j = 1, 2, \dots, k$, are extracted using the sample S and k iteration, and the posterior distribution is as follows:

$$\begin{aligned} \pi(\sigma_j \mid \xi, S, k) &\propto \prod_{t:S_t=j} f_N(\xi_t \mid 0, \sigma_j) G_0(d\sigma_j) \\ &\propto \sigma_j^{-(1+\nu/2)} \exp\left\{-\frac{c}{2\sigma_j}\right\} \cdot \prod_{t:S_t=j} \sigma_j^{-1/2} \exp\left\{-\frac{\xi_t^2}{2\sigma_j}\right\} \\ &= \sigma_j^{-(\sum_{t:S_t=j} 1+\nu)/2+1)} \exp\left\{-\frac{\sum_{t:S_t=j} \xi_t^2 + c}{2\sigma_j}\right\} \\ &\sim \text{IG}\left(\frac{\sum_{t:S_t=j} 1 + \nu}{2}, \frac{\sum_{t:S_t=j} \xi_t^2 + c}{2}\right). \end{aligned} \quad (18)$$

3.2. Sample ϕ and β . The sampling of $\pi(\phi \mid r, \beta, \{\Lambda_t\})$ and $\pi(\beta \mid r, \phi, \{\Lambda_t\})$ is based on the Metropolis-Hastings (MH) method [28]. The use of R language for this method is mentioned in a previous article [29], which proposed that the distribution is the conditional distribution of the parameters.

Let $p(\phi_i)$, $i = 1, 2, \dots, p$, denote the prior distribution of ϕ_i assuming known $S = (S_1, S_2, \dots, S_T)'$ and $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_k)$, $\xi_t/\sqrt{\sigma_{S_t}} \sim N(0, 1)$. In addition, write $h_t := 1 + \sum_{j=1}^q \beta_j r_{t-j}^2$. Then, $\xi_t = (r_t - \sum_{i=1}^p \phi_i r_{t-i})/\sqrt{h_t}$; therefore, the posterior distribution of ϕ_i is

$$\begin{aligned} \pi(\phi_i \mid r, S, \Theta \setminus \phi_i) &\propto \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma_{S_t} h_t}} \exp\left\{-\frac{(r_t - \sum_{j=1}^p \phi_j r_{t-j})^2}{2\sigma_{S_t} h_t}\right\} p(\phi_i) \\ &\propto \exp\left\{-\frac{(\phi_i - \mu_{\phi_i})^2}{2\sigma_{\phi_i}^2}\right\} \cdot p(\phi_i) \end{aligned} \quad (19)$$

where $\sigma_{\phi_i}^2 = (\sum_{i=1}^T (r_{t-i}^4/\sigma_{S_t} h_t))^{-1}$, $\mu_{\phi_i} = (\sum_{t=i}^T (r_{t-i}(r_t - \sum_{j \neq i} \phi_j r_{t-j}^2)/\sigma_{S_t} h_t))/\sigma_{\phi_i}^2$. Then, our proposed distribution for

ϕ_i is $N(\mu_{\phi_i}, \sigma_{\phi_i}^2)$, $i = 1, 2, \dots, p$. By MH sampling, the accepted probability of ϕ_i' , $\alpha(\phi_i', \phi_i) = \min\{g(\phi_i')f_N(\phi_i \mid \mu_{\phi_i}, \sigma_{\phi_i}^2)/g(\phi_i)f_N(\phi_i' \mid \mu_{\phi_i}, \sigma_{\phi_i}^2), 1\}$, where $g(\phi_i) \propto \pi(\phi_i \mid r, S, \Theta \setminus \phi_i)$. If the suggested value ϕ_i' has been refused, then the prior ϕ_i remains as the current value drawn from $\pi(\phi_i \mid r, S, \Theta \setminus \phi_i)$.

Sampling from $\pi(\beta \mid r, \phi, \{\Lambda_t\})$, set $p(\beta_i)$ as the prior of β_i , $i = 1, 2, \dots, q$, $\epsilon_t := r_t - \sum_{j=1}^p \phi_j r_{t-j} = \xi_t h_t^{1/2}$, and $\mathcal{F}_t := \sigma(r_s, s \leq t)$. Then, under \mathcal{F}_{t-1} ,

$$\begin{aligned} u_t &:= \epsilon_t^2 - \sigma_{S_t} h_t = \left(\frac{\epsilon_t^2}{\sigma_{S_t} h_t} - 1\right) \sigma_{S_t} h_t \\ &= (\chi^2(1) - 1) \sigma_{S_t} h_t. \end{aligned} \quad (20)$$

The conditional mean of u_t is $E[u_t \mid \mathcal{F}_{t-1}] = 0$, and the conditional variance is $\text{var}(u_t \mid \mathcal{F}_{t-1}) = 2(\sigma_{S_t} h_t)^2$. Nakatsuma [30] suggests replacing this u_t with $N(0, 2\sigma_{S_t}^2 h_t^2)$. Then, the posterior distribution of β_i is

$$\begin{aligned} \pi(\beta_i \mid r, S, \Theta \setminus \beta_i) &\propto \prod_{t=1}^T \frac{1}{\sqrt{2\pi \cdot 2\sigma_{S_t}^2 h_t^2}} \\ &\cdot \exp\left\{-\frac{(\epsilon_t^2 - (1 + \sum_{j=1}^q \beta_j r_{t-j}^2) \sigma_{S_t})^2}{4\sigma_{S_t}^2 h_t^2}\right\} \cdot p(\beta_i) \\ &\propto \exp\left\{-\frac{(\beta_i - \mu_{\beta_i})^2}{2\sigma_{\beta_i}^2}\right\} P(\beta_i) g(h_{[1:T]}) \end{aligned} \quad (21)$$

where $\sigma_{\beta_i}^2 = (\sum_{t=1}^T (r_{t-i}^4/2h_t^2))^{-1}$, $\mu_{\beta_i} = \sigma_{\beta_i}^2 \sum_{t=1}^T (r_{t-i}^2(\epsilon_t^2/\sigma_{S_t} - 1 - \sum_{j \neq i} \beta_j r_{t-j}^2)/2h_t^2)$. The suggested distribution of β_i is $N(\mu_{\beta_i}, \sigma_{\beta_i}^2)$, $i = 1, \dots, q$. Extracting β_i , the idea of rejecting or accepting is similar to ϕ_i .

3.3. Sample α . Drawing from the conditional posterior distribution of α can be performed, as explained in Escobar and West [5], by first sampling an auxiliary variable ζ from a beta distribution, $\zeta \sim \text{Beta}(\alpha + 1, T)$, and then sampling α from a gamma mixture, $\pi_\zeta G(a + S^*, b - \log \zeta) + (1 - \pi_\zeta) G(a + S^* - 1, b - \log \zeta)$, where $S^* = \max\{S_t\}_{t=1}^T$ and $\pi_\zeta/(1 - \pi_\zeta) = (a + S^* - 1)/(T(b - \log \zeta))$.

3.4. MCMC Algorithm. This part briefly discusses the steps of the MCMC used to fit the DAR-DPM model. Given the priors, we define the following MCMC algorithm:

Initialize $\phi, \alpha, \beta, k, S, a, b, \Lambda$
 Sample $\phi \mid \alpha, \beta, k, S, a, b, \Lambda, r$
 Sample $\beta \mid \phi, \alpha, k, S, a, b, \Lambda, r$
 Sample $\alpha \mid \phi, \beta, k, S, a, b, \Lambda, r$
 Sample $k, S \mid \phi, \alpha, \beta, a, b, \Lambda, r$
 Sample $a, b, \Lambda \mid \phi, \alpha, \beta, k, S, r$

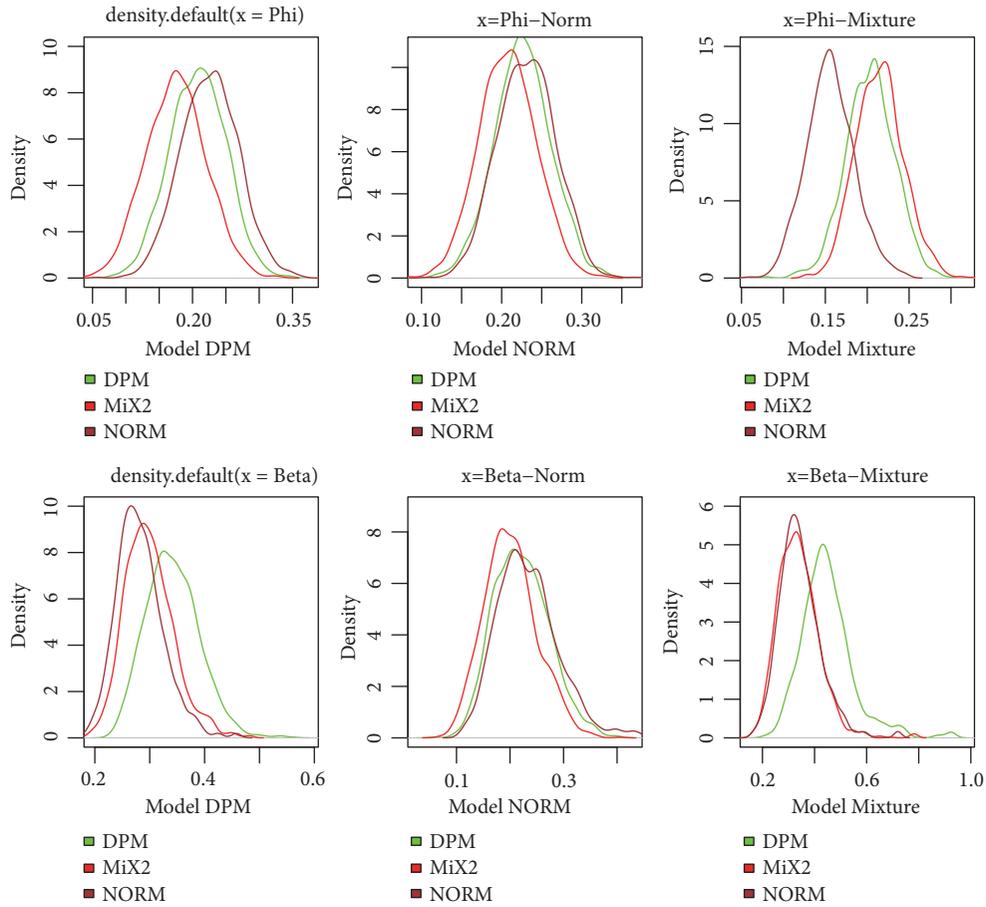


FIGURE 1: The postdensity plots of each parameter.

4. Simulation Study

In this section, the proposed methodology is illustrated using an artificial time series of size $T = 1000$. The simulated data are given by $y_t = 0.2y_{t-1} + \xi_t \sqrt{1 + 0.3y_{t-1}^2}$, where we set $\phi = 0.2$ and $\beta = 0.3$. The following three scenarios are considered: (1) $\xi_t \sim DP(\alpha, G_0)$; (2) $\xi_t \sim N(0, 1)$; and (3) $\xi_t \sim 0.9N(0, 0.6) + 0.1N(0, 4.6)$ (Ausín, 2014). The MCMC algorithm is run for a burn-in of 30000 iterations and subsequently a further 30000 iterations, storing one in 1500 iterations, thus obtaining a total of 1500 realizations from the posterior distribution.

Sensitivity Analysis, Comparison of Different Prior Effects. The set of hyperparameters has been set as $a = 2, b = 4, v = 10, c = 30$. Note that we set the prior $\alpha \sim \text{Gamma}(2, 4)$ initially because there is a fair degree of support for values near $\alpha = 1$ (see [5]), which leads to a rather noninformative prior.

Under the seven different priors in Table 1, we change only the hyperparameters a, b (Priors 2, 3, and 4), simply change the hyperparameters v, c (Prior 5 and 6), or change them all (Prior 7). Table 1 shows that all the estimated values of ϕ and β are closer to the true values of the parameters, with a smaller standard deviation. Therefore, we believe that the parameter estimation of DAR-DPM is not very sensitive to the setting of

the hyperparameters. In the following analysis, a significant loss of parameter estimation accuracy is not expected when using the hyperparameters of prior 7.

Normal and Finite Normal State Compared with Infinite Normal State. Table 2 shows that when estimating ϕ , the standard normal and finite mixture models behave well. However, both models are unable to estimate β . Clearly, the DPM method exhibits its comparative advantage in all cases and can accurately estimate the two parameters ϕ and β simultaneously, regardless of the data being from the standard Gaussian, finite mixture, or DPM mixed models.

We have drawn postdensity plots for each parameter based on simulation data for clarity; see Figure 1.

5. Empirical Application

This section applies the above method to actual financial data to fit the model. The utilized data are a series of US weekly observations of 3-month Treasury bill data: 1989/03-2019/03 (1566 observations, Data Source: Federal Reserve Economic Data). To make the data smooth, we perform some processing of the original data. Let x_t denote the 90-day T-bill weekly

TABLE 1: Sensitivity analysis of the choice of priors.

	Prior 1 $a = 2, b = 4$ $\nu = 10, c = 30$	Prior 2 $a = 6, b = 6$ $\nu = 10, c = 30$	Prior 3 $a = 0.1, b = 20$ $\nu = 10, c = 30$	Prior 4 $a = 2, b = 4$ $\nu = 3, c = 0.003$	Prior 5 $a = 2, b = 4$ $\nu = 10, c = 0.01$	Prior 6 $a = 2, b = 4$ $\nu = 10, c = 1$	Prior 7 $a = 1, b = 5$ $\nu = 3, c = 6$
ϕ	0.2008 0.0017	0.2008 0.0025	0.2033 0.0015	0.1999 0.0022	0.2042 0.0018	0.1949 0.0023	0.1919 0.0018
β	0.3164 0.0019	0.3151 0.0037	0.3092 0.0026	0.3089 0.0019	0.3319 0.0036	0.3407 0.0036	0.3146 0.0034
α	0.9625 0.0025	0.9780 0.0011	0.9289 0.0048	0.9187 0.0033	0.8949 0.0062	0.9096 0.0045	0.9492 0.0025
k	5.8512 1.3858	5.9677 1.4954	5.8917 2.1101	3.5473 0.6853	3.2172 0.4840	3.7628 0.4602	5.5979 1.3408

Note. This table presents the posterior mean and standard deviations of the model parameters for prior 1 sets $a = 2, b = 4, \nu = 10, c = 30$, and $k = 6$. The remaining prior uses the same choices but changes the values of the indicated hyperparameters in the first rows.

TABLE 2: Full sample estimates.

Model	Parameter	True values	DPM	Standard Normal	Finite mixture
			Post.Mean Post.std.dev	Post.Mean Post.std.dev	Post.Mean Post.std.dev
DPM	ϕ	0.2	0.1909 0.0364	0.2046 0.0441	0.1973 0.0492
			β	0.3	0.3095 0.0367
	α	1			0.9413 0.0472
			k		5.4095 1.1744
	Standard Normal	ϕ		0.2	0.1930 0.0367
β			0.3		0.2951 0.0532
		α			0.9124 0.0491
k				1.5485 0.1827	
		Finite mixture	ϕ	0.2	0.2088 0.0377
β	0.3				0.2957 0.0679
			α		0.9219 0.0370
k				2.4693 0.0774	

Note. This table displays the posterior means and standard deviations of the model parameters obtained from the DPM model, Gaussian distribution and mixture of two zero-mean Gaussians.

returns, and let y_t denote the rate of return of the order difference, i.e., $y_t = x_t - x_{t-1}$; see Figure 2.

To evaluate the performance of the different models, first, we give some descriptive statistics of the premodeling data and their autocorrelation function and partial autocorrelation function; see Table 3 and Figure 3.

We now illustrate the four models using the above data.

First, assume that $\xi_t \sim DP(\alpha, G)$; we may obtain the following fitted DAR(1, 1)-DPM model: $y_t = 0.2507y_{t-1} + \xi_t \sqrt{1 + 0.7043y_{t-1}^2}$, where the standard deviation of each coefficient is 0.0361 and 0.2460 and $AIC = -2.1361$. Then, assume that $\xi_t \sim N(0, \sigma^2)$; the fitted model is $y_t = 0.1776y_{t-1} + \xi_t \sqrt{1 + 0.2162y_{t-1}^2}$, the standard deviations of

TABLE 3: Descriptive statistics.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Variance	Sd
y	-0.590000	-0.030000	0.000000	0.004051	0.020000	0.980000	0.007110	0.084323

Note. This table shows the summary of the order difference.

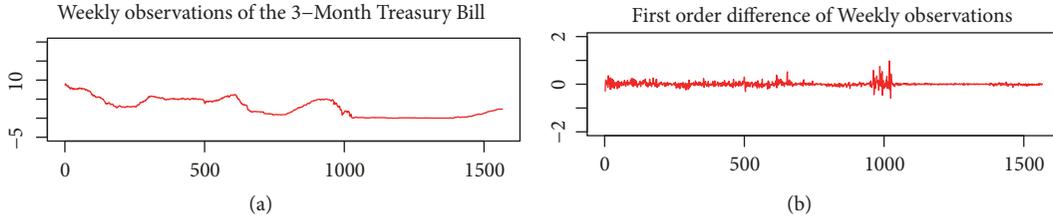


FIGURE 2: Plots of the observations (a) and the order difference of returns (b) of the weekly observations of American 3-month treasury bills. As shown in (b), the transformed data are smooth.

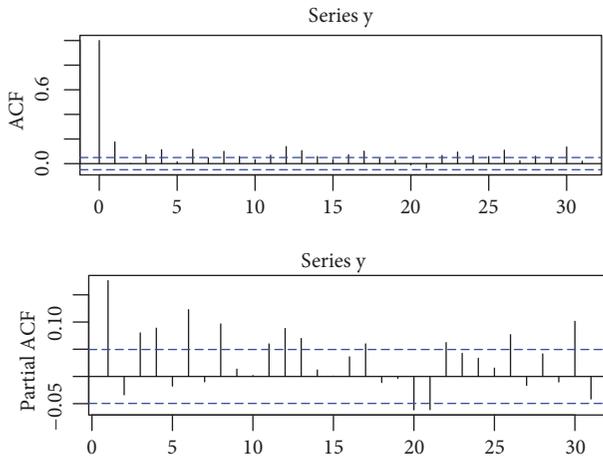


FIGURE 3: (a) is the ACF of y; the first-order lag autocorrelation coefficient is very significant. The PACF diagram (b) shows that we cannot build a simple autoregressive model.

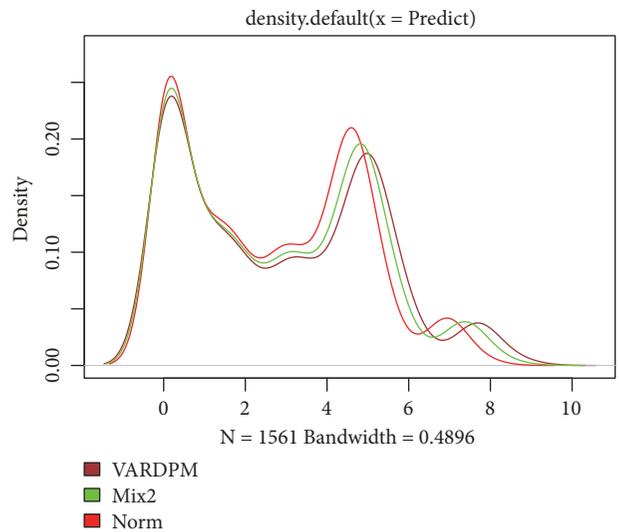


FIGURE 4: The predictive densities for both the parametric and nonparametric approaches.

the coefficients are 0.2162 and 0.1210, and $AIC = -1.7912$. Again, suppose that $\xi_t \sim \delta N(0, \sigma_1^2) + (1 - \delta)N(0, \sigma_2^2)$ to arrive at $y_t = 0.2517y_{t-1} + \xi_t \sqrt{1 + 0.6161y_{t-1}^2}$, where the standard deviations of the coefficients are 0.0385 and 0.2505 and $AIC = -2.0394$.

The predictive distribution of the series showcases the exhibity of the DPM. We computed and plotted the predictive densities for the above models; see Figure 4.

According to the three models obtained above, we can see that the parameter values with the DPM model have a smaller variance and minimum value of AIC . In the DPM model, we calculate that α is 0.8628 and that the mean of k is 2.7956. Obviously, under these figures, the disturbance tends to form several groups; therefore, using the DPM model to match these figures is satisfactory. To obtain the above results, we perform 20000 iterations; see Figure 5.

Finally, we also compare our model with that of Jensen and Maheu [19, 20] using the same datasets. The fitted GARCH(1,1)-DPM model is

$$y_t = h_t^{1/2} \xi_t \tag{22}$$

$$h_t = 0.2613 + 0.0619y_{t-1}^2 + 0.0462h_{t-1}$$

where $\xi_t \sim DP(\alpha, G)$. The standard deviations of the coefficients are 0.6421, 0.1288, and 0.1073, and $AIC = -2.0979$. This model was also applied in Ausin, Galeano, & Ghosh [22].

Based on this example, the results demonstrate that the DAR-DPM model developed in this paper is valuable and competitive. The support for u is always confined to values below 0.7, as is evident in the posterior density of u plotted in Figure 6 for the equity data, where $u = \alpha/(\alpha + 1)$, $BF(u = u_0) = \pi(u = u_0 | r, DPM)/\pi(u = u_0)$ [19, 20]. This indicates that the equity data strongly support a nonparametric specification with only a few.

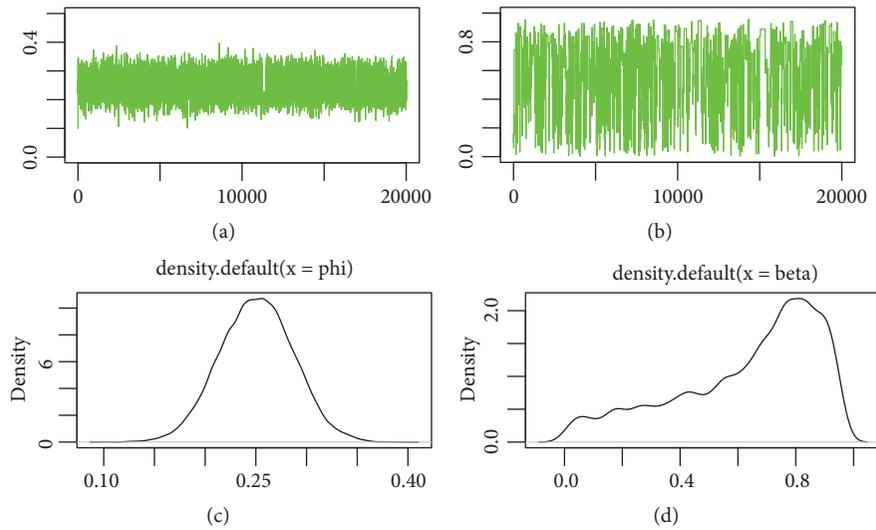


FIGURE 5: MCMC results for DAR-DPM: (a) the estimation results of ϕ and (b) the estimation of β ; (c) and (d) are their respective density maps. It can be observed that β is left-biased, so we take the median as the estimated value because the median is less affected by skewness.

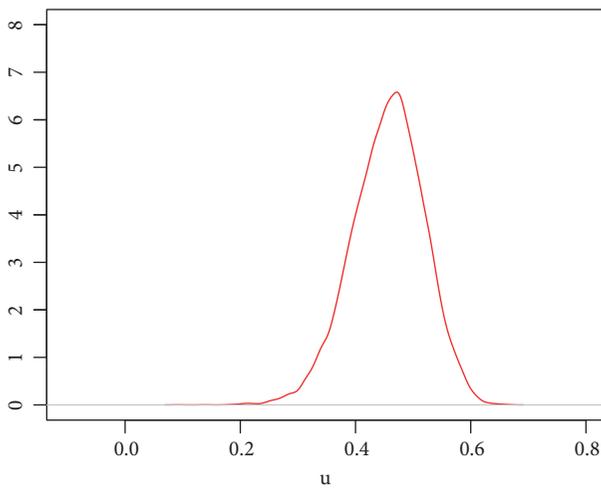


FIGURE 6: Posterior density of $u \equiv \alpha/(1 + \alpha)$.

6. Conclusion

In this article, a semiparametric Bayesian approach has been developed. The innovation distribution has been modeled using a scale mixture of a Gaussian model with a DP prior for the mixed distribution. An MCMC algorithm based on a combination of retrospective and slice sampling has been constructed to obtain samples from posterior distributions of the model parameters. The results that we achieved in each of our experiments in both a simulation study and a real data application are quite encouraging.

Data Availability

The text data used to support the findings of this study are included within the supplementary information file. Data Sources: Federal Reserve Economic Data can be obtained

by visiting the following websites: <https://www.federalreserve.gov/>. Additionally, the Treasury bill data of this study is available at “<https://fred.stlouisfed.org/categories/116>”.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was partially supported by NSFC (11071202& 11861025) & Science and Technology Foundation of Guizhou Province (LKS[2013]05), China.

References

- [1] D. V. Lindley, *Bayesian Statistics. A Review*, SIAM, Pennsylvania, Pa, USA, 1972.
- [2] S. H. Feizjavadian and R. Hashemi, “Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall-Olkin bivariate Weibull distribution,” *Computational Statistics and Data Analysis*, vol. 82, pp. 19–34, 2015.
- [3] J. E. Griffin, “An adaptive truncation method for inference in Bayesian nonparametric models,” *Statistics and Computing*, vol. 26, no. 1, pp. 1–19, 2016.
- [4] T. S. Ferguson, “A Bayesian analysis of some nonparametric problems,” *The Annals of Statistics*, vol. 1, pp. 209–230, 1973.
- [5] M. D. Escobar and M. West, “Bayesian density estimation and inference using mixtures,” *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 577–588, 1995.
- [6] T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [7] J. Davidson, “Moment and memory properties of linear conditional heteroscedasticity models, and a new model,” *Journal of Business and Economic Statistics*, vol. 22, no. 1, pp. 16–29, 2004.

- [8] J. Rombouts, L. Stentoft, and F. Violante, "The value of multivariate model sophistication: an application to pricing dow jones industrial average options," *International Journal of Forecasting*, vol. 30, no. 1, pp. 78–98, 2014.
- [9] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.
- [10] A. A. Weiss, "Asymptotic theory for ARCH models: estimation and testing," *Econometric Theory*, vol. 2, no. 1, pp. 107–131, 1986.
- [11] C. Francq and J. Zakoian, "Estimating linear representations of nonlinear processes," *Journal of Statistical Planning and Inference*, vol. 68, no. 1, pp. 145–165, 1998.
- [12] C. Francq and J. Zakoian, "Covariance matrix estimation for estimators of mixing weak ARMA models," *Journal of Statistical Planning and Inference*, vol. 83, no. 2, pp. 369–394, 2000.
- [13] S. Ling, "A double AR(p) model: structure and estimation," *Statistica Sinica*, vol. 17, no. 1, pp. 161–175, 2007.
- [14] E. Jacquier, N. G. Polson, and P. E. Rossi, "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, vol. 122, no. 1, pp. 185–212, 2004.
- [15] E.-I. Delatola and J. E. Griffin, "Bayesian nonparametric modelling of the return distribution with stochastic volatility," *Bayesian Analysis*, vol. 6, no. 4, pp. 901–926, 2011.
- [16] E.-I. Delatola and J. E. Griffin, "A Bayesian semiparametric model for volatility with a leverage effect," *Computational Statistics and Data Analysis*, vol. 60, pp. 97–110, 2013.
- [17] M. J. Jensen, "Semiparametric bayesian inference of long-memory stochastic volatility models," *Journal of Time Series Analysis*, vol. 25, no. 6, pp. 895–922, 2004.
- [18] M. J. Jensen and J. M. Maheu, "Bayesian semiparametric stochastic volatility modeling," *Journal of Econometrics*, vol. 157, no. 2, pp. 306–316, 2010.
- [19] M. J. Jensen and J. M. Maheu, "Bayesian semiparametric multivariate GARCH modeling," *Journal of Econometrics*, vol. 176, no. 1, pp. 3–17, 2013.
- [20] M. J. Jensen and J. M. Maheu, "Estimating a semiparametric asymmetric stochastic volatility model with a Dirichlet process mixture," *Journal of Econometrics*, vol. 178, pp. 523–538, 2013.
- [21] A. Virbickaite, M. C. Ausin, and P. Galeano, "A Bayesian non-parametric approach to asymmetric dynamic conditional correlation model with application to portfolio selection," *Computational Statistics & Data Analysis*, vol. 100, pp. 814–829, 2016.
- [22] M. C. Ausin, P. Galeano, and P. Ghosh, "A semiparametric Bayesian approach to the analysis of financial time series with applications to value at risk estimation," *European Journal of Operational Research*, vol. 232, no. 2, pp. 350–358, 2014.
- [23] J. F. Kingman, "Subadditive ergodic theory," *Annals of Probability*, vol. 1, no. 6, pp. 883–899, 1973.
- [24] Z. Lu and Z. Jiang, "L1 geometric ergodicity of a multivariate nonlinear AR model with an ARCH term," *Probability Letters*, vol. 51, no. 2, pp. 121–130, 2001.
- [25] A. Y. Lo, "On a class of Bayesian nonparametric estimates. I. density estimates," *The Annals of Statistics*, vol. 12, no. 1, pp. 351–357, 1984.
- [26] J. Sethuraman, "A constructive definition of Dirichlet priors," *Statistica Sinica*, vol. 4, no. 2, pp. 639–650, 1994.
- [27] S. N. Maceachern and P. Muller, "Estimating mixture of dirichlet process models," *Journal of Computational and Graphical Statistics*, vol. 7, no. 2, pp. 223–238, 1998.
- [28] T. Nakatsuma, "Bayesian analysis of ARMA-GARCH models: a Markov chain sampling approach," *Journal of Econometrics*, vol. 95, no. 1, pp. 57–69, 2000.
- [29] W. Ming and W. Yunshun, "Estimating GARCH modeling using metropolis-hastings method in R," *Open Journal of Statistics*, vol. 08, no. 06, pp. 931–938, 2018.
- [30] T. Nakatsuma, "A Markov-chain sampling algorithm for GARCH models," *Studies in Nonlinear Dynamics and Econometrics*, vol. 3, pp. 107–117, 1998.



Hindawi

Submit your manuscripts at
www.hindawi.com

