Brazilian Journal of Probability and Statistics
2016, Vol. 30, No. 3, 464–484
DOI: 10.1214/15-BJPS289

Statistical inference on restricted linear regression models with partial distortion measurement errors

Zhenghong Wei, Yongbin Fan and Jun Zhang

Shenzhen University

Abstract. We consider statistical inference for linear regression models when some variables are distorted with errors by some unknown functions of commonly observable confounding variables. The proposed estimation procedure is designed to accommodate undistorted as well as distorted variables. To test a hypothesis on the parametric components, a restricted least squares estimator is proposed for unknown parameters under some restricted conditions. Asymptotic properties for the estimators are established. A test statistic based on the difference between the residual sums of squares under the null and alternative hypotheses is proposed, and we also obtain the asymptotic properties of the test statistic. A wild bootstrap procedure is proposed to calculate critical values. Simulation studies are conducted to demonstrate the performance of the proposed procedure and a real example is analysed for an illustration.

1 Introduction

In some applications, variables may not be directly observed because of certain contamination, such as in health science and medicine research. It is well known that measurement errors in covariates may cause large bias, sometimes seriously, in the estimated regression coefficient if we ignore the measurement error. As such, measurement error models have received much attention. Some literature about measurement error models include Kneip, Simar and Van Keilegom (2015); Li and Xue (2008); Saleh and Shalabh (2014); Stefanski, Wu and White (2014); Xu and Zhu (2015); Yang, Li and Peng (2014); Zhang, Li and Xue (2011) and others. Carroll et al. (2006) systematically summarized some recent research developments of linear and non-linear models as well as in non-parametric and semi-parametric models. Our interest is in quantifying the following linear regression models with partial distortion measurement errors:

$$\begin{cases} Y = \boldsymbol{\beta}_0^\tau \mathbf{X} + \boldsymbol{\gamma}_0^\tau \mathbf{Z} + \varepsilon, \\ \tilde{Y} = \phi(U)Y, \qquad \tilde{\mathbf{X}} = \psi(U)\mathbf{X}, \end{cases} \tag{1.1}$$

where Y is an unobservable response, $\mathbf{X} = (X_1, X_2, \dots, X_p)^\tau$ is an unobservable continuous predictor vector (the superscript τ denotes the transpose operator throughout this paper), $\mathbf{Z} = (Z_1, Z_2, \dots, Z_q)^\tau$ is an observed predictor vector,

Key words and phrases. Distortion measurement errors, local linear smoothing, restricted estimator, bootstrap procedure.
Received February 2014; accepted March 2015.

$\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$ are the unknown coefficient parameters for the linear regression model of Y on $\mathbf{X}, \mathbf{Z}$. $\tilde{Y}$ and $\tilde{\mathbf{X}}$ are the distorted and observed response and predictors, $\psi(\cdot)$ is a $p \times p$ diagonal matrix $\operatorname{diag}(\psi_1(\cdot), \dots, \psi_p(\cdot))$, where $\phi(\cdot)$ and $\psi_r(\cdot)$ are unknown continuous distorting functions. The confounding variable $U \in \mathbb{R}^1$ is independent of $(\mathbf{X}^\tau, \mathbf{Z}^\tau, Y)^\tau$. The diagonal form of $\psi(\cdot)$ indicates that the confounding variable U distorts unobserved X_r, $r = 1, \dots, p$ in a multiplicative fashion. The model error ε is independent of $(\mathbf{X}^\tau, \mathbf{Z}^\tau, U)^\tau$.

This type of measurement error models is revealed by Şentürk and Nguyen (2009), who studied the Pima Indians diabetes data. It is interest to investigate which covariates have impact on the diabetes. To analyse this data, Şentürk and Nguyen (2009) suggested the covariate "body mass index" (BMI) to be a potential confounding variable. Kaysen et al. (2002) also treat BMI as the confounder on hemodialysis patients and they further realised that the fibrinogen level and serum transferrin level should be divided by BMI, in order to eliminate the contamination possibly caused by BMI. From a practical point of view, since the exact relationship between the confounder and primary variables is unknown, naively dividing BMI may be incorrect and may lead to an inconsistent estimator of the parameter. As a remedy, Şentürk and Müller (2005, 2006) introduced a flexible multiplicative adjustment by using unknown smooth distorting functions $\phi(\cdot), \psi_r(\cdot)$. Recently, there have been much literature on the statistical modelling of the distortion measurement errors data. Şentürk and Müller (2005, 2006) proposed parametric covariate-adjusted models with one-dimensional confounding variable in the setting in which the observed $\tilde{Y}$ and $\tilde{\mathbf{X}}$ are related through a varying coefficient model (Xu and Guo, 2013; Xu and Zhu, 2013), using the binning method (Fan and Zhang, 2000). Şentürk and Nguyen (2006) further proposed a local linear estimator to deal with partial distortion measurement error-in-variables. Nguyen, Şentürk and Carroll (2008) proposed an estimation procedure for linear mixed effects models with an application to longitudinal data. Cui et al. (2009) proposed a driect-plug-in estimation procedure. Later on, this direct-plug-in method has been developed to the case of multivariate confounders (Zhang, Zhu and Liang, 2012a, 2013a, 2014c) and some semi-parametric models, for example, partial linear models (Li, Lin and Cui, 2010), the partial linear single index models (Zhang et al., 2013b), the dimension reduction models (Zhang, Zhu and Zhu, 2012b). Zhang, Feng and Zhou (2014a) studied an efficient estimator for the correlation coefficient between two variables that are both observed with distortion measurement errors. Li et al. (2014) employs the smoothly clipped absolute deviation penalty (Liang and Li 2009; Fan and Li 2001, SCAD) least squares method to simultaneously select variables and estimate the coefficients for high-dimensional covariate adjusted linear regression models. Recently, Zhang, Li and Feng (2015) proposed a residuals based empirical process based test statistic to solve the problem of model checking on parametric distortion measurement error regression model.

In this article, we investigate the estimation and statistical inference for model (1.1). To estimate the unknown parameter $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$, we used the direct plug-in method proposed by Cui et al. (2009) and further investigate the asymptotic properties of the estimators. To make statistical inference of $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$, that is, testing whether the true parameter $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$ satisfies some linear restriction conditions or not, we further develop a restricted estimator by introducing Lagrange multipliers under the null hypothesis. The associated asymptotic properties of the restricted estimator are also revealed. Finally, a test statistic based on the difference between the residual sums of squares under the null and alternative hypotheses is proposed. The limiting distribution of the test statistic is shown to be a weighted sum of independent standard chi-squared distributions under the null hypothesis. To mimic the null distribution of the test statistic, a wild bootstrap procedure is proposed to define p-values. We conduct Monte Carlo simulation experiments to examine the performance of the proposed procedures. Our simulation results show that the proposed methods perform well both in estimation and hypothesis testing. We also apply our estimation and testing procedure to analyze the Pima Indians diabetes data set.

The paper is organized as follows. In Section 2, we describe the direct-plug-in estimation procedure for parameters $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$, and present the associated asymptotic results. In Section 3, we provide a test statistic for the testing problem and give a restricted estimator under the null hypothesis, associated their theoretical properties. A wild bootstrap procedure is also proposed to mimic the null distribution of the test statistic. In Section 4, we report the results of simulation studies and present the results of our statistical analysis of a diabetes study. All the technical proofs of the asymptotic results are given in the Appendix.

2 Estimation procedures for $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$

2.1 Covariate calibration

In this subsection, a calibration estimation procedure is proposed to estimate unobserved $Y, \mathbf{X}$. Using the observed i.i.d. samples $\{\tilde{Y}_i, \tilde{\mathbf{X}}_i, U_i\}_{i=1}^n$, we follow Cui et al.'s (2009) estimation procedure to estimate unknown distorting functions $\phi(\cdot)$, $\psi_r(\cdot)$. To ensure identifiability, Şentürk and Müller (2005, 2006) assumed that

$$E[\phi(U)] = 1, \qquad E[\psi_r(U)] = 1, \tag{2.1}$$

$r = 1, \dots, p$. By (1.1) and (2.1), we can have that $E[\tilde{Y}] = E[Y]$, $E[\tilde{X}_r] = E[X_r]$ and

$$E\left[\frac{\tilde{Y}}{E[\tilde{Y}]}\bigg|U\right] = \phi(U), \qquad E\left[\frac{\tilde{X}_r}{E[\tilde{X}_r]}\bigg|U\right] = \psi_r(U). \tag{2.2}$$

Using these equations, the unobserved $\{Y_i, \mathbf{X}_i\}_{i=1}^n$ can be estimated as

$$\hat{Y}_i = \frac{\tilde{Y}_i}{\hat{\phi}(U_i)}, \qquad \hat{X}_{ri} = \frac{\tilde{X}_{ri}}{\hat{\psi}_r(U_i)}, \tag{2.3}$$

$i = 1, \ldots, n$, $r = 1, \ldots, p$. Those estimators $\hat{\phi}(U_i)$, $\hat{\psi}_r(U_i)$ used in (2.3) are the Nadaraya–Watson estimators, which are defined as

$$\begin{aligned}\hat{\phi}(U_i) &= \frac{n^{-1}\sum_{j=1}^{n} K_h(U_j - U_i)\tilde{Y}_j}{n^{-1}\sum_{j=1}^{n} K_h(U_j - U_i)\bar{\tilde{Y}}},\\ \hat{\psi}_r(U_i) &= \frac{n^{-1}\sum_{j=1}^{n} K_h(U_j - U_i)\tilde{X}_{rj}}{n^{-1}\sum_{j=1}^{n} K_h(U_j - U_i)\bar{\tilde{X}}_r},\end{aligned} \tag{2.4}$$

$i = 1, \ldots, n$, where $\bar{\tilde{Y}} = \frac{1}{n}\sum_{i=1}^{n}\tilde{Y}_i$, $\bar{\tilde{X}}_r = \frac{1}{n}\sum_{i=1}^{n}\tilde{X}_{ri}$. Here $K_h(\cdot) = h^{-1}K(\cdot/h)$, $K(\cdot)$ denotes a kernel density function, and h is a positive-valued bandwidth.

2.2 A least squares estimator

In this subsection, we introduce the estimation procedure for the true parameter $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$. In what follows, $A^{\otimes 2} = AA^{\tau}$, $A^{\star 2} = A^{\tau}A$ any matrix or vector A. The estimators of $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$ are obtained by a least squares estimation method:

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} = \left(\frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{T}}_i\hat{Y}_i\right), \tag{2.5}$$

where $\hat{\mathbf{T}}_i = (\hat{\mathbf{X}}_i^{\tau}, \mathbf{Z}_i)^{\tau}$, $\hat{\mathbf{X}}_i = (\hat{X}_{1i}, \ldots, \hat{X}_{pi})^{\tau}$, $i = 1, \ldots, n$. We now present an asymptotic expression of $(\hat{\boldsymbol{\beta}}^{\tau}, \hat{\boldsymbol{\gamma}}^{\tau})^{\tau}$.

Theorem 2.1. *Under conditions* (A1)–(A6) *given in the Appendix*, *we can have*

$$\begin{aligned}&\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\\ &\quad = (E[\mathbf{T}^{\otimes 2}])^{-1}\frac{1}{n}\sum_{i=1}^{n}(\tilde{Y}_i - Y_i)\frac{E[\mathbf{T}Y]}{E[Y]} + (E[\mathbf{T}^{\otimes 2}])^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{T}_i\varepsilon_i\\ &\qquad - (E[\mathbf{T}^{\otimes 2}])^{-1}\sum_{r=1}^{p}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{E[\mathbf{T}X_r]}{E[X_r]}(\tilde{X}_{ri} - X_{ri})\right)\boldsymbol{\beta}_{0,r} + o_P(n^{-1/2}).\end{aligned} \tag{2.6}$$

Remark 1. In this asymptotic expression, the second term $(E[\mathbf{T}^{\otimes 2}])^{-1}\frac{1}{n} \times \sum_{i=1}^{n}\mathbf{T}_i\varepsilon_i$ is the usual asymptotic expression for the least squares estimator when data are observed without errors. The others in the (2.6) are caused by the distorting functions $\phi(\cdot)$, $\psi_r(\cdot)$'s and the confounding variable U.

The proposed procedure involves the bandwidth, h, to be selected. It is worthwhile to point out that under-smoothing is necessary for estimating $\hat{\mathbf{X}}_i$, $\hat{Y}_i$. This is because, condition (A6) requires the bandwidth h to satisfy $nh^4 \to 0$. To meet condition (A6), Carroll et al. (1997) suggested that the bandwidth h can be chosen as an order of $O(n^{-1/5}) \times n^{-2/15} = O(n^{-1/3})$. Thus, a useful and reasonable choice for h is the rule of thumb suggested by Silverman (1986), namely, $h = \hat{\sigma}_U n^{-1/3}$, where $\hat{\sigma}_U$ is the sample deviation of U. See also in Zhang et al. (2014b); Zhou and Liang (2009).

3 A Hypothesis test for $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$

In the previous section, we discuss the estimation of $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$ rather than testing. Another interesting problem is to evaluate if certain explanatory variables in the parametric components influence the response significantly. As in many important statistical applications, in addition to sample information, we may have some prior information on the regression parametric vector which can be used to improve the parametric estimators. Specially, this study can be used to select unobserved underlying variable $\mathbf{X}$ for distortion measurement errors data. We consider the following linear hypothesis:

$$\mathcal{H}_0 : \mathbf{A}\boldsymbol{\theta}_0 = \mathbf{b} \quad \text{vs} \quad \mathcal{H}_1 : \mathbf{A}\boldsymbol{\theta}_0 \neq \mathbf{b}, \tag{3.1}$$

where $\boldsymbol{\theta}_0 = (\boldsymbol{\beta}_0^\tau, \boldsymbol{\gamma}_0^\tau)^\tau$, $\mathbf{A}$ is a $k \times (p+q)$ matrix of known constants and $\mathbf{b}$ is a k-vector of known constants. We shall also assume that $\text{rank}(\mathbf{A}) = k \leq p+q$.

3.1 A restricted estimator and its asymptotic normality

Under the null hypothesis $\mathcal{H}_0$, the restriction conditions $\mathbf{A}\boldsymbol{\theta}_0 = \mathbf{b}$ should be used to obtain an estimator of $\boldsymbol{\theta}_0$ without losing the restriction information involved in the null hypothesis $\mathcal{H}_0$. The information for regression coefficients involved in this null hypothesis may improve the efficiency of the estimator. Following Wei and Wang (2012), we can construct a restricted estimator by using Lagrange multiplier technique:

$$\mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \sum_{i=1}^{n} [\hat{Y}_i - \hat{\mathbf{T}}_i^\tau \boldsymbol{\theta}]^2 + 2\boldsymbol{\lambda}^\tau (\mathbf{A}\boldsymbol{\theta} - \mathbf{b}), \tag{3.2}$$

where $\boldsymbol{\lambda}$ is a $k \times 1$ vector of the Lagrange multipliers. Differentiating $\mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\lambda})$ with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$, it is easily seen that

$$\begin{cases} \dfrac{\partial \mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\theta}} = -2\displaystyle\sum_{i=1}^{n} \hat{\mathbf{T}}_i [\hat{Y}_i - \hat{\mathbf{T}}_i^\tau \boldsymbol{\theta}] + 2\mathbf{A}^\tau \boldsymbol{\lambda} = 0, \\ \dfrac{\partial \mathcal{S}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = 2(\mathbf{A}\boldsymbol{\theta} - \mathbf{b}) = 0. \end{cases} \tag{3.3}$$

Solving (3.3) with respective to $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$, the restricted least squares estimator of $\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0$ can be obtained as

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_R \\ \hat{\boldsymbol{\gamma}}_R \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} - \left[\sum_{i=1}^{n} \hat{\mathbf{T}}_i^{\otimes 2} \right]^{-1} \mathbf{A}^\tau \left\{ \mathbf{A} \left[\sum_{i=1}^{n} \hat{\mathbf{T}}_i^{\otimes 2} \right]^{-1} \mathbf{A}^\tau \right\}^{-1} \left[\mathbf{A} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\gamma}} \end{pmatrix} - \mathbf{b} \right], \tag{3.4}$$

where the estimator $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}$ used in (3.4) is obtained from (2.5). In the following, we present the asymptotic normality for the restricted estimator $\hat{\boldsymbol{\beta}}_R, \hat{\boldsymbol{\gamma}}_R$.

Theorem 3.1. *Under conditions* (A1)–(A6) *given in the Appendix, we can have that*

$$\sqrt{n}\left(\begin{pmatrix}\hat{\boldsymbol{\beta}}_R\\ \hat{\boldsymbol{\gamma}}_R\end{pmatrix}-\begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\right)\xrightarrow{\mathcal{L}} N\big(\mathbf{0}, \boldsymbol{\Omega}_{\mathbf{A}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Omega}_{\mathbf{A}}^{\tau}\sigma_{\varepsilon}^2+\boldsymbol{\Omega}_{\mathbf{A}}\mathbf{Q}_{\boldsymbol{\beta}_0}\boldsymbol{\Omega}_{\mathbf{A}}^{\tau}\big), \tag{3.5}$$

where $\boldsymbol{\Omega}_{\mathbf{A}}=\mathbf{I}_{p+q}-\boldsymbol{\Gamma}^{-1}\mathbf{A}^{\tau}[\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^{\tau}]^{-1}\mathbf{A}$, $\mathbf{I}_{p+q}$ *is an identity matrix of size* $p+q$, *and*

$$\begin{aligned}\mathbf{Q}_{\boldsymbol{\beta}_0}&=\sum_{l=1}^{p}\sum_{r=1}^{p}\boldsymbol{\beta}_{0,l}\boldsymbol{\beta}_{0,r}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda}_{X_l}\boldsymbol{\Lambda}_{X_r}^{\tau}\boldsymbol{\Gamma}^{-1}\operatorname{Cov}\big(\psi_l(U),\psi_r(U)\big)\frac{E[X_lX_r]}{E[X_l]E[X_r]}\\&\quad-\sum_{l=1}^{p}\boldsymbol{\beta}_{0,l}\boldsymbol{\Gamma}^{-1}[\boldsymbol{\Lambda}_{X_l}\boldsymbol{\Lambda}_Y^{\tau}+\boldsymbol{\Lambda}_Y\boldsymbol{\Lambda}_{X_l}^{\tau}]\boldsymbol{\Gamma}^{-1}\operatorname{Cov}\big(\psi_l(U),\phi(U)\big)\frac{E[X_lY]}{E[X_l]E[Y]}\\&\quad+\boldsymbol{\Gamma}^{-1}\boldsymbol{\Lambda}_Y^{\otimes 2}\boldsymbol{\Gamma}^{-1}\operatorname{Var}\big(\phi(U)\big)\frac{E[Y^2]}{(E[Y])^2}\end{aligned}$$

and $\boldsymbol{\Lambda}_{X_r}=E[\mathbf{T}X_r]$,$r=1,\dots,p$, $\boldsymbol{\Lambda}_Y=E[\mathbf{T}Y]$ *and* $\sigma_{\varepsilon}^2=\operatorname{Var}(\varepsilon)$.

Remark 2. The first term $\boldsymbol{\Omega}_{\mathbf{A}}\boldsymbol{\Gamma}^{-1}\boldsymbol{\Omega}_{\mathbf{A}}^{\tau}\sigma_{\varepsilon}^2$ in the asymptotic variance (3.5) is the usual asymptotic covariance matrix of the restricted least squares estimator under the null hypothesis $\mathcal{H}_0$ when the data can be directly observed, that is, $\phi(\cdot)\equiv 1$ and $\psi_r(\cdot)\equiv 1$. The second term $\boldsymbol{\Omega}_{\mathbf{A}}\mathbf{Q}_{\boldsymbol{\beta}_0}\boldsymbol{\Omega}_{\mathbf{A}}^{\tau}$ is an extra term caused by the distortion in the covariates.

3.2 A test statistic and its asymptotic result

After obtaining the restricted least squares estimator $\hat{\boldsymbol{\beta}}_R, \hat{\boldsymbol{\gamma}}_R$, we can define the residual sum of squares under the null hypothesis $\mathcal{H}_0$:

$$\operatorname{RSS}(\mathcal{H}_0)=\sum_{i=1}^{n}\big[\hat{Y}_i-\hat{\mathbf{X}}_i^{\tau}\hat{\boldsymbol{\beta}}_R-\mathbf{Z}_i^{\tau}\hat{\boldsymbol{\gamma}}_R\big]^2. \tag{3.6}$$

Similarly, we can define the residual sum of squares under the alternative hypothesis:

$$\operatorname{RSS}(\mathcal{H}_1)=\sum_{i=1}^{n}\big[\hat{Y}_i-\hat{\mathbf{X}}_i^{\tau}\hat{\boldsymbol{\beta}}-\mathbf{Z}_i^{\tau}\hat{\boldsymbol{\gamma}}\big]^2. \tag{3.7}$$

Our test statistic for the null hypothesis $\mathcal{H}_0$ is based on the difference between the residual sums of squares under the null and alternative hypotheses:

$$\mathcal{T}_n=\operatorname{RSS}(\mathcal{H}_0)-\operatorname{RSS}(\mathcal{H}_1). \tag{3.8}$$

Intuitively, if the null hypothesis $\mathcal{H}_0$ is true, the value of $\mathcal{T}_n$ is small. If the null hypothesis $\mathcal{H}_0$ is not true, there should be significant difference between $\mathrm{RSS}(\mathcal{H}_0)$ and $\mathrm{RSS}(\mathcal{H}_1)$, and the value of $\mathcal{T}_n$ is large. Large value of $\mathcal{T}_n$ further indicates that the null hypothesis $\mathcal{H}_0$ should be rejected. The following theorem gives the asymptotic distribution of $\mathcal{T}_n$ under the null hypothesis $\mathcal{H}_0$.

Theorem 3.2. *Under conditions* (A1)–(A6) *given in the Appendix, we can have that*:

- *Under the null hypothesis* $\mathcal{H}_0 : \mathbf{A}\boldsymbol{\theta}_0 = \mathbf{b}$,
$$\mathcal{T}_n \xrightarrow{L} \ell_1\chi^2_{1,1} + \ell_2\chi^2_{2,1} + \cdots + \ell_k\chi^2_{k,1},$$
where $\chi^2_{i,1}$, $i = 1, \ldots, k$ *are independent standard chi-squared random variables with one degree of freedom, and* ℓ_i*'s are the eigenvalues of matrix* $(\mathbf{A}[\boldsymbol{\Gamma}^{-1}\sigma^2_\varepsilon + \mathbf{Q}_{\boldsymbol{\beta}_0}]\mathbf{A}^\tau)(\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau)^{-1}$. *Here, matrices* $\boldsymbol{\Gamma}$ *and* $\mathbf{Q}_{\boldsymbol{\beta}_0}$ *are defined in Theorem* 3.1.
- *Under the local alternative hypothesis* $\mathcal{H}_{1n} : \mathbf{A}\boldsymbol{\theta}_0 = \mathbf{b} + n^{-1/2}\mathbf{c}$, *the test statistic* $\mathcal{T}_n$ *follows a generalised chi-squared distribution* (Sheil and O'Muircheartaigh, 1977), *namely*,
$$\mathcal{T}_n \xrightarrow{L} (\mathfrak{M} + \boldsymbol{\Gamma}^{-1/2}\mathbf{A}^\tau[\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau]^{-1}\mathbf{c})^\tau(\mathfrak{M} + \boldsymbol{\Gamma}^{-1/2}\mathbf{A}^\tau[\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau]^{-1}\mathbf{c}), \quad (3.9)$$
where $\mathfrak{M}$ *is a random vector follows a multivariate normal distribution* $N(\mathbf{0}, \mathbf{D})$ *with*
$$\mathbf{D} = \boldsymbol{\Gamma}^{-1/2}\mathbf{A}^\tau(\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau)^{-1}\mathbf{A}(\boldsymbol{\Gamma}^{-1}\sigma^2_\varepsilon + \mathbf{Q}_{\boldsymbol{\beta}_0})\mathbf{A}^\tau(\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau)^{-1}\mathbf{A}\boldsymbol{\Gamma}^{-1/2}.$$

3.3 A wild bootstrap procedure

To apply Theorem 3.2 under the null hypothesis $\mathcal{H}_0$, one way is to estimate the unknown weights ℓ_i, $1 \le i \le k$ by estimating $\boldsymbol{\Gamma}$ and $\mathbf{Q}_{\boldsymbol{\beta}_0}$ consistently, then one can use the distribution of $\hat{\ell}_1\chi^2_{1,1} + \hat{\ell}_2\chi^2_{2,1} + \cdots + \hat{\ell}_k\chi^2_{k,1}$ or Welch–Satterthwaite approximation to define the p-values. However, as the asymptotic variances obtained in Theorem 3.1 are very complex. Especially for $\mathbf{Q}_{\boldsymbol{\beta}_0}$, its estimator may not be precise in finite samples. To overcome this problem, we suggest to apply a wild bootstrap (Stute, González Manteiga and Presedo Quindimil, 1998; Escanciano, 2006; Wu, 1986) technique to mimic the distribution of the test statistic $\mathcal{T}_n$ under the null hypothesis $\mathcal{H}_0$.

The procedure for calculating the critical values based on the bootstrap test statistic is proposed as follows:

Step 1. Compute the test statistic $\mathcal{T}_n$ defined in (3.8).

Step 2. Generate N times i.i.d. variables ς_{ib}, $i = 1, \ldots, n$, $b = 1, \ldots, B$ with a Bernoulli distribution which respectively take values at $\frac{1\mp\sqrt{5}}{2}$ with probability $\frac{5\pm\sqrt{5}}{10}$. Let $\hat{\varepsilon}_i = \hat{Y}_i - \hat{\mathbf{X}}_i^\tau\hat{\boldsymbol{\beta}} - \mathbf{Z}_i^\tau\hat{\boldsymbol{\gamma}}$ and $\hat{\varepsilon}_i^{(b)} = \hat{\varepsilon}_i\varsigma_{ib}$, and further obtain
$$\hat{Y}_i^{(b)} = \hat{\mathbf{X}}_i^\tau\hat{\boldsymbol{\beta}} + \mathbf{Z}_i^\tau\hat{\boldsymbol{\gamma}} + \hat{\varepsilon}_i^{(b)}.$$

Step 3. For each b, using bootstraps $(\hat{Y}_i^{(b)}, \hat{\mathbf{X}}_i, \mathbf{Z}_i)$, we re-calculate the bootstrap estimators $\hat{\boldsymbol{\beta}}^{(b)}, \hat{\boldsymbol{\gamma}}^{(b)}$ and $\hat{\boldsymbol{\beta}}_R^{(b)}, \hat{\boldsymbol{\gamma}}_R^{(b)}$ by

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}^{(b)} \\ \hat{\boldsymbol{\gamma}}^{(b)} \end{pmatrix} = \left(\frac{1}{n}\sum_{i=1}^{n} \hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} \hat{\mathbf{T}}_i \hat{Y}_i^{(b)}\right),$$

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_R^{(b)} \\ \hat{\boldsymbol{\gamma}}_R^{(b)} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\beta}}^{(b)} \\ \hat{\boldsymbol{\gamma}}^{(b)} \end{pmatrix} - \left[\sum_{i=1}^{n} \hat{\mathbf{T}}_i^{\otimes 2}\right]^{-1} \mathbf{A}^{\tau} \left\{\mathbf{A}\left[\sum_{i=1}^{n} \hat{\mathbf{T}}_i^{\otimes 2}\right]^{-1} \mathbf{A}^{\tau}\right\}^{-1} \left[\mathbf{A}\begin{pmatrix} \hat{\boldsymbol{\beta}}^{(b)} \\ \hat{\boldsymbol{\gamma}}^{(b)} \end{pmatrix} - \mathbf{b}\right],$$

and we then calculate the bootstrap test statistic

$$\mathcal{T}_n^{(b)} = \mathrm{RSS}^{(b)}(\mathcal{H}_0) - \mathrm{RSS}^{(b)}(\mathcal{H}_1),$$

$$\mathrm{RSS}^{(b)}(\mathcal{H}_0) = \sum_{i=1}^{n} [\hat{Y}_i^{(b)} - \hat{\mathbf{X}}_i^{\tau} \hat{\boldsymbol{\beta}}_R^{(b)} - \mathbf{Z}_i^{\tau} \hat{\boldsymbol{\gamma}}_R^{(b)}]^2,$$

$$\mathrm{RSS}^{(b)}(\mathcal{H}_1) = \sum_{i=1}^{n} [\hat{Y}_i^{(b)} - \hat{\mathbf{X}}_i^{\tau} \hat{\boldsymbol{\beta}}^{(b)} - \mathbf{Z}_i^{\tau} \hat{\boldsymbol{\gamma}}^{(b)}]^2.$$

Step 4. We calculate the $1 - \kappa$ quantile of the bootstrap test statistics $\mathcal{T}_n^{(b)}$, $b = 1, \dots, N$ as the κ-level critical value.

4 A simulation study

In this section, we present some numerical results to evaluate the performance of our proposed estimators. In the following, the Epanechnikov kernel $K(t) = 0.75(1 - t^2)^+$ is used. As noted in Remark 1, the rule of thumb bandwidth $h = \hat{\sigma}_U n^{-1/3}$ is used. In Example 1, a comparison is made between the direct-plug-in (DPI) estimation method and the local linear approximation (LLA) method proposed in Şentürk and Nguyen (2006). The simulation results is reported in Table 1. In Example 2, a comparison is made between the usual least squares estimator $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}$ and the restricted estimator $\hat{\boldsymbol{\beta}}_R, \hat{\boldsymbol{\gamma}}_R$. The simulation results in reported in Table 2. We also investigate the performance of the test statistic $\mathcal{T}_n$ and its bootstrap estimators. The values of power calculations are reported in Table 3.

Example 1. We generate 500 realizations from the following model (4.1) and the sample size is chosen as $n = 500$ and 1000:

$$Y = \boldsymbol{\beta}_{0,1} X_1 + \boldsymbol{\beta}_{0,2} X_2 + \boldsymbol{\gamma}_{0,1} Z_1 + \varepsilon, \tag{4.1}$$

where $(\boldsymbol{\beta}_{0,1}, \boldsymbol{\beta}_{0,2}, \boldsymbol{\gamma}_{0,1}) = (1, 3, -2)$. Those predictors are independently generated from normal distributions: $X_1 \sim N(2, 1.2^2)$, $X_2 \sim N(2, 0.5^2)$ and $Z_1 \sim$

Table 1 *Simulation results of Example 1. "Mean" is the simulation mean; "SD" is the standard deviation*

	Sample size			$\beta_{0,1}=1$	$\beta_{0,2}=3$	$\gamma_{0,1}=-2$
$\sigma=0.50$	DPI	$n=500$	Mean	0.9858	2.9835	−1.9843
			SD	0.0352	0.0692	0.1042
			MSE	0.0014	0.0050	0.0111
		$n=1000$	Mean	0.9900	2.9888	−1.9815
			SD	0.0248	0.0452	0.0818
			MSE	0.0007	0.0022	0.0070
	LLA	$n=500$	Mean	0.9911	2.9855	−2.0017
			SD	0.0386	0.0715	0.1112
			MSE	0.0016	0.0053	0.0123
		$n=1000$	Mean	0.9925	2.9889	−1.9964
			SD	0.0259	0.0471	0.0894
			MSE	0.0007	0.0023	0.0080
$\sigma=1.00$	DPI	$n=500$	Mean	0.9883	2.9837	−1.9920
			SD	0.0609	0.1336	0.2056
			MSE	0.0038	0.0180	0.0422
		$n=1000$	Mean	0.9915	2.9879	−1.9939
			SD	0.0410	0.0889	0.1369
			MSE	0.0018	0.0080	0.0187
	LLA	$n=500$	Mean	0.9878	2.9727	−2.0181
			SD	0.0624	0.1348	0.2183
			MSE	0.0040	0.0188	0.0478
		$n=1000$	Mean	0.9901	2.9769	−2.0093
			SD	0.0455	0.0933	0.1487
			MSE	0.0022	0.0092	0.0221

$N(0.5, 0.25^2)$, and the covariance matrix of (X_1, X_2, Z_1) is chosen as

$$\begin{pmatrix} 1 & -0.5 & 0.25 \\ -0.5 & 1 & -0.5 \\ 0.25 & -0.5 & 1 \end{pmatrix}.$$

The model error ε is independent with (X_1, X_2, Z_1), $\varepsilon \sim N(0, \sigma^2)$ with $\sigma = 0.5$ and 1.0. The confounding covariate U is generated from the Uniform (0, 1) distribution, independent with $(X_1, X_2, Z_1, \varepsilon)$. The distorting functions for X_1 and X_2 are chosen as $\psi_1(U) = 1 + 0.3\sin(2\pi U)$ and $\psi_1(U) = 1 + 3(U - 0.5)^3$, respectively. The distorting function for Y is chosen as $\phi(U) = 0.5 + U$.

In Table 1, we report the simulation results of the sample mean (Mean), the sample standard deviation (SD) and the sample mean squared error (MSE). We can see that both the DPI estimator and LLA estimator are close to the true value. As the sample size increases or σ decreases, their means are generally closer to the true values, the SD and MSE of both the estimators decrease, and the biases of DPI estimator is much smaller than those of LLA estimator. In addition, by com-

Table 2 *Simulation results of Example 2. Restricted condition (I):* $\boldsymbol{\beta}_{0,1} + \boldsymbol{\beta}_{0,2} + 2\boldsymbol{\gamma}_{0,1} = 0$; *Restricted condition (II):* $-\boldsymbol{\beta}_{0,1} + \boldsymbol{\beta}_{0,2} + \boldsymbol{\gamma}_{0,1} = 0$. *"Mean" is the simulation mean; "SD" is the standard deviation*

	Sample size			$\boldsymbol{\beta}_{0,1} = 1$	$\boldsymbol{\beta}_{0,2} = 3$	$\boldsymbol{\gamma}_{0,1} = -2$
$\sigma = 0.50$	(I)	$n = 500$	Mean	0.9826	2.9778	−1.9802
			SD	0.0388	0.0611	0.0469
			MSE	0.0018	0.0042	0.0026
		$n = 800$	Mean	0.9902	2.9841	−1.9872
			SD	0.0308	0.0510	0.0384
			MSE	0.0010	0.0028	0.0016
		$n = 1000$	Mean	0.9939	2.9910	−1.9924
			SD	0.0263	0.0436	0.0325
			MSE	0.0007	0.0020	0.0011
	(II)	$n = 500$	Mean	0.9851	2.9760	−1.9910
			SD	0.0365	0.0593	0.0352
			MSE	0.0015	0.0041	0.0013
		$n = 800$	Mean	0.9913	2.9833	−1.9920
			SD	0.0292	0.0492	0.0292
			MSE	0.0009	0.0027	0.0009
		$n = 1000$	Mean	0.9947	2.9904	−1.9957
			SD	0.0247	0.0422	0.0265
			MSE	0.0006	0.0018	0.0007
$\sigma = 0.75$	(I)	$n = 500$	Mean	0.9841	2.9776	−1.9809
			SD	0.0497	0.0841	0.0633
			MSE	0.0027	0.0076	0.0044
		$n = 800$	Mean	0.9864	2.9818	−1.9841
			SD	0.0402	0.0669	0.0497
			MSE	0.0018	0.0048	0.0027
		$n = 1000$	Mean	0.9919	2.9889	−1.9904
			SD	0.0361	0.0597	0.0451
			MSE	0.0013	0.0036	0.0021
	(II)	$n = 500$	Mean	0.9859	2.9764	−1.9905
			SD	0.0477	0.0813	0.0476
			MSE	0.0024	0.0071	0.0023
		$n = 800$	Mean	0.9872	2.9817	−1.9945
			SD	0.0385	0.0640	0.0401
			MSE	0.0016	0.0044	0.0016
		$n = 1000$	Mean	0.9926	2.9886	−1.9960
			SD	0.0348	0.0573	0.0342
			MSE	0.0012	0.0034	0.0012

Table 2 (*Continued*)

	Sample size			$\beta_{0,1} = 1$	$\beta_{0,2} = 3$	$\gamma_{0,1} = -2$
$\sigma = 1.00$	(I)	$n = 500$	Mean	0.9846	2.9881	−1.9864
			SD	0.0659	0.1108	0.0819
			MSE	0.0045	0.0124	0.0069
		$n = 800$	Mean	0.9924	2.9858	−1.9891
			SD	0.0499	0.0883	0.0646
			MSE	0.0025	0.0080	0.0043
		$n = 1000$	Mean	0.9969	2.9958	−1.9964
			SD	0.0468	0.0763	0.0579
			MSE	0.0022	0.0058	0.0034
	(II)	$n = 500$	Mean	0.9872	2.9864	−1.9993
			SD	0.0630	0.1060	0.0672
			MSE	0.0041	0.0114	0.0045
		$n = 800$	Mean	0.9931	2.9852	−1.9921
			SD	0.0481	0.0847	0.0530
			MSE	0.0023	0.0074	0.0029
		$n = 1000$	Mean	0.9975	2.9953	−1.9977
			SD	0.0451	0.0734	0.0438
			MSE	0.0020	0.0053	0.0019

Table 3 *The simulation results of the power calculations*

Significant level	0.01	0.025	0.05	0.10
$\mathbf{c} = 0.000$	0.0094	0.0243	0.0517	0.0975
$\mathbf{c} = -0.100$	0.0488	0.1057	0.1220	0.1951
$\mathbf{c} = -0.200$	0.1789	0.2114	0.2846	0.3902
$\mathbf{c} = -0.300$	0.2927	0.4228	0.5203	0.6504
$\mathbf{c} = -0.400$	0.6260	0.7236	0.7805	0.8374
$\mathbf{c} = -0.500$	0.8537	0.8943	0.9187	0.9512
$\mathbf{c} = -0.600$	0.9837	0.9919	1.0000	1.0000
$\mathbf{c} = 0.100$	0.0732	0.1138	0.1789	0.2602
$\mathbf{c} = 0.200$	0.1870	0.2602	0.3496	0.4959
$\mathbf{c} = 0.300$	0.4472	0.5691	0.6423	0.7236
$\mathbf{c} = 0.400$	0.7236	0.8049	0.8780	0.9268
$\mathbf{c} = 0.500$	0.9268	0.9674	0.9919	1.0000
$\mathbf{c} = 0.600$	1.0000	1.0000	1.0000	1.0000

paring the MSE, we can see that our DPI estimator is slightly more efficient than LAP estimator, which suggests that DPI estimation procedure is indeed worthy of recommendation.

Example 2. We first consider the restricted estimator under two restricted conditions $\mathbf{A}^{[1]} = (1, 1, 2)^\tau$ (i.e., $\boldsymbol{\beta}_{0,1} + \boldsymbol{\beta}_{0,2} + 2\boldsymbol{\gamma}_{0,1} = 0$) and $\mathbf{A}^{[2]} = (-1, 1, 1)^\tau$ (i.e., $-\boldsymbol{\beta}_{0,1} + \boldsymbol{\beta}_{0,2} + \boldsymbol{\gamma}_{0,1} = 0$). The covariates $X_1, X_2, Z_1, \varepsilon, U$ and the choices of distorting functions are the same as Example 1. We generate 500 realizations from the following model (4.1) and the sample size is chosen as $n = 500$, 800 and 1000, and the σ for the error term ε is chosen as $\sigma = 0.25, 0.5$ and 1.0 in this example.

In Table 2, we can see that the values of MSE for the restricted estimator of $\boldsymbol{\beta}_{0,1}$ is slightly larger than those obtained in Table 1. However, the restricted estimator of $\boldsymbol{\beta}_{0,2}$ is generally smaller than those in Table 1, the restricted estimator $\hat{\boldsymbol{\gamma}}_R$ are much smaller than those in Table 1, which indicates both $\mathbf{A}^{[1]}$ and $\mathbf{A}^{[2]}$ can improve the estimation efficiency for $\boldsymbol{\beta}_{0,2}, \boldsymbol{\gamma}_{0,1}$ without losing much estimation efficiency for $\boldsymbol{\beta}_{0,1}$.

Next, we consider a hypothesis test problem by considering

$$\mathcal{H}_0 : \boldsymbol{\gamma}_{0,1} = 0 \quad \text{vs} \quad \mathcal{H}_1 : \boldsymbol{\gamma}_{0,1} = \mathbf{c}, \tag{4.2}$$

where $\mathbf{c} = 0, \pm 0.1, \pm 0.2, \pm 0.3, \pm 0.4, \pm 0.5, \pm 0.6$. The parameter $(\boldsymbol{\beta}_{0,1}, \boldsymbol{\beta}_{0,2})$ is set to be $(1, 3)$. In this test problem (4.2), it is noted that the alternative hypothesis $\mathcal{H}_1$ becomes the null hypothesis $\mathcal{H}_0$ when $\mathbf{c} = 0$. 1000 bootstrap samples are generated in each simulation to calculate its powers. In Table 3, all empirical levels obtained by the bootstrap test statistics are close to the four nominal levels when $\mathbf{c} = 0$, which indicates that the bootstrap statistic $\mathcal{T}_n^{(b)}$ gives proper Type I errors. As the absolute value of $\mathbf{c}$ increases, the powers increases rapidly and increases to one.

5 A real data analysis

We apply our method to analyse the Pima Indian diabetes data for an illustration. The dataset is available on the web site http://www.ics.uci.edu/~mlearn/databases/. This data has been analysed by Şentürk and Nguyen (2009). They suggested the body mass index (BMI) as the potential confounding variable and further investigated a linear regression model between the unobserved covariate "plasma glucose concentration" (Y), the covariate "diastolic blood pressure" (X), the observed covariate "triceps skin fold thickness" (Z_1) and the observed covariate "age" (Z_2):

$$Y = \boldsymbol{\beta}_{0,0} + \boldsymbol{\beta}_{0,1} X + \boldsymbol{\gamma}_{0,1} Z_1 + \boldsymbol{\gamma}_{0,2} Z_2 + \varepsilon. \tag{5.1}$$

We first present the patterns of $\hat{\phi}(u)$ and $\hat{\psi}(u)$ in Figure 1 by using the local linear smoothing procedure (Fan and Gijbels, 1996). Those two plots indicate that $\phi(u)$ and $\psi(u)$ are not constants (i.e., $\phi(u) \not\equiv 1$, $\psi(u) \not\equiv 1$), which suggests that the variable "BMI" is the potential confounder for the response Y and covariate X. Next, we use our proposed direct-plug-in estimation procedure and obtain that $(\hat{\boldsymbol{\beta}}_{0,0}, \hat{\boldsymbol{\beta}}_{0,1}, \hat{\boldsymbol{\gamma}}_{0,1}, \hat{\boldsymbol{\gamma}}_{0,2}) = (84.0284, 0.2118, 0.0642, 0.6375)$. 1000 bootstrap samples are conducted for the testing problem, the results are presented in Table 4.

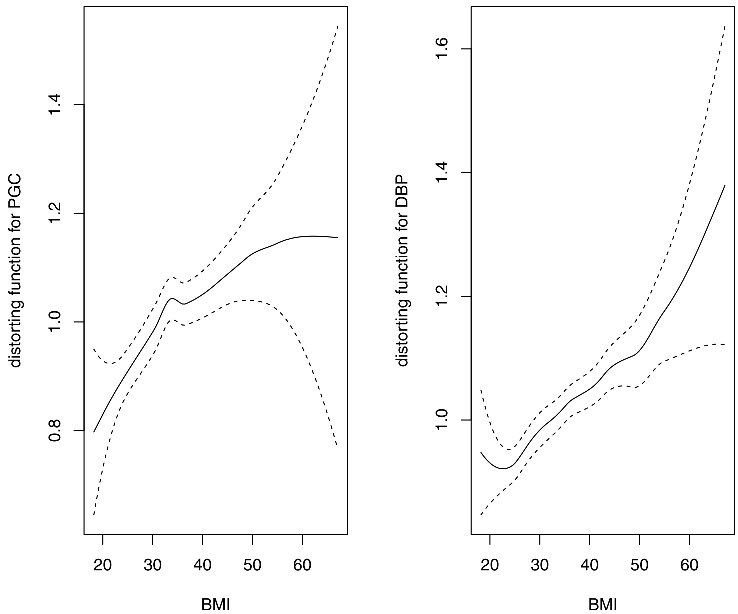

Figure 1 *Plots for distorting functions.*

Table 4 *The results of the hypothesis testing for the Pima Indian diabetes data*

Null hypothesis	Alternative hypothesis	p-values
$\mathcal{H}_0^{[1]} : \boldsymbol{\beta}_{0,1} = 0.2$	$\mathcal{H}_1^{[1]}, \boldsymbol{\beta}_{0,1} \neq 0.2$	0.922
$\mathcal{H}_0^{[2]} : \boldsymbol{\gamma}_{0,2} = 0.6$	$\mathcal{H}_1^{[2]}, \boldsymbol{\gamma}_{0,2} \neq 0.6$	0.790
$\mathcal{H}_0^{[3]} : \boldsymbol{\gamma}_{0,2} - 3\boldsymbol{\beta}_{0,1} = 0.0$	$\mathcal{H}_1^{[3]}, \boldsymbol{\gamma}_{0,2} - 3\boldsymbol{\beta}_1 \neq 0.0$	0.994
$\mathcal{H}_0^{[4]} : \boldsymbol{\gamma}_{0,1} = 0.0$	$\mathcal{H}_1^{[4]}, \boldsymbol{\gamma}_{0,1} \neq 0.0$	0.636

Based on the estimate $(\hat{\boldsymbol{\beta}}_{0,0}, \hat{\boldsymbol{\beta}}_{0,1}, \hat{\boldsymbol{\gamma}}_{0,1}, \hat{\boldsymbol{\gamma}}_{0,2})$, we are interested in the null hypothesis $\mathcal{H}_0^{[s]}$, $s = 1, 2, 3, 4$ in Table 4. From this table, we cannot reject the hypothesis $\mathcal{H}_0^{[1]} : \boldsymbol{\beta}_{0,1} = 0.2$ as the associated p-value is 0.922, if we use the significant level $\alpha = 0.05$. Similarly, we cannot reject the hypothesis $\mathcal{H}_0^{[3]} : \boldsymbol{\gamma}_{0,2} = 0.6$ either. We consider the hypothesis $\mathcal{H}_0^{[4]} : \boldsymbol{\gamma}_{0,2} - 3\boldsymbol{\beta}_{0,1} = 0.0$, the associated p-value is 0.994, which indicate that the relationship in the null hypothesis $\mathcal{H}_0^{[4]}$ can be true. For the null hypothesis $\mathcal{H}_0^{[4]}$, we cannot reject the null hypothesis $\mathcal{H}_0^{[4]}$, as its p-value is 0.636, much larger than critical value 0.05. This implies that the variable "Triceps skin fold thickness" may be excluded from model (5.1).

Appendix

In this appendix, we present the conditions, prepare several preliminary lemmas, and give the proofs of the main results.

A.1 Conditions

(A1) The matrix $\mathbf{\Gamma}$ defined in Theorem 3.1 is positive-definite, $0 < E[\varepsilon^2] = \sigma_\varepsilon^2 < +\infty$.

(A2) The density function $f_U(u)$ of the confounding variable U is bounded away from 0 and satisfies the Lipschitz condition of order 1 on $\mathfrak{U}$, where $\mathfrak{U}$ stands for a compact support set of U. Moreover, $\inf_{u\in\mathfrak{U}} f_U(u) \geq c_0$, $0 < c_0 < +\infty$.

(A3) $\phi(\cdot)$, $\psi_r(\cdot)$ have three bounded and continuous derivatives. Moreover, $\phi(u)$ and $\psi_r(u)$ are non-zero on $\mathfrak{U}$.

(A4) $E[Y]$ and $E[X_s]$ are bounded away from 0; moreover, $E[|Y|^3] < \infty$, $E[|X_s|^3] < \infty$, $s = 1, \ldots, p$.

(A5) The kernel function $K(\cdot)$ is a density function with a compact support, symmetric about zero, satisfying a Lipschitz condition and having bounded derivatives. Furthermore, $\int_{-\infty}^{\infty} u^2 K(u)\,du \neq 0$, $\int_{-\infty}^{\infty} |u|^j K(u)\,du < \infty$, for $j = 1, 2, 3$.

(A6) As $n \to \infty$, the bandwidth h satisfies: $\frac{\log^2 n}{nh^2} \to 0$, $nh^2 \to \infty$, and $nh^4 \to 0$.

A.2 Technical lemmas

Lemma A.1. *Suppose $E(W|U = u) = m_W(u)$ and its derivatives up to second order are bounded for all $u \in \mathfrak{U}$, where $\mathfrak{U}$ is defined in condition* (A1), *and that $E|W|^3$ exists, and $\sup_{u\in\mathfrak{U}} \int |w|^{s_0} f(u, w)\,dw < \infty$, where $f(u, w)$ is the joint density of $(U, W)^\tau$. Let $(U_i, W_i)^\top$, $i = 1, 2, \ldots, n$ be independent and identically distributed* (*i.i.d.*) *samples from $(U, W)^\tau$. If* (A1)–(A3) *hold, and $n^{2\delta-1}h \to \infty$ for $\delta < 1 - s_0^{-1}$, then*

$$\sup_{u\in\mathfrak{U}} \left| \frac{1}{n} \sum_{i=1}^{n} K_h(U_i - u) W_i - f_U(u) m_W(u) - \frac{h^2}{2} [f_U(u) m_W(u)]'' \int K(u) u^2\,du \right| = O(\tau_{n,h}), \qquad a.s.,$$

where let $\tau_{n,h} = h^3 + \sqrt{\frac{\log n}{nh}}$.

Proof. Lemma A.1 can be immediately proved from the result obtained by Mack and Silverman (1982). □

Lemma A.2. *Suppose that conditions* (A1)–(A6) *hold. Let* $S(x)$ *be a continuous function satisfying* $ES^2(X) < \infty$. *Then, for* $l = 1, \dots, p$,

$$n^{-1}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)S(X_i) = n^{-1}\sum_{i=1}^{n}(\tilde{Y}_i - Y_i)\frac{E[YS(X)]}{EY} + o_P(n^{-1/2}), \tag{A.2}$$

$$n^{-1}\sum_{i=1}^{n}(\hat{X}_{li} - X_{li})S(X_i) = n^{-1}\sum_{i=1}^{n}(\tilde{X}_{li} - X_{li})\frac{E[X_l S(X)]}{EX_l} + o_P(n^{-1/2}). \tag{A.3}$$

Proof. Lemma A.2 is the direct result of Lemma B.2 in Zhang, Zhu and Liang (2012a). □

A.3 Proofs of Theorems 2.1, 3.1 and 3.2

Proof of Theorem 2.1. *Step* 1. Denote $\kappa_n = h^2 + \sqrt{\frac{\log n}{nh}}$.

Using Lemma A.1, it is easily seen that

$$\sup_{u\in\mathfrak{U}}\left|\frac{1}{n}\sum_{i=1}^{n}K_h(U_i - u) - f_U(u)\right| = O_P(\kappa_n), \qquad \text{a.s.,} \tag{A.4}$$

$$\sup_{u\in\mathfrak{U}}\left|\frac{1}{n}\sum_{i=1}^{n}K_h(U_i - u)\tilde{Y}_i - \phi(u)f_U(u)E[Y]\right| = O_P(\kappa_n), \qquad \text{a.s.} \tag{A.5}$$

Together with (A.4) and condition (A2), we can have that

$$\begin{aligned}\inf_{u\in\mathfrak{U}}\left|\frac{1}{n}\sum_{i=1}^{n}K_h(U_i - u)\right| &\geq \inf_{u\in\mathfrak{U}} f_U(u) - \sup_{u\in\mathfrak{U}}\left|\frac{1}{n}\sum_{i=1}^{n}K_h(U_i - u) - f_U(u)\right| \\ &\geq c_0 + O_P(\kappa_n).\end{aligned} \tag{A.6}$$

Then, using (A.4), (A.6), $\bar{\tilde{Y}} - E[Y] = O_P(n^{-1/2})$, we can have that

$$\begin{aligned}&\sup_{u\in\mathfrak{U}}|\hat{\phi}(u) - \phi(u)| \\ &\quad\leq \frac{\sup_{u\in\mathfrak{U}}|1/n\sum_{i=1}^{n}K_h(U_i - u)\tilde{Y}_i - 1/n\sum_{i=1}^{n}K_h(U_i - u)\phi(u)\bar{\tilde{Y}}|}{\inf_{u\in\mathfrak{U}}|1/n\sum_{i=1}^{n}K_h(U_i - u)\bar{\tilde{Y}}|} \\ &\quad\leq \frac{\sup_{u\in\mathfrak{U}}|1/n\sum_{i=1}^{n}K_h(U_i - u)\tilde{Y}_i - \phi(u)f_U(u)E[Y]|}{c_0 + O_P(\kappa_n)} \\ &\qquad + \frac{\sup_{u\in\mathfrak{U}}|\phi(u)f_U(u)||E[Y] - \bar{\tilde{Y}}|}{c_0 + O_P(\kappa_n)} \\ &\qquad + \frac{\sup_{u\in\mathfrak{U}}|\phi(u)|\sup_{u\in\mathfrak{U}}|1/n\sum_{i=1}^{n}K_h(U_i - u) - f_U(u)|}{c_0 + O_P(\kappa_n)}|\bar{\tilde{Y}}| \\ &\quad= O_P(\kappa_n + n^{-1/2}).\end{aligned} \tag{A.7}$$

Similar to (A.7), we can show that

$$\sup_{u\in\mathfrak{U}}\left|\hat{\psi}_r(u)-\psi_r(u)\right| = O_P\left(\kappa_n + n^{-1/2}\right), \tag{A.8}$$

$r = 1, \ldots, p$.

Step 2. Note that

$$\begin{aligned}
&\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix} \\
&\quad= \left(\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i\left[\hat{Y}_i - \hat{\mathbf{X}}_i^\tau \boldsymbol{\beta}_0 - \mathbf{Z}_i^\tau \boldsymbol{\gamma}_0\right]\right) \\
&\quad= \left(\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1}\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i(\hat{Y}_i - Y_i) \\
&\qquad+ \left(\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1}\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i\varepsilon_i \\
&\qquad+ \left(\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right)^{-1}\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i(\mathbf{X}_i - \hat{\mathbf{X}}_i)^\tau \boldsymbol{\beta}_0.
\end{aligned}\tag{A.9}$$

Using (A.7)–(A.8) and Lemma A.1, we have

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \hat{X}_{ri}(\hat{Y}_i - Y_i) \\
&\quad= \frac{1}{n}\sum_{i=1}^n X_{ri}(\hat{Y}_i - Y_i) \\
&\qquad+ \frac{1}{n}\sum_{i=1}^n X_{ri}Y_i\frac{(\hat{\psi}_r(U_i)-\psi_r(U_i))(\hat{\phi}(U_i)-\phi(U_i))}{\phi(U_i)\psi_r(U_i)} \\
&\quad= \frac{1}{n}\sum_{i=1}^n (\tilde{Y}_i - Y_i)\frac{E[X_rY]}{E[Y]} + o_P\left(n^{-1/2}\right) + O_P\left(\kappa_n^2 + n^{-1}\right).
\end{aligned}\tag{A.10}$$

As $nh^8 \to 0$, $\frac{\log^2 n}{nh^2} \to 0$, it yields that $\kappa_n^2 = o(n^{-1/2})$. Using this fact, similar to (A.10), we can have that

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i(\hat{Y}_i - Y_i) &= \frac{1}{n}\sum_{i=1}^n \mathbf{T}_i(\hat{Y}_i - Y_i) + \frac{1}{n}\sum_{i=1}^n \begin{pmatrix}\hat{\mathbf{X}}_i - \mathbf{X}_i\\ \mathbf{0}\end{pmatrix}(\hat{Y}_i - Y_i) \\
&= \frac{1}{n}\sum_{i=1}^n (\tilde{Y}_i - Y_i)\frac{E[\mathbf{T}Y]}{E[Y]} + o_P\left(n^{-1/2}\right),
\end{aligned}$$

where $E[\mathbf{T}Y] = (E[X_1 Y], \dots, E[X_p Y], E[Z_1 Y], \dots, E[Z_q Y])^\tau$, and

$$\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i(\hat{\mathbf{X}}_i - \mathbf{X}_i)^\tau \boldsymbol{\beta}_0 = \sum_{r=1}^p \left(\frac{1}{n}\sum_{i=1}^n \frac{E[\mathbf{T}X_r]}{E[X_r]}(\tilde{X}_{ri} - X_{ri})\right)\boldsymbol{\beta}_{0,r} + o_P(n^{-1/2}).$$

By using $E[\varepsilon \mathbf{T}] = 0$, we can have that $\frac{1}{n}\sum_{i=1}^n (\hat{\mathbf{T}}_i - \mathbf{T}_i)\varepsilon_i = o_P(n^{-1/2})$. Moreover, similar to (A.10), it is easily seen that $\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2} = \frac{1}{n}\sum_{i=1}^n \mathbf{T}_i^{\otimes 2} + o_P(1) = E[\mathbf{T}^{\otimes 2}] + o_P(1)$. As a result, (A.9) can be represented as

$$\begin{aligned}
&\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\\
&\quad = \left(E[\mathbf{T}^{\otimes 2}]\right)^{-1}\frac{1}{n}\sum_{i=1}^n (\tilde{Y}_i - Y_i)\frac{E[\mathbf{T}Y]}{E[Y]} + \left(E[\mathbf{T}^{\otimes 2}]\right)^{-1}\frac{1}{n}\sum_{i=1}^n \mathbf{T}_i\varepsilon_i\\
&\qquad - \left(E[\mathbf{T}^{\otimes 2}]\right)^{-1}\sum_{r=1}^p \left(\frac{1}{n}\sum_{i=1}^n \frac{E[\mathbf{T}X_r]}{E[X_r]}(\tilde{X}_{ri} - X_{ri})\right)\boldsymbol{\beta}_{0,r}\\
&\qquad + o_P(n^{-1/2}).
\end{aligned} \tag{A.11}$$

We complete the proof of Theorem 2.1. □

Proof of Theorem 3.1. Using (3.5) and the null hypothesis $\mathcal{H}_0 : \mathbf{A}\boldsymbol{\theta}_0 = \mathbf{b}$, we can have that

$$\begin{aligned}
&\begin{pmatrix}\hat{\boldsymbol{\beta}}_R\\ \hat{\boldsymbol{\gamma}}_R\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\\
&\quad = \left\{\mathbf{I}_{p+q} - \left[\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right]^{-1}\mathbf{A}^\tau\left\{\mathbf{A}\left[\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\right]^{-1}\mathbf{A}^\tau\right\}^{-1}\mathbf{A}\right\}\\
&\qquad \times \left\{\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\right\}\\
&\quad = \{\mathbf{I}_{p+q} - \boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol{\Gamma}^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{A}\}\left\{\begin{pmatrix}\hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}\end{pmatrix} - \begin{pmatrix}\boldsymbol{\beta}_0\\ \boldsymbol{\gamma}_0\end{pmatrix}\right\}\\
&\qquad + o_P(n^{-1/2}).
\end{aligned} \tag{A.12}$$

Together with the asymptotic expression (A.11), we complete the proof of Theorem 3.1. □

Proof of Theorem 3.2. *Step* 3.1. Under the null hypothesis $\mathcal{H}_0 : \mathbf{A}\boldsymbol{\theta}_0 = b$, and Using the fact that $\sum_{i=1}^n [\hat{Y}_i - \hat{\mathbf{X}}_i^\tau \hat{\boldsymbol{\beta}} - \mathbf{Z}_i^\tau \hat{\boldsymbol{\gamma}}]\binom{\hat{\mathbf{X}}_i}{\mathbf{Z}_i} = 0$, we can have that

$$\mathcal{T}_n = \begin{pmatrix}\hat{\boldsymbol{\beta}}_R - \hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}_R - \hat{\boldsymbol{\gamma}}\end{pmatrix}^\tau \sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2}\begin{pmatrix}\hat{\boldsymbol{\beta}}_R - \hat{\boldsymbol{\beta}}\\ \hat{\boldsymbol{\gamma}}_R - \hat{\boldsymbol{\gamma}}\end{pmatrix}. \tag{A.13}$$

Using (A.12), recalling the definition of $\mathbf{\Omega_A}$ used in Theorem 3.1, we know that

$$\begin{pmatrix} \hat{\boldsymbol\beta}_R - \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma}_R - \hat{\boldsymbol\gamma} \end{pmatrix} = -\{\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{A}\}\left\{\begin{pmatrix} \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma} \end{pmatrix} - \begin{pmatrix} \boldsymbol\beta_0 \\ \boldsymbol\gamma_0 \end{pmatrix}\right\} + o_P(n^{-1/2}). \quad \text{(A.14)}$$

Together with (A.11) and that of $\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2} = \boldsymbol\Gamma + o_P(1)$, we have

$$\begin{aligned}
\mathcal{T}_n &= \left\{\sqrt{n}[\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}]^{-1/2}\left[\begin{pmatrix} \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma} \end{pmatrix} - \begin{pmatrix} \boldsymbol\beta_0 \\ \boldsymbol\gamma_0 \end{pmatrix}\right]\right\}^\tau \\
&\quad \times [\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}]^{1/2}[\mathbf{I}_{p+q} - \mathbf{\Omega_A}]^\tau \boldsymbol\Gamma[\mathbf{I}_{p+q} - \mathbf{\Omega_A}][\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}]^{1/2} \\
&\quad \times \left\{\sqrt{n}[\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}]^{-1/2}\left[\begin{pmatrix} \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma} \end{pmatrix} - \begin{pmatrix} \boldsymbol\beta_0 \\ \boldsymbol\gamma_0 \end{pmatrix}\right]\right\} \\
&\xrightarrow{L} \sum_{i=1}^k \ell_i \chi_{i,1}^2,
\end{aligned}$$

where $\chi_{i,1}^2$, $i = 1, \dots, k$ are independent standard chi-squared random variables with one degree of freedom, and ℓ_i's are the eigenvalues of matrix $[\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}][\mathbf{I}_{p+q} - \mathbf{\Omega_A}]^\tau\boldsymbol\Gamma[\mathbf{I}_{p+q} - \mathbf{\Omega_A}]$. Using the fact that $[\mathbf{I}_{p+q} - \mathbf{\Omega_A}]^\tau\boldsymbol\Gamma[\mathbf{I}_{p+q} - \mathbf{\Omega_A}] = \mathbf{A}^\tau(\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau)^{-1}\mathbf{A}$, we complete the proof of Theorem 3.2. □

Step 3.2. Under the local hypothesis $\mathcal{H}_{1n} : \mathbf{A}\boldsymbol\theta_0 = \mathbf{b} + n^{-1/2}\mathbf{c}$, similar to (A.14), using (3.5), we have that

$$\begin{aligned}
&\sqrt{n}\boldsymbol\Gamma^{1/2}\begin{pmatrix} \hat{\boldsymbol\beta}_R - \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma}_R - \hat{\boldsymbol\gamma} \end{pmatrix} \\
&\quad = -\boldsymbol\Gamma^{1/2}(\mathbf{I}_{p+q} - \mathbf{\Omega_A})\sqrt{n}\left\{\begin{pmatrix} \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma} \end{pmatrix} - \begin{pmatrix} \boldsymbol\beta_0 \\ \boldsymbol\gamma_0 \end{pmatrix}\right\} \\
&\qquad - \boldsymbol\Gamma^{-1/2}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{c} + o_P(1) \qquad \text{(A.15)} \\
&\quad \xrightarrow{L} N\big(-\boldsymbol\Gamma^{-1/2}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{c}, \\
&\qquad\quad \boldsymbol\Gamma^{1/2}(\mathbf{I}_{p+q} - \mathbf{\Omega_A})[\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0}](\mathbf{I}_{p+q} - \mathbf{\Omega_A})^\tau\boldsymbol\Gamma^{1/2}\big).
\end{aligned}$$

Using (A.16) and that of $\frac{1}{n}\sum_{i=1}^n \hat{\mathbf{T}}_i^{\otimes 2} = \boldsymbol\Gamma + o_P(1)$, we can have that

$$\begin{aligned}
\mathcal{T}_n &= \left[\sqrt{n}\boldsymbol\Gamma^{1/2}\begin{pmatrix} \hat{\boldsymbol\beta}_R - \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma}_R - \hat{\boldsymbol\gamma} \end{pmatrix}\right]^\tau \left[\sqrt{n}\boldsymbol\Gamma^{1/2}\begin{pmatrix} \hat{\boldsymbol\beta}_R - \hat{\boldsymbol\beta} \\ \hat{\boldsymbol\gamma}_R - \hat{\boldsymbol\gamma} \end{pmatrix}\right] + o_P(1) \\
&\xrightarrow{L} (\mathfrak{M} + \boldsymbol\Gamma^{-1/2}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{c})^\tau(\mathfrak{M} + \boldsymbol\Gamma^{-1/2}\mathbf{A}^\tau\{\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau\}^{-1}\mathbf{c}),
\end{aligned} \qquad \text{(A.16)}$$

where $\mathfrak{M}$ follows multivariate normal distribution $N(\mathbf{0}, \mathbf{D})$ with

$$\mathbf{D} = \boldsymbol\Gamma^{-1/2}\mathbf{A}^\tau(\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau)^{-1}\mathbf{A}(\boldsymbol\Gamma^{-1}\sigma_\varepsilon^2 + \mathbf{Q}_{\boldsymbol\beta_0})\mathbf{A}^\tau(\mathbf{A}\boldsymbol\Gamma^{-1}\mathbf{A}^\tau)^{-1}\mathbf{A}\boldsymbol\Gamma^{-1/2}.$$

Acknowledgments

The authors thank the editor, the associate editor for their constructive suggestions that helped them to improve the early manuscript. The research work of Wei Zhenghong is supported by the Project of Department of Education of Guangdong Province of China (Grant No. 2014KTSCX112). The research work of Zhang Jun is supported by the National Natural Sciences Foundation of China (Grant No. 11401391) and the Project of Department of Education of Guangdong Province of China (Grant No. 2014KTSCX112).

References

Carroll, R. J., Fan, J., Gijbels, I. and Wand, M. P. (1997). Generalized partially linear single-index models. *Journal of the American Statistical Association* **92**, 477–489. MR1467842

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Nonlinear Measurement Error Models. A Modern Perspective*, 2nd ed. New York: Chapman and Hall. MR2243417

Cui, X., Guo, W., Lin, L. and Zhu, L. (2009). Covariate-adjusted nonlinear regression. *The Annals of Mathematical Statistics* **37**, 1839–1870. MR2533473

Escanciano, J. C. (2006). A consistent diagnostic test for regression models using projections. *Econometric Theory* **22**, 1030–1051. MR2328527

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. London: Chapman and Hall. MR1383587

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association* **96**, 1348–1360. MR1946581

Fan, J. and Zhang, J. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **62**, 303–322. MR1749541

Kaysen, G. A., Dubin, J. A., Müller, H.-G., Mitch, W. E., Rosales, L. M. and Nathan W. Levin the Hemo Study Group (2002). Relationships among inflammation nutrition and physiologic mechanisms establishing albumin levels in hemodialysis patients. *Kidney International* **61**, 2240–2249.

Kneip, A., Simar, L. and Van Keilegom, I. (2015). Frontier estimation in the presence of measurement error with unknown variance. *Journal of Econometrics* **184**, 379–393. MR3291009

Li, F., Lin, L. and Cui, X. (2010). Covariate-adjusted partially linear regression models. *Communications in Statistics. Theory and Methods* **39**, 1054–1074. MR2745361

Li, G. and Xue, L. (2008). Empirical likelihood confidence region for the parameter in a partially linear errors-in-variables model. *Communications in Statistics. Theory and Methods* **37**, 1552–1564. MR2440453

Li, X., Du, J., Li, G. and Fan, M. (2014). Variable selection for covariate adjusted regression model. *Journal of Systems Science and Complexity* **27**, 1227–1246. MR3284348

Liang, H. and Li, R. (2009). Variable selection for partially linear models with measurement errors. *Journal of the American Statistical Association* **104**, 234–248. MR2504375

Mack, Y. P. and Silverman, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **61**, 405–415. MR0679685

Nguyen, D. V., Şentürk, D. and Carroll, R. J. (2008). Covariate-adjusted linear mixed effects model with an application to longitudinal data. *Journal of Nonparametric Statistics* **20**, 459–481. MR2446438

Saleh, A. E. and Shalabh (2014). A ridge regression estimation approach to the measurement error model. *Journal of Multivariate Analysis* **123**, 68–84. MR3130421

Şentürk, D. and Müller, H.-G. (2005). Covariate adjusted correlation analysis via varying coefficient models. *Scandinavian Journal of Statistics* **32**, 365–383. MR2204625

Şentürk, D. and Müller, H.-G. (2006). Inference for covariate adjusted regression via varying coefficient models. *The Annals of Mathematical Statistics* **34**, 654–679. MR2281880

Şentürk, D. and Nguyen, D. V. (2006). Estimation in covariate-adjusted regression. *Computational Statistics & Data Analysis* **50**, 3294–3310. MR2239669

Şentürk, D. and Nguyen, D. V. (2009). Partial covariate adjusted regression. *Journal of Statistical Planning and Inference* **139**, 454–468. MR2474018

Sheil, J. and O'Muircheartaigh, I. (1977). Algorithm as106: The distribution of non-negative quadratic forms in normal variables. *Applied Statistics* **26**, 92–98.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. *Monographs on Statistics and Applied Probability.* London: Chapman and Hall. MR0848134

Stefanski, L., Wu, Y. and White, K. (2014). Variable selection in nonparametric classification via measurement error model selection likelihoods. *Journal of the American Statistical Association* **109**, 574–589. MR3223734

Stute, W., González Manteiga, W. and Presedo Quindimil, M. (1998). Bootstrap approximations in model checks for regression. *Journal of the American Statistical Association* **93**, 141–149. MR1614600

Wei, C. and Wang, Q. (2012). Statistical inference on restricted partially linear additive errors-in-variables models. *Test* **21**, 757–774. MR2992091

Wu, C.-F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *The Annals of Mathematical Statistics* **14**, 1261–1350. MR0868303

Xu, W. and Guo, X. (2013). Nonparametric checks for a varying coefficient model with missing response at random. *Metrika* **76**, 459–482. MR3049497

Xu, W. and Zhu, L. (2013). Testing the adequacy of varying coefficient models with missing responses at random. *Metrika* **76**, 53–69. MR3018357

Xu, W. and Zhu, L. (2015). Nonparametric check for partial linear errors-in-covariables models with validation data. *Annals of the Institute of Statistical Mathematics* **67**, 793–815. MR3357939

Yang, Y., Li, G. and Peng, H. (2014). Empirical likelihood of varying coefficient errors-in-variables models with longitudinal data. *Journal of Multivariate Analysis* **127**, 1–18. MR3188873

Zhang, J., Feng, Z. and Zhou, B. (2014a). A revisit to correlation analysis for distortion measurement error data. *Journal of Multivariate Analysis* **124**, 116–129. MR3147315

Zhang, J., Gai, Y. and Wu, P. (2013a). Estimation in linear regression models with measurement errors subject to single-indexed distortion. *Computational Statistics & Data Analysis* **59**, 103–120. MR3000045

Zhang, J., Li, G. and Feng, Z. (2015). Checking the adequacy for a distortion errors-in-variables parametric regression model. *Computational Statistics & Data Analysis* **83**, 52–64. MR3281795

Zhang, J., Wang, X., Yu, Y. and Gai, Y. (2014b). Estimation and variable selection in partial linear single index models with error-prone linear covariates. *Statistics. A Journal of Theoretical and Applied Statistics* **48**, 1048–1070. MR3259875

Zhang, J., Yu, Y., Zhou, B. and Liang, H. (2014c). Nonlinear measurement errors models subject to additive distortion. *Journal of Statistical Planning and Inference* **150**, 49–65. MR3206720

Zhang, J., Yu, Y., Zhu, L. and Liang, H. (2013b). Partial linear single index models with distortion measurement errors. *Annals of the Institute of Statistical Mathematics* **65**, 237–267. MR3011622

Zhang, J., Zhu, L. and Liang, H. (2012a). Nonlinear models with measurement errors subject to single-indexed distortion. *Journal of Multivariate Analysis* **112**, 1–23. MR2957282

Zhang, J., Zhu, L. and Zhu, L. (2012b). On a dimension reduction regression with covariate adjustment. *Journal of Multivariate Analysis* **104**, 39–55. MR2832185

Zhang, W., Li, G. and Xue, L. (2011). Profile inference on partially linear varying-coefficient errors-in-variables models under restricted condition. *Computational Statistics & Data Analysis* **55**, 3027–3040. MR2813063

Zhou, Y. and Liang, H. (2009). Statistical inference for semiparametric varying-coefficient partially linear models with error-prone linear covariates. *The Annals of Mathematical Statistics* **37**, 427–458. MR2488358

Z. Wei
College of Mathematics and Computational Science
Institute of Statistical Sciences
Shenzhen University
Shenzhen
China
E-mail: zhhwei@szu.edu.cn

Y. Fan
College of Mathematics and Computational Science
Shenzhen University
Shenzhen
China
E-mail: fanybsta@gmail.com

J. Zhang
Institute of Statistical Sciences
Shen Zhen–Hong Kong Joint Research Center for Applied Statistical Sciences
College of Mathematics and Computational Science
Shenzhen University
Shenzhen
China
E-mail: zhangjunstat@gmail.com