JOURNAL OF APPLIED STATISTICS
2020, VOL. 47, NO. 1, 117–131
https://doi.org/10.1080/02664763.2019.1633285

Likelihood-based quantile autoregressive distributed lag models and its applications

Yuzhu Tian[a,b], Liyong Wang[c], Manlai Tang[d], Yanchao Zang[a] and Maozai Tian[e]

[a]School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, People's Republic of China; [b]School of Mathematics and Statistics, Northwest Normal University, LanZhou, People's Republic of China; [c]School of Statistics and Mathematics, The Central University of Finance and Economics, Beijing, People's Republic of China; [d]Department of Mathematics and Statistics, The Hang Seng University of Hong Kong, Hong Kong, Hong Kong; [e]School of Statistics, Renmin University of China, Beijing, People's Republic of China

ABSTRACT

Time lag effect exists widely in the course of economic operation. Some economic variables are affected not only by various factors in the current period but also by various factors in the past and even their own past values. As a class of dynamical models, autoregressive distributed lag (ARDL) models are frequently used to conduct dynamic regression analysis. In this paper, we are interested in the quantile regression (QR) modeling of the ARDL model in a dynamic framework. By combining the working likelihood of asymmetric Laplace distribution (ALD) with the expectation–maximization (EM) algorithm into the considered ARDL model, the iterative weighted least square estimators (IWLSE) are derived. Some Monte Carlo simulations are implemented to evaluate the performance of the proposed estimation method. A dataset of the consumption of electricity by residential customers is analyzed to illustrate the application.

ARTICLE HISTORY
Received 22 November 2018
Accepted 13 June 2019

KEYWORDS
Dynamic regression model; QR analysis; EM algorithm; information criterion; electricity consumption

1. Introduction

Time lag effect exists extensively in the running process of the economic system. Some economic indicators are affected not only by various factors in the current period but also by various factors in the past and even their own past values. Generally, the models with lagging variables are termed lagging variable model. Taking time factor into account in the lagging variable model can make statistical analysis possible to become dynamic analysis. Dynamic regression model is a kind of main model to establish a dynamic economic system. Ordinary time series models, distribution lag models and error autocorrelation models belong to dynamic regression models. One eminent feature of this kind of model is that the left side of the regression model depicts the current value of the interpreted variable, while the right side contains the historical value of the interpreted variable, the current value and historical value of the interpreting variables, and even the autocorrelation errors.

CONTACT Yuzhu Tian pole1999@163.com School of Mathematics and Statistics, Northwest Normal University, LanZhou, People's Republic of China
This article has been republished with minor changes. These changes do not impact the academic content of the article.

As a main class of dynamic regression models, linear regression model with correlated errors has widely emerged in statistical literatures. For example, Lin *et al.* [19] studied the regression model with time series errors, Lee and Lund [16] proposed a linear regression with stationary autocorrelated errors, Yang [32] studied linear regression model with serially correlated errors, Wu and Wang [31] studied shrinkage estimation for linear regression with ARMA errors, Yoon *et al.* [34] studied penalized regression model with autoregressive errors, Rosadi and Peiris [26] discussed second-order least-squares regression model with autocorrelated errors, Yoon *et al.* [33] studied penalized linear regression model with ARMA-GARCH errors.

Another kind of important dynamic regression model is the ARDL model. The ARDL model has widely been used in the fields of income consumption, investment output, macro and micro econometric model analysis. The introduction of lag variables can describe economic phenomena more comprehensively and accurately to improve the accuracy of modeling. The related theory of the ARDL model has been studied and applied to the dynamic economic system modeling. For example, Dufour and Kiviet [5] studied the exact inference method of the first-order ARDL model, Hassler and Wolters [10] discussed the ARDL model and cointegration, Lauridsen [15] discussed the estimation and application of spatial ARDL model, Buss [2] studied Bayesian ARDL model and application of economic downturn forecast.

The traditional statistical modeling method of these two kinds of dynamic regression models is to transform them into independent and identically distributed error models by difference or mathematical transformation, and then conduct conditional mean modeling using the least squares estimation (LSE) of mean regression or the maximum likelihood estimation (MLE) under the assumption of normal errors. However, the regression analysis method based on mean feature is not robust for non-normal errors and anomalous data. The analysis results are often biased and can only react the conditional mean information of the interpreted variable. In fact, the effects of explanatory variables on the response variable at different quantiles are not identical. As a natural alternative of mean regression, QR approach proposed by Koenker and Bassett [13] can build different regression relationships between the response variable and explanatory variables over the given quantiles. As a popular statistical analysis tool, the QR method has many advantages over the traditional mean regression, which makes its analysis results more robust and not affected by abnormal data and extreme data. This makes QR method widely used in financial insurance, income consumption, environmental science and clinical medicine, etc.

Some related studies on the QR analysis based on ARDL models and linear models with autocorrelation errors have emerged in recent literatures. For instance, Lim and Oh [18] studied variable selection in linear QR with autoregressive errors, Jiang and Li [11] discussed penalized weighted composite QR of the linear regression model with heavy-tailed autocorrelated errors. Li *et al.* [17] studied quantile ARDL model and applied it analyze the international tourism demand of Korea, Antonio *et al.* [1] discussed quantile ARDL model and applied it to the analysis of housing price regulation, Jin *et al.* [12] studied quantile cointegration theory of the ARDL model, Pal and Mitra [23] used the ARDL models to analyze the relationship between diesel and soybean prices.

From a Bayesian viewpoint, Yu and Moyeed [35] proposed the Bayesian QR approach based on the working likelihood of the asymmetric Laplace distribution (ALD) error, Geraci and Bottai [7] studied QR longitudinal data, Zhao and Lian [36] considered

Bayesian Tobit QR single index model, Tian *et al.* [28] investigated Bayesian quantile mixed effects model with censoring and errors in covariates. Based on the working likelihood of the ALD, Liu and Bottai [20] studied conditional quantile mixed effects model for longitudinal data using MCEM algorithm, Tian *et al.* [30] developed the EM algorithm for linear QR model, Tian *et al.* [29] studied the EM algorithm estimation for linear QR model with autoregressive errors.

In this paper, we focus on the QR analysis of ARDL models using the EM algorithm. Other sections of this article are organized as follows. In Section 2, we give the hierarchical working likelihood of QR ARDL models. In Section 3, we employ the EM algorithm to obtain the MLE of the considered model. In Section 4, simulation studies are conducted to illustrate the finite sample performance of the proposed method. In Section 5, a dataset of the electricity consumption is analyzed to illustrate the application. Section 6 summarizes the full text.

2. Preliminary description

2.1. The model

Consider the following ARDL model:

$$y_t = \mu + \sum_{i=0}^{d} x_{t-i}^T \theta_i + \sum_{j=1}^{q} \phi_j y_{t-j} + \varepsilon_t, \quad t = 1, \ldots, n, \tag{1}$$

where d and q are lag lengths of lagged response variable y_t and $p \times 1$ dimensional lagged covariate x_t, y_{t-j} is the j-th order lag of lagged response variable y_t, x_{t-i} is the i-th order lag of lagged covariate x_t, $\theta_i = (\theta_{i1}, \ldots, \theta_{ip})^T$ are regression coefficient vectors of lagged covariates x_{t-i} for $i = 0, \ldots, d$, ε_t is independent and identically distributed (i.i.d) error terms with finite first-order moment. The above model (1) is denoted as ARDL(d, q). In model (1), we assume that each variable x_t have the same lag truncation d for convenience. The case of different lag truncation for each variable is immediate.

Denote $Z_t = (1, x_t^T, \ldots, x_{t-d}^T)^T$, $Y_{t-1} = (y_{t-1}, \ldots, y_{t-q})^T$, $\boldsymbol{\beta} = (\mu, \theta_0^T, \ldots, \theta_d^T)^T$, $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_q)^T$, ARDL model (1) can be represented as

$$y_t = Z_t^T \boldsymbol{\beta} + Y_{t-1}^T \boldsymbol{\phi} + \varepsilon_t, \quad t = s+1, \ldots, n, \tag{2}$$

where $s = \max(d, q)$.

For model (1), the conditional τth quantile of response variable can be specified as

$$Q_\tau(y_t | \mathcal{F}_{t-1}) = Z_t^T \boldsymbol{\beta}_\tau + Y_{t-1}^T \boldsymbol{\phi}_\tau, \quad \tau \in (0, 1), \tag{3}$$

where $\mathcal{F}_{t-1}$ is a sigma algebra of information up to time $t-1$, the subscript τ in both $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\phi}_\tau$ can be omitted for simplicity.

2.2. The hierarchical working likelihood

Based on Koenker and Bassett [13], the τth quantile estimator of regression coefficient $\boldsymbol{\beta}$ and autoregression coefficient $\boldsymbol{\phi}$ can be obtained by minimizing the following objective loss function:

$$\arg\min_{\beta,\phi} \sum_{t=s+1}^{n} \rho_\tau \left(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi} \right), \tag{4}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ is the quantile check function. From Yu and Moyeed [35], minimizing the objective loss function of quantile regression is equivalent to maximizing the working likelihood of the asymmetric Laplace distribution (ALD) errors. And they argued that empirical results are robust by forcing the ALD on errors even if it is a misspecification of the true errors. Sriram *et al.* [27] studied posterior consistency of Bayesian QR based on the ALD specification. The probability density function (pdf) of ALD is

$$f(y|\mu,\sigma,\tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{ -\rho_\tau \left(\frac{y-\mu}{\sigma} \right) \right\}, \tag{5}$$

where μ is the location, σ is the scale, and $0 < \tau < 1$ is the skewness.

Using the working likelihood of ALD, the working likelihood of model (2) is

$$\mathrm{L}(\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma | x, y) = \prod_{t=s+1}^{n} \frac{\tau(1-\tau)}{\sigma} \exp\left\{ -\rho_\tau \left[\frac{y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi}}{\sigma} \right] \right\},$$

where $y = (y_1, \ldots, y_n)^T$, $x = \{x_1, \ldots, x_n\}$.

Additionally, based on the mixture representation of ALD developed by Reed and Yu [25] and Kozumi and Kobayashi [14], model (2) can be hierarchically represented as

$$y_t = Z_t^T \boldsymbol{\beta} + Y_{t-1}^T \boldsymbol{\phi} + \kappa_1 \upsilon_t + \sqrt{\kappa_2 \sigma \upsilon_t} \cdot e_t, \quad t = s+1, 2, \ldots, n, \tag{6}$$

where $\upsilon_t \sim \mathrm{Exp}(1/\sigma), e_t \sim N(0,1), \kappa_1 = 1 - 2\tau/\tau(1-\tau), \kappa_2 = 2/\tau(1-\tau)$, e_t and υ_t are independent of each other.

From (6), the joint hierarchical working likelihood of the complete data $\{y, x, \upsilon\}$ is

$$\mathrm{HL}(\boldsymbol{\beta}, \Phi, \sigma | y, x, v) = \prod_{t=s+1}^{n} \left[\frac{1}{\sqrt{\kappa_2 \sigma \upsilon_t}} \exp\left\{ -\frac{(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi} - \kappa_1 \upsilon_t)^2}{2\kappa_2 \sigma \upsilon_t} \right\} \cdot \frac{1}{\sigma} \exp\left\{ -\frac{1}{\sigma} \upsilon_t \right\} \right]. \tag{7}$$

3. The estimation methodology

3.1. EM algorithm procedure

The EM algorithm is an iterative method which has widely been used to find MLE of parameters in statistical models, where the model depends on incomplete data such as missing data or unobserved latent variables. EM algorithm has been broadly applied in statistical learning, mixture models, image processing, data mining and other applied fields. There are two key steps in EM algorithm including an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. One can refer to Dempster *et al.* [4], McLachlan and Krishnan [21] as well as Gupta and Chen [8] for some comprehensive summaries of EM algorithm. In the following contents, we employ the EM algorithm to conduct the QR estimation of the ARDL model.

From the joint hierarchical likelihood (7), up to a constant, we derive it working log-likelihood function as follows:

$$\begin{aligned}\log \mathrm{HL}(\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma | y, x, v) = & -\frac{3(n-s)}{2} \log \sigma - \frac{1}{2} \sum_{t=s+1}^{n} \log \upsilon_t \\ & - \sum_{t=s+1}^{n} \frac{1}{\sigma} \left[\frac{(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi})^2}{2\kappa_2} \upsilon_t^{-1} \right. \\ & \left. + \frac{\kappa_1^2 + 2\kappa_2}{2\kappa_2} \upsilon_t - \frac{\kappa_1 (y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi})}{\kappa_2} \right]. \end{aligned} \tag{8}$$

Set initial value of parameters set $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma\}$ to be $\hat{\Theta}^{(0)}$, the EM algorithm can be implemented as follows:

E-step: Assume the hth iteration value $\hat{\Theta}^{(h)}$ has been derived, the optimized objective loss function of the $(h+1)$th iteration is

$$\begin{aligned} & Q(\Theta | \Theta^{(h)}) \\ & \triangleq E[\log \mathrm{HL}(\boldsymbol{\beta}, \Phi, \sigma | y, x, v) | \hat{\Theta}^{(h)}] \\ & = -\frac{3(n-s)}{2} \log \sigma - \frac{1}{2} \sum_{t=s+1}^{n} E\left(\log \upsilon_t | \hat{\Theta}^{(h)} \right) \\ & \quad - \frac{1}{\sigma} \sum_{t=s+1}^{n} \left[\frac{\left(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi}\right)^2}{2\kappa_2} E(\upsilon_t^{-1} | \hat{\Theta}^{(h)}) \right. \\ & \quad \left. + \frac{\kappa_1^2 + 2\kappa_2}{2\kappa_2} E(\upsilon_t | \hat{\Theta}^{(h)}) - \frac{\kappa_1 \left(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi}\right)}{\kappa_2} \right]. \end{aligned} \tag{9}$$

For deriving parameter estimator $\hat{\Theta}^{(h+1)}$ of the $(h+1)$th iteration, it is essential for us to calculate the conditional expectation values for variables $\log(\upsilon_t)$, υ_t^{-1} and υ_t. It is noticed that the conditional pdf of υ_t van be easy to get as follows:

$$\begin{aligned} f(\upsilon_t | y, x) & \propto \frac{1}{\sqrt{\upsilon_t}} \exp \left\{ -\frac{1}{2} \left[\frac{\left(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi}\right)^2}{\kappa_2 \sigma} \upsilon_t^{-1} + \frac{\kappa_1^2 + 2\kappa_2}{\kappa_2 \sigma} \upsilon_t \right] \right\} \\ & \sim \mathrm{GIG} \left(\frac{1}{2}, \frac{\left(y_t - Z_t^T \boldsymbol{\beta} - Y_{t-1}^T \boldsymbol{\phi}\right)^2}{\kappa_2 \sigma}, \frac{\kappa_1^2 + 2\kappa_2}{\kappa_2 \sigma} \right), \quad t = s+1, \ldots, n, \end{aligned}$$

where $\mathrm{GIG}(\lambda, \chi, \psi)$ denotes the generalized inverse Gaussian (GIG) distribution. Hence, we have

$$E(\upsilon_t^\alpha) = \left(\frac{\chi}{\psi} \right)^{\alpha/2} \frac{K_{\lambda+\alpha}(\sqrt{\chi \psi})}{K_\lambda(\sqrt{\chi \psi})}, \quad \alpha \in R, E(\log \upsilon_t) = \frac{dE(\upsilon_t^\alpha)}{d\alpha}|_{\alpha=0}.$$

Hence, we obtain

$$\hat{\delta}_t^{(h)} \triangleq E(\upsilon_t^{-1}|\hat{\Theta}^{(h)}) = \frac{\sqrt{\kappa_1^2 + 2\kappa_2}}{\left|y_t - Z_t^T\hat{\boldsymbol{\beta}}^{(h)} - Y_{t-1}^T\hat{\boldsymbol{\phi}}^{(h)}\right|},$$

$$\hat{\gamma}_t^{(h)} \triangleq E(\log \upsilon_t|\hat{\Theta}^{(h)}) = \frac{dE(\upsilon_t^{\alpha}|\hat{\Theta}^{(h)})}{d\alpha}|_{\alpha=0},$$

$$\hat{\xi}_t^{(h)} \triangleq E(\upsilon_t|\hat{\Theta}^{(h)}) = \frac{\kappa_2\hat{\sigma}^{(h)}}{\kappa_1^2 + 2\kappa_2} + \frac{\left|y_t - Z_t^T\hat{\boldsymbol{\beta}}^{(h)} - Y_{t-1}^T\hat{\boldsymbol{\phi}}^{(h)}\right|}{\sqrt{\kappa_1^2 + 2\kappa_2}}.$$

M-step: Maximize the Q function of E-step with respect to $\Theta = \{\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma\}$ and obtain updated estimator $\hat{\Theta}^{(h+1)}$. For regression parameter $\boldsymbol{\beta}$, let

$$\frac{\partial Q}{\partial \boldsymbol{\beta}} = -\frac{1}{\sigma\kappa_2}\left[\sum_{t=s+1}^{n} Z_t\hat{\delta}_t^{(h)}Z_t^T\right]\beta + \frac{1}{\sigma\kappa_2}\sum_{t=s+1}^{n}\left[\hat{\delta}_t^{(h)}\left(y_t - Y_{t-1}^T\boldsymbol{\phi}\right)Z_t - \kappa_1 Z_t\right] = 0. \tag{10}$$

Represent the regularized equation (10) as the following linear equation group:

$$\left[\sum_{t=s+1}^{n} Z_t\hat{\delta}_t^{(h)}Z_t^T\right]\cdot\boldsymbol{\beta} = \sum_{t=s+1}^{n} Z_t\left[\hat{\delta}_t^{(h)}\left(y_t - Y_{t-1}^T\boldsymbol{\phi}\right) - \kappa_1\right]. \tag{11}$$

Denote

$$\begin{cases} Y = (y_{s+1} - Y_s^T\boldsymbol{\phi}, \ldots, y_n - Y_{n-1}^T\boldsymbol{\phi})^T,\ Z = (Z_{s+1}, \ldots, Z_n)^T, \\ \hat{W}^{(h)} = \begin{pmatrix} \hat{\delta}_{s+1}^{(h)} & & \\ & \ddots & \\ & & \hat{\delta}_n^{(h)} \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} \kappa_1 \\ \vdots \\ \kappa_1 \end{pmatrix}_{(n-s)\times 1}, \end{cases}$$

Equation (11) can be represented to be matrix equation as follows:

$$Z^T\hat{W}^{(h)}Z\cdot\boldsymbol{\beta} = Z^T\hat{W}^{(h)}Y - Z^T\boldsymbol{\eta}. \tag{12}$$

Furthermore, we replace the autoregressive parameter vector $\boldsymbol{\phi}$ in Y of Equation (3.1) as the hth iteration value $\hat{\boldsymbol{\phi}}^{(h)}$, and denote

$$\hat{Y} = \left(y_{s+1} - Y_s^T\hat{\boldsymbol{\phi}}^{(h)}, \ldots, y_n - Y_{n-1}^T\hat{\boldsymbol{\phi}}^{(h)}\right)^T.$$

Therefore, the $(h+1)$th iteration QR estimator $\hat{\boldsymbol{\beta}}^{(h+1)}$ is obtained as follows:

$$\hat{\boldsymbol{\beta}}^{(h+1)} = \left(Z^T\hat{W}^{(h)}Z\right)^{-1}Z^T\hat{W}^{(h)}\bar{Y}, \tag{13}$$

where $\bar{Y} = \hat{Y} - [\hat{W}^{(h)}]^{-1}\boldsymbol{\eta}$, $\hat{W}^{(h)}$ is a weight matrix which is associated with quantile τ. From (13), additionally, we find that the derived estimator $\hat{\boldsymbol{\beta}}^{(h+1)}$ has the form of the IWLSE.

For autoregressive parameter $\boldsymbol{\phi}$, from the Q function (9), we obtain its regularized equation group as follows:

$$\frac{\partial Q}{\partial \phi_k} = -\frac{1}{\sigma\kappa_2}\sum_{t=s+1}^{n}\left[\hat{\delta}_t^{(h)}\left(y_t - Z_t^T\boldsymbol{\beta} - \sum_{j=1}^{q}\phi_j y_{t-j}\right)\cdot(-y_{t-k}) + \kappa_1 y_{t-k}\right] = 0,$$
$$k = 1,\ldots,q. \tag{14}$$

Equation (14) can be represented as the following linear equation group:

$$\sum_{j=1}^{q}\phi_j \sum_{t=s+1}^{n}\hat{\delta}_t^{(h)} y_{t-j}y_{t-k} = \sum_{t=s+1}^{n} y_{t-k}\left[\hat{\delta}_t^{(h)}\left(y_t - Z_t^T\boldsymbol{\beta}\right) - \kappa_1\right], \quad k = 1,\ldots,q. \tag{15}$$

To obtain the updated estimator $\hat{\Phi}^{(h+1)}$, we use the Gauss–Seidel method [9]. Replacing regression parameter $\boldsymbol{\beta}$ in Equation (15) as the $(h+1)$-th updated value $\hat{\boldsymbol{\beta}}^{(h+1)}$ in Equation (13), and denote

$$\hat{e}_t = y_t - Z_t^T\hat{\beta}^{(h+1)}, \quad t = s+1,\ldots,n. \tag{16}$$

From Equation (15), we see $(h+1)$th iteration estimator $\hat{\Phi}^{(h+1)}$ is the solution of linear equation group $A\Phi = b$, where A and b are given respectively by

$$A = (a_{lk})_{q\times q}, \quad a_{lk} = \begin{cases} \sum_{t=s+1}^{n}\hat{\delta}_t^{(h)} y_{t-l}^2, & l = k, \\ \sum_{t=s+1}^{n}\hat{\delta}_t^{(h)} y_{t-l}y_{t-k}, & l \neq k, \end{cases}$$
$$b = \left(\sum_{t=s+1}^{n} y_{t-1}\left(\hat{\delta}_t^{(h)}\hat{e}_t - \kappa_1\right),\ldots,\sum_{t=s+1}^{n} y_{t-q}\left(\hat{\delta}_t^{(h)}\hat{e}_t - \kappa_1\right)\right)^T.$$

Furthermore, denote

$$E = \begin{pmatrix} y_s & y_{s-1} & \cdots & y_{s+1-q} \\ y_{s+1} & y_s & \cdots & y_{s+2-q} \\ \vdots & \vdots & & \vdots \\ y_{n-1} & y_{n-2} & \cdots & y_{n-q} \end{pmatrix}.$$

Matrix A can be represented as $A = E^T\hat{W}^{(h)}E$ and vector b is expressed as $b = E^T\hat{W}^{(h)}(\hat{\mathbf{e}} - (\hat{W}^{(h)})^{-1}\boldsymbol{\eta})$, where $\hat{\mathbf{e}} = (\hat{e}_{s+1},\ldots,\hat{e}_n)^T$.

Thus, the $(h+1)$th iteration value $\hat{\Phi}^{(h+1)}$ is

$$\hat{\Phi}^{(h+1)} = A^{-1}b = \left(E^T\hat{W}^{(h)}E\right)^{-1}\cdot\left[E^T\hat{W}^{(h)}\left(\hat{\mathbf{e}} - (\hat{W}^{(h)})^{-1}\boldsymbol{\eta}\right)\right]. \tag{17}$$

From (17), it can be seen that $\hat{\Phi}^{(h+1)}$ has the form of the IWLSE with the same weighted matrix $\hat{W}^{(h)}$ with $\hat{\boldsymbol{\beta}}^{(h+1)}$.

For σ, let

$$\frac{\partial Q}{\partial \sigma} = -\frac{3(n-s)}{2\sigma} + \frac{1}{\sigma^2}\sum_{t=s+1}^{n}\left[\frac{\left(y_t - Z_t^T\boldsymbol{\beta} - Y_{t-1}^T\boldsymbol{\phi}\right)^2}{2\kappa_2}\delta_t^{(h)} + \frac{\kappa_1^2 + 2\kappa_2}{2\kappa_2}\xi_t^{(h)} \right.$$
$$\left. - \frac{\kappa_1\left(y_t - Z_t^T\boldsymbol{\beta} - Y_{t-1}^T\boldsymbol{\phi}\right)}{\kappa_2}\right] = 0. \tag{18}$$

Replacing parameter $\boldsymbol{\beta}$ and Φ in Equation (17) as the $(h+1)$th iteration values $\hat{\boldsymbol{\beta}}^{(h+1)}$ and $\hat{\Phi}^{(h+1)}$, and denote

$$\hat{\omega}_t^{(h+1)} = y_t - Z_t^T\hat{\boldsymbol{\beta}}^{(h+1)} - Y_{t-1}^T\hat{\boldsymbol{\phi}}^{(h+1)}, \quad t = s+1,\ldots,n.$$

We obtain the $(h+1)$th iteration estimator of scale parameter σ as follows:

$$\hat{\sigma}^{(h+1)} = \frac{\sum_{t=s+1}^{n}\left[\hat{\delta}_t^{(h)}(\hat{\omega}_t^{(h+1)})^2 - 2\kappa_1\hat{\omega}_t^{(h+1)} + (\kappa_1^2 + 2\kappa_2)\cdot\xi_t^{(h)}\right]}{3(n-s)\kappa_2}. \tag{19}$$

From (13), (17), (21), we repeat E-step and M-step to update $\{\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma\}$ continually till the total error of the estimates attains the predetermined convergence constraint. In practice, in order to make the algorithm fully converge, we can start the EM iterations from several different initial points. Additionally, Oakes [22] provided a calculation formula to derive the confidence intervals of parameters for EM algorithm. However, it is essential to estimate complex variance–covariance matrix. As a kind of widespread random simulation algorithm, Bootstrap resampling method provide an efficient and simple alternative to construct confidence intervals of unknown parameters. A large number of empirical studies demonstrated that Bootstrap method generally outperform the direct confidence intervals based on asymptotic variance, especially for small sample cases. A detailed introduction about Bootstrap method can refer to Efron and Tibshirani [6] and Davison and Hinkley [3], etc.

3.2. Choices of orders

In subsection 3.1, we discussed the QR estimators of ARDL model (1) under prefixed lag order d and regressive order q. In practical application, both two orders are unknown and need to estimate. AIC (Akaike information criterion) or BIC (Bayesian information criterion) can be used to select appropriate lag orders. It is well-known that BIC criterion is consistent. We suggest to adopt BIC criteria in real applications. The BIC criterion based on quantile regression estimates is defined as

$$\mathrm{BIC}(d,q) = \log(\hat{\sigma}) + \frac{1 + q + (1+d)\times\dim(x)}{2n}\log(n), \tag{20}$$

where $\hat{\sigma} = 1/n - s\sum_{t=s+1}^{n}\rho_\tau(y_t - Z_t^T\hat{\boldsymbol{\beta}} - Y_{t-1}^T\hat{\boldsymbol{\phi}})$, $\dim(x)$ is the dimension of covariate x_t. The appropriate d and p can be selected by $\min_{d,q}\mathrm{BIC}(d,q)$.

4. Numerical simulations

In this section, some Monte Carlo simulation experiments are conducted to evaluate the finite sample performance of the proposed estimation procedure. We generate 1000 data sets from the following three ARDL(1, 1) models with sample size $n = 200$:

$$y_t = \mu + x_t^T\theta_0 + x_{t-1}^T\theta_1 + \phi y_{t-1} + \varepsilon_t. \tag{21}$$

Model 1: $\mu = 0.5, \theta_0 = 2, \theta_1 = 1, \phi = 0.5$.

Model 2: $\mu = 0.5, \theta_0 = (2, 1.5)^T, \theta_1 = (1, 0.5)^T, \phi = 0.5$.

Model 3: $\mu = 0.5, \theta_0 = (2, 1.5, 0, 0)^T, \theta_1 = (1, 0.5, 0, 0)^T, \phi = 0.5$.

For each model, we employ the following three schemes to generate the errors ε_t: standard normal distribution ($N(0, 1)$); standard Laplace distribution ($L(0, 1)$); mixture normal distribution ($0.5N(0, 1) + 0.5N(0, 9)$). Three quantiles 0.25, 0.5 and 0.75 are considered for each models. For all cases, we set $y_0 = 0$ to generate y_t for $t = 1, \ldots, n$ based

Table 1. Estimation results of model 1.

Error	τ	Parameter	$\mu = 0.5$	$\theta_0 = 2$	$\theta_1 = 1$	$\phi = 0.5$
N(0,1)	0.25	Bias	0.081	−0.002	0.062	−0.023
		RMSE	0.158	0.099	0.139	0.043
		95% C.L.	0.324	1.793	0.814	0.405
		95% C.U.	0.868	2.186	1.323	0.544
	0.50	Bias	0.002	0.001	0.001	0.002
		RMSE	0.085	0.089	0.106	0.032
		95% C.L.	0.336	1.836	0.787	0.436
		95% C.U.	0.676	2.185	1.210	0.567
	0.75	Bias	−0.011	−0.006	−0.019	0.005
		RMSE	0.089	0.102	0.122	0.034
		95% C.L.	0.313	1.789	0.753	0.437
		95% C.U.	0.653	2.196	1.203	0.572
L(0,1)	0.25	Bias	0.087	−0.025	0.027	−0.018
		RMSE	0.336	0.261	0.278	0.056
		95% C.L.	0.044	1.480	0.402	0.381
		95% C.U.	1.238	2.451	1.564	0.588
	0.50	Bias	−0.014	0.005	−0.025	0.004
		RMSE	0.147	0.165	0.181	0.040
		95% C.L.	0.207	1.682	0.615	0.420
		95% C.U.	0.779	2.342	1.307	0.579
	0.75	Bias	−0.013	0.006	−0.025	0.002
		RMSE	0.240	0.240	0.265	0.054
		95% C.L.	0.028	1.553	0.432	0.390
		95% C.U.	0.983	2.478	1.481	0.602
Mixture	0.25	Bias	0.098	−0.010	0.037	−0.021
		RMSE	0.251	0.209	0.221	0.050
		95% C.L.	0.150	1.605	0.593	0.387
		95% C.U.	1.047	2.348	1.462	0.558
	0.50	Bias	−0.003	0.000	−0.014	0.004
		RMSE	0.125	0.138	0.159	0.039
		95% C.L.	0.263	1.723	0.679	0.425
		95% C.U.	0.741	2.270	1.326	0.576
	0.75	Bias	−0.005	−0.005	−0.027	0.002
		RMSE	0.182	0.186	0.205	0.050
		95% C.L.	0.177	1.630	0.572	0.400
		95% C.U.	0.838	2.338	1.366	0.594

Table 2. Estimation results of model 2.

Error	τ	Parameter	$\mu = 0.5$	$\theta_{01} = 2$	$\theta_{02} = 1.5$	$\theta_{11} = 1$	$\theta_{12} = 0.5$	$\phi = 0.5$
N(0,1)	0.25	Bias	0.063	−0.001	−0.004	0.044	0.024	−0.017
		RMSE	0.143	0.102	0.101	0.127	0.113	0.035
		95% C.L.	0.318	1.803	1.298	0.817	0.314	0.423
		95% C.U.	0.830	2.193	1.690	1.264	0.748	0.542
	0.50	Bias	−0.002	0.000	0.001	−0.012	−0.010	0.004
		RMSE	0.085	0.090	0.088	0.111	0.095	0.028
		95% C.L.	0.324	1.820	1.328	0.779	0.303	0.449
		95% C.U.	0.661	2.169	1.673	1.198	0.672	0.560
	0.75	Bias	−0.013	−0.003	−0.001	−0.015	−0.017	0.008
		RMSE	0.089	0.098	0.105	0.114	0.107	0.031
		95% C.L.	0.316	1.806	1.302	0.761	0.279	0.450
		95% C.U.	0.665	2.182	1.694	1.195	0.690	0.565
L(0,1)	0.25	Bias	0.082	−0.023	−0.009	0.025	0.030	−0.018
		RMSE	0.318	0.278	0.267	0.264	0.254	0.050
		95% C.L.	0.013	1.492	0.964	0.510	0.019	0.388
		95% C.U.	1.191	2.488	1.991	1.550	1.032	0.572
	0.50	Bias	−0.011	−0.007	−0.013	−0.032	−0.029	0.009
		RMSE	0.150	0.178	0.173	0.184	0.175	0.040
		95% C.L.	0.200	1.643	1.155	0.605	0.108	0.430
		95% C.U.	0.785	2.339	1.821	1.303	0.792	0.585
	0.75	Bias	0.018	0.002	−0.009	−0.026	−0.023	0.007
		RMSE	0.253	0.253	0.262	0.267	0.249	0.050
		95% C.L.	0.041	1.487	0.986	0.449	0.010	0.403
		95% C.U.	1.058	2.483	2.008	1.487	0.953	0.607
Mixture	0.25	Bias	0.050	−0.005	−0.016	0.022	0.036	−0.015
		RMSE	0.244	0.214	0.205	0.214	0.214	0.045
		95% C.L.	0.095	1.648	1.098	0.615	0.109	0.398
		95% C.U.	1.010	2.358	1.849	1.434	0.952	0.566
	0.50	Bias	−0.008	0.003	0.000	−0.024	−0.017	0.009
		RMSE	0.130	0.138	0.137	0.154	0.150	0.034
		95% C.L.	0.224	1.734	1.222	0.688	0.184	0.442
		95% C.U.	0.752	2.269	1.774	1.278	0.768	0.570
	0.75	Bias	0.008	−0.010	−0.005	−0.029	−0.023	0.008
		RMSE	0.175	0.200	0.189	0.211	0.196	0.046
		95% C.L.	0.174	1.629	1.099	0.518	0.081	0.415
		95% C.U.	0.864	2.329	1.864	1.340	0.846	0.594

on model (21). For exogenous covariate x_t, we generate them using the standard normal distribution. We set the ordinary LSEs as the initial values of regression coefficient $\boldsymbol{\beta}$ and autoregressive parameter Φ for simplicity. Specifically, we first obtain the LSE of $\boldsymbol{\beta}$ by omitting the lag terms of model (21). Then, we derive the LSE of autoregressive parameter Φ based on the residual of the LSE of regression coefficient $\boldsymbol{\beta}$. The initial value of scale parameter σ is simply set as 1. The estimation biases, root mean square errors (RMSE) and 95% equal-tailed confidence intervals (CI) of parameters based on 1000 simulations are reported in Tables 1–3, where 95% C.L. and 95% C.U. are 95% confidence lower limit and confidence upper limit, respectively.

From Tables 1 and 2, it is evident that the EM algorithm procedure performs very well for the considered error distributions over three quantile values. The EM algorithm can achieve a fast convergence in simulations for both regression coefficients and autoregressive parameters. Generally, autoregressive parameters are estimated more accurately with smaller RMSEs than regressive parameters. Additionally, we find that the estimation

Table 3. Estimation results of model 3.

Error	τ	Parameter	$\mu = 0.5$	$\theta_{01} = 2$	$\theta_{02} = 1.5$	$\theta_{03} = 0$	$\theta_{04} = 0$	$\theta_{11} = 1$	$\theta_{12} = 0.5$	$\theta_{13} = 0$	$\theta_{14} = 0$	$\phi = 0.5$
N(0,1)	0.25	Bias	0.070	−0.007	−0.006	0.004	−0.000	0.049	0.039	0.001	−0.002	−0.023
		RMSE	0.150	0.112	0.101	0.102	0.100	0.133	0.124	0.107	0.104	0.040
		95% C.L.	0.323	1.775	1.285	−0.198	−0.203	0.810	0.313	−0.204	−0.203	0.415
		95% C.U.	0.832	2.195	1.678	0.205	0.193	1.287	0.762	0.201	0.207	0.541
	0.50	Bias	−0.005	−0.005	−0.004	−0.001	−0.001	−0.009	−0.009	−0.001	−0.000	0.005
		RMSE	0.092	0.102	0.093	0.089	0.094	0.108	0.102	0.095	0.093	0.029
		95% C.L.	0.318	1.803	1.314	−0.167	−0.185	0.776	0.282	−0.192	−0.191	0.446
		95% C.U.	0.676	2.174	1.687	0.175	0.183	1.198	0.689	0.170	0.177	0.560
	0.75	Bias	−0.018	−0.015	−0.010	0.000	0.002	−0.031	−0.017	0.002	−0.004	0.013
		RMSE	0.098	0.130	0.128	0.097	0.105	0.129	0.116	0.103	0.106	0.032
		95% C.L.	0.292	1.768	1.278	−0.194	−0.201	0.740	0.275	−0.192	−0.191	0.458
		95% C.U.	0.651	2.199	1.704	0.182	0.202	1.194	0.687	0.195	0.195	0.569
L(0,1)	0.25	Bias	0.061	−0.057	−0.044	−0.004	−0.005	0.020	0.021	0.004	−0.003	−0.021
		RMSE	0.321	0.282	0.272	0.242	0.247	0.276	0.257	0.249	0.255	0.050
		95% C.L.	−0.058	1.366	0.932	−0.472	−0.486	0.457	−0.017	−0.513	−0.491	0.390
		95% C.U.	1.225	2.451	1.953	0.469	0.486	1.533	1.001	0.497	0.502	0.566
	0.50	Bias	−0.019	0.001	−0.005	−0.009	−0.003	−0.039	−0.040	0.003	0.003	0.020
		RMSE	0.154	0.181	0.173	0.167	0.174	0.189	0.192	0.172	0.180	0.043
		95% C.L.	0.179	1.616	1.179	−0.351	−0.345	0.617	0.069	−0.347	−0.350	0.443
		95% C.U.	0.784	2.339	1.838	0.332	0.354	1.323	0.806	0.347	0.373	0.590
	0.75	Bias	0.040	−0.027	−0.033	−0.012	0.021	−0.053	−0.007	0.004	0.007	0.016
		RMSE	0.246	0.258	0.265	0.236	0.244	0.271	0.261	0.229	0.244	0.051
		95% C.L.	0.064	1.487	0.967	−0.488	−0.453	0.404	−0.057	−0.454	−0.475	0.422
		95% C.U.	1.024	2.462	1.966	0.422	0.501	1.465	0.996	0.459	0.495	0.611
Mixture	0.25	Bias	0.041	−0.039	−0.032	0.001	0.003	0.032	0.030	0.004	0.000	−0.021
		RMSE	0.246	0.218	0.204	0.193	0.199	0.207	0.205	0.201	0.197	0.045
		95% C.L.	0.081	1.537	1.098	−0.354	−0.383	0.615	0.130	−0.382	−0.395	0.401
		95% C.U.	1.017	2.351	1.851	0.385	0.395	1.441	0.921	0.404	0.379	0.554
	0.50	Bias	−0.021	0.000	−0.011	0.009	−0.004	−0.036	−0.033	0.001	−0.002	0.019
		RMSE	0.137	0.151	0.152	0.139	0.139	0.167	0.161	0.149	0.141	0.040
		95% C.L.	0.237	1.719	1.212	−0.259	−0.281	0.651	0.164	−0.284	−0.288	0.447
		95% C.U.	0.741	2.284	1.778	0.268	0.276	1.286	0.768	0.309	0.264	0.585
	0.75	Bias	0.030	−0.037	−0.020	0.000	0.016	−0.057	−0.043	0.002	−0.007	0.021
		RMSE	0.190	0.221	0.210	0.189	0.198	0.216	0.210	0.184	0.194	0.047
		95% C.L.	0.173	1.520	1.096	−0.375	−0.350	0.546	0.075	−0.364	−0.402	0.442
		95% C.U.	0.905	2.334	1.864	0.374	0.434	1.331	0.839	0.363	0.359	0.606

results under $N(0,1)$ error have smaller estimation RMSEs and shorter confidence intervals than $t3$ error and mixture normal error in most cases. From Table 3, it can be seen the EM algorithm procedure also performs well for sparse model 3. The estimation values of unknown parameters are close to the true values for both nonzero coefficients and zero-valued coefficients. For zero-valued parameters, the estimated 95% confidence intervals cover zeros for all cases. Those results show that the proposed estimation procedures have good estimation effects even for sparse models.

5. Real data example

The dataset can be found from Ramanathan [24] which presented the consumption of electricity served by San Diego Gas and Electric Company. 87 quarterly observations from February 1974 to April 1994 were recorded. This dataset has also been analyzed by Lim and Oh [18], Jiang and Li [11] as well as Tian *et al.*[29] based on the linear QR model with autoregressive errors. Lim and Oh [18] and Jiang and Li [11] considered the penalized variable selection, while Tian *et al.* [29] presented a EM algorithm estimation. In this section, we use the proposed QR ARDL model to fit this dataset. The response variable is the electricity consumption measured by the logarithm of the kwh sales per residential customer (LKWH, y). The explanatory covariates include the logarithm of per-capita income (LIncome, x_1), the logarithm of price of electricity (LPrice, x_2), cooling degree days (CCD, x_3) and heat degree days (HDD, x_4). We centralize the response variable, standardize the covariates (minus the sample mean and divided by the sample standard deviation), and then construct the first-order ARDL$(1,1)$ model as follows:

$$y_t = \mu + x_t^T\theta_0 + x_{t-1}^T\theta_1 + \phi y_{t-1} + \varepsilon_t, \quad t = 1,\dots,n, \tag{22}$$

where $x_t = (x_{1t}, x_{2t}, x_{3t}, x_{4t})^T, \theta_0 = (\theta_{01},\dots,\theta_{04})^T, \theta_1 = (\theta_{11},\dots,\theta_{14})^T$.

The responses y_t in model (22) has an increasing linear trend as time t increases. The mean and standard deviation are 0 and 0.099, the quantiles at 0.25th, 0.50th and 0.75th are -0.060, -0.004 and 0.080, respectively. In order to test the prediction efficiency of the proposed estimation method, we split 87 observations into two parts. The first 85 observations are used to fit the model, and the last two observations are used to make a prediction. We employ the proposed EM algorithm procedure to fit the first 85 observations over the 0.25th, 0.5th and 0.75th quantiles. In order to derive the estimated standard errors (SE) and 95% confidence intervals, 1000 residual bootstrap replications are employed. The initial values of parameters are set as the same to Section 4. The average estimation values (Est.), SE, 95% confidence lower limits and confidence upper limits are listed in Table 4.

From the estimation results of Table 4, we see four explanatory covariates LIncome, LPrice, CCD and HDD have the expected signs as Ramanathan [24], where LIncome, CCD and HDD have positive effects on LKWH while LPrice exhibits negative effect over three quantiles. In addition, LIncome has 1st-order positive lag effect on LKWH while LPrice, CCD and HDD have 1st-order negative lag effects. Furthermore, we find 95% Bootstrap confidence intervals of parameters θ_{01}, θ_{02} and θ_{11} cover zero over all three quantiles, which indicate that LIncome and LPrice have no significant effect and LIncome has no significant 1st-order lag effect on current LKWH over three given quantiles. 95% Bootstrap confidence intervals of parameters θ_{12} and θ_{13} cover zero at the 0.50th and 0.75th quantiles but do not cover zero at 0.25th quantile, which indicate that LPrice and CDD have significant

Table 4. Estimation results for the data of electricity consumption.

τ	Parameter	μ	θ_{01}	θ_{02}	θ_{03}	θ_{04}	θ_{11}	θ_{12}	θ_{13}	θ_{14}	ϕ
0.25	Est.	−0.009	0.074	−0.008	0.015	0.030	0.036	−0.039	−0.007	−0.018	0.136
	SE	0.003	0.074	0.020	0.003	0.003	0.072	0.020	0.003	0.003	0.013
	95% C.L.	−0.019	−0.010	−0.025	0.009	0.025	−0.221	−0.080	−0.013	−0.025	0.111
	95% C.U.	−0.005	0.335	0.047	0.022	0.036	0.120	−0.009	−0.001	−0.012	0.163
0.50	Est.	0.001	0.073	−0.015	0.017	0.031	0.033	−0.019	−0.006	−0.021	0.156
	SE	0.004	0.089	0.020	0.003	0.003	0.087	0.020	0.004	0.004	0.015
	95% C.L.	−0.009	−0.057	−0.057	0.010	0.025	−0.208	−0.063	−0.013	−0.028	0.126
	95% C.U.	0.007	0.316	0.028	0.024	0.038	0.158	0.024	0.002	−0.013	0.186
0.75	Est.	0.013	0.073	−0.015	0.010	0.025	0.033	−0.027	−0.007	−0.026	0.174
	SE	0.004	0.103	0.020	0.004	0.004	0.101	0.021	0.004	0.004	0.019
	95% C.L.	0.001	−0.113	−0.052	0.002	0.018	−0.265	−0.070	−0.015	−0.034	0.132
	95% C.U.	0.020	0.375	0.028	0.017	0.033	0.217	0.012	0.001	−0.018	0.211

Table 5. Prediction values of the data of electricity consumption.

Prediction model	τ	The 86th response True value: 0.144	The 87th response True value: 0.140
QR ARDL model	0.25	0.121	0.138
	0.50	0.126	0.148
	0.75	0.136	0.164
QR model with AR errors	0.25	0.098	0.131
	0.50	0.136	0.126
	0.75	0.139	0.145

1st-order lag effects on current LKWH only at low quantile. 95% Bootstrap confidence intervals of parameters θ_{03}, θ_{04}, θ_{14} and ϕ don't cover zero at three given quantiles, which show CCD, HDD, 1st-order lagged HDD and LKWH have significant effects on current LKWH. In addition, the estimation values of ϕ have an increasing trend as quantiles range from 0.25 to 0.75, which indicates 1st-order lagged LKWH has a bigger impact at higher quantile. From the above analysis results, we know that CCD, HDD and 1st-order lagged HDD and 1st-order lagged LKWH are main influence factors on current LKWH.

Furthermore, we conduct prediction studies for the 86th and 87th responses of model (22) over three quantiles using the proposed method. The prediction values of the 86th response and the 87th response over the 0.25th, 0.50th and 0.75th quantiles are presented in Table 5. We see that the proposed method performs well. To compare the prediction efficiency with QR model with autoregressive errors, we also provide model prediction values using the proposed method by Tian *et al.* [29] for the 86th and the 87th responses over three quantiles in Table 5. From Table 5, we see both two prediction models have good prediction effects, but the QR ARDL model has better prediction results which are closer to true values for most of the cases.

6. Conclusion

We study the EM algorithm QR estimation of the ARLD models. Using the proposed estimation procedure, the IWLSE forms of the QR ARLD models are derived. Some simulations and a real data example are implemented to illustrate the proposed procedures. High-dimensional regression modeling is a hot issue in recent statistical literatures. In future work, we will incorporate the penalized likelihood method into the QR ARDL models to conduct variable selection or sparse estimation in the high-dimensional framework.

Acknowledgments

The authors thank reviewers for their constructive comments and valuable suggestions which have greatly improved the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work is partly supported by China Postdoctoral Science Foundation (No. 2017M610156), National Natural Science Foundation of China (Nos. 11501167, 71473280), Young academic leaders project of Henan University of Science and Technology (No. 13490008) and Major project of the National Social Science Foundation of China (No. 15ZDA009).

References

[1] J.R. Antonio, G. Montes-Rojas and S.Y Park, *Quantile autoregressive distributed lag model with an application to house price returns*, Oxf. Bull. Econ. Stat. 75 (2013), pp. 307–321.
[2] G. Buss, *Forecasting Economy with Bayesian Autoregressive Distributed Lag Model: Choosing Optimal Prior in Economic Downturn*, Mpra Paper, 2009, pp. 100–105.
[3] A.C. Davison and D.V Hinkley, *Bootstrap Methods and their Application*, Cambridge University Press, Cambridge, 1997.
[4] A.P. Dempster, N.M. Laird and D.B Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, J. R. Stat. Soc. B 39 (1977), pp. 1–38.
[5] J.M. Dufour and J.F Kiviet, *Exact inference methods for first-order autoregressive distributed lag models*, Econometrica 66 (1998), pp. 79–104.
[6] B. Efron and R.J Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.
[7] M. Geraci and M. Bottai, *Quantile regression for longitudinal data using the asymmetric Laplace distribution*, Biostatistics 8 (2007), pp. 140–154.
[8] M.R. Gupta and Y Chen, *Theory and use of the EM algorithm*, Found Trends Signal Process. 4 (2010), pp. 223–296.
[9] L. Hageman and D Young, *Applied Iterative Methods*, Academic Press, New York, 1981.
[10] U. Hassler and J Wolters, *Autoregressive distributed lag models and cointegration*, AStA – Adv. Stat. Anal. 90 (2006), pp. 59–74.
[11] Y. Jiang and H Li, *Penalized weighted composite quantile regression in the linear regression model with heavy-tailed autocorrelated errors*, J. Korean Stat. Soc. 43 (2014), pp. 531–543.
[12] S.C. Jin, T.H. Kim and Y Shin, *Quantile cointegration in the autoregressive distributed-lag modeling framework*, J. Econom. 188 (2015), pp. 281–300.
[13] R. Koenker and J Bassett, *Regression quantiles*, Econometrica 46 (1978), pp. 33–50.
[14] H. Kozumi and G Kobayashi, *Gibbs sampling methods for Bayesian quantile regression*, J. Stat. Comput. Simul. 81 (2011), pp. 1565–1578.
[15] J Lauridsen, *Spatial autoregressively distributed lag models: Equivalent forms, estimation, and an illustrative commuting model*, Ann. Reg. Sci. 40 (2006), pp. 297–311.
[16] J. Lee and R Lund, *Revisiting simple linear regression with autocorrelated errors*, Biometrika 91 (2004), pp. 240–245.
[17] H.Q. Li, P. Sungyong and J.H Seo, *Quantile elasticity of international tourism demand for South Korea using the quantile autoregressive distributed lag model*, Tourism Econ. 17 (2011), pp. 997–1015.

[18] Y. Lim and H Oh, *Variable selection in quantile regression when the models have autoregression errors*, J. Korean Stat. Soc. 43 (2014), pp. 513–530.
[19] T. Lin, M. Pourahmadi, and A. Schick, *Regression models with time series errors*, J. Time Ser. Anal. 20 (1999), pp. 425–433.
[20] Y. Liu and M Bottai, *Mixed-effects models for conditional quantiles with longitudinal data*, Int. J. Biostat.2009. DOI: 10.2202/1557-4679.1186.
[21] G.J. McLachlan and T Krishnan, *The EM Algorithm and Extensions*, 2nd ed. Wiley, Hoboken, 2008.
[22] D Oakes, *Direct calculation of the information matrix via the EM algorithm*, J. R. Stat. Soc. B 61 (1999), pp. 479–482.
[23] D. Pal and S.K Mitra, *Diesel and soybean price relationship in the USA: Evidence from a quantile autoregressive distributed lag model*, Emp. Econom. 52 (2017), pp. 1609–1626.
[24] R Ramanathan, *Introductory Econometrics with Applications*, Harcourt Brace Jovanovich, Orlando, 1998.
[25] C. Reed and K Yu, Efficient Gibbs sampling for Bayesian quantile regression. Technical report, Brunel University, UK, 2009.
[26] D. Rosadi and S Peiris, *Second-order least-squares estimation for regression models with autocorrelated errors*, Comput. Stat. 29 (2015), pp. 931–943.
[27] K. Sriram, R.V. Ramamoorthi and P Ghosh, *Posterior consistency of Bayesian quantile regression based on the misspecified asymmetric Laplace density*, Bayesian Anal. 8 (2013), pp. 479–504.
[28] Y.Z. Tian, E.Q. LI and M.Z Tian, *Bayesian joint quantile regression for mixed effects models with censoring and errors in covariates*, Comput. Stat. 31 (2016), pp. 1–27.
[29] Y.Z. Tian, M.L. Tang, Y.C. Zang and M.Z Tian, *Quantile regression for linear models with autoregressive errors using EM algorithm*, Comput. Stat. 33 (2018), pp. 1605–1625.
[30] Y.Z. Tian, M.Z. Tian and Q.Q Zhu, *Linear quantile regression based on EM algorithm*, Commun. Stat. Theor. Methods 43 (2014), pp. 3464–3484.
[31] R. Wu and Q Wang, *Shrinkage estimation for linear regression with ARMA errors*, J. Stat. Plann. Infer. 142 (2012), pp. 2136–2148.
[32] C.Y Yang, *Estimation of linear regression models with serially correlated errors*, J. Data Sci. 10 (2012), pp. 723–755.
[33] Y.J. Yoon, S. Lee and T Lee, *Adaptive LASSO for linear regression models with ARMA-GARCH errors*, Commun. Stat. Simul. C 46 (2017), pp. 3479–3490.
[34] Y.-J. Yoon, C. Park and T Lee, *Penalized regression models with autoregressive error terms*, J. Stat. Comput. Simul. 83 (2013), pp. 1756–1772.
[35] K. Yu and R.A. Moyeed, *Bayesian quantile regression*, Stat. Probab. Lett. 54 (2001), pp. 437–447.
[36] K.F. Zhao and H Lian, *Bayesian Tobit quantile regression with single-index models*, J. Stat. Comput. Simul. 85 (2015), pp. 1247–1263.