Statistica Sinica **33** (2023), 2257-2280
doi:https://doi.org/10.5705/ss.202021.0246

PARTIALLY FUNCTIONAL LINEAR QUANTILE REGRESSION WITH MEASUREMENT ERRORS

Mengli Zhang[1], Lan Xue[2], Carmen D. Tekwe[3], Yang Bai[1] and Annie Qu[4]

[1]*Shanghai University of Finance and Economics,* [2]*Oregon State University,*
[3]*Indiana University Bloomington and* [4]*University of California, Irvine*

Abstract: Ignoring measurement errors in conventional regression analyses can lead to biased estimation and inference results. Reducing such bias is challenging when the error-prone covariate is a functional curve. In this paper, we propose a new corrected loss function for a partially functional linear quantile model with function-valued measurement errors. We establish the asymptotic properties of both the functional coefficient and the parametric coefficient estimators. We also demonstrate the finite-sample performance of the proposed method using simulation studies, and illustrate its advantages by applying it to data from a children obesity study.

Key words and phrases: Corrected score, functional measurement error, functional principle component, physical activity, quantile regression, wearable devices.

1. Introduction

Wearable devices are increasingly used in health research to monitor health and wellbeing, and can be combined with lifestyle interventions to reduce obesity. Wearable devices provide continuous granulated measurements of physical activity (PA), with raw PA measured at a sub-second level. These measurements can be viewed as functions or curves, rather than as vectors. Although recent developments in functional data analysis (FDA) can be applied to such data, analytical challenges need to be solved in order to draw more accurate inferences in obesity prevention studies. First, the standard analytical methods employed in obesity studies apply regression approaches to model the mean BMI, which could potentially limit the detection of intervention effects among participants whose weights differ from the mean (Koenker (2005); Geraci and Bottai (2014)). Second, because the true patterns of PA behavior are not directly observable, wearable devices are used to monitor PA, thus introducing measurement errors including variability in predictions at various PA intensity levels and errors associated with

Corresponding author: Lan Xue, Department of Statistics, Oregon State University, Corvallis, OR 97331-4501, USA. E-mail: Lan.Xue@oregonstate.edu.

the prediction equations (Bassett (2012); Crouter, Churilla and Bassett (2006); Jacobi et al. (2007); Warolin et al. (2012); Rothney et al. (2008)). Failure to account for such errors when assessing the effects of PA measures on health outcomes often leads to bias and an underestimation of these effects (Tekwe et al. (2019)). Third, the data collected from wearable devices are typically not discrete vectors, but rather curvilinear functions of time, and require using the functional data approaches of Ramsay and Silverman (2005).

However, while there is extensive research on measurement errors in traditional multivariate data analyses (Carroll et al. (2006); Fuller (1987)), few studies have examined such errors in the context of FDA. When conducting an FDA, it is usually assumed that measurement errors are independent and identically distributed (i.i.d.). However, this assumption is not feasible, because the functional measurement errors tend to be correlated over time in practice. Therefore, there is a need to consider functional measurement errors with more complex error structures in FDAs. In this paper, we address these critical needs and analytical challenges. The proposed method provides a better understanding of how measurement errors affect the evaluation of wearable-device-based PA patterns in obesity studies.

Our research is motivated by a childhood obesity study in which stand-biased desks were introduced to schools as an intervention to increase school day PA. The original research goal was to assess the association between daily energy expenditure (DEE) and subsequent progression toward obesity. Because DEE is not directly observable, the investigators measured it using accelerometer armbands provided to the children. In this project, we develop statistical methods to model the DEE measurements obtained from these wearable devices as functional data that are prone to measurement errors, and reduce the effects of DEE measurement errors when predicting obesity. In addition, our methods are based on a quantile regression and focus on the entire distribution of the body mass index (BMI), rather than simply using the average BMI values.

Quantile regression (Koenker and Bassett (1978)) has emerged as an important statistical method that offers a systematic approach for examining the effects of covariates on the entire distribution of the response variable, in contrast to a traditional mean regression. Researchers have recently applied quantile regression to functional data. Kato (2012) studied estimation in a functional linear quantile model, and derived the rate of convergence for the functional coefficient estimator. Yao, Sue-Chee and Wang (2017) proposed a regularized partially linear functional quantile regression for the simultaneous estimation and selection of the important covariates. Both of the aforementioned works assume that the functional covariate is exactly observed at subject-specific sampling points.

When the covariates are not directly observable and are instead measured using error-prone proxies, calibrating the measurement error in a quantile regression framework is challenging, for two reasons. First, a parametric assumption of the regression error distribution is often unavailable in quantile regressions. Second, the quantile function, unlike the mean function, does not inherit the additive property (Wang, Stefanski and Zhu (2012)). As a result, only a few studies have examined measurement errors in quantile regression settings, including the works of Wang, Stefanski and Zhu (2012),Wei and Carroll (2009), He and Liang (2000), and Hu and Schennach (2008). However, the aforementioned methods are all restricted to cases in which the covariates are scalars of finite dimensions. In particular, He and Liang (2000) introduced a consistent estimation procedure based on orthogonal residuals under the assumption that the model error and the measurement error follow a common distribution that is spherically symmetric. Wei and Carroll (2009) introduced joint estimating equations that hold simultaneously for all quantile levels, although they require a general linear regression structure assumption for all conditional quantiles. In addition, their method requires estimating the conditional density of the responses, given the independent variables. Firpo, Galvao and Song (2017) proposed a semiparametric two-step estimator to improve the strong linearity assumption on all quantile levels and the iterative algorithm in Wei and Carroll (2009). However, this requires nonparametric estimations of the conditional density of the true variables, given a response and other error-free covariates. It is difficult to obtain such estimates when the dimensionality of the error-prone covariates is large. In contrast, Wang, Stefanski and Zhu (2012) constructed a corrected-loss estimation only at specific quantile levels, thus avoiding the strong assumption of a mutual symmetric error and the nonparametric estimation of the conditional densities.

The idea of a correction for loss/score functions in the context of covariate measurement errors was first introduced by Nakamura (1990), who constructed the corrected log-likelihood/score function as an unbiased estimator of the conditional expectation of the original log likelihood/score function, given measurement errors. Estimating equations were also derived for generalized linear models with normal measurement errors. However, the nondifferentiable loss function in a quantile regression makes it difficult to construct such equations. Wang, Stefanski and Zhu (2012) addressed this issue by using a smooth function to approximate the indicator function, deriving the corrected loss function under multivariate normal or Laplace distributions of the measurement errors.

In this study, we solve a more challenging problem in which the covariate of interest is functional and the associated unknown regression coefficient is nonparametric and of infinite dimension. In this scenario, it is difficult to develop algorithms and derive the asymptotic properties of the proposed methods. To the best of our knowledge, this study is the first to address function-valued measurement errors in functional quantile regression models. We propose a corrected loss approach, as in Wang, Stefanski and Zhu (2012), for a partially functional linear quantile regression when the functional covariate is contaminated with functional measurement errors. In particular, our method identifies the measurement error model by assuming a parametric form for the measurement error distribution, and constructs a corrected objective function using a class of smoothed quantile objective functions. The proposed method provides a consistent estimator of the functional regression coefficient, and asymptotically normal estimates of the parametric coefficients. In addition, our method does not require a specification of the distribution of the regression error.

The rest of the paper is organized as follows. Section 2 introduces the partially functional quantile model. The corrected loss and a practical implementation are provided in Section 3. Section 4 establishes the asymptotic property of our proposed estimators. Simulation studies in Section 5 demonstrate the finite-sample performance of the proposed method. Section 6 contains a real-data application from a children obesity study. Section 7 concludes the paper. The technical lemmas and proofs are included in the Supplementary Material.

2. Background and Notation

Suppose $(Y_i, X_i, \mathbf{Z}_i)_{i=1}^n$ are independent realizations from the distribution of $(Y, X, \mathbf{Z})$, where Y is a scalar random variable, and $X = \{X(t), t \in \mathcal{T}\}$ is a random function assumed to be square integrable on a bounded closed interval $\mathcal{T}$ in $\mathbb{R}$ and possibly contaminated with measurement errors. Without loss of generality, we assume $\mathcal{T} = [0, 1]$, and X is centered with $E[X(t)] = 0$, for $t \in \mathcal{T}$. Here, $\mathbf{Z}$ is a p-dimensional vector of error-free covariates, including the intercept term. For a given $\tau \in (0, 1)$, the τth conditional quantile function $Q_{Y|X,\mathbf{Z}}(\tau)$ is defined as $F^{-1}_{Y|X,\mathbf{Z}}(\tau)$, where $F_{Y|X,\mathbf{Z}}(y) = P(Y \leq y|X, \mathbf{Z})$ is the cumulative distribution function of Y conditional on X and $\mathbf{Z}$. For $i = 1, \ldots, n$, we assume

$$Q_{Y_i|X_i,\mathbf{Z}_i}(\tau) = \mathbf{Z}_i^T \boldsymbol{\theta}_0(\tau) + \int_0^1 \beta_0(t, \tau) X_i(t) dt, \tag{2.1}$$

where $\boldsymbol{\theta}_0(\tau)$, including a scalar-valued intercept term, represents the vector of coefficients associated with the error-free covariates for the τth quantile, and

$\beta_0(t,\tau) \in L_2[0,1]$ is a functional coefficient that quantifies the effect of the functional covariate. For each individual i, we assume the functional covariate X_i is not directly observable. Instead, it is approximated using a surrogate W_i as

$$W_i(t) = X_i(t) + U_i(t). \tag{2.2}$$

Here, W_i serves as an unbiased measure of X_i subject to a functional measurement error $U_i(t)$, where $E[U_i(t)] = 0$ for $t \in \mathcal{T}$ and U_i is independent of $X_i, \mathbf{Z}_i$, and Y_i. Studies on functional data often assume that the measurement errors $U_i(t)$ are i.i.d. with a common variance over time t. Here, we consider a functional measurement error, and allow $U_i(t)$ to have an unstructured covariance function $\Sigma_U(t,s)$. When $U_i(t)$ are i.i.d., its common variance can be estimated using $\{W_i(t)\}$ under the smoothness of the covariance function of $X_i(t)$, which is similar to the estimation of the nugget effect in spatial statistics. However, when $\Sigma_U(t,s)$ is unstructured, additional information is needed to identify the covariance structure of $U_i(t)$, and measurement errors need to be taken into account to ensure a consistent estimation and inference of the partially functional linear quantile regression model (2.1).

The model in (2.1) is a useful generalization of both the classical linear quantile regression model and the functional linear quantile regression model. For mean regression models, recent studies (Shin (2009); Lu, Du and Sun (2014)) have considered estimating partially functional linear regression models when the covariates are measured without a classical measurement error. Kong et al. (2016) developed a penalized estimation procedure for variable selection in a high-dimensional partially functional linear regression model. In terms of quantiles, Kato (2012) studied the estimation of the model when the covariates are measured without errors. When the functional covariates are measured with errors, the aforementioned methods are no longer applicable. Ignoring measurement errors can lead to biased estimations and misleading inferences. When the functional covariate is contaminated by errors, a consistent estimation of the model in (2.1) is technically challenging. Existing methods require a complete specification of the conditional distribution of the response given the true functional covariates, which is often impractical. In addition, most existing methods focus on a fixed number of scalar covariates measured with errors, whereas the model in (2.1) involves a functional covariate of infinite dimension, which poses both computational and theoretical difficulties.

3. The Proposed Method

Functional principal component analysis (FPCA) plays an important role in functional data analysis, because it provides an efficient mechanism to represent random functions as a linear combination of basis functions. Our method is based on the FPCA of the covariate process $X(t)$. Suppose that $\int_0^1 \mathrm{E}[X^2(t)]dt < \infty$. Denote the covariance kernel of $X(t)$ as $K_x(s,t) = \mathrm{Cov}(X(s), X(t))$. Then, the Hilbert-Schmidt theorem entails that $K_x(s,t)$ can be represented as $K_x(s,t) = \sum_{j=1}^{\infty} \kappa_j \phi_j(s)\phi_j(t)$, where $\kappa_1 \geq \kappa_2 \geq \cdots \geq 0$ are ordered eigenvalues, and $\{\phi_j\}_{j=1}^{\infty}$ is an orthonormal basis of $L_2[0,1]$. Thus, we have the following expansions in $L_2[0,1]$, $X_i(t) = \sum_{j=1}^{\infty} X_{ij}\phi_j(t)$, $U_i(t) = \sum_{j=1}^{\infty} U_{ij}\phi_j(t)$, $\beta_0(t,\tau) = \sum_{j=1}^{\infty} b_{0j}(\tau)\phi_j(t)$, where X_{ij}, U_{ij} and $b_{0j}(\tau)$ are defined as $X_{ij} = \int_0^1 X_i(t)\phi_j(t)dt$, $U_{ij} = \int_0^1 U_i(t)\phi_j(t)dt$, and $b_{0j}(\tau) = \int_0^1 \beta_0(t,\tau)\phi_j(t)dt$. Thus, model (2.1) can be represented by a linear quantile regression model with an infinite number of "regressors",

$$Q_{Y_i|X_i,\mathbf{Z}_i}(\tau) = \sum_{j=1}^{\infty} b_{0j}(\tau)X_{ij} + \mathbf{Z}_i^T \boldsymbol{\theta}_0(\tau). \tag{3.1}$$

If $X(t)$ is observed without errors, a truncated version of (3.1) is often considered, where $\sum_{j=1}^{\infty} b_{0j}(\tau)X_{ij}$ is truncated by $\sum_{j=1}^{m} b_{0j}(\tau)X_{ij}$ with a large integer m, and the unknown coefficients are estimated by

$$\left(\tilde{\mathbf{b}}(\tau), \tilde{\boldsymbol{\theta}}(\tau)\right) = \underset{(\mathbf{b},\boldsymbol{\theta})}{\mathrm{argmin}} \sum_{i=1}^{n} \rho_\tau \left(Y_i - \mathbf{X}_i^T \mathbf{b}(\tau) - \mathbf{Z}_i^T \boldsymbol{\theta}(\tau)\right), \tag{3.2}$$

where $\rho_\tau(\epsilon) = \{\tau - I(\epsilon \leq 0)\}\,\epsilon$ is the check loss at the τth quantile. Similar estimation procedures have been considered in quantile regressions involving functional covariates without classic measurement errors. Kato (2012) systematically investigated the asymptotic property of such an estimator in a functional linear quantile model. Yao, Sue-Chee and Wang (2017) considered a regularized procedure for simultaneous variable selection and estimation.

We consider a scenario in which the functional covariate is not fully observed, but is instead contaminated by a classic measurement error. This means that the measurement error is additive and independent of the responses and covariates in the regression model. Our goal is to construct consistent estimators of the regression coefficients in (2.1) in the presence of a classic measurement error. Applying similar basis expansions on both sides of the measurement error equation (2.2), we have, for $j = 1, \ldots, m$,

$$W_{ij} = X_{ij} + U_{ij}, \tag{3.3}$$

where $W_{ij} = \int_0^1 W_i(t)\phi_j(t)dt$.

When $X_i(t)$ is contaminated by a measurement error, naively replacing X_{ij} with W_{ij} in the objective function (3.2) can lead to biased and inconsistent estimators of the unknown parameters. Instead, we adopt the corrected loss approach proposed in Wang, Stefanski and Zhu (2012) for linear quantile regression models, and extend it to partially functional quantile regression models. Our estimation strategy is based on a corrected objective or loss function, which leads to an unbiased estimation of a smooth approximation of $\rho_\tau(Y_i - \mathbf{X}_i^T\mathbf{b}(\tau) - \mathbf{Z}_i^T\boldsymbol{\theta}(\tau))$. The corrected objective function uses a property of Gaussian random variables (Stefanski and Cook (1995)) that $\mathrm{E}\{\mathrm{E}[\phi(Z_1 + \mathrm{i}\sigma Z_2|Z_1)]\} = \phi(\mu)$, where $\mathrm{i} = \sqrt{-1}$, for two independent Gaussian random variables $Z_1 \sim N(\mu, \sigma^2)$ and $Z_2 \sim N(0,1)$, as long as $\phi(\cdot)$ is a sufficiently smooth function. Another key element in the construction of the corrected loss function is that we use an infinitely smooth function $\rho_h(\cdot)$ to approximate the check loss $\rho_\tau(\cdot)$ in quantile regression with $\rho_h(\epsilon) = \epsilon[\tau - 1/2 + \pi^{-1}\int_0^{\epsilon/h}\sin(t)/tdt]$. Here, h controls the goodness of the approximation with $\lim_{h\to 0}\rho_h(\epsilon) = \rho_\tau(\epsilon)$.

Motivated by Wang, Stefanski and Zhu (2012), when the measurement errors follow a normal distribution, we consider the following corrected quantile objective function:

$$\left(\hat{\mathbf{b}}(\tau), \hat{\boldsymbol{\theta}}(\tau)\right) = \underset{(\mathbf{b},\theta)}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_h^*(Y_i - \mathbf{W}_i^T\mathbf{b}(\tau) - \mathbf{Z}_i^T\boldsymbol{\theta}(\tau), \mathbf{b}^T\Sigma_u\mathbf{b}), \tag{3.4}$$

where $\rho_h^*(\epsilon, \sigma^2) = \pi^{-1}\int_0^{1/h}\left\{y^{-1}\epsilon\sin(y\epsilon) - \sigma^2\cos(y\epsilon)\right\}\exp\left(y^2\sigma^2/2\right)dy + \epsilon(\tau - 1/2)$ and $\Sigma_u = \mathrm{Cov}(\mathbf{U}_i)$. The function ρ_h^* is constructed so that it is an unbiased estimator of ρ_h. In particular, Wang, Stefanski and Zhu (2012) showed that $\mathrm{E}\left[\rho_h^*(Y - \mathbf{W}^T\mathbf{b}(\tau)\ -\mathbf{Z}^T\boldsymbol{\theta}(\tau), \mathbf{b}^T\Sigma_u\mathbf{b})|Y, X, \mathbf{Z}\right] = \rho_h\left(Y - \mathbf{X}^T\mathbf{b}(\tau) - \mathbf{Z}^T\boldsymbol{\theta}(\tau)\right) \approx \rho_\tau\left(Y - \mathbf{X}^T\mathbf{b}(\tau)\ -\mathbf{Z}^T\boldsymbol{\theta}(\tau)\right)$, where the parameter h controls the goodness of the approximation, with a better approximation for a smaller h. Thus, consistent estimators of the coefficients $\theta(\tau)$ and $\beta(t,\tau)$ can be obtained from (3.4), with $\hat{\beta}(t,\tau) = \sum_{j=1}^{m}\hat{b}_j(\tau)\hat{\phi}_j(t)$ and $\{\hat{\phi}_j(t)\}_{j=1}^{m}$ an orthonormal basis of the estimated $K_x(s,t)$, as in subsection 3.1.

3.1. Implementation

Note that $(\hat{\mathbf{b}}(\tau), \hat{\boldsymbol{\theta}}(\tau))$ in (3.4) is not feasible in practice, because it relies on $K_x(s,t)$, which is generally unavailable when the functional covariate X is not fully observed. Therefore, additional data are needed to estimate and identify $K_x(s,t)$. In this study, we use repeated observations to estimate

$K_x(s,t)$ and $K_u(s,t)$ due to measurement errors. Suppose we have R repeated observations of $W_i(t)$, denoted as $W_i^r(t)$, for $r = 1,\ldots,R$ and $i = 1,\ldots,n$. After a simple interpolation rule as in Kato (2012), $K_x(s,t)$ is estimated as $\hat{K}_x(s,t) = \hat{K}_w(s,t) - \hat{K}_u(s,t)$, where $\hat{K}_w(s,t) = \sum_{i=1}^n [W_i(s) - \bar{W}(s)][W_i(t) - \bar{W}(t)]/(n-1)$, $\hat{K}_u(s,t) = \sum_{i=1}^n \sum_{r=1}^R [W_i^r(s) - \bar{W}_i(s)][W_i^r(t) - \bar{W}_i(t)]/(nR-n)$, with $\bar{W}_i(t) = (1/R)\sum_{r=1}^R W_i^r(t)$ and $\bar{W}(t) = (1/n)\sum_i \bar{W}_i(t)$. Let the empirical eigen-decomposition of $\hat{K}_x(s,t)$ be $\hat{K}_x(s,t) = \sum_{j=1}^m \hat{\kappa}_{xj}\hat{\phi}_j(s)\hat{\phi}_j(t)$. Let $\hat{W}_{ij} = \int_0^1 W_i(t)\hat{\phi}_j(t)dt$ and $\hat{W}_{ij}^r = \int_0^1 W_i^r(t)\hat{\phi}_j(t)dt$ be the empirical projections of $W_i(t)$ and $W_i^r(t)$, respectively, onto $\hat{\phi}_j(t)$ for $j = 1,\ldots,m$. Furthermore, let $\hat{\mathbf{W}}_i^r = (\hat{W}_{i1}^r,\ldots,\hat{W}_{im}^r)^T$ be the m-dimensional vector composed of the empirical components of $W_i^r(t)$. Then, Σ_u can be estimated using the within-subject covariance of the repeated empirical components $\hat{\mathbf{W}}_i^r$. Specifically, $\hat{\Sigma}_u = \sum_{i=1}^n \sum_{r=1}^R (\hat{\mathbf{W}}_i^r - \bar{\mathbf{W}}_i^*)(\hat{\mathbf{W}}_i^r - \bar{\mathbf{W}}_i^*)^T/(nR-n)$, with $\bar{\mathbf{W}}_i^* = (1/R)\sum_{r=1}^R \hat{\mathbf{W}}_i^r$. Now, we can obtain our estimators using

$$\big(\check{\mathbf{b}}(\tau), \check{\boldsymbol{\theta}}(\tau)\big) = \operatorname{argmin} \sum_{i=1}^n \rho_h^*(Y_i - \hat{\mathbf{W}}_i^T \mathbf{b}(\tau) - \mathbf{Z}_i^T \boldsymbol{\theta}(\tau), \mathbf{b}^T \hat{\Sigma}_u \mathbf{b}, h), \tag{3.5}$$

and the functional coefficient is estimated as $\check{\beta}(t,\tau) = \sum_{j=1}^m \check{b}_j(\tau)\hat{\phi}_j(t)$.

Another practical issue is to determine m, the number of eigen functions, and the tuning parameter h. In our simulation, we use the Bayesian information criterion (BIC) to determine m, which is more stable, as suggested in (Kato (2012)). When there is no measurement error, the oracle estimation gives $BIC(m) = \log[(1/n)\sum_{i=1}^n \rho_\tau(Y_i - Z_i^T\hat{\theta}(\tau) - \sum_{j=1}^m \hat{b}_j(\tau)\hat{X}_{ij})] + (m+p)\log(n)/n$. The BIC can be defined similarly for the naive method by replacing $\hat{X}_{ij}$ with $\hat{W}_{ij}$. As shown in Table 1 in the Supplementary Material, the number of selected scores using the naive and oracle methods are close. Thus, we propose using the naive method to choose m, which is also used for the proposed corrected loss method.

For the tuning parameter h in the corrected quantile objective function, we apply the simulation and extrapolation method (SIMEX), as in Wang, Stefanski and Zhu (2012) and Delaigle and Hall (2008). For a given m, let $\hat{\gamma}(h) = (\hat{\mathbf{b}}^T(h), \hat{\theta}(h))$ be the corrected loss estimator associated with h. Here, we omit τ for notational simplicity. An optimal h_0 minimizes the mean squared error of $\hat{\gamma}(h)$, defined as $\mathrm{E}[(\hat{\gamma}(h) - \gamma_0)^T \Sigma_{\hat{\gamma}}^{-1}(\hat{\gamma}(h) - \gamma_0)]$, where γ_0 denotes the true regression coefficients and $\Sigma_{\hat{\gamma}}$ is the covariance matrix of $\hat{\gamma}(h)$. Because the mean squared error depends on an unknown covariate $X(t)$, it cannot be calculated directly. Instead, we estimate it using the simulation and extrapolation method. In the simulation step, we generate additional independent error terms $\{\mathbf{U}_{i,1}\}_{i=1}^n$ and $\{\mathbf{U}_{i,2}\}_{i=1}^n$ from $\mathrm{N}(\mathbf{0}, \hat{\Sigma}_u)$, and obtain new surrogate variables $\{\mathbf{W}^*_{i,1} = \mathbf{W}_i + \mathbf{U}_{i,1}\}_{i=1}^n$ and $\{\mathbf{W}^*_{i,2} = \mathbf{W}^*_{i,1} + \mathbf{U}_{i,2}\}_{i=1}^n$, with increasing levels of measurement errors. Based on the simulated data sets $\{Y_i, \mathbf{Z}_i, \mathbf{W}^*_{i,1}\}_{i=1}^n$ and $\{Y_i, \mathbf{Z}_i, \mathbf{W}^*_{i,2}\}_{i=1}^n$, we obtain the corrected loss estimators $\hat{\gamma}_1(h)$ and $\hat{\gamma}_2(h)$, respectively. We repeat the simulation step N_s times. Using $\{\hat{\gamma}_{1,s}(h)\}_{s=1}^{N_s}$ and $\{\hat{\gamma}_{2,s}(h)\}_{s=1}^{N_s}$, we can estimate the mean squared error of $\hat{\gamma}_1(h)$ and $\hat{\gamma}_2(h)$, respectively, as

$$M_1(h) = N_s^{-1}\sum_{s=1}^{N_s}[\hat{\gamma}_{1,s}(h) - \hat{\gamma}(h)]^T \hat{\Sigma}_{\hat{\gamma}_1}^{-1}[\hat{\gamma}_{1,s}(h) - \hat{\gamma}(h)],$$
$$M_2(h) = N_s^{-1}\sum_{s=1}^{N_s}[\hat{\gamma}_{2,s}(h) - \hat{\gamma}_{1,s}(h)]^T \hat{\Sigma}_{\hat{\gamma}_2}^{-1}[\hat{\gamma}_{2,s}(h) - \hat{\gamma}_{1,s}(h)],$$

where $\hat{\Sigma}_{\hat{\gamma}_1}$ and $\hat{\Sigma}_{\hat{\gamma}_2}$ are the sample covariance matrices of $\{\hat{\gamma}_{1,s}(h) - \hat{\gamma}(h)\}_{s=1}^{N_s}$ and $\{\hat{\gamma}_{2,s}(h) - \hat{\gamma}_{1,s}(h)\}_{s=1}^{N_s}$, respectively. Let $\hat{h}_1 = \mathrm{argmin}_h M_1(h)$ and $\hat{h}_2 = \mathrm{argmin}_h M_2(h)$. For the extrapolation step, note that $\{\mathbf{W}_{i,2}\}_{i=1}^n$ measures $\{\mathbf{W}_{i,1}\}_{i=1}^n$ in the same way that $\{\mathbf{W}_{i,1}\}_{i=1}^n$ measures $\{\mathbf{W}_i\}_{i=1}^n$, and $\{\mathbf{W}_i\}_{i=1}^n$ measures $\{\mathbf{X}_i\}_{i=1}^n$. Therefore, the relationship between $\hat{h}_1$ to $\hat{h}_2$ is similar to that between $\hat{h}_0$ to $\hat{h}_1$. Here, $\hat{h}_0 = \hat{h}$ is the optimal tuning parameter under the observed surrogate $\{W_i\}_{i=1}^n$. In particular, Delaigle and Hall (2008) considered a linear back-extrapolation with $\log(\hat{h}_1) - \log(\hat{h}_2) \approx \log(\hat{h}_0) - \log(\hat{h}_1)$. Therefore h_0 can be approximated by $\hat{h} = \hat{h}_1^2/\hat{h}_2$.

4. Theoretical Properties

We first introduce the following notation. For any $z \in R^p$, let $\|z\|$ and $\|z\|_\infty$ be the vector L_2 and the supremum norm of z, respectively. For any $K : [0,1]^2 \to R$, let $\|K\|^2 = \int_0^1\int_0^1 K^2(s,t)dsdt$. For any two positive sequences r_n and s_n, $r_n \asymp s_n$ denotes that r_n/s_n is bounded away from zero and infinity. In addition, E_n denotes the sample mean operator. Moreover, we use the same letters c and C for any positive constants, without distinction in each case. To establish the asymptotic results, we need the following assumptions.

Let $K_x(s,t)$ be the covariance kernel of $X(t)$ with $\{\kappa_j\}_{j=1}^\infty$, and $\{\phi_j\}_{j=1}^\infty$ be the eigenvalue and eigen function sequences. Then, $\{\phi_j\}_{j=1}^\infty$ forms an orthonormal basis of $L_2([0,1])$. For $j \geq 1$, let $X_j = \int X(t)\phi_j(t)dt$ be the projection scores. Similarly, define $\{\kappa_{uj}\}_{j=1}^\infty$, $\{\phi_{uj}\}_{j=1}^\infty$ and $\{U_{uj}\}_{j=1}^\infty$ for $U(t)$. To establish the asymptotic results, we need the following assumptions.

(A1) $\{Y_i, X_i(t), Z_i, U_i(t)\}_{i=1}^n$ are i.i.d. copies of $\{Y, X(t), Z, U(t)\}$, in which the functional measurement error $\{U(t)\}$ is a zero-mean Gaussian process on $[0,1]$ and independent of $\{Y, X(t), Z\}$.

(A2) The functional covariate satisfies $\int_0^1 \mathrm{E}[X^4(t)]dt \leq c$ and $\mathrm{E}[X_j^4] \leq c\kappa_j^2$, where $c^{-1}j^{-\alpha_x} \leq \kappa_j \leq cj^{-\alpha_x}$ and $\kappa_j - \kappa_{j+1} \geq c^{-1}j^{-\alpha_x-1}$, for some $\alpha_x > 1$ and all $j \geq 1$. Similarly, assume $\int_0^1 \mathrm{E}[U^4(t)]dt \leq c$ and $\mathrm{E}[U_{uj}^4] \leq c\kappa_{uj}^2$ with $c^{-1}j^{-\alpha_u} \leq \kappa_{uj} \leq cj^{-\alpha_u}$, for some $\alpha_u > 1$ and all $j \geq 1$.

(A3) There exist constants $\nu_1, \nu_2 \in (0, 2]$ such that $\mathrm{E}[(X(t) - X(s))^2] \leq c|t-s|^{\nu_1}$ and $\mathrm{E}[(U(t) - U(s))^2] \leq c|t-s|^{\nu_2}$, for all $s, t \in [0,1]$.

(A4) Let $\mathbf{A} = (\tilde{\mathbf{X}}^T, \mathbf{Z}^T)$, where $\tilde{X}_j = \kappa_j^{-1/2} X_j$, for each j. Assume $c^{-1} \leq \lambda_{min}(E(\mathbf{A}^T\mathbf{A})) \leq \lambda_{max}(E(\mathbf{A}^T\mathbf{A})) \leq c$ for all n, where λ_{max} and λ_{max} are the smallest and largest eigenvalues.

(A5) For some $\beta > \alpha_x/2 + 1$, $\sup_{\tau\in(0,1)} |b_{0j}(\tau)| \leq cj^{-\beta}$, for all $j \geq 1$.

(A6) For $i = 1, \ldots, n$, the functional curve is observed only at discrete points $0 = t_{i1} \leq t_{i2} \leq \cdots \leq t_{i,L_i+1} = 1$. Define $\Delta_n = \max_{1\leq i\leq n} \max_{1\leq l\leq L_i}(t_{i,l+1} - t_{il})$. Assume $\Delta_n \to 0$, $nm^{2\alpha_x}\Delta_n^{\nu_0} = O(1)$ as $n \to \infty$, where $\nu_0 = \min(\nu_1, \nu_2)$.

(A7) The vector θ_0 is an interior point of the parameter space Θ, which is a compact subset of R^p.

(A8) Let $\varepsilon_0(\tau) = Y - Q_\tau(Y|X(t), Z)$. The conditional density $f(\varepsilon_0(\tau)|X(t), Z)$ is continuously differentiable and bounded away from zero almost surely. In addition, $\mathrm{E}(\varepsilon_0^4(\tau)|X(t), Z)$ is bounded as a function of τ.

Assumptions on the response and functional covariate similar to (A1)-(A3), (A5)-(A6), and (A8) can also be found in Kato (2012), which are needed to establish the estimation consistency of the functional coefficient function when there is no measurement error. In particular, (A2)-(A3) determine the smoothness of the random functions $X(t)$ and $U(t)$, while (A5) controls the smoothness of the regression function. Similar assumptions can also be found in Hall and Horowitz (2007). (A6) implies that the sampling points are dense in $[0,1]$ as the sample size increases. The positive-definite matrix assumption in (A4) is similar to that in Kong et al. (2016). For identifiability purposes, we assume that the measurement error is a Gaussian process. The upper bound for $\{\mathrm{E}[(U_{uj})^4]\}_{j=1}^\infty$ in assumption (A3) is a sufficient condition for bounded moments of $\{U_j\}_{j=1}^\infty$.

Theorem 1. *Under assumptions* (A1)-(A8)*, if* $m^{2\alpha_x+2}/n \to 0$*, then as* $n \to \infty$ *and* $h \to 0$*,*

$$\int_0^1 \{\hat{\beta}(t,\tau) - \beta_0(t,\tau)\}^2 dt = O_p\left[\frac{h^{-1}\exp(ch^{-2})m^{\alpha_x+1/2}[\log(m+n)]^{1/2}}{n^{1/2}} + hm^{\alpha_x}\right] + O_p(mn^{-1}) + O_p(m^{-2\beta+1}).$$

Remark 1. Theorem 1 shows that our functional estimator is consistent in probability. The rate of convergence is composed of three terms, which arise from the truncation and estimation errors associated with approximating the functional coefficient $\beta_0(t,\tau)$ with the orthonormal basis $\{\phi_j\}_{j=1}^m$, and the approximation of the check function in the quantile regression using a smooth differentiable function to incorporate the measurement errors. Often, the second term is dominated by the third term for some m, of which $m \asymp n^{1/(\alpha_x+2\beta)}$ in Kato (2012) is such an example. Let $h = c(\log n)^{-\delta}$ and $m = c(\log n)^{\delta/(2\beta+\alpha_x-1)}$, for some $0 < \delta < 1/2$. Then, the rate of convergence simply reduces to $O_p(m^{-2\beta+1})$. This rate of convergence can also be found in Kato (2012), although it requires a higher order of the number of eigen functions with $m \asymp n^{1/(\alpha_x+2\beta)}$.

To establish the asymptotic normality of $\hat{\boldsymbol{\theta}}$, we require additional assumptions.

(A9) For $l = 1,\ldots,p$, there exists $g_l(t)$ such that $Z_l^* = Z_l - \mathrm{E}(Z_l) - \int_0^1 X(t)g_l(t)dt$ satisfying $\mathrm{E}[Z_l^*|X(t)] = 0$ and $|g_{lj}| \le cj^{-\beta}$ with $g_{lj} = \int_0^1 g_l(t)\phi_j(t)dt$, for each l and $j \ge 1$. In addition, denote $\tilde{Z}_l = Z_l - \mathrm{E}(Z_l) - \int_0^1 W(t)g_l(t)dt$.

(A10) Denote $\mathbf{Z}^* = (Z_1^*,\ldots,Z_p^*)^T$ and $\tilde{\mathbf{Z}} = (\tilde{Z}_1,\ldots,\tilde{Z}_p)^T$. There exist positive definite matrices $\mathbf{B}$ and $\mathbf{D}$ such that $\mathrm{E}\{\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^T[(\partial\rho_h^*/\partial\varepsilon)(\varepsilon_0 - \mathbf{U}^T\mathbf{b}_0, \mathbf{b}_0^T\Sigma_u\mathbf{b}_0)]^2\} \to \mathbf{B}$ and $\mathrm{E}\left[\mathbf{Z}^*\mathbf{Z}^{*T}(\partial^2\rho_h^*/\partial\varepsilon^2)(\varepsilon_0 - \mathbf{U}^T\mathbf{b}_0, \mathbf{b}_0^T\Sigma_u\mathbf{b}_0)\right] \to \mathbf{D}$ as $n \to \infty$ and $h \to 0$. Furthermore, we assume $(\partial^2\rho_h^*/\partial\varepsilon\partial\sigma^2)(\varepsilon_0 - \mathbf{U}^T\mathbf{b}_0, \mathbf{b}_0^T\Sigma_u\mathbf{b}_0)$ and $[(\partial^2\rho_h/\partial\varepsilon^2)(\varepsilon_0)]^2$ have bounded expectations conditional on $\mathbf{Z}$ and $X(t)$.

Remark 2. Assumptions similar to (A9) can be found in Lu, Du and Sun (2014) and Shin (2009), and it is introduced to adjust the dependence between $\mathbf{Z}$ and $X(t)$. The first part of assumption (A10) is the same as Assumption 7 in Wang, Stefanski and Zhu (2012) in the case of error-free covariates. A bounded expectation of the second-order derivatives is need for the partially linear functional model.

Theorem 2. *Under the same assumptions as Theorem* 1 *and Assumptions* (A9) *and* (A10)*, we have* $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \to N(\mathbf{0}, \mathbf{D^{-1}BD^{-1}})$ *in distribution.*

5. Simulation Studies

In this section, we conduct simulation studies to assess the finite-sample performance of the proposed corrected-loss method (CL), and compare it with that of the naive method (NAIVE), which ignores measurement errors by simply replacing X_{ij} with W_{ij} in (3.2). As a benchmark, we also consider the oracle estimator (ORACLE), which assumes that the functional covariate $X(t)$ is fully observed. The data are generated independently from the models $Y = \int_0^1 b(t)X(t) + (Z_1, Z_2)\boldsymbol{\theta} + \epsilon$ and $W(t) = X(t) + U(t)$, where $b(t) = \sum_{j=1}^{50} b_j\phi_j(t)$, with $b_1 = 0.3$, $b_j = 5(-1)^{j+1}/j^{3.5}$ for $j \geq 2$, $X(t) = \sum_{j=1}^{50} \gamma_j Z_j \phi_j(t)$ with $\gamma_j = (-1)^{j+1} j^{-\alpha/2}$, $Z_j \sim U[-\sqrt{3}, \sqrt{3}]$, $U(t) = \sum_{j=1}^{50} \nu_j U_j \phi_j(t)$ with $\nu_j = (-1)^{j+1} j^{-\alpha/2}$, $U_j \sim N(0, 1)$, and the regression error $\epsilon = \{1 + \eta \int_0^1 \cos(\pi t)X(t)dt\}N(0, 1)$. We consider the two values $\eta = 0$ and 0.5 in Case 1 and Case 2, respectively, corresponding to homogeneous and heteroscedastic regression models. In Case 3, ϵ is generated from a t distribution with five degrees of freedom, but is normalized to have mean zero and variance one. We consider the basis functions $\phi_j(t) = \sqrt{2}\cos(\pi j t)$ and the parameter $\alpha = 1.1$ or 2, which controls the effective number of basis functions needed for functional data. For the parametric part, we consider the regression coefficients $\boldsymbol{\theta} = (0.3, 0.5)^T$ and the covariates $Z_1 \sim Binomial(1, 0.6)$ and $Z_2 \sim N(\mu, 0.25)$, where μ depends on the functional covariate $X(t)$ by $\mu = 1 + \int_0^1 \sum_{j=1}^{5} (-1)^{j+1}(1/j^2)\phi_j(t)X(t)dt$. We set $n = 200, 400, 600$, and repeat the experiment 100 times. In the following, we consider the proposed estimation method at both the 50th and the 75th quantiles.

For our method, we assume $X(t)$ is unobserved. Instead, several replications of $W(t)$ are observed to estimate the covariance function $K_U(s,t)$ of the measurement error. Here, we estimate $K_U(s,t)$ based on three replications of $W(t)$, and the average of three replications is taken as the observed $W(t)$. Then, for any two time points s and t, the covariance function of $X(t)$ can be estimated as $\hat{K}_X(s,t) = \hat{K}_W(s,t) - \hat{K}_U(s,t)$. We then calculate $\{\hat{\phi}_j(t)\}_{j=1}^m$ as the first m eigen functions of $\hat{K}_X(s,t)$. The number of eigen functions m is selected using the BIC in subsection 3.1. For the selection of h, the SIMEX method in subsection 3.1 is used with $N_s = 20$ in the simulation step. To save computational time, the SIMEX selection is performed on only 30 experiments, and the average of 30 values $\hat{h}_{simex}$ is used for each of the 100 replications. The average number of selected eigen functions, with standard deviations and the tuning parameter h, are reported in Tables 1 and 2 of the Supplementary Material.

Figure 1 shows box plots of $\hat{\theta}_k(\tau) - \theta_k(\tau)$ $(k = 0, 1, 2)$, and Table 1 summarizes the MSE values of $\hat{\theta}_k(\tau)$ $(k = 0, 1, 2)$ at $\tau = 0.75$ from three different methods.

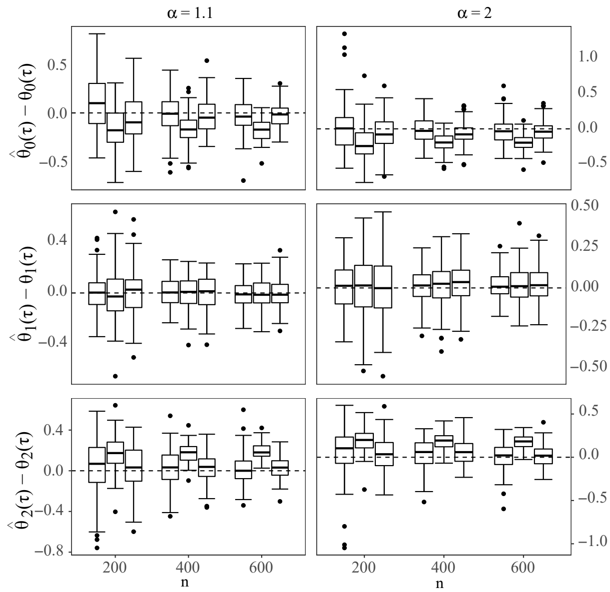

Figure 1. Box plots of $\hat{\theta}_k(\tau) - \hat{\theta}_{k0}(\tau), k = 0, 1, 2$ at $\tau = 0.75$, but for different α values and sample sizes from the data generated in Case 1. In each figure, the box plots from left to right show CL, NAIVE, and ORACLE respectively.

Table 1. MSEs of the parametric coefficient estimators from the proposed method (CL), naive method (NAIVE), and oracle method (ORACLE) at $\tau = 0.75$ in Case 1.

		MSE($\hat{\theta}_0$)			MSE($\hat{\theta}_1$)			MSE($\hat{\theta}_2$)		
α	n	CL	NAIVE	ORACLE	CL	NAIVE	ORACLE	CL	NAIVE	ORACLE
1.1	200	0.0944	0.0767	0.0554	0.0209	0.0389	0.0331	0.0793	0.0567	0.0424
	400	0.0385	0.0507	0.0322	0.0124	0.0184	0.0180	0.0328	0.0387	0.0209
	600	0.0268	0.0416	0.0165	0.0096	0.0158	0.0143	0.0256	0.0429	0.0116
2	200	0.1049	0.0955	0.0668	0.0195	0.0360	0.0379	0.0867	0.0556	0.0389
	400	0.0272	0.0538	0.0304	0.0104	0.0181	0.0163	0.0276	0.0431	0.0240
	600	0.0311	0.0485	0.0239	0.0080	0.0141	0.0134	0.0255	0.0376	0.0159

The data are generated with regression errors in Case 1 and $\alpha = 1.1$ or 2. Table 1 shows that the CL, NAIVE, and ORACLE approaches all give similar estimation results for θ_1, with negligible bias under all scenarios, because Z_1 is independent of the functional covariate that is contaminated with errors. However, because the distribution of Z_2 is affected by measurement errors, the naive estimates of θ_0 and θ_2 have obvious biases. In contrast, our method is effective in reducing the estimation bias, and results in much smaller biases than that of the NAIVE method. The NAIVE approach gives the smallest estimation variances, because of the extra variability in the observed covarites due to measurement errors.

Table 2. Bias, Variance, and IMSE of the functional coefficient estimators at $\tau = 0.75$ in Case 1.

		CL			NAIVE			ORACLE		
α	n	$Bias^2$	Var	IMSE	$Bias^2$	Var	IMSE	$Bias^2$	Var	IMSE
1.1	200	0.0439	0.2015	0.2454	0.1240	0.0660	0.1900	0.0136	0.1176	0.1312
	400	0.0124	0.0812	0.0936	0.0862	0.0280	0.1142	0.0106	0.0486	0.0592
	600	0.0100	0.0594	0.0693	0.0770	0.0173	0.0943	0.0086	0.0296	0.0382
2	200	0.1135	0.1841	0.2976	0.1802	0.0603	0.2405	0.0357	0.1818	0.2175
	400	0.0420	0.1172	0.1592	0.1114	0.0543	0.1656	0.0197	0.0913	0.1109
	600	0.0159	0.0831	0.0990	0.0840	0.0322	0.1162	0.0139	0.0428	0.0567

However, Table 1 shows that the CL approach has better overall performance, with a smaller MSE than that of the NAIVE method when the sample size is larger ($n = 400, 600$), and the overall performance of CL is comparable with that of the ORACLE approach. In addition, the MSEs for both CL and ORACLE decrease as n increases, supporting our asymptotic results. However, for $\hat{\theta}_0$ and $\hat{\theta}_2$, the NAIVE method has nondiminishing MSEs due to nonignorable bias.

The integrated mean squared error (IMSE) is used to assess the accuracy of the functional coefficient estimators. Let $\hat{b}_l(t, \tau)$ be an estimated coefficient from the lth experiment, for $l = 1, \ldots, 100$. Let $\bar{b}(t, \tau) = \sum_{l=1}^{100} \hat{b}_l(t, \tau)/100$. We define $\text{Bias}^2 = \int_0^1 [\bar{b}(t, \tau) - b(t, \tau)]^2 dt$, $\text{Var} = \int_0^1 \sum_{l=1}^{100} [\hat{b}_l(t, \tau) - \bar{b}(t, \tau)]^2/100dt$, and $\text{IMSE} = \int_0^1 \sum_{l=1}^{100} [\hat{b}_l(t, \tau) - b(t, \tau)]^2/100dt$.

Table 2 reports the Bias^2, Var, and IMSE of the three functional coefficient estimates. It again shows the effectiveness of the proposed method in reducing the estimation bias due to the measurement error. The size of the Bias^2 under the CL is smaller than that under the NAIVE approach, and is close to the ORACLE estimator, particularly when n is large. For larger n, our proposed method obtains smaller IMSE values than those of the naive method. In addition, when α increases from 1.1 to 2, the IMSE of $\hat{b}(t, \tau)$ increases which is consistent with Remark 1. The same phenomenon was also observed in Kato (2012). Figure 2 plots the average of the functional coefficient estimates over 100 replications for each of three methods. The curves of the oracle estimator and our proposed estimator are closer to the true line than is the naive curve at each α and sample size n. Other results, not provided here, show the same conclusions.

To investigate the performance of the proposed estimator under different regression error distributions, we generate data from the above model with $n = 400$ and $\alpha = 1.1$, but with three different regression error distributions, as described in cases 1, 2, and 3. We also compare the performance of the proposed method

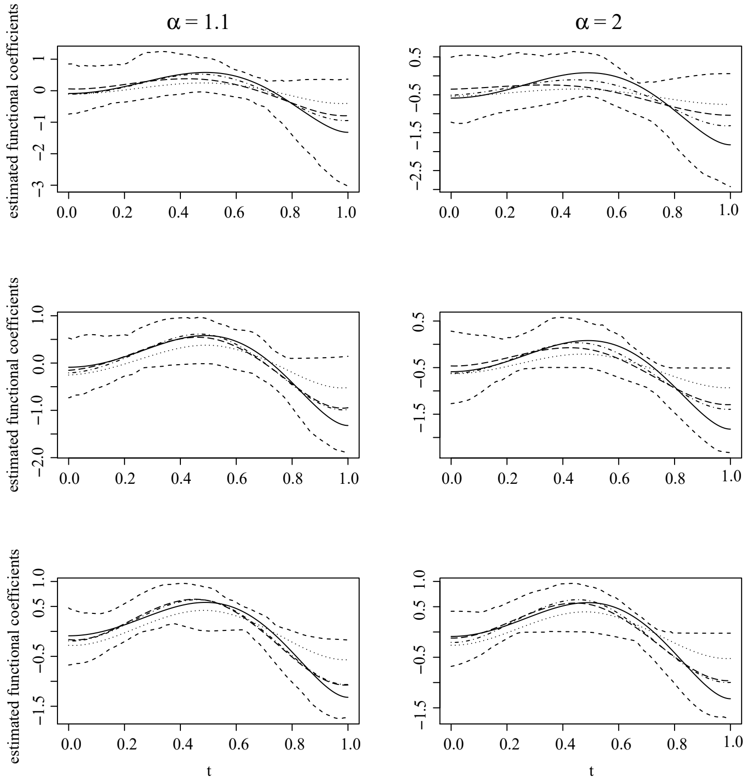

Figure 2. Estimated functional coefficients at $\tau = 0.75$ for $\alpha = 1.1$ or 2 with the data generated in Case 1. Top to bottom show $n = 200, 400, 600$ for each α. In each plot, the solid line represents the true functional coefficient, the dotted and dashed line shows ORACLE, the longer dashed line shows CL, the dotted line shows NAIVE, and the shorter dashed lines show the 95% interval of CL.

at the 50th and 75th quantiles. Tables 3 and 4 in the Supplementary Material report the estimation results. Under all error distributions and both quantile levels, the NAIVE approach gives nonignorable biases when estimating both parametric and nonparametric regression coefficients. In contrast, the proposed method is effective in reducing the estimation bias under all scenarios.

We also use a simulation to verify the asymptotic variance formula for the parametric part given in Theorem 2. Data of sizes $n = 200, 400, 600$ are generated from Case 1, with $\tau = 0.5$ and $\alpha = 2$. Table 3 compares the empirical standard errors with the theoretical values calculated using the asymptotic formula in Theorem 2 for $\hat{\theta}_0$, $\hat{\theta}_1$, and $\hat{\theta}_2$. The results show that the empirical and theoretical standard errors are close for all three parametric regression coefficients, thus supporting the asymptotic results in Theorem 2.

Table 3. Empirical and theoretical standard errors for the data generated in Case 1.

	n	Empirical SE	Theoretical SE
θ_0	200	0.279	0.248
	400	0.172	0.177
	600	0.166	0.147
θ_1	200	0.152	0.150
	400	0.109	0.106
	600	0.074	0.086
θ_2	200	0.258	0.220
	400	0.158	0.158
	600	0.147	0.131

Table 4. Empirical coverage probabilities (in %) of the bootstrap confidence intervals with a nominal level of 95% for Case 1 at $\tau = 0.5$, $\alpha = 2$.

n	θ_0	θ_1	θ_2	$b(t)$
200	97	99	94	95
400	96	96	96	95
600	95	98	96	93

However, it is challenging to use the theoretical standard errors to make an inference on the regression coefficients, because the calculation of former relies on the true data-generating process, which is unavailable and needs to be estimated in practice. In addition, it involves projecting error free covariates on $\{X(t)\}$, which is not feasible in practice because $\{X(t)\}$ is unobserved due to the measurement error. Therefore, we propose using the bootstrap method to construct confidence intervals for the unknown regression coefficients. Table 4 provides the empirical coverage probabilities of the 95% confidence intervals based on 500 bootstrap samples for $\tau = 0.5$ and $\alpha = 2$ of Case 1. For the functional coefficient, the coverage probability of the pointwise confidence interval is calculated by averaging the coverage probabilities at 201 observed time points in the interval $[0, 1]$. Table 4 shows that the bootstrap approach performs reasonably well, with empirical coverage probabilities close to 95%.

6. Real Data Analysis

In this section, we apply the proposed method to data from a childhood obesity study. About 20% of children in the United States are obese, and the prevalence of childhood obesity has more than tripled in the last 40 years. Childhood obesity negatively affects children's physiological, behavioral, and psychological development. To combat the obesity epidemic, researchers have begun implementing behavioral school-based interventions aimed at increasing physical activity in children. In this section, we apply our proposed method to a data set from a children obesity study conducted by Dr. Mark Benden and colleagues from 2012 to 2014.

In this study, a total of 230 students from three different elementary schools in the College Station Independent School District were enrolled and followed over a six-month period. The students were randomly assigned to receive either stand-biased desks (treatment) or traditional desks (control) in their classrooms. The purpose of the study was to evaluate the effect of stand-biased desks as an intervention aimed at increasing energy expenditure and reducing obesity in elementary school-aged children. In our application, daily energy expenditure, $X(t)$, is defined as the total number of calories or energy used by the body to perform everyday bodily functions. The true values of $X(t)$ are not directly observable. Instead a surrogate measure for daily energy expenditure, $W(t)$, was collected per minute using the Sense Wear Armband® (BodyMedia, Pittsburgh, PA) for students who wore accelerometers while in school for one week (five days) at baseline. When energy expenditure measurements have missing values, cubic splines are applied to smooth each individual energy expenditure curve, which is then used to impute the missing values. In addition to accelerometry-based energy expenditure data, other covariates collected at baseline include each subject's school, age, sex, ethnicity, height, and treatment. The covariate treatment is a binary variable used to indicate desk assignment, taking the value one for a stand-biased desks and zero for a traditional desk. BMI values are measured both at baseline and after six months for all 230 students in the study. We take the average of the two BMI values to reflect the overall BMI level during the six-month period for each student. Table 5 provides a summary of the variables.

Students from different schools enrolled in the study at different time points. To eliminate any potential effect of enrollment time on students' daily energy expenditure patterns, we center the measured energy expenditure data by subtracting the daily averaged energy expenditure from each individual curve. In addition, the daily energy expenditure curve is averaged for every five minutes to reduce the variability, which leaves 52 data points in each curve. Furthermore, the time interval is scaled to $[0, 1]$. Because some students have only three days' measurements of energy expenditure, we randomly pick three days' data to estimate the covariance kernel of $X(t)$ and variance of the measurement errors. Figure 3 provides plots of the three-day data, where $W_1(t)$, $W_2(t)$, and $W_3(t)$ are plotted against time for all subjects. The gray lines are individual energy expenditure observations, and the black solid line is the mean energy expenditure.

Table 5. Descriptive statistics for the children obesity study; "Other" = Hispanics/Asians/Native Americans.

Variable	mean(sd)/N(%)
log(BMI)	2.846(0.159)
Age	8.378(0.754)
Treatment	134(58.261)
Control	96(41.026)
Blacks	25(10.870)
Other	42(18.260)
Whites	163(70.870)
Boys	120(52.174)
Girls	110(47.826)
School 1	57(24.783)
School 2	89(38.696)
School 3	84(36.522)

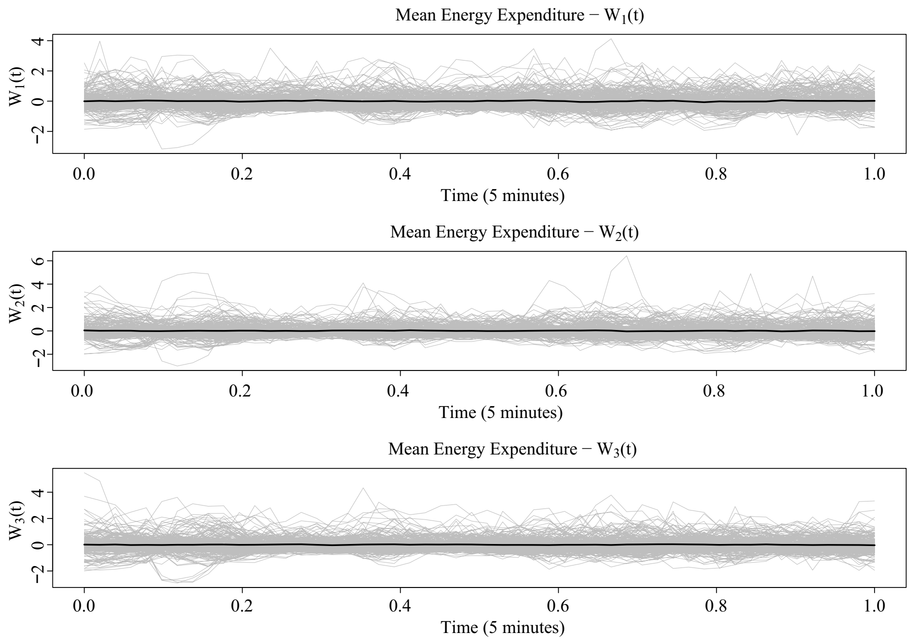

Figure 3. Plots of observed energy expenditure $\{W_1(t), W_2(t), W_3(t)\}$ versus time for all subjects at baseline.

We consider the following models for Q_τ, the τth conditional quantile of log(BMI), and the measurement errors, $Q_\tau = \theta_0 + \theta_1 S_1 + \theta_2 S_2 + \theta_3 Trt + \theta_4 Sex + \theta_5 R_1 + \theta_6 R_2 + \theta_7 Age + \int X(t)\beta(t)dt$, and $W_j(t) = X(t) + U_j(t)$, for $j = 1, 2, 3$,

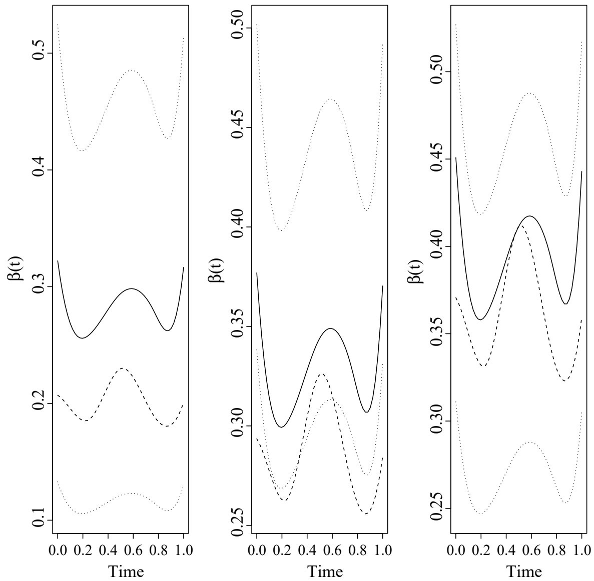

Figure 4. Plots of proposed (solid line) and naive (dashed line) estimates of $\beta(t)$ at $\tau = 0.2,\ 0.5,\ 0.8$ from left to right. The two dotted lines are the 95% pointwise bootstrap confidence intervals with 500 boostrap samples.

where S_1 and S_2 are binary indicators for School 1 and School 2, respectively, and School 3 is set as the baseline. In addition, Trt is a binary treatment indicator for desk assignment, and Sex is a binary variable, taking the value one for male, and zero for female students. For ethnicity, R_1 and R_2 are binary indicators for Other Ethnicity and Black respectively and White is the reference category. For illustration, we consider $\tau = 0.2,\ 0.5,\ 0.8$, and compare our proposed method with the naive method, which ignores measurement errors. The number of bases used to represent the functional data is selected using the BIC criterion based on the naive method. It selected one basis ($m = 1$) for all three levels of τ. Using SIMEX, we choose h as 2.116, 3.462,1.900 at $\tau = 0.2,\ 0.5,\ 0.8$, respectively, for our method.

The estimated functional coefficients are provided in Figure 4. For each quantile level, the 95% nonparametric bootstrap confidence intervals do not contain the zero line, indicating that energy expenditure has a significant effect on the 20th, 50th, and 80th quantiles of log(BMI). In addition, the estimated coefficient functions are positive, indicating more daily energy expenditures associated with higher BMI values. This may be because a higher energy cost is needed to perform weight-bearing activities for individuals with higher BMI values. This positive association was also observed in Maffeis et al. (1996). In addition, Figure

Table 6. Parametric regression coefficient estimates by the proposed (CL) and naive methods at $\tau = 0.2,\ 0.5,\ 0.8$. Standard errors from 500 bootstrap samples are included in parentheses, and significance at 5% level is shown in bold face.

τ	Method	Inter	S1	S2	Trt	Sex	R1	R2	Age
0.2	CL	**1.788**	-0.098	-0.071	-0.052	-0.035	0.031	0.075	0.014
		(0.267)	(0.061)	(0.040)	(0.029)	(0.034)	(0.046)	(0.074)	(0.033)
	NAIVE	**2.493**	**-0.054**	-0.019	-0.013	-0.025	0.010	-0.009	**0.035**
		(0.092)	(0.022)	(0.018)	(0.015)	(0.014)	(0.018)	(0.042)	(0.011)
0.5	CL	**2.539**	**-0.069**	-0.023	0.005	**-0.044**	0.039	0.049	**0.040**
		(0.146)	(0.024)	(0.019)	(0.019)	(0.018)	(0.023)	(0.035)	(0.017)
	NAIVE	**2.562**	**-0.069**	-0.038	-0.000	-0.032	0.020	0.048	**0.036**
		(0.147)	(0.028)	(0.028)	(0.026)	(0.026)	(0.029)	(0.044)	(0.017)
0.8	CL	**3.058**	**-0.143**	-0.043	0.000	-0.037	0.017	-0.029	**0.093**
		(0.191)	(0.040)	(0.027)	(0.027)	(0.024)	(0.033)	(0.044)	(0.024)
	NAIVE	**2.785**	-0.080	-0.049	-0.024	**-0.071**	0.004	-0.011	0.029
		(0.163)	(0.044)	(0.031)	(0.023)	(0.027)	(0.047)	(0.072)	(0.021)

4 shows that the naive estimates attenuate its effects toward zero compared with the proposed estimates. Table 6 reports the estimation results of the parametric regression coefficients. As expected, the estimates of the interceptor increase with the quantile level τ, and all are significantly different from zero. Our proposed estimates indicate significant school effects at the 50th and 80th quantiles, but no statistically significant school effect is found at the 20th quantile due to the larger variability of the proposed method. For desk assignment (Trt), there are no statistically significant findings at any quantile levels for both methods. This is possibly because its effect is mediated by the inclusion of energy expenditure in the model. For Sex, our proposed estimates show significant effects at the 50th quantile, rather than at the 80th quantile indicated by the naive estimates. Both methods reveal no significant effects of ethnicity on log(BMI) at the three quantile levels. For age, both methods find a significant positive effect for median BMI. However, the proposed method finds age to be significant at the 80th quantile, while naive method found it to be significant at the 20th quantile.

7. Discussion

We have established consistent estimators for a partially functional linear quantile model when the functional covariate is contaminated by a function-valued error. The nondifferentiable check loss in the quantile regression and the functional measurement error impose additional challenges. Our method does not require a specification of the conditional distribution of the response given the true covariates, and does not assume independence between measurement errors at different time points. We assume the functional measurement error is a Gaussian process, and develop the corrected loss function based on a smooth function of the check loss function, where a smoothing parameter needs to be determined for a bias and variance trade-off. We show the consistency and asymptotic normality of the proposed functional estimator and the parametric estimator. Simulations and a real-data analysis show that our method outperforms the naive method.

The proposed model allows for only one functional predictor. However, it can be extended to allow multiple functional predictors. With the basis representation, it leads to more terms in either $\mathbf{Z}$ and/or $\mathbf{X}$, depending on whether the functional predictors are measured with or without errors. While the algorithm proposed in this paper can be extended to obtain resulting estimators, the proofs for establishing the asymptotic results require additional work, owing to the extra functional predictors.

In our implementation, we determine the number of eigen functions by simply ignoring the existence of measurement errors. A more precise method should be investigated to determine the number of eigen functions in the presence of measurement errors. Furthermore, we could consider a more general measurement error model, such as, the multiplicative measurement error, or relaxing the Gaussian assumption of the measurement error used in the proposed method.

In addition, the Laplace distribution is another popular distribution for the measurement errors. In fact, Wang, Stefanski and Zhu (2012) considered the corrected score approach for measurement error models with a Laplace distribution in a linear quantile regression. However, one challenge to extending it to the functional covariate case is that the Laplace distribution does not have the linear additive property. That is, a linear combination of Laplace random variables does not necessarily follow the Laplace distribution. Therefore, if the functional measurement errors are assumed to follow a multivariate Laplace distribution, the individual scores may no longer follow the same distribution. Thus, an extension to Laplace functional measurement errors is not straightforward, and is left to future research.

Supplementary Material

The online Supplementary Material includes additional simulation results and detailed proofs of the main theorems and necessary lemmas.

Acknowledgments

The authors thank Mark Benden for providing the data used in the application section. The authors also thank the reviewers, associate editor, and co-editor for their helpful suggestions and comments. Research reported in this publication was supported by the National Institute Of Diabetes And Digestive And Kidney Diseases of the National Institutes of Health under Award Number R01DK132385. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. In addition, Xue's research was supported by National Science Foundation (DMS-1812258). Tekwe's research was supported by National Cancer Institute Supplemental Award U01-CA057030-29S2. Bai's research was supported by the Natural Science Foundation of China (11771268). Qu's research was supported by the National Science Foundation (DMS-1821198).

References

Bassett, D. R. (2012). Device-based monitoring in physical activity and public health research. *Physiological Measurement* **33**, 1769–1783.

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective.* 2nd Edition. Chapman and Hall/CRC, New York.

Crouter, S. E., Churilla, J. R. and Bassett, D. R. (2006). Estimating energy expenditure using accelerometers. *European Journal of Applied Physiology* **98**, 601–612.

Delaigle, A. and Hall, P. (2008). Using simex for smoothing-parameter choice in errors-in-variables problems. *Journal of the American Statistical Association* **103**, 280–287.

Firpo, S., Galvao, A. F. and Song, S. (2017). Measurement errors in quantile regression models. *Journal of Econometrics* **198**, 146–164.

Fuller, W. (1987). *Measurement Error Models.* John Wiley & Sons, New York.

Geraci, M. and Bottai, M. (2014). Linear quantile mixed models. *Statistics and Computing* **24**, 461–479.

Hall, P. and Horowitz, J. L. (2007). Methodology and convergence rates for functional linear regression. *The Annals of Statistics* **35**, 70–91.

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.

Hu, Y. and Schennach, S. (2008). Instrumental variable treatment of nonclassical measurement error models. *Econometrica* **76**, 195–216.

Jacobi, D., Perrin, A.-E., Grosman, N., Doré, M.-F., Normand, S., Oppert, J.-M. et al. (2007). Physical activity-related energy expenditure with the RT3 and TriTrac accelerometers in overweight adults. *Obesity* **15**, 950–956.

Kato, K. (2012). Estimation in functional linear quantile regression. *The Annals of Statistics* **40**, 3108–3136.

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

Koenker, R. and Bassett, G., Jr. (1978). Regression quantiles. *Econometrica* **46**, 33–50.

Kong, D., Xue, K., Yao, F. and Zhang, H. (2016). Partially functional linear regression in high dimensions. *Biometrika* **103**, 147–159.

Lu, Y., Du, J. and Sun, Z. (2014). Functional partially linear quantile regression model. *Metrika* **77**, 317–332.

Maffeis, C., Zaffanello, M., Pinelli, L. and Schutz, Y. (1996). Total energy expenditure and patterns of activity in 8-10-year-old obese and nonobese children. *Journal of Pediatric Gastroenterology and Nutrition* **23**, 256–261.

Nakamura, T. (1990). Corrected score function for errors-in-variables models: Methodology and application to generalized linear models. *Biometrika* **77**, 127–137.

Ramsay, J. O. and Silverman, B. W.(2005). *Functional Data Analysis.* 2nd Edition. Springer, New York.

Rothney, M. P., Schaefer, E. V., Neumann, M. M., Choi, L. and Chen, K. Y. (2008). Validity of physical activity intensity predictions by actigraph, actical, and RT3 accelerometers. *Obesity* **16**, 1946–1952.

Shin, H. (2009). Partial functional linear regression. *Journal of Statistical Planning and Inference* **139**, 3405–3418.

Stefanski, L. A. and Cook, J. R. (1995). Simulation-extrapolation: The measurement error jackknife. *Journal of the American Statistical Association* **90**, 1247–1256.

Tekwe, C., Zoh, R., Yang, M., Carroll, R., Honvoh, G., Allison, D. et al. (2019). Instrumental variable approach to estimating the scalar-on-function regression model with measurement error with application to energy expenditure assessment in childhood obesity. *Statistics in Medicine* **38**, 3764–3781.

Wang, H. J., Stefanski, L. A. and Zhu, Z. (2012). Corrected-loss estimation for quantile regression with covariate measurement errors. *Biometrika* **99**, 405–421.

Warolin, J., Carrico, A., Whitaker, L., Wang, L., Chen, K., Acra, S. et al. (2012). Effect of bmi on prediction of accelerometry-based energy expenditure in youth. *Medicine and Science in Sports and Exercise* **44**, 2428.

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

Yao, F., Sue-Chee, S. and Wang, F. (2017). Regularized partially functional quantile regression. *Journal of Multivariate Analysis* **156**, 39–56.

Mengli Zhang

School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China.

E-mail: zhangmengli_sufe@163.com

Lan Xue
Department of Statistics, Oregon State University, Corvallis, OR 97331-4501, USA.
E-mail: Lan.Xue@oregonstate.edu
Carmen D. Tekwe
School of Public Health, Indiana University Bloomington, Bloomington, IN 47405, USA.
E-mail: ctekwe@iu.edu
Yang Bai
School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China.
E-mail: statbyang@mail.shufe.edu.cn
Annie Qu
Department of Statistics, University of California, Irvine, Irvine, CA 92697-1250, USA.
E-mail: aqu2@uci.edu

(Received July 2021; accepted January 2022)