Journal of the Korean Data & Information Science Society
2015, **26**(1), 209–216

http://dx.doi.org/10.7465/jkdi.2015.26.1.209
한국데이터정보과학회지

Partially linear support vector orthogonal quantile regression with measurement errors[†]

Changha Hwang[1]

[1]Department of Statistics, Dankook University

Received 24 October 2014, revised 11 November 2014, accepted 21 November 2014

Abstract

Quantile regression models with covariate measurement errors have received a great deal of attention in both the theoretical and the applied statistical literature. A lot of effort has been devoted to develop effective estimation methods for such quantile regression models. In this paper we propose the partially linear support vector orthogonal quantile regression model in the presence of covariate measurement errors. We also provide a generalized approximate cross-validation method for choosing the hyperparameters and the ratios of the error variances which affect the performance of the proposed model. The proposed model is evaluated through simulations.

Keywords: Errors-in-variables, generalized approximate cross-validation, measurement error, orthogonal regression, partially linear, quantile regression, support vector quantile regression.

1. Introduction

Standard regression models assume that covariates have been measured exactly without error. Therefore, these models account only for errors in the dependent variables or responses. In contrast, measurement error models or errors-in-variables (EIV) models are regression models that account for measurement errors in the covariates. The covariates of interest are often not observable and instead are measured with errors. Disregarding these measurement errors often leads to bias in estimating the mean and quantile functions (Fuller, 1987; Carroll *et al.*, 2006; Wei and Carroll, 2009; Ma and Yin, 2011; Wang *et al.*, 2012). This is the motivation for investigating measurement error models, which are frequently encountered by researchers conducting empirical studies in the social and natural sciences.

The measurement error problem has a long history and a vast literature. In the last three decades, the linear measurement error model for mean regression has frequently been used in practice and has attracted considerable attention in the statistical literature. Recently there have been quite a lot of works on nonlinear and nonparametric measurement error models for mean regression. Carroll *et al.* (2006) and Chen *et al.* (2011) summarize the literature on nonlinear models. During the last two decades, partially linear measurement errors models

[†] The present research was conducted by the research fund of Dankook University in 2014.
[1] Professor, Department of Applied Statistics, Dankook University, Gyeonggido 448-701, Korea. E-mail: chwang@dankook.ac.kr

have received much attention in the literature as a generalization of the linear measurement errors model. However, less attention has been paid to quantile regression (QR) than to mean regression with covariate measurement errors because of two main difficulties for correcting the bias in QR caused by measurement error (Wang *et al.*, 2012). One is that a parametric regression-error likelihood is usually not specified in QR. The other is that the quantile of the sum of two random variables is not necessarily the sum of the two marginal quantiles.

There have been some works on measurement error in QR (He and Liang, 2000; Chesher, 2001; Schennach, 2008; Wei and Carroll, 2009; Wang *et al.*, 2012; Ma and Yin, 2011). In this paper we propose the partially linear support vector orthogonal QR (PLSVOQR) model with covariate measurement errors by applying quantile loss function of orthogonal residuals to the formulation of semiparametric support vector QR (SSVQR) of Shim *et al.* (2012). This PLSVOQR model utilizes iterative reweighted least squares (IRWLS) procedure. Unlike He and Liang (2000), this method avoids the assumption that the random errors in the response and the measurement errors in the covariates follow the same symmetric distribution. This is the first paper that utilizes the idea of support vector machine (SVM) of Vapnik (1995) when some of the covariates are linear and have measurement errors. For other applications of SVM see Cho *et al.* (2010), Hwang and Shim (2012) and Shim and Hwang (2013, 2014).

The rest of this paper is organized as follows. Section 2 outlines a partially linear orthogonal QR (PLOQR) with measurement errors in some covariates that affect linearly the response. Section 3 illustrates the IRWLS-based PLSVOQR with measurement errors in covariates in parametric component. Section 4 and Section 5 present simulation study and conclusion, respectively.

2. PLOQR with measurement errors

For the purpose of deriving PLSVOQR, we now consider estimating QR function of the partially linear measurement error model

$$\begin{cases} y_i = b + \boldsymbol{\beta}^t \boldsymbol{x}_i^* + f(\boldsymbol{z}_i) + \epsilon_i, \ i = 1, \cdots, n, \\ \boldsymbol{x}_i = \boldsymbol{x}_i^* + \boldsymbol{u}_i, \end{cases} \tag{2.1}$$

where b is the intercept term, $\boldsymbol{x}_i^*$ and $\boldsymbol{u}_i$ are $d \times 1$ vectors of unobservable random covariates and measurement errors, respectively, $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_d)^t$ is a $d \times 1$ vector of unknown parameters, $\boldsymbol{z}_i$ is an $m \times 1$ vector of non-random covariates, $f(\cdot)$ is an unknown regression function, y_i is a scalar response, the random errors ϵ_i's are independent and identically distributed with zero mean and finite variance σ_ϵ^2, ϵ_i and $\boldsymbol{u}_i$ are uncorrelated, and $\boldsymbol{x}_i^*$ and $(\epsilon_i, \boldsymbol{u}_i^t)^t$ are independent. The measurement errors $\boldsymbol{u}_i$'s are independent and identically distributed with with mean zero and covariance matrix $\Sigma_{\boldsymbol{u}}$. In order to identify model (2.1), we usually assume $\Sigma_{\boldsymbol{x}^*} = Cov(\boldsymbol{x}_i^*)$ is a positive definite matrix and $\Sigma_{\boldsymbol{u}}/\sigma_\epsilon^2$ is a known $d \times d$ positive definite matrix. Another way to identify model (2.1) is to assume that $\Sigma_{\boldsymbol{u}}$ is a known $d \times d$ positive definite matrix or estimator of $\Sigma_{\boldsymbol{u}}$ is available.

We now outline the basic principle of PLOQR with measurement errors, which is based on Ma and Yin (2011). In the ordinary mean regression, we minimize the vertical standardized distance $D\left\{\left(y - b - \boldsymbol{\beta}^t \boldsymbol{x}^* - f(\boldsymbol{z})\right) \,/\, SD\left(y - b - \boldsymbol{\beta}^t \boldsymbol{x}^* - f(\boldsymbol{z})\right)\right\}$, where D stands for a desirable distance measure and SD stands for the standard deviation, because only the vertical y direction has errors. However, in the measurement error situation, errors also occur along

the horizontal $\boldsymbol{x}^*$ direction. Thus, we should consider a distance containing both vertical and horizontal components. In fact, the minimization of the same standardized distance with $\boldsymbol{x}^*$ replaced by $\boldsymbol{x}$ automatically corrects for this. Thus, we have

$$\frac{\big(y - b - \boldsymbol{\beta}^t\boldsymbol{x} - f(\boldsymbol{z})\big)}{SD\left(y - b - \boldsymbol{\beta}^t\boldsymbol{x} - f(\boldsymbol{z})\right)} = \frac{\big(y - b - \boldsymbol{\beta}^t\boldsymbol{x} - f(\boldsymbol{z})\big)}{\sqrt{\sigma_\epsilon^2 + \boldsymbol{\beta}^t\Sigma_{\boldsymbol{u}}\boldsymbol{\beta}}}, \tag{2.2}$$

which is proportional to $\big(y - b - \boldsymbol{\beta}^t\boldsymbol{x} - f(\boldsymbol{z})\big)/\sqrt{1 + \boldsymbol{\beta}^t\boldsymbol{\beta}}$ under the assumption that ϵ and $\boldsymbol{u}$ have a joint distribution that is spherically symmetric. This correction was also used in He and Liang (2000) for the L_1 distance.

By the way, the recent literature became aware of the inadequacy of assumption that ϵ and $\boldsymbol{u}$ have a joint distribution that is spherically symmetric, and proposes the use of general orthogonal regression method that accounts for different uncertainties of the two types of errors. For general orthogonal regression, it is usually assumed that $\Sigma_{\boldsymbol{u}}$ is a diagonal matrix, i.e., $\Sigma_{\boldsymbol{u}} = \mathrm{diag}\{\sigma_{u_1}^2, \cdots, \sigma_{u_d}^2\}$, and $\sigma_{u_i}^2$'s are different from σ_ϵ^2. Then, the distance (2.2) is proportional to $\big(y - b - \boldsymbol{\beta}^t\boldsymbol{x} - f(\boldsymbol{z})\big)\,/\,\sqrt{1 + \sum_{i=1}^d \lambda_i\beta_i^2}$, where λ_i's are the error variances' ratios defined as $\lambda_i = \sigma_{u_i}^2/\sigma_\epsilon^2$. Thus, the PLOQR problem can be expressed as

$$\min \ \sum_{i=1}^n \rho_\tau\left(\frac{y_i - b - \boldsymbol{\beta}^t\boldsymbol{x}_i - f(\boldsymbol{z}_i)}{\sqrt{1 + \sum_{i=1}^d \lambda_i\beta_i^2}}\right), \tag{2.3}$$

where $\rho_\tau(\cdot)$ is the quantile loss function called as check function and is defined as

$$\rho_\tau(r) = \tau r I(r \geq 0) + (\tau - 1) r I(r < 0), \tag{2.4}$$

where $I(\cdot)$ is the indicator function. The PLOQR will be an acceptable method as long as λ_i's are specified correctly.

3. PLSVOQR with measurement errors

In this section we derive the PLSVOQR with measurement errors in covariates by applying the basic principle of SVM to the PLOQR problem (2.3). According to the basic principle of SVM, we can consider the following representation of $f(\boldsymbol{z})$ in feature space

$$f(\boldsymbol{z}) = \boldsymbol{w}^t\boldsymbol{\phi}(\boldsymbol{z}), \tag{3.1}$$

where the nonlinear function $\boldsymbol{\phi}(\cdot) : R^m \to R^{m_h}$ maps the input space to a so-called higher dimensional feature space. If we use the kernel trick of the SVM, then the function $f(\boldsymbol{z})$ becomes

$$f(\boldsymbol{z}) = \sum_{i=1}^n \alpha_i K(\boldsymbol{z}, \boldsymbol{z}_i), \tag{3.2}$$

where α_i's are the dual parameters, and $K(\cdot,\cdot) = \boldsymbol{\phi}(\cdot)^t\boldsymbol{\phi}(\cdot)$ is a kernel function obtained from the Mercer's condition (Mercer, 1909). Several choices of the kernel function are possible. One popular choice of kernel function in practice is Gaussian kernel

$$K(\boldsymbol{z}_i, \boldsymbol{z}_j) = \exp\left(-\|\boldsymbol{z}_i - \boldsymbol{z}_j\|^2/2\kappa\right), \ i, j = 1, \cdots, n,$$

where $\kappa > 0$ is prespecified kernel parameter. In this paper we use this Gaussian kernel.

Incorporating the function (3.2) into the PLOQR problem (2.3) and utilizing the regularization technique of SVM, given a training data set we can define the following PLSVOQR optimization problem to get optimal $(b, \boldsymbol{\beta}, \boldsymbol{\alpha})$:

$$\min \ \frac{1}{2}\boldsymbol{\alpha}^t \boldsymbol{K}\boldsymbol{\alpha} + C\sum_{i=1}^{n} \rho_\tau \left(\frac{y_i - b - \boldsymbol{\beta}^t \boldsymbol{x}_i - \boldsymbol{K}_i\boldsymbol{\alpha}}{\sqrt{1+\sum_{i=1}^{d}\lambda_i\beta_i^2}} \right), \tag{3.3}$$

where $\boldsymbol{\alpha} = (\alpha_1 \cdots, \alpha_n)^t$ and $\boldsymbol{K}_i$ is the ith row of the $n \times n$ kernel matrix $\boldsymbol{K}$ with elements $K_{ij} = K(\boldsymbol{z}_i, \boldsymbol{z}_j)$. Solving the above PLSVOQR problem (3.3), we can obtain the τ-th QR function estimate of covariate vectors $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ as follows:

$$\hat{Q}(\tau|\boldsymbol{x}_i, \boldsymbol{z}_i) = b_\tau + \boldsymbol{\beta}_\tau^t \boldsymbol{x}_i + \boldsymbol{K}_i\boldsymbol{\alpha}_\tau, \tag{3.4}$$

where $b_\tau, \boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ are the estimates of $b, \boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ at quantile level τ, respectively.

In this section we propose an IRWLS procedure to estimate PLSVOQR with measurement errors, in which b_τ, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ are estimated simultaneously. This procedure makes it possible to derive the generalized approximate cross-validation (GACV) function for selecting λ_i's and hyperparameters of PLSVOQR. The IRWLS procedure basically uses differentiation of the check function $\rho_\tau(\cdot)$. To overcome the nondifferentiability of $\rho_\tau(\cdot)$ at 0, we employ the modified check function $\rho_{\tau,\delta}(\cdot)$ (Nychka *et al.*, 1995) which differs from $\rho_\tau(\cdot)$ only in the small interval $(-\delta, \delta)$ such as,

$$\rho_{\tau,\delta}(r) = \tau \ \frac{r^2}{\delta} \ I(r \geq 0) + (1-\tau) \ \frac{r^2}{\delta} \ I(r < 0). \tag{3.5}$$

By setting δ small enough, we can get a good approximate solution to the optimal problem (3.3). Now (3.3) becomes the one of obtaining b_τ, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ to minimize

$$L(b, \boldsymbol{\beta}, \boldsymbol{\alpha}) = \frac{1}{2}\boldsymbol{\alpha}^t \boldsymbol{K}\boldsymbol{\alpha} + aC\sum_{i=1}^{n} \rho_{\tau,\delta}(y_i - b - \boldsymbol{\beta}^t \boldsymbol{x}_i - \boldsymbol{K}_i\boldsymbol{\alpha}) \text{ for } \tau \in (0,1), \tag{3.6}$$

where $a = \frac{1}{\sqrt{1+\sum_{i=1}^{d}\lambda_i\beta_i^2}}$.

Taking partial derivatives of (3.6) with regard to b, $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, respectively, leads to the optimal values of b_τ, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ to be the solution to

$$\begin{pmatrix} \mathbf{1}^t\boldsymbol{W}\mathbf{1} & \mathbf{1}^t\boldsymbol{W}\boldsymbol{X} & \mathbf{1}^t\boldsymbol{W}\boldsymbol{K} \\ \boldsymbol{X}^t\boldsymbol{W}\mathbf{1} & \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{X} & \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{K} \\ \boldsymbol{W}\mathbf{1} & \boldsymbol{W}\boldsymbol{X} & \boldsymbol{W}\boldsymbol{K} + \boldsymbol{I}/(aC) \end{pmatrix} \begin{pmatrix} b_\tau \\ \boldsymbol{\beta}_\tau \\ \boldsymbol{\alpha}_\tau \end{pmatrix} = \begin{pmatrix} \mathbf{1}^t\boldsymbol{W}\boldsymbol{y} \\ \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{y} \\ \boldsymbol{W}\boldsymbol{y} \end{pmatrix}, \tag{3.7}$$

where $\boldsymbol{X}$ is an $n \times d$ matrix consisting of $\boldsymbol{x}_i^t$ for $i = 1, 2, \cdots, n$, and $\boldsymbol{W}$ is a diagonal matrix with the ith diagonal element w_{ii} obtained from the derivative of the modified check function as

$$w_{ii} = \begin{cases} 2\tau/\delta, & 0 < r_i < \delta \\ \tau/r_i, & r_i > \delta \\ 2(1-\tau)/\delta, & -\delta < r_i < 0 \\ (\tau-1)/r_i, & r_i < -\delta \end{cases}, \tag{3.8}$$

where $r_i = y_i - b_\tau - \boldsymbol{\beta}_\tau^t \boldsymbol{x}_i - \boldsymbol{K}_i \boldsymbol{\alpha}_\tau$. We know that the solution to the equation (3.7) cannot be obtained in a single step since a contains λ_i's and $\boldsymbol{\beta}_\tau$, and $\boldsymbol{W}$ contains b_τ, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ therein.

We now illustrate the model selection method which chooses the optimal values of λ_i's and hyperparameters of the PLSVOQR using IRWLS procedure. The functional structure of the PLSVOQR is characterized by the error variances' ratios λ_i's, the regularization parameter C and the kernel parameter. To choose the hyperparameters and λ_i's of PLSVOQR we use GACV function (Yuan, 2006),

$$GACV(\boldsymbol{\lambda}) = \frac{\sum_{i=1}^n \rho_{\tau,\delta}(y_i - \hat{Q}(\tau|\boldsymbol{x}_i, \boldsymbol{z}_i))}{n - \text{trace}(\boldsymbol{H})}, \tag{3.9}$$

where $\boldsymbol{\lambda}$ is the set of hyperparameters and λ_i's, and $\boldsymbol{H}$ is the hat matrix such that

$$\left(\hat{Q}(\tau|\boldsymbol{x}_1, \boldsymbol{z}_1), \cdots, \hat{Q}(\tau|\boldsymbol{x}_n, \boldsymbol{z}_n)\right)^t = \boldsymbol{H}\boldsymbol{y} \tag{3.10}$$

with the ijth element $h_{ij} = \partial \hat{Q}(\tau|\boldsymbol{x}_i, \boldsymbol{z}_i)/\partial y_j$. The hat matrix $\boldsymbol{H}$ can be obtained from (3.4) and (3.7) as follows:

$$\boldsymbol{H} = (\boldsymbol{1}, \boldsymbol{X}, \boldsymbol{K}) \begin{pmatrix} \boldsymbol{1}^t\boldsymbol{W}\boldsymbol{1} & \boldsymbol{1}^t\boldsymbol{W}\boldsymbol{X} & \boldsymbol{1}^t\boldsymbol{W}\boldsymbol{K} \\ \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{1} & \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{X} & \boldsymbol{X}^t\boldsymbol{W}\boldsymbol{K} \\ \boldsymbol{W}\boldsymbol{1} & \boldsymbol{W}\boldsymbol{X} & \boldsymbol{W}\boldsymbol{K} + \boldsymbol{I}/(aC) \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{1}^t\boldsymbol{W} \\ \boldsymbol{X}^t\boldsymbol{W} \\ \boldsymbol{W} \end{pmatrix}. \tag{3.11}$$

Thus we need to apply IRWLS procedure which starts with initialized values of b_τ, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\alpha}_\tau$ for given λ_i's and hyperparameters at quantile level τ as follows:

1. Set the initial values $b_\tau^{(0)}$, $\boldsymbol{\beta}_\tau^{(0)}$, $\boldsymbol{\alpha}_\tau^{(0)}$ and $a_\tau^{(0)} = \sqrt{1 + \sum_{i=1}^d \lambda_i \beta_{i\tau}^{(0)2}}$.
2. (a) Given $b_\tau^{(k)}$, $\boldsymbol{\beta}_\tau^{(k)}$, $\boldsymbol{\alpha}^{(k)}$ and $a_\tau^{(k)}$, calculate $\boldsymbol{W}^{(k)}$ from (3.8) using δ, τ and $r_t = y_t - b_\tau^{(k)} - \boldsymbol{\beta}_\tau^{(k)'} \boldsymbol{z}_t - \boldsymbol{K}_t \boldsymbol{\alpha}_\tau^{(k)}$.
   (b) Calculate $b_\tau^{(k+1)}$, $\boldsymbol{\beta}_\tau^{(k+1)}$ and $\boldsymbol{\alpha}^{(k+1)}$ from

$$\begin{pmatrix} b_\tau^{(k+1)} \\ \boldsymbol{\beta}_\tau^{(k+1)} \\ \boldsymbol{\alpha}_\tau^{(k+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{1}'\boldsymbol{W}^{(k)}\boldsymbol{1} & \boldsymbol{1}'\boldsymbol{W}^{(k)}\boldsymbol{X} & \boldsymbol{1}'\boldsymbol{W}^{(k)}\boldsymbol{K} \\ \boldsymbol{X}'\boldsymbol{W}^{(k)}\boldsymbol{1} & \boldsymbol{X}'\boldsymbol{W}^{(k)}\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{W}^{(k)}\boldsymbol{K} \\ \boldsymbol{W}^{(k)}\boldsymbol{1} & \boldsymbol{W}^{(k)}\boldsymbol{X} & \boldsymbol{W}^{(k)}\boldsymbol{K} + \boldsymbol{I}/(aC) \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{1}'\boldsymbol{W}^{(k)}\boldsymbol{y} \\ \boldsymbol{X}'\boldsymbol{W}^{(k)}\boldsymbol{y} \\ \boldsymbol{W}^{(k)}\boldsymbol{y} \end{pmatrix}.$$

   (c) Update $a_\tau^{(k+1)} = \sqrt{1 + \sum_{i=1}^d \lambda_i \beta_{i\tau}^{(k+1)2}}$ with and $\boldsymbol{\beta}_\tau^{(k+1)}$.
3. Iterate Step 2 until convergence.
4. Compute $\hat{Q}(\tau|\boldsymbol{x}_i, \boldsymbol{z}_t) = b_\tau + \boldsymbol{\beta}_\tau^t \boldsymbol{x}_i + \boldsymbol{K}_i \boldsymbol{\alpha}_\tau$ at quantile level τ.

The algorithm is iterated until the following stop criterion is satisfied:

$$\frac{1}{1+d+n} \|(b_\tau^{(k)}, \boldsymbol{\beta}_\tau^{(k)t}, \boldsymbol{\alpha}_\tau^{(k)t})^t - (b_\tau^{(k+1)}, \boldsymbol{\beta}_\tau^{(k+1)t}, \boldsymbol{\alpha}_\tau^{(k+1)t})^t\|^2 < \epsilon,$$

where $1 + d + n$ is the number of parametres and $\epsilon = 10^{-6}$ is taken as the tolerance level.

4. Simulation study

In this section we perform simulation study to investigate the effects of measurement errors and to demonstrate the finite sample behavior of the PLSVOQR under different error distributions and quantile levels. We are concerned with the PLSVOQR in which $\boldsymbol{x}^*$ is 2-dimensional vector and $\boldsymbol{z}$ is scalar or 2-dimensional vector. We choose the same variance for the errors of the covariates, $\sigma_{u_1}^2 = \sigma_{u_2}^2 = \sigma_u^2$, in order to simplify slightly the reporting of results. Thus we have $\lambda_1 = \lambda_2 = \lambda$. For λ we only consider $\lambda = 0.2, 0.5, 1.0, 1.25$ or 1.5. For quantile levels we will focus $\tau = 0.1, 0.5$ and 0.9. We compare the proposed PLSVOQR with SSVQR of Shim *et al.* (2012). This competing model does not consider measurement errors in covariates.

We generate 100 data sets of size 50 from each of the following two partially linear measurement error models:

- **Model 1**: $y_i = 1 + 0.5x_{1i}^* + x_{2i}^* + 2\sin(2\pi z_i) + \epsilon_i$, $\boldsymbol{x}_i = \boldsymbol{x}_i^* + \boldsymbol{u}_i$, $x_{1i}^*, x_{2i}^* \sim$ *i.i.d.* $N(2,1)$, $z_i \sim$ *i.i.d.* $U(0,1)$
- **Model 2**: $y = -1 - 0.4x_{1i}^* + 0.6x_{2i}^* + z_{1i}\exp(-z_{2i}^2) + \epsilon_i$, $\boldsymbol{x}_i = \boldsymbol{x}_i^* + \boldsymbol{u}_i$, $x_{1i}^*, x_{2i}^* \sim$ *i.i.d.* $U(2,4)$, $z_{1i}, z_{2i} \sim$ *i.i.d.* $U(0,1)$

Here we assume ϵ and $\boldsymbol{u}$ are independent of $\boldsymbol{x}^*$. For distribution of ϵ, we consider normal distribution $N(0, \sigma_\epsilon^2)$, t-distribution $\frac{\sigma_\epsilon}{\sqrt{2}} t_4$, where t_4 stands for the t-distribution with 4 degree of freedom, and Laplace distribution $L(0, \sigma_\epsilon^2)$ with mean 0 and variance σ_ϵ^2, where $\sigma_\epsilon^2 = 0.25$ for Model 1 and $\sigma_\epsilon^2 = 0.1$ for Model 2. For distributions of two components of $\boldsymbol{u}$, we also consider normal distribution $N(0, \sigma_u^2)$, t-distribution $\frac{\sigma_u}{\sqrt{2}} t_4$ and Laplace distribution $L(0, \sigma_u^2)$ with mean 0 and variance σ_u^2, where σ_u^2 is determined such that $\lambda = \sigma_u^2/\sigma_\epsilon^2$ for given λ. Although there are 9 combinations of distributions for two errors σ_u^2 and σ_ϵ^2, we here consider only three combinations $\big(N(0,\sigma_u^2), N(0,\sigma_\epsilon^2)\big)$, $\left(\frac{\sigma_u}{\sqrt{2}} t_4, \frac{\sigma_\epsilon}{\sqrt{2}} t_4\right)$ and $\big(L(0,\sigma_u^2), L(0,\sigma_\epsilon^2)\big)$ because the results are very similar.

For two partially linear measurement error models the τth QR functions of y are given as follows:

$$\begin{aligned} Q(\tau|\boldsymbol{x}^*, z) &= 1 + 0.5x_1^* + x_2^* + 2\sin(2\pi z) + \sigma_\epsilon F_\epsilon^{-1}(\tau), \\ Q(\tau|\boldsymbol{x}^*, \boldsymbol{z}) &= -1 - 0.4x_1^* + 0.6x_2^* + z_1\exp(-z_2^2) + \sigma_\epsilon F_\epsilon^{-1}(\tau), \end{aligned}$$

where $F_\epsilon^{-1}(\cdot)$ is the quantile function of ϵ. For 100 data sets of size 50, we compare the proposed PLSVOQR with SSVQR. For comparison we calculate the mean and standard deviation of 100 mean square errors (MSEs) for each estimated QR function as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\Big(Q(\tau|\boldsymbol{x}_i^*, z_i) - \hat{Q}(\tau|\boldsymbol{x}_i, z_i)\Big)^2. \tag{4.1}$$

Tables 4.1 and 4.2 show the results for the mean and standard deviation of 100 MSEs for each estimated QR function. Standard deviations are given in parenthesis. Boldfaced values indicate best performance/result in the particular categories of τ, λ and error distributions. We use a grid-search approach with GACV function (3.9) to find the optimal values of three parameters λ, C, κ of the PLSVOQR. As seen from Tables 4.1 and 4.2, the proposed PLSVOQR yields the smaller means of MSEs and smaller standard deviations of MSEs for all cases. Therefore, the PLSVOQR performs better than SVQR in estimating QR function

when the input variables are contaminated with noise. In fact, we have found that the PLSVOQR obtains the estimate of λ that is pretty close to the true value of λ, although we do not report here.

Table 4.1 Comparison of MSEs for 100 $\hat{Q}_\tau$'s for Model 1

λ	Distribution (u, ϵ)	Method	$\hat{Q}_\tau$ 0.1	0.5	0.9
0.5	(N, N)	PLSVOQR	**0.2593** (0.0099)	**0.1854** (0.0049)	**0.2515** (0.0080)
		SSVQR	0.8878 (0.0636)	0.7360 (0.0615)	0.9017 (0.0616)
	(t_4, t_4)	PLSVOQR	**0.3305** (0.0203)	**0.1649** (0.0072)	**0.2958** (0.0178)
		SSVQR	0.8878 (0.0636)	0.7360 (0.0615)	0.9017 (0.0616)
	(L, L)	PLSVOQR	**0.3003** (0.0142)	**0.1685** (0.0045)	**0.3167** (0.0150)
		SSVQR	1.0845 (0.0977)	0.6668 (0.0500)	1.2660 (0.1355)
1.0	(N, N)	PLSVOQR	**0.4065** (0.0149)	**0.2961** (0.0076)	**0.4149** (0.0145)
		SSVQR	1.1070 (0.0876)	0.9561 (0.0653)	1.0779 (0.0627)
	(t_4, t_4)	PLSVOQR	**0.4890** (0.0260)	**0.2728** (0.0111)	**0.4587** (0.0241)
		SSVQR	1.2737 (0.0093)	0.7748 (0.0050)	1.4591 (0.0094)
	(L, L)	PLSVOQR	**0.4745** (0.0201)	**0.2803** (0.0080)	**0.4966** (0.0207)
		SSVQR	1.4298 (0.1027)	0.7605 (0.0553)	1.3194 (0.1070)
1.25	(N, N)	PLSVOQR	**0.4766** (0.0169)	**0.3396** (0.0087)	**0.4847** (0.0169)
		SSVQR	1.1844 (0.0881)	1.0014 (0.0715)	1.2462 (0.0845)
	(t_4, t_4)	PLSVOQR	**0.5573** (0.0281)	**0.3207** (0.0125)	**0.5465** (0.0278)
		SSVQR	1.3400 (0.1029)	0.7901 (0.0502)	1.4725 (0.1046)
	(L, L)	PLSVOQR	**0.5417** (0.0232)	**0.3235** (0.0092)	**0.5527** (0.0223)
		SSVQR	1.4849 (0.1023)	0.7871 (0.0512)	1.4526 (0.1303)

Table 4.2 Comparison of MSEs for 100 $\hat{Q}_\tau$'s for Model 2

λ	Distribution (u, ϵ)	Method	$\hat{Q}_\tau$ 0.1	0.5	0.9
0.2	(N, N)	PLSVOQR	**0.0901** (0.0028)	**0.0418** (0.0016)	**0.0944** (0.0030)
		SSVQR	0.1357 (0.0137)	0.2266 (0.0309)	0.1805 (0.0193)
	(t_4, t_4)	PLSVOQR	**0.1091** (0.0114)	**0.0287** (0.0013)	**0.0832** (0.0037)
		SSVQR	0.1752 (0.0241)	0.1611 (0.0297)	0.1630 (0.0184)
	(L, L)	PLSVOQR	**0.0962** (0.0042)	**0.0321** (0.0016)	**0.0996** (0.0052)
		SSVQR	0.1855 (0.0205)	0.1314 (0.0160)	0.1592 (0.0115)
1.0	(N, N)	PLSVOQR	**0.1163** (0.0029)	**0.0675** (0.0018)	**0.1241** (0.0035)
		SSVQR	0.1762 (0.0158)	0.2534 (0.0336)	0.2538 (0.0288)
	(t_4, t_4)	PLSVOQR	**0.1363** (0.0117)	**0.0558** (0.0015)	**0.1100** (0.0046)
		SSVQR	0.2346 (0.0265)	0.1899 (0.0225)	0.2388 (0.0215)
	(L, L)	PLSVOQR	**0.1224** (0.0047)	**0.0604** (0.0018)	**0.1260** (0.0054)
		SSVQR	0.2414 (0.0249)	0.1844 (0.0219)	0.2144 (0.0249)
1.5	(N, N)	PLSVOQR	**0.1279** (0.0031)	**0.0783** (0.0018)	**0.1339** (0.0037)
		SSVQR	0.2154 (0.0194)	0.2670 (0.0393)	0.2938 (0.0343)
	(t_4, t_4)	PLSVOQR	**0.1478** (0.0118)	**0.0677** (0.0018)	**0.1214** (0.0045)
		SSVQR	0.2467 (0.0269)	0.2311 (0.0293)	0.2454 (0.0211)
	(L, L)	PLSVOQR	**0.1333** (0.0050)	**0.0721** (0.0019)	**0.1391** (0.0058)
		SSVQR	0.2611 (0.0237)	0.1939 (0.0238)	0.2482 (0.0286)

5. Concluding remarks

In this paper, we dealt with estimating QR function of the partially linear measurement error model with the PLSVOQR. We found that the PLSVOQR provides good results in estimating QR function for the given examples. The proposed PLSVOQR also make the model selection easier and faster than a leave-one-out cross validation or k-fold cross validation technique by using GACV function. In general, the orthogonal regression analysis requires knowledge of λ_i's. As long as the error variances' ratios λ_i's are specified correctly, the orthogonal regression fitting method is an acceptable method. By the way, the model selection process of PLSVOQR makes it possible to obtain the estimates of λ_i's that are very close to the true values. Thus, the proposed PLSVOQR appears to be useful in estimating QR function of the partially linear measurement errors model.

References

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: A modern perspective*, Chapman & Hall/CRC, New York.

Chen, X., Hong, H. and Nekipelov, D. (2011). Nonlinear models of measurement errors. *Journal of Economic Literature*, **49**, 901-37.

Chesher, A. (2001). *Parameter approximations for quantile regressions with measurement error*, Working Paper CWP02/01, University College London, Dept. of Economics.

Cho, D. H., Shim, J. and Seok, K. H. (2010). Doubly penalized kernel method for heteroscedastic autoregressive data. *Journal of the Korean Data & Information Science Society*, **21**, 155-162.

Fuller, W. A. (1987). *Measurement error models*, Wiley, New York.

He, X. and Liang, H. (2000). Quantile regression estimates for a class of linear and partially linear errors-in-variables models. *Statistica Sinica*, **10**, 129-140.

Hwang, C. and Shim, J. (2012). Smoothing Kaplan-Meier estimate using monotone suppot vector regression. *Journal of the Korean Data & Information Science Society*, **23**, 1045-1054.

Ma, Y. and Yin, G. (2011). Censored quantile regression with covariate measurement errors. *Statistica Sinica*, **21**, 949-971.

Mercer, J. (1909). Function of positive and negative type and their connection with theory of integral equations. *Philosophical Transactions of Royal Society A*, 415-446.

Nychka, D., Gray, G., Haaland, P., Martin, D. and O'Connell, M. (1995). A nonparametric regression approach to syringe grading for quality improvement. *Journal of the American Statistical Association*, **90**, 1171-1178.

Schennach, S. M. (2008). Quantile regression with mismeasured covariates. *Econometric Theory*, **24**, 1010-1043.

Shim, J. and Hwang, C. (2013). Expected shortfall estimation using kernel machines. *Journal of the Korean Data & Information Science Society*, **24**, 625-636.

Shim, J. and Hwang, C. (2014). Estimating small area proportions with kernel logistic regressions models. *Journal of the Korean Data & Information Science Society*, **25**, 941-949.

Shim, J., Kim, Y., Lee, J. and Hwang, C. (2012). Estimating value at risk with semiparametric support vector quantile regression. *Computational Statistics*, **27**, 685-700.

Vapnik, V. N. (1995). *The nature of statistical learning theory*, Springer, New York.

Wang, H. J., Stefanski, L. A. and Zhu, Z. (2012). Corrected-loss estimation for quantile regression with covariate measurement errors. *Biometrika*, **99**, 405–21.

Wei, Y. and Carroll, R. J. (2009). Quantile regression with measurement error. *Journal of the American Statistical Association*, **104**, 1129-1143.

Yuan, M. (2006). GACV for quantile smoothing splines. *Computational Statistics and Data Analysis*, **50**, 813-829.