Quantile Double AR Time Series Models for Financial Returns

Yuzhi Cai [*] , *Swansea University*

Gabriel Montes-Rojas, *City University London*

Jose Olmo, *Centro Universitario de la Defensa de Zaragoza*

Abstract

In this paper we develop a novel quantile double AR model for modelling financial time series. This is done by specifying a generalized lambda distribution to the quantile function of the location-scale double autoregressive model developed in Ling (2004, 2007). Model parameter estimation uses MCMC Bayesian methods. A novel simulation technique is introduced for forecasting the conditional distribution of financial returns m-periods ahead, and hence any predictive quantities of interest. The application to forecasting Value-at-Risk at different time horizons and coverage probabilities for Dow Jones Industrial Average shows that our method works very well in practice.

Key words: Bayesian methods, density forecasts, generalized lambda distribution, quantile function, quantile forecasts.

---

[*]Address for correspondence: Dr Yuzhi Cai, School of Business and Economics, Swansea University, Swansea, SA2 8PP, United Kingdom. Email: y.cai@swansea.ac.uk

1

1 Introduction

Consider Ling (2004, 2007) double AR(p) model defined by

$$y_t = a_0 + a_1 y_1 + \cdots a_p y_{t-p} + \eta_t \sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_p y_{t-p}^2}, \qquad (1)$$

where $b_i > 0$ $(i = 0, \ldots, p)$, η_t is an independent random sequence and $\eta_t \sim N(0, 1)$, and y_s is independent of η_t for $s < t$. This model is a special case of ARMA-ARCH models proposed by Weiss (1984), but it is different from the ARCH models proposed by Engle (1982) if $a_i \neq 0$. This model encompasses a large proportion of applications in empirical economics and finance where volatility plays an important role in modeling autoregressive series (further discussion on the motivation for the double AR(p) models can be found in Weiss (1984) and Ling (2004, 2007)).

It is worth mentioning that model (1) has only been investigated for the conditional mean. Moreover, the normality requirement on the error term η_t of the model is quite restrictive as many economic and financial time series are non-Gaussian. This motivated us to develop a novel quantile double AR model corresponding to model (1) that also allows us to deal with general non-Gaussian time series easily. In order to illustrate our approach we apply the developed model to the Dow Jones Industrial Average (DJIA).

In many areas of research studying extreme quantiles is of fundamental importance. An example is Value-at-Risk (VaR) in economics and finance. Statistical inference on extreme quantiles can be made once the probability distribution or density function of the innovations η_t is known. However, a direct quantile approach to statistical modelling has recently become more popular. One of the methods for estimating conditional quantiles of y_t is to use the quantile regression techniques (see Koenker and Bassett (1978) and Koenker (2005)), which allow us to obtain a sequence of conditional quantiles by using a semi-parametric model, that is, without imposing strong distributional assumptions on η_t. The development in this area is rapid. For example, Koenker and Zhao (1996) extended

quantile regression to linear ARCH models and Engle and Manganelli (2004) developed a different conditional autoregressive VaR model. Xiao and Koenker (2009) developed a two-step approach of quantile regression estimation for linear GARCH time series. Taylor (2008) proposed the exponentially weighted quantile regression for estimating time-varying quantiles, and Giot and Laurent (2003) model VaR using ARCH models based on skewed t-distribution. Galvao (2009, 2011) considered unit root quantile autoregression testing and quantile regression dynamic panel model with fixed effects. Cai and Stander (2008) proposed a quantile self-exciting autoregressive time series model and developed a Bayesian approach for parameter estimation. Cai (2007, 2010c) also proposed forecasting methods for such models.

However, one of the problems associated with the above models is that the extreme quantiles (corresponding to extreme risks) may not be properly estimated, for example, the estimated quantile curves may cross over (non-monotonicity). This is because when the probability τ approaches the extremes (i.e. 0 or 1), the estimated τth conditional quantile becomes less and less precise.

One way to deal with the crossing-over problem is to specify a parametric conditional quantile function model (see Gilchrist (2000) for an excellent introduction to this parametric approach). This procedure allows to estimate the whole conditional quantile function of y_t directly using a wider class of distributions for η_t, including those which are defined only via their quantile functions and that may not have closed mathematical expressions for their density or distribution functions. Our quantile double AR model follows this procedure and it enables us to obtain valid estimation of extreme quantiles. Furthermore, quantile function properties allow us to construct the distribution of η_t by combining several quantile functions in a proper way (see Gilchrist, 2000), leading to an appropriate model for capturing important features of economic and financial time series, including the occurrence of extremes and volatility clustering. The flexibility of the generalized lambda distribution (see Freimer et al. (1988)) motivated us to use this distribution in the construction of the

3

proposed quantile double AR model and this distribution because also enables us to study the conditional quantile function of y_t directly.

We also propose a Bayesian method to estimate the model parameters, which is a moderate extension of the Bayesian method proposed by Cai (2009, 2010a, b). It will become clear later in the paper that, unlike some other estimation methods, our Bayesian approach also plays an important role in our proposed forecasting method which allows us to take the uncertainty of the estimated model parameters into account when forecasting. Our forecasting method can be used to obtain m-step ahead ($m > 0$) out-of-sample forecasts, not just point forecasts but also the whole predictive distributions via the quantile function models. Little work can be found on this in the literature.

Our results show that volatility clustering phenomenon observed in many financial returns is reflected more parsimoniously in this model by the generalized lambda distribution parameters. This indicates that despite the prominence in the literature of models to forecast conditional volatility, it can be the case that current volatility is not so instrumental for forecasting the conditional distribution of returns and researchers/practitioners need to look at other parameters driving the behavior of the distribution tails. Furthermore, the flexibility of the quantile double AR model permits asymmetries in the dynamics of the tails of the predictive conditional distribution of returns. This implies that this model provides a better understanding of the impact of large negative/positive returns in the likelihood of future losses/gains.

Therefore, the main contributions of this paper are: (a) It proposes a quantile double AR model for economic and financial time series. Different from model (1), our model study the conditional quantile function of y_t, which allows us to model the extreme quantiles directly and to study non-Gaussian time series easily. (b) Combined with our Bayesian method we also propose a forecasting method for quantile function models, which enables us to obtain m-step ahead out-of-sample predictive distributions, and hence any predictive quantities of interest, including extreme quantiles.

The article is structured as follows. In Section 2, we propose the model and briefly discuss the Bayesian estimation method for model fitting. Section 3 discusses an out-of-sample forecasting method that also exploits features of MCMC Bayesian methods. Section 4 applies these techniques to modeling and forecasting m-periods ahead the conditional distribution of the Dow Jones Industrial Average (DJIA). Further discussion and comments are found in Section 5.

2   The model and parameter estimation

Let $y_1, \ldots, y_n$ be an observed time series. The proposed quantile double AR(k_1, k_2) time series model takes the form

$$Q_{y_t}(\tau \mid \boldsymbol{\beta}, \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_{k_1} y_{t-k_1} + \sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_{k_2} y_{t-k_2}^2} \, Q(\tau, \boldsymbol{\gamma}),$$
(2)

where (k_1, k_2) is the order of the model, $\boldsymbol{\beta} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\gamma})$ is a vector of model parameters, where $\boldsymbol{\eta}_1 = (a_0, \ldots, a_{k_1})$ and $\boldsymbol{\eta}_2 = (b_0, \ldots, b_{k_2})$, in which $b_0 > 0$, $b_j \geq 0$, $j = 1, \ldots, k_2$ and $\mathbf{y}_{t-1} = (y_1, \ldots, y_{t-1})^\top$. Finally, $Q(\tau, \boldsymbol{\gamma})$ is a well defined quantile function used to describe the distribution of the error term of the model.

Note that model (2) is equivalent to the quantile process of model (1) if $Q(\tau, \boldsymbol{\gamma})$ is the quantile function of $N(0, 1)$. If we let

$$Q(\tau, \boldsymbol{\gamma}) = \frac{\tau^{\gamma_1} - 1}{\gamma_1} - \frac{(1 - \tau)^{\gamma_2} - 1}{\gamma_2}, \quad \gamma_1 < 0, \ \gamma_2 < 0,$$
(3)

then our proposed double AR quantile function model becomes

$$Q_{y_t}(\tau \mid \boldsymbol{\beta}, \mathbf{y}_{t-1}) = a_0 + a_1 y_{t-1} + \cdots + a_{k_1} y_{t-k_1}$$
(4)

$$+ \sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_{k_2} y_{t-k_2}^2} \left(\frac{\tau^{\gamma_1} - 1}{\gamma_1} - \frac{(1 - \tau)^{\gamma_2} - 1}{\gamma_2} \right),$$

which in fact defines the generalized lambda distribution (see Freimer et al. (1988)), for which, both the mean and the spread of the distribution depend on the past observed data.

This quantile function model is appropriate for modeling the sequence of financial log returns for the following reasons: (a) The generalized lambda distribution can provide a very accurate approximation to most common distributions such as normal, lognormal and Weibull distributions as well as others (see Fournier et al. (2007)), and γ_1 and γ_2 jointly determine the shape of the left and right tails of the distribution respectively. (b) When $\gamma_1 < 0$ and $\gamma_2 < 0$ the support of the distribution is $(-\infty, \infty)$. As the log returns of daily financial time series can be positive or negative, they are well defined on the support of the generalized lambda distribution. (c) Generally we would expect that both the center and the scale of the conditional distribution of log returns will depend on the past log returns. The proposed model suggests a way to deal with these issues. (d) Quantile models are not sensitive to outliers compared with the models for the conditional mean, hence model (4) is robust with respect to the modeling of extreme log returns. (e) We will see later in the paper that the proposed quantile model also has the ability to deal with the clustering of financial log returns.

Now let us consider parameter estimation. The conditional likelihood of the observed data is given by

$$L(y_{k+1}, \ldots, y_n \mid \mathbf{y}_k, \boldsymbol{\beta}) = \prod_{t=k+1}^{n} \left\{ \sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_{k_2} y_{t-k_2}^2} \left(\tau_t^{\gamma_1 - 1} + (1 - \tau_t)^{\gamma_2 - 1} \right) \right\}^{-1},$$

where $k = \max\{k_1, k_2\}$, $\mathbf{y}_k = \{y_1, \ldots, y_k\}$ and τ_t $(t = k+1, \ldots, n)$ satisfy

$$
\begin{aligned}
y_t &= a_0 + a_1 y_{t-1} + \cdots + a_{k_1} y_{t-k_1} \\
&\quad + \sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_{k_2} y_{t-k_2}^2} \left(\frac{\tau_t^{\gamma_1} - 1}{\gamma_1} - \frac{(1-\tau_t)^{\gamma_2} - 1}{\gamma_2} \right).
\end{aligned}
\tag{5}
$$

Let $\pi(\boldsymbol{\beta})$ be the prior density function of the parameters. Then the posterior distribution

6

of the parameters is given by

$$\pi(\boldsymbol{\beta} \mid \mathbf{y}_n) \propto L(y_{k+1}, \ldots, y_n \mid \mathbf{y}_k, \boldsymbol{\beta}) \pi(\boldsymbol{\beta}).$$

It can be shown that under certain conditions the posterior distribution is well defined. Specifically, we can use the following prior density functions:

$$\pi(a_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-a_i^2/2\sigma_i^2}, \quad \pi(b_j) = \frac{1}{\sqrt{2\pi}s_j b_j} e^{-(\log(b_j))^2/2s_j^2}, \ b_j > 0$$

(6)

$$\pi(\gamma_\ell) = \frac{1}{\sqrt{2\pi}\lambda_\ell(-\gamma_\ell)} e^{-(\log(-\gamma_\ell))^2/2\lambda_\ell^2}, \ \gamma_\ell < 0,$$

where $i = 0, \ldots, k_1$, $j = 0, \ldots, k_2$, $\ell = 1, 2$.

Note that in (6) we have used the usual default for the prior means (see, for example, Chipman et al. (2001)), i.e. the means of the prior distributions are all 0, which is a neutral choice reflecting indifference between positive and negative values. In practice, if we have some prior knowledge about those parameters, then we could set the prior means at other proper values. On the other hand, large values of σ_i, s_j and λ_ℓ imply weak prior information, while small values of them reflect strong prior information. Therefore, we believe these prior distributions are useful in practice and we will use them in this paper.

The posterior samples of the model parameters of a Markov chain Monte Carlo (MCMC) method are useful for taking into account the uncertainty around the estimated model parameters. For this reason, we generalize the MCMC method developed by Cai (2009, 2010a, b) with the aim of constructing forecasts of the quantile process that do not suffer from estimation risk. The outline of the generalized MCMC method is given below.

Let $\boldsymbol{\beta}$ be the current value of the parameters and τ_t $(t = k + 1, \ldots, n)$ the associated probabilities. Let $\boldsymbol{\beta}'$ be the proposed value, whose components are independently simulated as follows: $a_i' \sim N(a_i, \sigma_{a_i}^2)$ for $i = 0, \ldots, k_1$, $b_j' \sim N(b_j, \sigma_{b_j}^2)$ such that $b_0' > 0$ and $b_j' \geq 0$ for $j = 1, \ldots, k_2$, and $\gamma_\ell' \sim N(\gamma_\ell, \sigma_{\gamma_\ell}^2)$ such that $\gamma_\ell' < 0$ for $\ell = 1, 2$. Then we obtain τ_t'

$(t = k+1, \ldots, n)$ by solving equation (5). We accept the proposed value with a probability $\min\{a, 1\}$, where

$$a = \frac{\pi(\boldsymbol{\beta}' \mid \mathbf{y}_n) \prod_{j=0}^{k_2}(1 - \Phi\left(-b_j/\sigma_{b_j}\right)) \prod_{\ell=1}^{2} \Phi(-\gamma_\ell/\sigma_{\gamma_\ell})}{\pi(\boldsymbol{\beta} \mid \mathbf{y}_n) \prod_{j=0}^{k_2}(1 - \Phi\left(-b_j'/\sigma_{b_j}\right)) \prod_{\ell=1}^{2} \Phi(-\gamma_\ell'/\sigma_{\gamma_\ell})},$$

the proof of which is available upon request.

The fitted model can be checked through the standardized residual series defined by

$$r_t = \frac{y_t - (a_0 + a_1 y_{t-1} + \cdots + a_{k_1} y_{t-k_1})}{\sqrt{b_0 + b_1 y_{t-1}^2 + \cdots + b_{k_2} y_{t-k_2}^2}}.$$

If model (1) is an appropriate description of the time series y_t, then r_t is an independently and identically distributed (iid) random sequence following a generalized lambda distribution as defined in (3).

In Bayesian analysis, posterior odds may be used to compare several fitted models:

$$\frac{p((k_1, k_2) \mid \mathbf{y}_n)}{p((k_1', k_2') \mid \mathbf{y}_n)} = \frac{p(\mathbf{y}_n \mid (k_1, k_2))}{p(\mathbf{y}_n \mid (k_1', k_2'))} \times \frac{p((k_1, k_2))}{p((k_1', k_2'))},$$

where (k_1, k_2) and (k_1', k_2') represent two models with different orders, $p((k_1, k_2))$ is the prior probability for model of order (k_1, k_2). The best model corresponds to the largest $p((k_1, k_2) \mid \mathbf{y}_n)$. In this paper, we assume a uniform prior for each model of order (k_1, k_2). Hence the posterior odds comparisons reduce to Bayes factor comparisons

$$\frac{p((k_1, k_2) \mid \mathbf{y}_n)}{p((k_1', k_2') \mid \mathbf{y}_n)} = \frac{p(\mathbf{y}_n \mid (k_1, k_2))}{p(\mathbf{y}_n \mid (k_1', k_2'))}.$$

Therefore, equivalently, the best model now corresponds to the largest $p(\mathbf{y}_n \mid (k_1, k_2))$. Various methods have been developed to estimate the Bayes factor in the literature. For example, Newton and Raftery (1994) proposed a method that was based on a mixture of the samples obtained from the prior and posterior distribution of the parameters. Lewis and

Raftery (1997) studied the approximation to $p(\mathbf{y}_n \mid (k_1, k_2))$ based on Laplace-Metropolis estimators. In this paper, we use the method proposed by Gelfand and Dey (1994) to estimate $p(\mathbf{y}_n \mid (k_1, k_2))$ and the formula is given by

$$\{p(\mathbf{y}_n \mid (k_1, k_2))\}^{-1} \approx (U - U_0)^{-1} \sum_{u=U_0}^{U} g(\boldsymbol{\beta}^{(u)})\{p(\mathbf{y}_n \mid (k_1, k_2), \boldsymbol{\beta}^{(u)})p(\boldsymbol{\beta}^{(u)} \mid (k_1, k_2))\}^{-1},$$

where $\boldsymbol{\beta}^{(u)}$ are posterior samples and $g(\cdot)$ is an arbitrary density function. Note that the above approximation is unstable if g has tails thicker than $p(\mathbf{y}_n \mid (k_1, k_2), \boldsymbol{\beta}^{(u)})p(\boldsymbol{\beta}^{(u)} \mid (k_1, k_2))$. In this paper $g(\cdot)$ is taken as the product of the density functions which have the same form as the prior density functions but with smaller parameter values, leading to a stable approximation to $\{p(\mathbf{y}_n \mid (k_1, k_2))\}^{-1}$. Also note that $p(\mathbf{y}_n \mid (k_1, k_2), \boldsymbol{\beta}^{(u)})p(\boldsymbol{\beta}^{(u)} \mid (k_1, k_2))$ is just the product of the likelihood of the data and the priors of the parameters for model (k_1, k_2), both of which are evaluated at the uth posterior sample.

3   Forecasting

An important and empirically relevant issue in modeling financial and economic time series is forecasting ability. Cai (2010c) in a similar setting proposed a forecasting method for quantile self-exciting autoregressive (QSETAR) time series models. This method is, however, semi-parametric and cannot be used then for the proposed quantile double AR models. In this section we propose a forecasting method based on the quantile function model (4), which enables us to obtain multi-step ahead predictive density functions, and hence any quantities of interest. The methodology also takes the uncertainty of the parameters into account in the forecasts. Note that the forecasting method is based on the assumption that the estimated parameter values remain valid even when $t > n$.

Specifically, when $m = 1$, i.e. for one step ahead forecasting, we have

$$
Q_{y_{n+1}}(\tau \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n) = a_0^{(u)} + a_1^{(u)} y_n + \cdots + a_{k_1}^{(u)} y_{n+1-k_1}
$$
$$
+ \sqrt{b_0^{(u)} + b_1^{(u)} y_n + \cdots + b_{k_2}^{(u)} y_{n+1-k_2}} \left(\frac{\tau^{\gamma_1^{(u)}} - 1}{\gamma_1^{(u)}} - \frac{(1-\tau)^{\gamma_2^{(u)}} - 1}{\gamma_2^{(u)}} \right),
$$

(7)

where $\boldsymbol{\beta}^{(u)}$ is the uth posterior sample. Let $f(y_{n+1} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n)$ be the corresponding density function of y_{n+1} given $\boldsymbol{\beta}^{(u)}$ and $\mathbf{y}_n$, then

$$
f(y_{n+1} \mid \mathbf{y}_n) = \int_{\boldsymbol{\beta}} f(y_{n+1} \mid \boldsymbol{\beta}, \mathbf{y}_n) \pi(\boldsymbol{\beta} \mid \mathbf{y}_n) d\boldsymbol{\beta} \approx \frac{1}{U-U_0} \sum_{u=U_0}^{U} f(y_{n+1} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n).
$$

(8)

Expression (8) suggests that a sample of y_{n+1} given $\mathbf{y}_n$ can be obtained by simulating $y_{n+1}^{(u,i_1)}$ ($i_1 = 1, \ldots, I_1$) from each $f(y_{n+1} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n)$ ($u = U_0, \ldots, U$). Then $\{y_{n+1}^{(u,i_1)} : u = U_0, \ldots, U, \ i_1 = 1, \ldots, I_1\}$ forms a sample from $f(y_{n+1} \mid \mathbf{y}_n)$. This simulation procedure is valid because each of the $f(y_{n+1} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n)$ makes equal contributions to $f(y_{n+1} \mid \mathbf{y}_n)$. Note that $y_{n+1}^{(u,i_1)}$ can be simulated by using (7). These simulated samples can then be used for estimating $f(y_{n+1} \mid \mathbf{y}_n)$ and for forecasting when $m > 1$.

For $m = 2$, we have

$$
f(y_{n+2} \mid \mathbf{y}_n) = \int_{y_{n+1}} \int_{\boldsymbol{\beta}} f(y_{n+2} \mid \boldsymbol{\beta}, y_{n+1}, \mathbf{y}_n) f(y_{n+1} \mid \boldsymbol{\beta}, \mathbf{y}_n) \pi(\boldsymbol{\beta} \mid \mathbf{y}_n) d\boldsymbol{\beta} dy_{n+1}
$$
$$
\approx \frac{1}{U-U_0} \sum_{u=U_0}^{U} \int_{y_{n+1}} f(y_{n+2} \mid \boldsymbol{\beta}^{(u)}, y_{n+1}, \mathbf{y}_n) f(y_{n+1} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_n) dy_{n+1}
$$
$$
\approx \frac{1}{U-U_0} \sum_{u=U_0}^{U} \frac{1}{I_1} \sum_{i_1=1}^{I_1} f(y_{n+2} \mid \boldsymbol{\beta}^{(u)}, y_{n+1}^{(u,i_1)}, \mathbf{y}_n)
$$

Therefore, a sample $y_{n+2}^{(u,i_1,i_2)}$ of y_{n+2} given $\mathbf{y}_n$ can also be obtained by using (7) conditional on $y_{n+1}^{(u,i_1)}$, $\mathbf{y}_n$ and $\boldsymbol{\beta}^{(u)}$. These samples can then be used for estimating $f(y_{n+2} \mid \mathbf{y}_n)$ and for forecasting when $m > 2$.

Generally, we have

$$f(y_{n+m} \mid \mathbf{y}_n) \approx \{(U - U_0)I_1 \cdots I_{m-1}\}^{-1} \sum_{u=U_0}^{U} \sum_{i_1=1}^{I_1} \cdots$$
$$\sum_{i_{m-1}=1}^{I_{m-1}} f(y_{n+m} \mid \boldsymbol{\beta}^{(u)}, \mathbf{y}_{n+m-1}^{(u,i_1,\ldots,i_{m-1})}, \mathbf{y}_n)$$

where $\mathbf{y}_{n+m-1}^{(u,i_1,\ldots,i_{m-1})} = (y_{n+m-1}^{(u,i_1,\ldots,i_{m-1})}, y_{n+m-2}^{(u,i_1,\ldots,i_{m-2})}, \ldots, y_{n+m-k_2}^{(u,i_1,\ldots,i_{m-k_2})})$, and $y_{n+m-j}^{(u,i_1,\ldots,i_{m-j})} = y_{n+m-j}$ if $m - j \leq 0$ $(j = 1, \ldots, k_2)$. Hence, a sample of y_{n+m} given $\mathbf{y}_n$ can also be obtained.

Our results in the next section show that the above forecasting procedure works very well in practice.

4 Quantile double AR models for Financial Returns

In this section, we present some interesting results on applying the above developed methodology to the log returns of the Dow Jones Industrial Average. The data covers the period between 2 January 2004 and 8 October 2010. The length of the series is 1705. The time series plots of the observed series and its log returns, denoted by y_t $(t = 1, \ldots, n = 1704)$ are given in Figure 1. As expected, the observed DJIA series shows occurrence of extremes and volatility clustering. We will see that the proposed quantile function model can cope with these features very well.

We fitted several quantile double AR models to the log returns with different orders: $k_1 = 1, 2$ and $k_2 = 1, 2$. The initial values required by the MCMC method for each model were taken as $a_0^{(0)} = \bar{y}$, $a_i^{(0)} = 0$ $(i = 1, \ldots, k_1)$, $b_0^{(0)} = \bar{s}$, $b_j^{(0)} = 0$ $(j = 1, \ldots, k_2)$; $\gamma_1^{(0)}$ and $\gamma_2^{(0)}$ are two random samples from negative exponential distributions with rates 3 and 4, respectively. This specification about the initial parameter values guarantees that it is in the support of the posterior distribution. In this application, we have $\bar{y} = 0.003271$, $\bar{s} = 1.635145$, $\gamma_1^{(0)} = -0.258775$ and $\gamma_2^{(0)} = -0.332507$. The prior distributions of the model parameters are given by (6). To reflect the fact that we do not have any prior
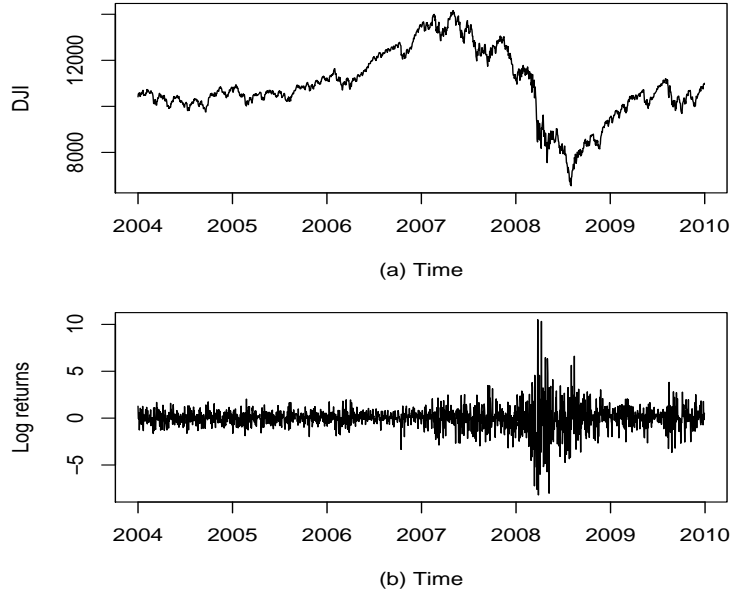
11

Figure 1: (a) Time series plot of the DJIA between 2/1/2004-8/10/2010. (b) Time series plot of the DJIA log returns.

Order	$k_1 = 1, k_2 = 1$	$k_1 = 2, k_2 = 1$	$k_1 = 1, k_2 = 2$	$k_1 = 2, k_2 = 2$
$\{p(\mathbf{y}_n \mid (k_1, k_2))\}^{-1}$	6.4431	13.6356	19.0237	17.1192

Table 1: $\{p(\mathbf{y}_n \mid (k_1, k_2))\}^{-1}$ values of the fitted models for the DJIA log returns.

information about the true parameter values, we let σ_i, s_j and λ_ℓ equal 5 for all possible values of i, j and ℓ.

For each model of order (k_1, k_2), a Markov chain of length 200000 steps was run. Time series plots of the posterior samples suggest that it is appropriate to adopt a burn-in of 10000 values, after which the simulated parameter values were saved once every 100 steps. The Bayesian estimates of the model parameters were taken to be the sample means of the posterior samples. Figure 2 shows the results for $k_1 = 1$ and $k_2 = 1$, where the vertical lines on the histograms give the locations of the Bayesian estimates of the model parameters.

Table 1 shows the estimated $\{p(\mathbf{y}_n \mid (k_1, k_2))\}^{-1}$ values for each fitted model. Therefore, by using the Bayes factor, we see that the best fitted model is the model of order $k_1 = 1$

12

and $k_2 = 1$, which is further supported by the QQ-plots of the four fitted models given in Figure 3.
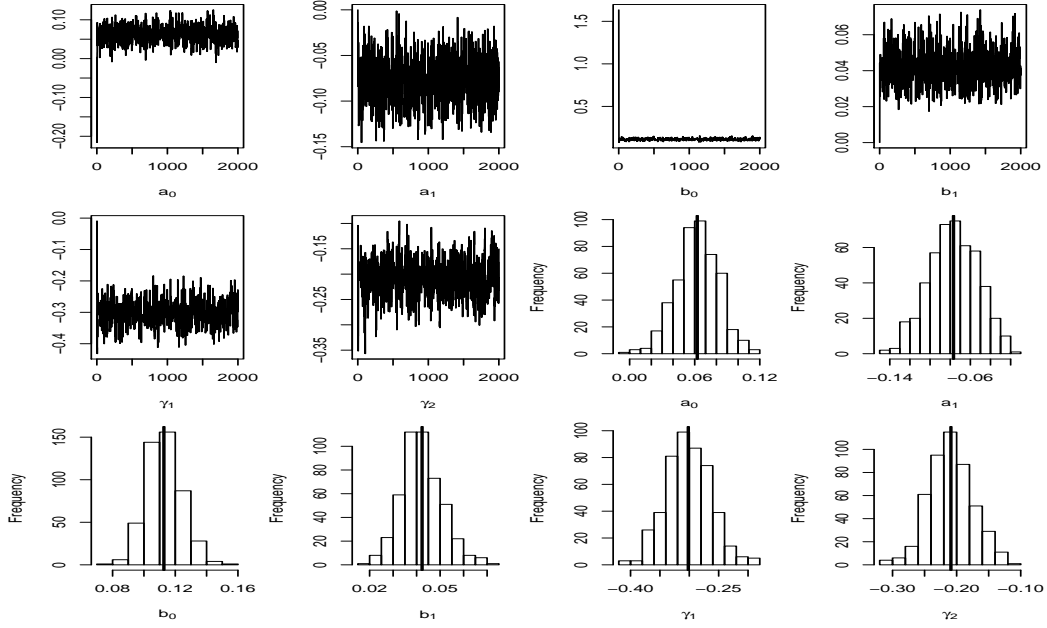

Figure 2: Time series plots and histograms of the posterior samples for the model with $k_1 = 1$ and $k_2 = 1$ for the DJIA log returns.

If we use the mean of the posterior samples as the Bayesian estimate of the model parameters, then the fitted model is given by

$$Q_{y_t}(\tau \mid \hat{\boldsymbol{\beta}}, \mathbf{y}_{t-1}) = 0.0623 - 0.077 y_{t-1} + \sqrt{0.113 + 0.042 y_{t-1}^2} \left(\frac{\tau^{-0.301} - 1}{-0.301} - \frac{(1-\tau)^{-0.209} - 1}{-0.209} \right).$$
(9)

We also fitted several ARMA-GARCH models to the same data set by using the package fGarch in the statistical software R. As the log returns show large variations, we used the t-distribution to model the error term of the model, see Bollerslev (1987). Table 2 shows the BIC values of the fitted models, which suggests that the best fitted model is AR(1)-GARCH(1,1) and which is further supported by Figure 4 as none of other QQ-plots shows
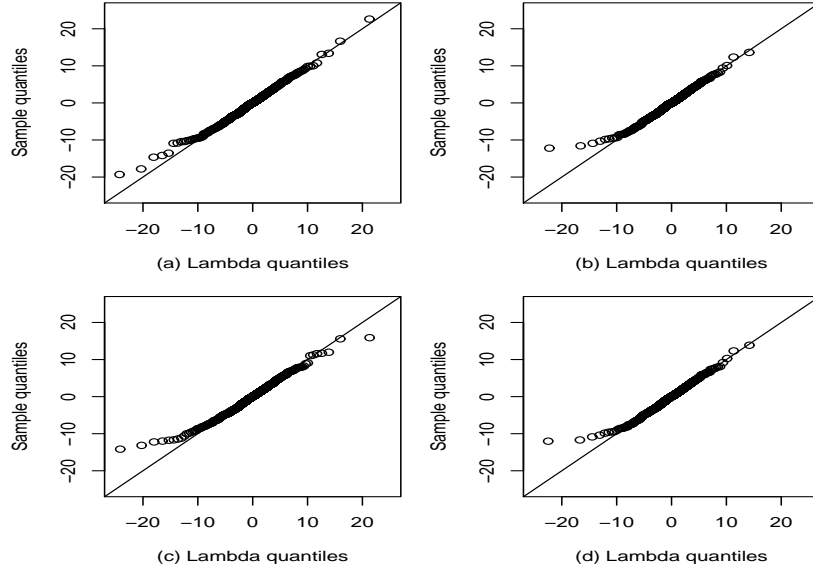
13

Figure 3: QQ-plots of the four fitted models of orders (a) $k_1 = 1, k_2 = 1$, (b) $k_1 = 1, k_2 = 2$, (c) $k_1 = 2, k_2 = 1$ and (d) $k_1 = 2, k_2 = 2$ for the DJIA log returns.

a better fit. The fitted AR(1)-GARCH(1,1) model is given by

$$
\begin{aligned}
y_t &= \underset{(0.0184)}{0.0327} - \underset{(0.0242)}{0.0557} y_{t-1} + \sqrt{h_t} \varepsilon_t \,,
\end{aligned}
\tag{10}
$$

where

$$
h_t = \underset{(0.0036)}{0.0085} + \underset{(0.0133)}{0.0808} v_{t-1}^2 + \underset{(0.0133)}{0.9143} h_{t-1}, \quad v_t = y_t - 0.0327 + 0.0557 y_{t-1},
$$

ε_t follows the t-distribution with 7.2150(1.3335) degrees of freedom, and the numbers in brackets are the standard errors of the estimated parameter values.

A sequence of fitted one-step ahead predictive density functions of the DJIA log returns conditional on different information sets have been obtained by using the developed forecasting method. Figure 5 shows the one-step ahead predictive density functions during the period from 23 December 2008 to 19 May 2009. It is worth mentioning that the differences

ARMA(1,0)-GARCH(1,1)	ARMA(1,1)-GARCH(1,1)	ARMA(0,1)-GARCH(1,1)
-2.6466	-2.6406	-2.6462
ARMA(1,0)-GARCH(1,2)	ARMA(1,1)-GARCH(1,2)	ARMA(0,1)-GARCH(1,2)
-2.6425	-2.6361	-2.6422
ARMA(1,0)-GARCH(2,1)	ARMA(1,1)-GARCH(2,1)	ARMA(0,1)-GARCH(2,1)
-2.6346	-2.6283	-2.6343

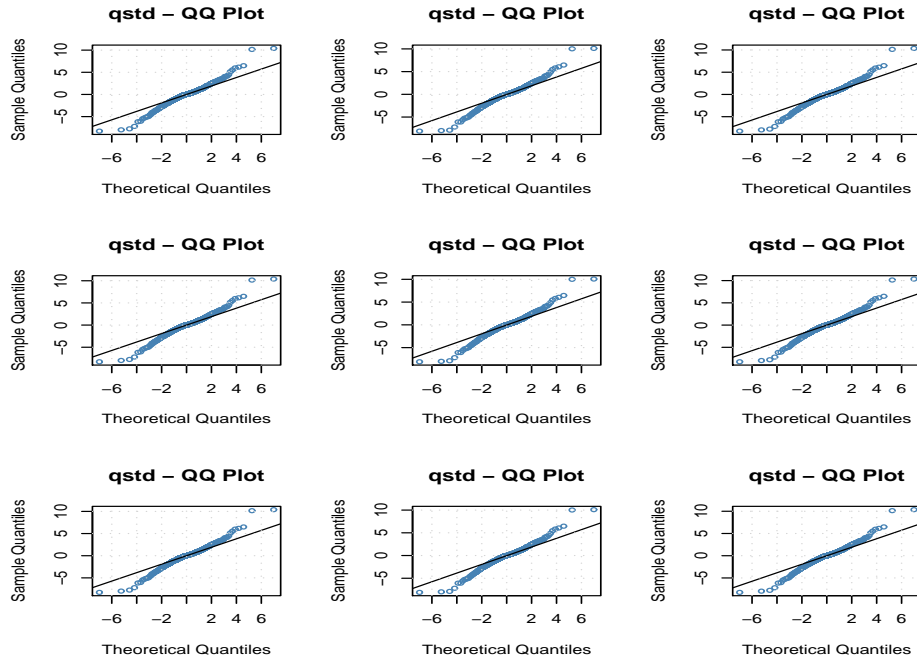Table 2: BIC values of the fitted ARMA-GARCH models

Figure 4: QQ-plots of the fitted ARMA-GARCH models for the DJIA log returns.

between the predictive density functions indicate the effects of the differences in information sets. For example, in this period of time the minimum closing DJIA was obtained on 9 March 2009 and on 10 March 2009 the closing DJIA had a significant recovery leading to a large log return (5.633%) shown by the darker continuous vertical line. The large log return on this day implies a higher level of uncertainty leading to a very flat predictive distribution of the log returns on the next day. On the other hand, for example, on 26 February 2009, we had a relatively small negative log return (-1.09%) shown by the darker dashed vertical line, leading to a relatively high peaked predictive distribution.
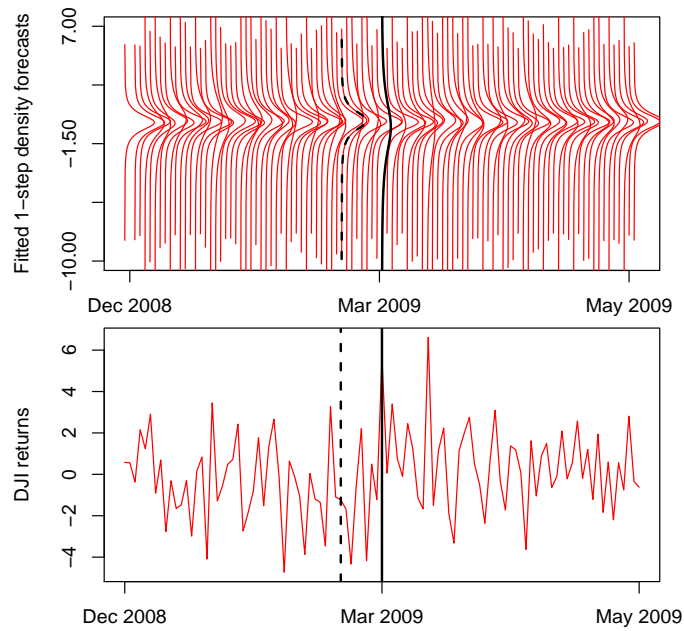

Figure 5: One-step ahead density forecasts for the DJIA log returns between 23 December 2008-19 May 2009.

We also estimated one-step ahead quantile forecasts at $\tau = 0.025, 0.25, 0.5, 0.75, 0.975$ levels by using the posterior samples. That is, for each posterior sample, we obtain a corresponding predictive quantile, and the final quantile is the average of those estimated quantiles. Table 3 shows the number of the observed returns and the corresponding percentage in different quantile ranges. A simple goodness-of-fit Pearson's test shows that our

Quantile ranges	< 0.025	$(0.025, 0.25)$	$(0.25, 0.5)$	$(0.5, 0.75)$	$(0.75, 0.975)$	> 0.975
True probabilities	0.025	0.225	0.25	0.25	0.225	0.025
Number of observed returns	43	384	422	449	359	46
Percentages	0.0252	0.2255	0.2478	0.2637	0.2108	0.0270

Table 3: The coverage of the fitted 1-step ahead predicted quantiles for the DJIA log returns.

τ	0.025	0.25	0.5	0.75	0.975
$\hat{\beta}_0^\tau$	-2.6622	-0.4731	0.0553	0.5406	2.3286
$SE(\hat{\beta}_0^\tau)$	0.2297	0.0351	0.0188	0.0290	0.1581
$\hat{\beta}_1^\tau$	0.0016	-0.0337	-0.0833	-0.1239	-0.2996
$SE(\hat{\beta}_0^\tau)$	0.1620	0.0274	0.0148	0.0224	0.1226

Table 4: Estimated parameter values of model (11).

one-step ahead predicted quantiles have the right coverage. The test does not reject the null hypothesis at 5% significance level.

Figure 6 shows the five fitted one-step ahead quantile curves (darker curves), corresponding to quantile levels at $\tau = 0.025, 0.25, 0.5, 0.75, 0.975$ from the bottom to the top respectively, where the lighter curve is the observed log returns. So it becomes clear that if we want to forecast the conditional VaR in risk management, then we could focus on the specific feature of the distribution at a quantile level that we are interested in.

We further compared our approach with the semi-parametric approach. A semi-parametric model

$$q_{y_t}^\tau = \beta_0^\tau + \beta_1^\tau y_{t-1} \tag{11}$$

was fitted to the DJIA log returns for $\tau = 0.025, 0.25, 0.5, 0.75, 0.975$. The estimated model parameters are given in Table 4 and the one-step ahead predictive quantiles for DJIA log returns are shown in Figure 7. It is seen that the one-step ahead predictive quantiles obtained from model (11) are quite different from those obtained from model (9), especially
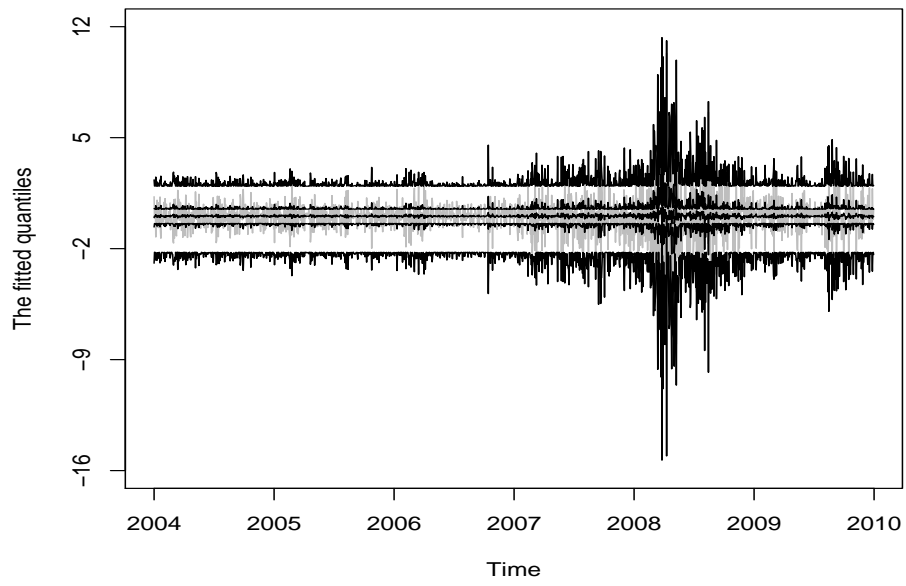
17

Figure 6: One-step ahead predictive quantiles for the DJIA log returns using our model.
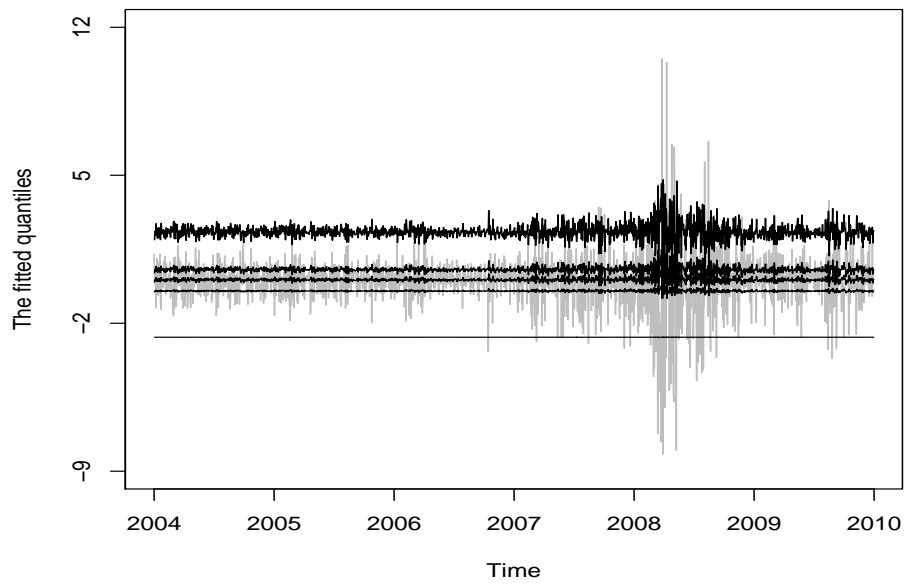

Figure 7: One-step ahead predictive quantiles for the DJIA log returns using the semi-parametric model (11).

for the extreme quantiles when $\tau = 0.025$ and 0.975. In particular, the 0.025th one-step ahead predictive quantiles obtained from model (11) is almost equal to a constant, which is not reasonable. Furthermore, the 0.975th and 0.75th one-step ahead predictive quantiles cross over at several time points between 2008 and 2009, leading to a failure in estimating the coverage of the fitted one-step ahead predicted quantiles for the DJIA log returns.

Forecasting for several periods ahead is particularly relevant in a risk management framework where risk managers can be interested in applying these methods not only for monitoring risk with regulatory purposes, but also for stress-testing exercises. In these cases it is very convenient to assume elliptical distributions for the error term of the model along with simple ARCH type parametric structures for the conditional volatility model, that allow one to forecast the conditional forecast a few periods into the future. Estimation in these cases is usually obtained using rolling windows of data. In contrast, our method provides a very convenient technique for obtaining reliable forecasts of the risk measures, VaR, several periods ahead. Figure 8 reports multi-step ahead out-of-sample predictive density functions (continuous curves) and predictive quantiles (i.e. predictive values at risk). The dashed vertical lines in Figure 8 show the 97.5% and 2.5% multi-step ahead predictive quantiles. Note that we only showed the forecasts up to 15-steps ahead, although any finite steps-ahead forecasts can be obtained. The continuous vertical lines are the actual observed log returns from 11 October 2010 to 29 October 2010. The other two dotted vertical lines in each plot give a 95% prediction interval obtained from model (10) by using the statistical software R.

Table 5 shows the point forecasts obtained from the AR-GARCH and quantile double AR models under study; the second column reports the actual observed log returns, columns three to five report the predicted mean returns and the 2.5% and 97.5% predicted quantiles from our quantile double AR model, while the columns six to eight give the predicted mean returns and associated 95% prediction intervals from the fitted AR-GARCH model using R.
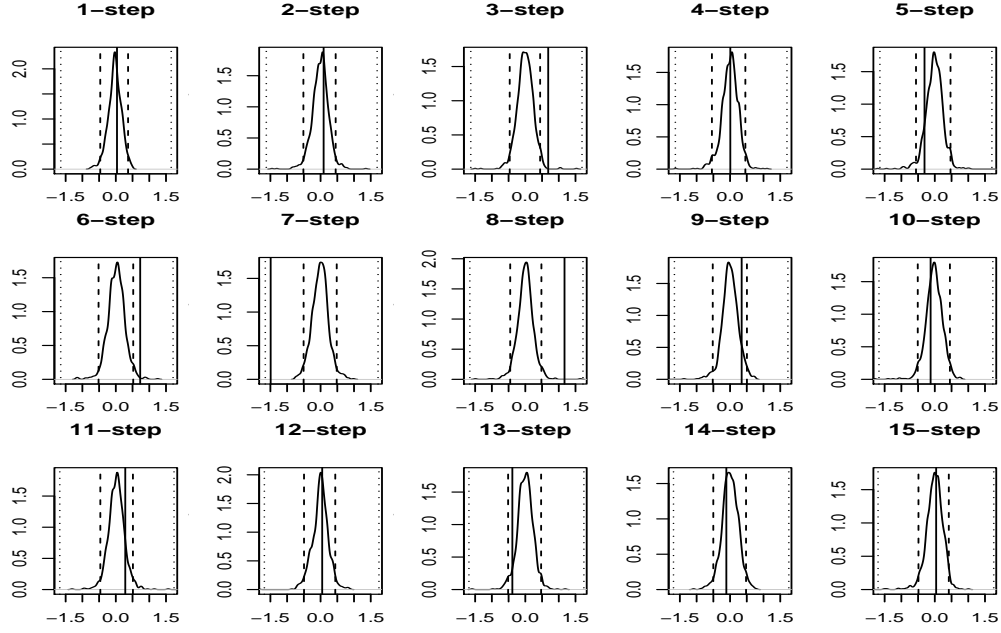
Figure 8: Up to 15-steps ahead out-of-sample forecasts for the DJIA log returns.

Steps	Observed	Predicted (quantile double AR)	2.5% quantile	97.5% quantile	Predicted (AR-GARCH)	Lower CI	Upper CI
1	0.035	-0.034	-0.465	0.366	0.005	-1.646	1.656
2	0.091	-0.008	-0.514	0.451	0.034	-1.622	1.691
3	0.684	-0.004	-0.466	0.443	0.033	-1.630	1.695
4	0.008	-0.007	-0.543	0.457	0.033	-1.635	1.701
5	-0.308	-0.013	-0.562	0.470	0.033	-1.641	1.707
6	0.729	0.007	-0.516	0.520	0.033	-1.647	1.712
7	-1.492	-0.014	-0.507	0.483	0.033	-1.653	1.718
8	1.171	-0.002	-0.451	0.481	0.033	-1.658	1.724
9	0.347	0.004	-0.517	0.504	0.033	-1.664	1.729
10	-0.126	-0.018	-0.510	0.452	0.033	-1.669	1.735
11	0.282	0.023	-0.463	0.511	0.033	-1.675	1.740
12	0.048	-0.010	-0.494	0.440	0.033	-1.680	1.746
13	-0.387	-0.001	-0.514	0.473	0.033	-1.686	1.751
14	-0.111	0.004	-0.501	0.478	0.033	-1.691	1.756
15	0.040	-0.002	-0.4898	0.419	0.033	-1.696	1.762
MSE		0.3346521			0.3357334		

Table 5: Out-of-sample point forecasts for the log returns of the Dow Jones Industrial Average.

It is seen that in most cases the actual observed log returns are well within the two predicted quantiles. The results show that, for example, the probability of the log return on 15 October 2010, conditional on the history up to 8 October 2010, to be less than -0.562 or greater than 0.470 is 0.05. Therefore, the predicted quantiles also provide an estimate of a future value at risk at a given level. It is worth mentioning that our point forecasts are usually not in the center of the interval formed by the two predicted quantiles, while the point forecasts from the AR-GARCH model are always in the middle of the estimated prediction interval. We also noticed that the 95% prediction intervals obtained from the fitted AR-GARCH model are much wider compared with those determined by the two predicted quantiles. The Mean Squared Error (MSE) between the observed and the predicted values can be found in the last row of Table 5, which shows that the MSE for our model is slightly smaller than that for the fitted AR-GARCH model. However, the differences between the two types of models are certainly worth further investigations in the future.

5   Comments and conclusions

In this paper we proposed a quantile double AR model that can estimated using MCMC Bayesian methods. Our methodology accommodates very easily a forecasting method for multi-step ahead prediction of the conditional quantile process. In this way, the proposed quantile double AR model allows us to study the whole conditional distribution of financial returns and to obtain the corresponding multi-step ahead conditional predictive distributions. The generalized lambda distribution is proposed to construct the quantile double AR model. We have found that this specific quantile double AR model is appropriate to deal with extreme quantiles, the crossing-over problem and the stylized fact of the non-Gaussianity of financial returns. It is worth mentioning that the developed methodologies can also be easily generalized to other quantile function models beyond the generalized

lambda distribution. In fact, we believe the optimal choice of a quantile double AR model is data dependent. Further investigations are required in the future.

The model can be further generalized by taking both the location and the scale of the quantile double AR model as a more general function of the past values of the response variable or by including additional regressors, which may lead to other useful models in practice.

We illustrated our methodologies by applying them to the DJIA log returns. However, extensive comparisons with other existing flexible methodologies for modeling the conditional distribution of returns have not been covered by this paper and certainly further investigations are required in the future.

References

[1] Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, 69, 542-547.

[2] Cai, Y. (2010a). Multivariate quantile function models. *Statistica Sinica*, 20, 481-496.

[3] Cai, Y. (2010b). Polynomial power-Pareto quantile function models. *Extremes*, 13, 291-314.

[4] Cai, Y. (2010c). Forecasting for quantile self exciting threshold autoregressive time series models. *Biometrika*. 97, 199 - 208.

[5] Cai, Y. (2009). Autoregression with non-Gaussian Innovations. *Journal of Time Series Econometrics*, Vol.1, Iss.2, Article 2. DOI: 10.2202/1941-1928.1016.

[6] Cai, Y. (2007). A quantile approach to US GNP. *Economic Modelling*. 24, 969-979.

[7] Cai, Y.i and Stander, J. (2008). Quantile self-exciting threshold autoregressive time series models. *Journal of Time Series Analysis*, 29, 186-202.

[8] Chipman, H. George, E.I. and McCulloch, R.E. (2001). The practical implementation of Bayesian model selection. *Model Selection: IMS Lecture Notes*, 38, 66-134.

[9] Engle, R.F. and Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression quantiles. *Journal of Business & Economic Statistics*, 22, 367-381.

[10] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50, 987-1007.

[11] Fournier, B., Rupin, N., Bigerelle, M., Najjar, D., Iost, A. and Wilcox, R. (2007). Estimating the parameters of a generalized lambda distribution. *Computational Statistics & Data Analysis*, 51, 2813-2835.

[12] Freimer, M., Mudholkar, G.S., Kollia, G., and Lin, C.T. (1988). A study of the generalised Tukey lambda distribution. Communications of Statistics, *Theory Math.*, 17, 3547-3567.

[13] Galvao, A. (2009). Unit Root Quantile Autoregression Testing Using Covariates. *Journal of Econometrics*, 152, 165-178.

[14] Galvao, A. (2011). Quantile Regression for Dynamic Panel Data with Fixed Effects. *Journal of Econometrics*, 164, 142-157.

[15] Gelfand, A.E. and Dey, D. (1994). Bayesian model choice: Asymptotics and exact calcula- tions. *Journal of the Royal Statistical Society B* 56, 501-514.

[16] Gilchrist, W.G. (2000). *Statistical Modelling with Quantile Functions*. Chapman & Hall/CRC.

[17] Giot, P. and Laurent, S. (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics*, 18, 641-663.

[18] Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

[19] Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46, (1) 33-50.

[20] Koenker, R. and Zhao, Q. (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory*, 12, 793-813.

[21] Lewis, S.M. and Raftery, A.E. (1997). Estimating Bayes' factors via posterior simulation with the Laplace-Metropolis estimator. *Journal of the American Statistical Association*, 92, 648-655.

[22] Ling, S. (2004). Estimation and testing stationarity for double-autoregressive models. *Journal of the Royal Statistical Society B*, 66, 63-78.

[23] Ling, S. (2007). A double AR(p) model: structure and estimation. *Statistica Sinica*, 17, 161-175.

[24] Newton, M.A. and Raftery, A.E. (1994). Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society B*, 56, 3-48.

[25] Taylor, J.W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, 6, 382-406.

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

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