

Research Article

Profile Inferences on Restricted Additive Partially Linear EV Models

Xiuli Wang

School of Mathematical Science, Shandong Normal University, Jinan 250014, China

Correspondence should be addressed to Xiuli Wang; wxlmath@163.com

Received 5 June 2013; Revised 1 August 2013; Accepted 7 August 2013

Academic Editor: Qian Guo

Copyright © 2013 Xiuli Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider the testing problem for the parameter and restricted estimator for the nonparametric component in the additive partially linear errors-in-variables (EV) models under additional restricted condition. We propose a profile Lagrange multiplier test statistic based on modified profile least-squares method and two-stage restricted estimator for the nonparametric component. We derive two important results. One is that, without requiring the undersmoothing of the nonparametric components, the proposed test statistic is proved asymptotically to be a standard chi-square distribution under the null hypothesis and a noncentral chi-square distribution under the alternative hypothesis. These results are the same as the results derived by Wei and Wang (2012) for their adjusted test statistic. But our method does not need an adjustment and is easier to implement especially for the unknown covariance of measurement error. The other is that asymptotic distribution of proposed two-stage restricted estimator of the nonparametric component is asymptotically normal and has an oracle property in the sense that, though the other component is unknown, the estimator performs well as if it was known. Some simulation studies are carried out to illustrate relevant performances with a finite sample. The asymptotic distribution of the restricted corrected-profile least-squares estimator, which has not been considered by Wei and Wang (2012), is also investigated.

1. Introduction

To balance the modeling bias and the “curse of dimensionality,” different kinds of semiparametric models, such as partially linear model, partially linear varying coefficient model, partially linear single-index model, and additive partially linear model, have been proposed and investigated. In this paper, we consider the semiparametric additive partially linear model, which can be written as

$$Y = X'\beta + \sum_{d=1}^D f_d(Z_d) + \varepsilon, \quad (1)$$

where Y is the response, X and $Z = (Z_1, \dots, Z_D)'$ are covariates on R^p and R^D , respectively, f_1, \dots, f_D are unknown smooth functions, $\beta = (\beta_1, \dots, \beta_p)'$ is a p -dimensional vector of unknown parameters, and ε is the random error with conditional mean zero given X and Z . Model (1), inheriting the interpretability of the linear model and flexibility of the additive model, has been studied by Opsomer and Ruppert [1], Li [2], and Jiang et al. [3], among others. Two

important models, partially linear model, which is the case when $D = 1$ for model (1), and additive model, when $\beta = 0$, can be regarded as its special models.

However in many practical applications, it may be difficult or impossible to measure some explanatory variables accurately, and people usually can only observe its surrogate. For the literature about errors in variables (EV) the reader can resort to Fuller [4], Cheng and Van Ness [5], Carroll et al. [6], Liang et al. [7], You and Chen [8], Li and Greene [9], Wang et al. [10], and the references therein. In this paper, we consider model (1) with the case that only covariate X is measured with additive error; that is, X cannot be observed, but an unbiased measure of X , denoted by W , can be obtained such that

$$W = X + U, \quad (2)$$

where U is the measurement error with mean zero and independent of (X, Z, Y) . Like Liang et al. [11], we only consider the case of $D = 2$ and assume that the covariance matrix of measurement error U , say Σ_{uu} , is known. Otherwise, we

can estimate it by repeatedly measuring W as mentioned by Liang et al. [7, 11].

In many practical applications such as production and consumption studies, in addition to the sample information, people always have some prior information on regression parametric vector which can be formed into some constraint conditions and used to improve the parametric estimators. For more detailed discussion see Jorgenson [12]. It is well known that for general linear models and nonlinear regression models there are restricted estimation and corresponding results, but for semi-parametric regression models, the results of restricted estimation are very few. Recently, Wei and Wang [13] considered the testing problem under the following linear hypothesis on the parametric component:

$$H_0 : A\beta = \mathbf{b} \quad \text{versus} \quad H_1 : A\beta \neq \mathbf{b}, \quad (3)$$

where A is a $k \times p$ matrix of known constants, $\text{rank}(A) = k$, and \mathbf{b} is a k -vector of known constants. They first proposed the restricted corrected-profile least-squares estimator $\hat{\beta}_r$ of β under the restriction $A\beta = \mathbf{b}$. Based on the difference between the corrected residual sums of squares under the null and alternative hypotheses, a test statistic is suggested, but the limiting distribution is a weighted sum of independent standard $\chi^2(1)$, so adjustment is needed, and the adjusted test statistic is also suggested. Though they proposed the restricted corrected-profile least-squares estimator $\hat{\beta}_r$ of β under the condition $A\beta = \mathbf{b}$, they did not investigate the asymptotic property of $\hat{\beta}_r$. In particular they did not give the estimator of nonparametric component. So in this paper, we will investigate the above problems. Once the asymptotic distribution of $\hat{\beta}_r$ is obtained, for a different matrix A , we can easily derive the asymptotic distribution of different constrained estimates $A\hat{\beta}_r$ with different purposes. In addition to the just-mentioned problems, our other aim is to suggest the profile Lagrange multiplier test statistic for the unknown parameter β on testing problem and show that its limiting distribution is a standard chi-squared distribution under the null hypothesis and a noncentral chi-square distribution under the alternative hypothesis. These results are the same as the results derived by Wei and Wang [13] for their adjusted test statistic, but our method does not need an adjustment and is easier to implement especially when the covariance Σ_{uu} is unknown. For the Lagrange multiplier test method, see Wei and Wu [14] and Zhang et al. [15].

The rest of this paper is organized as follows. In Section 2, we first review the restricted corrected-profile least-squares estimator $\hat{\beta}_r$ of β and then study the asymptotic distribution of $\hat{\beta}_r$. After that, we construct the modified profile Lagrange multiplier test statistic and derive its asymptotic distribution under the null and alternative hypotheses. In Section 3, two-stage restricted estimators for the nonparametric components are proposed and their asymptotic distribution are presented. Some simulation studies are carried out to assess the performance of the derived results in Section 4. Proofs of the main results are given in Section 5.

2. Asymptotic Results of the Restricted-Profile Least-Squares Estimator and Modified Profile Lagrange Multiplier Test

2.1. The Restricted-Profile Least-Squares Estimator and Its Asymptotic Property. For the case of $D = 2$, like Liang et al. [11] and Wang et al. [10], we assume $E\{f_1(Z_1)\} = E\{f_2(Z_2)\} = 0$ to ensure identifiability of the nonparametric functions, and X and Y are centered for simplicity. Suppose the observable data $\{Y_i, Z_{i1}, Z_{i2}, W_i = (W_{i1}, W_{i2}, \dots, W_{ip})', i = 1, \dots, n\}$ are independent and identically distributed (i.i.d.) and are generated from the following model:

$$\begin{aligned} Y_i &= X_i' \beta + f_1(Z_{i1}) + f_2(Z_{i2}) + \varepsilon_i, \\ W_i &= X_i + U_i, \end{aligned} \quad (4)$$

where the model errors $\{\varepsilon_i, i = 1, \dots, n\}$ are i.i.d. with $E(\varepsilon_i | X_i, Z_{i1}, Z_{i2}) = 0$ almost surely and the measurement errors U_i are i.i.d. with mean zero and independent of $\{Y_i, X_i = (X_{i1}, X_{i2}, \dots, X_{ip})', Z_{i1}, Z_{i2}\}$. Besides, we also have some prior information on regression parametric vector β that can be presented by the following restricted condition:

$$A\beta = \mathbf{b} \quad (5)$$

with A being a $k \times p$ matrix of known constants and $\text{rank}(A) = k$ and \mathbf{b} being a k -vector of known constants.

First, we need introduce the restricted-profile least-squares estimator of Wei and Wang [13]. The notations are like those of Wang et al. [10]. If β is known, the first equation of model (4) can be rewritten as

$$Y_i - X_i' \beta = f_1(Z_{i1}) + f_2(Z_{i2}) + \varepsilon_i, \quad i = 1, \dots, n. \quad (6)$$

Denote

$$\begin{aligned} \mathbf{Y} &= (Y_1, \dots, Y_n)', \\ \boldsymbol{\varepsilon} &= (\varepsilon_1, \dots, \varepsilon_n)', \\ \mathbf{X} &= (X_1, \dots, X_n)', \\ \mathbf{Z}_1 &= (Z_{11}, \dots, Z_{n1})', \\ \mathbf{Z}_2 &= (Z_{12}, \dots, Z_{n2})', \\ F_1 &= (f_1(Z_{11}), \dots, f_1(Z_{n1}))', \\ F_2 &= (f_2(Z_{12}), \dots, f_2(Z_{n2}))'. \end{aligned} \quad (7)$$

Hence, (6) can be written as

$$\mathbf{Y} - \mathbf{X}\beta = F_1 + F_2 + \boldsymbol{\varepsilon}. \quad (8)$$

Let s'_{1,z_1}, s'_{2,z_2} denote the equivalent kernels for the local linear regression at z_1, z_2 :

$$\begin{aligned} s'_{1,z_1} &= e'_1 (\mathcal{Z}'_1 \Omega_1 \mathcal{Z}_1)^{-1} \mathcal{Z}'_1 \Omega_1, \\ s'_{2,z_2} &= e'_1 (\mathcal{Z}'_2 \Omega_2 \mathcal{Z}_2)^{-1} \mathcal{Z}'_2 \Omega_2, \end{aligned} \quad (9)$$

where $e_1 = (1, 0)'$,

$$\begin{aligned}\Omega_1 &= \text{diag} \left(\frac{1}{h_1} K \left(\frac{Z_{11} - z_1}{h_1} \right), \dots, \frac{1}{h_1} K \left(\frac{Z_{n1} - z_1}{h_1} \right) \right), \\ \Omega_2 &= \text{diag} \left(\frac{1}{h_2} K \left(\frac{Z_{12} - z_2}{h_2} \right), \dots, \frac{1}{h_2} K \left(\frac{Z_{n2} - z_2}{h_2} \right) \right),\end{aligned}\quad (10)$$

for a kernel function $K(\cdot)$ and bandwidths h_1, h_2 , and

$$\mathcal{X}_1 = \begin{pmatrix} 1 & Z_{11} - z_1 \\ \vdots & \vdots \\ 1 & Z_{n1} - z_1 \end{pmatrix}, \quad \mathcal{X}_2 = \begin{pmatrix} 1 & Z_{12} - z_2 \\ \vdots & \vdots \\ 1 & Z_{n2} - z_2 \end{pmatrix} \quad (11)$$

are $n \times 2$ design matrices. Let S_1 and S_2 represent the smoother matrices whose rows are the equivalent kernels at the observations \mathbf{Z}_1 and \mathbf{Z}_2 , respectively:

$$S_1 = \begin{pmatrix} s'_{1,Z_{11}} \\ \vdots \\ s'_{1,Z_{n1}} \end{pmatrix}, \quad S_2 = \begin{pmatrix} s'_{1,Z_{12}} \\ \vdots \\ s'_{2,Z_{n2}} \end{pmatrix}. \quad (12)$$

Using the backfitting algorithm, we get the backfitting estimators of F_1 and F_2 , say \hat{F}_1 and \hat{F}_2 , which can be expressed as

$$\begin{aligned}\hat{F}_1 &= \{I - (I - S_1^c S_2^c)^{-1} (I - S_1^c)\} (\mathbf{Y} - \mathbf{X}\beta), \\ \hat{F}_2 &= \{I - (I - S_2^c S_1^c)^{-1} (I - S_2^c)\} (\mathbf{Y} - \mathbf{X}\beta),\end{aligned}\quad (13)$$

where $S_i^c = (I - \mathbf{1}\mathbf{1}'/n)S_i$ is the centered smoothing matrix corresponding to $S_i, i = 1, 2$. Substituting (13) into (8), we have

$$(I - S) \mathbf{Y} = (I - S) \mathbf{X}\beta + \varepsilon, \quad (14)$$

where $S = \{I - (I - S_1^c S_2^c)^{-1} (I - S_1^c)\} + \{I - (I - S_2^c S_1^c)^{-1} (I - S_2^c)\}$.

Since X_i cannot be observed, the corrected-profile least-squares estimator of β is defined as

$$\begin{aligned}\hat{\beta} &= \arg\min_n [(\mathbf{Y} - \mathbf{W}\beta)' (I - S)' (I - S) (\mathbf{Y} - \mathbf{W}\beta) \\ &\quad - n\beta' \Sigma_{uu} \beta] \\ &= [\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu}]^{-1} \widetilde{\mathbf{W}}' (I - S) \mathbf{Y},\end{aligned}\quad (15)$$

where \mathbf{W} is defined similar to \mathbf{X} and $\widetilde{\mathbf{W}} = (I - S)\mathbf{W}$.

When considering the restrictions $A\beta = \mathbf{b}$, Wei and Wang [13] defined the corrected Lagrange function as

$$\begin{aligned}F(\beta, \lambda) &= [(\mathbf{Y} - \mathbf{W}\beta)' (I - S)' (I - S) (\mathbf{Y} - \mathbf{W}\beta) - n\beta' \Sigma_{uu} \beta] \\ &\quad + 2\lambda' (A\beta - \mathbf{b}),\end{aligned}\quad (16)$$

where λ is a $k \times 1$ Lagrange multipliers vector. Differentiating $F(\beta, \lambda)$ with respect to β and λ and setting them to zero, we obtain

$$\begin{aligned}\frac{\partial F(\beta, \lambda)}{\partial \beta} &= -2\widetilde{\mathbf{W}}' (I - S) \mathbf{Y} + 2\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} \beta - 2n\Sigma_{uu} \beta \\ &\quad + 2A' \lambda = 0, \\ \frac{\partial F(\beta, \lambda)}{\partial \lambda} &= 2(A\beta - \mathbf{b}) = 0.\end{aligned}\quad (17)$$

Solving (17), we have

$$\begin{aligned}\hat{\beta}_r &= \hat{\beta} - (\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A' [A(\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A']^{-1} \\ &\quad \times (A\hat{\beta} - \mathbf{b}), \\ \hat{\lambda} &= [A(\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A']^{-1} (A\hat{\beta} - \mathbf{b}),\end{aligned}\quad (18)$$

where $\hat{\beta}$ is given by (15).

According to Wei and Wang [13], $\hat{\beta}_r$ is called the modified restricted profile least-squares estimator of β , but the asymptotic distribution of $\hat{\beta}_r$ is vacant. In the following, we will give the asymptotic distribution of $\hat{\beta}_r$. The following assumptions will be required to derive the main results. These assumptions are common and can be found in the works of Liang et al. [11], Wang et al. [10], and Wei and Wang [13].

(A1) The density functions of Z_1 and Z_2 are bounded away from 0 and have bounded continuous second partial derivatives.

(A2) Let $\check{X}_i = X_i - E(X_i | Z_{i1}) - E(X_i | Z_{i2})$. The matrix $\Gamma_{X|Z} \triangleq E(\check{X}_1 \check{X}_1')$ is positive-definite, $E(\varepsilon_i | X_i, Z_{i1}, Z_{i2}) = 0$, and $E(|\varepsilon_i|^3 | X_i, Z_{i1}, Z_{i2}) < \infty$.

(A3) The bandwidths h_1 and h_2 are of order $n^{-1/5}$.

(A4) The function $K(\cdot)$ is a symmetric density function with compact support and satisfies $\int K(u) du = 1$, $\int uK(u) du = 0$, $\int u^2 K(u) du = 1$, and $\int u^4 K(u) du < \infty$.

(A5) $EU_i = 0$ and $E(\|U_i\|^3) < \infty$.

Under these assumptions, we give the following theorem that states the asymptotic distribution of $\hat{\beta}_r$.

Theorem 1. Suppose that conditions (A1)–(A5) hold and ε_i is homoscedastic with variance σ^2 and independent of U_i . Then the modified restricted profile least-squares estimator $\hat{\beta}_r$ is asymptotically normal. Namely,

$$\sqrt{n}(\hat{\beta}_r - \beta) \xrightarrow{\mathcal{D}} N(0, \Sigma), \quad n \rightarrow \infty, \quad (19)$$

where " $\xrightarrow{\mathcal{D}}$ " denotes the convergence in distribution and

$$\begin{aligned}\Sigma &= \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} - \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} \Xi \Gamma_{X|Z}^{-1} - \Gamma_{X|Z}^{-1} \Xi \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} \\ &\quad + \Gamma_{X|Z}^{-1} \Xi \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} \Xi \Gamma_{X|Z}^{-1}, \\ \Lambda &= E(\varepsilon_1 - U_1' \beta)^2 \Gamma_{X|Z} + E\{(U_1 U_1' - \Sigma_{uu}) \beta\}^{\otimes 2} \\ &\quad + \Sigma_{uu} \sigma^2, \quad \Xi = A' (A \Gamma_{X|Z}^{-1} A')^{-1} A, \quad R^{\otimes 2} = R R'.\end{aligned}\quad (20)$$

Remark 2. From the previous theorem, it is easy to check that when X is observed exactly the asymptotic distribution of modified restricted profile least-squares estimator $\hat{\beta}_r$ is the same as the asymptotic distribution obtained by Theorem 3.1 which appeared in the works of Wei and Liu [16].

To apply Theorem 1 to inference, we need to estimate $\Gamma_{X|Z}$ and Λ . Using plug-in method, a consistent estimator $\hat{\Sigma}$ of Σ can be obtained. Wei and Wang [13] proposed to estimate $\Gamma_{X|Z}$ and Λ , respectively, by

$$\hat{\Gamma}_{X|Z} = \frac{1}{n} \sum_{i=1}^n \bar{W}_i \bar{W}_i' - \Sigma_{uu}, \quad (21)$$

$$\hat{\Lambda} = \frac{1}{n} \sum_{i=1}^n \{\bar{W}_i (\bar{Y}_i - \bar{W}_i' \hat{\beta}) + \Sigma_{uu} \hat{\beta}\}^{\otimes 2}. \quad (22)$$

The asymptotic distribution of $B \hat{\beta}_r$, where B is an $s \times p$ matrix with $\text{rank}(B) = s$, can be given by the following result.

Corollary 3. Suppose that conditions (A1)–(A5) hold and ε_i is homoscedastic with variance σ^2 and independent of U_i . Then as $n \rightarrow \infty$, one has

$$\sqrt{n} (B \hat{\beta}_r - B \beta) \xrightarrow{\mathcal{D}} N(0, B \Sigma B'). \quad (23)$$

Using Corollary 3, confidence regions of any linear combination of the parameter components given in advance can be constructed. In particular, for any $0 < \alpha < 1$, let $\chi_{1-\alpha}^2(p)$ denote the $(1 - \alpha)$ -quantile of $\chi^2(p)$. Then

$$\begin{aligned}\left\{ B \beta \in R^s : (B \beta - B \hat{\beta}_r)' \left(\frac{1}{n} B \hat{\Sigma} B' \right)^{-1} (B \beta - B \hat{\beta}_r) \right. \\ \left. \leq \chi_{1-\alpha}^2(p) \right\}\end{aligned}\quad (24)$$

constitute a confidence region of $B \beta$ with asymptotic coverage $1 - \alpha$.

2.2. Modified Profile Lagrange Multiplier Test and Its Asymptotic Properties. Using the estimation method described in Section 2.1, we now consider the following linear hypothesis:

$$H_0 : A \beta = \mathbf{b}, \quad H_1 : A \beta = \mathbf{b} + \delta, \quad (25)$$

where δ is a $k \times 1$ positive constant vector. Wei and Wang [13] constructed a test statistic based on the difference

between the corrected residual sums of squares under the null and alternative hypotheses and showed that the asymptotic null distribution of proposed statistic is a weighted sum of independent standard $\chi^2(1)$, so adjustment is needed. Notice that the statistics proposed in Wei and Wu [14] and Zhang et al. [15] can achieve the standard chi-squared limit, so in this paper, we use the similar idea to construct the profile Lagrange multiplier test statistic.

Using the estimator of Lagrange multiplier defined in (18), the modified profile Lagrange multiplier test statistic can be constructed by

$$T_n = \frac{1}{n} \hat{\lambda}' (C_0 L_0 C_0')^{-1} \hat{\lambda}, \quad (26)$$

where $C_0 = (A \hat{\Gamma}_{X|Z}^{-1} A')^{-1}$, $L_0 = A \hat{\Gamma}_{X|Z}^{-1} \hat{\Lambda} \hat{\Gamma}_{X|Z}^{-1} A'$, and $\hat{\Gamma}_{X|Z}$ and $\hat{\Lambda}$ are estimators of $\Gamma_{X|Z}$ and Λ defined by (21) and (22), respectively.

The following theorem gives the asymptotic distribution of the modified profile Lagrange multiplier test statistic T_n .

Theorem 4. Suppose that conditions (A1)–(A5) hold. Then

(1) under the null hypothesis H_0 of testing problem (25),

$$T_n \xrightarrow{\mathcal{D}} \chi^2(k), \quad \text{as } n \rightarrow \infty, \quad (27)$$

(2) under the alternative hypothesis H_1 of testing problem (25), T_n follows the asymptotic noncentral $\chi^2(k, \zeta)$ distribution with k degrees of freedom, and the noncentral parameter is

$$\zeta = \lim_{n \rightarrow \infty} n (A \beta - \mathbf{b})' (CLC')^{-1} (A \beta - \mathbf{b}), \quad (28)$$

where $C = (A \Gamma_{X|Z}^{-1} A')^{-1}$, $L = A \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} A'$.

Remark 5. From the above theorem, we known that our results are the same to the results of adjusted test statistic derived by Wei and Wang [13], but our proposed test statistic is more easier to perform, especially for the case when the covariance matrix Σ_{uu} of measurement error is unknown. From Wei and Wang [13] we know that, when Σ_{uu} is unknown, it is difficult to estimate $RSS(H_0)$ and $RSS(H_1)$; thus their proposed test statistic is unapplicable in this case. So our proposed profile Lagrange multiplier test statistic is more attractive.

3. Two-Stage Restricted Estimator for the Nonparametric Component

According to $\hat{\beta}_r$, the corrected backfitting estimators, say $\tilde{F}_1 = (\tilde{f}_1(Z_{11}), \dots, \tilde{f}_1(Z_{n1}))'$ and $\tilde{F}_2 = (\tilde{f}_2(Z_{12}), \dots, \tilde{f}_2(Z_{n2}))'$, of F_1 and F_2 can be defined by

$$\begin{aligned}\tilde{F}_1 &= \{I - (I - S_1^c S_2^c)^{-1} (I - S_1^c)\} (\mathbf{Y} - \mathbf{W} \hat{\beta}_r), \\ \tilde{F}_2 &= \{I - (I - S_2^c S_1^c)^{-1} (I - S_2^c)\} (\mathbf{Y} - \mathbf{W} \hat{\beta}_r).\end{aligned}\quad (29)$$

From (29), we know that \tilde{F}_1 and \tilde{F}_2 are only the estimators of $f_1(\cdot)$ and $f_2(\cdot)$ at the observations $(Z_{11}, \dots, Z_{n1})'$ and $(Z_{12}, \dots, Z_{n2})'$, respectively. We next give the two-stage restricted estimator of $f_1(\cdot)$.

Noting the fact that $f_1(z_1) = E[Y - W'\beta - f_2(Z_2) | Z_1 = z_1]$ and using the local linear expansion, we define a two-stage restricted estimator of $(f_1(z_1), f_1'(z_1))'$ by

$$\begin{aligned} & (\hat{f}_1^R(z_1), \hat{f}_1'^R(z_1))' \\ &= \arg \min_{a_0, a_1} \sum_{i=1}^n \left[Y_i - W_i' \hat{\beta}_r - \tilde{f}_2(Z_{i2}) - a_0 \right. \\ & \quad \left. - a_1 (Z_{i1} - z_1) \right]^2 K_h(Z_{i1} - z_1), \end{aligned} \quad (30)$$

where $K_h(x) = (1/h)K(x/h)$ and h is a bandwidth. Then the two-stage restricted estimator $\hat{f}_1^R(z_1)$ of $f_1(z_1)$ can be written as

$$\begin{aligned} \hat{f}_1^R(z_1) &= (1, 0) (\mathcal{X}_1' \Omega \mathcal{X}_1)^{-1} \mathcal{X}_1' \Omega (I - S_2^c S_1^c)^{-1} (I - S_2^c) \\ & \quad \times (\mathbf{Y} - \mathbf{W} \hat{\beta}_r), \end{aligned} \quad (31)$$

where $\Omega = \text{diag}((1/h)K((Z_{11} - z_1)/h), \dots, (1/h)K((Z_{n1} - z_1)/h))$.

Let $p_1(z_1)$ denote the density function of Z_1 , $v_0 = \int K^2(u)du$, $\mu_2 = \int u^2 K(u)du$. The following theorem gives the asymptotic normality of the two-stage restricted estimator.

Theorem 6. Suppose that conditions (A1)–(A5) hold and $nh^5 = O(1)$. If $\sigma^2(z_1) = E(\varepsilon^2 | Z_1 = z_1)$ has continuous derivative, then as $n \rightarrow \infty$, one has

$$\sqrt{nh} \left(\hat{f}_1^R(z_1) - f_1(z_1) - \frac{h^2}{2} \mu_2 f_1''(z_1) \right) \xrightarrow{\mathcal{D}} N(0, \Sigma(z_1)), \quad (32)$$

where $\Sigma(z_1) = ((\sigma^2(z_1) + \beta' \Sigma_{uu} \beta) / p_1(z_1)) v_0$.

Remark 7. From Theorem 6, we known that the two-stage restricted estimator $\hat{f}_1^R(z_1)$ is asymptotically normal and has an oracle property; that is, the estimator performs as if the other nonparametric component $f_2(z_2)$ was known though it was unknown. Simulation studies further confirm our theory result. Similarly, the two-stage restricted estimator of $f_2(z_2)$ and its asymptotic result can be obtained.

4. Simulation

In this section, some simulations are carried out to evaluate the finite sample performance of the testing procedure and the proposed two-stage restricted estimator for nonparametric components.

In our simulation studies, consider the following additive partially linear EV model:

$$\begin{aligned} Y_i &= X_{i1}\beta_1 + X_{i2}\beta_2 + f_1(Z_{i1}) + f_2(Z_{i2}) + \varepsilon_i, \\ W_{ij} &= X_{ij} + U_{ij}, \quad j = 1, 2, \end{aligned} \quad (33)$$

where X_{i1} , X_{i2} , Z_{i1} , and Z_{i2} are mutually independent variables and $\beta_1 = 3, \beta_2 = 2$, $X_{i1} \sim N(0, 1)$, and $X_{i2} \sim N(2, 1)$. $f_1(Z_{i1}) = Z_{i1} \exp(-Z_{i1}^2)$, $f_2(Z_{i2}) = 2 \sin(Z_{i2})$, and Z_{i1}, Z_{i2} are generated from a uniform distribution $U[-1, 1]$. The model error $\varepsilon_i \sim N(0, 0.2^2)$ and measure error $U_{ij} \sim N(0, \sigma_u^2)$. The kernel function is taken as the Epanechnikov kernel $K(t) = (3/4)(1 - t^2)I_{(|t| \leq 1)}$.

We first study the performance of the proposed testing procedure. Consider that the null hypothesis is $H_0 : A\beta = 5$ with $A = (1, 1)$ and the corresponding alternative hypothesis is $H_1 : A\beta = 5 + \delta$ with δ being a series of positive constants. For $\delta = 0$, the alternative hypothesis becomes the null hypothesis. We take the bandwidths $h_1 = h_2 = 0.8 \cdot n^{-1/5}$, the sample sizes $n = 200, 300$, and $\sigma_u^2 = 0.25, 0.4$. To illustrate the effectiveness of the proposed test statistics, the estimated power function curves with the significance level $\alpha = 0.05$ are plotted for $\sigma_u^2 = 0.4$ and $\sigma_u^2 = 0.25$. For each case, we repeat 1000 times. The results are given in Figure 1.

From Figure 1, we can see the following results.

- (1) Our proposed test statistic is sensitive to the corresponding alternative hypothesis. This can be seen from the fact that if we increase the constant δ a little, the estimated power function increases rapidly.
- (2) The measurement errors affect the power function. For the same sample sizes, increasing the variance of the measurement error results in the decrease of the estimated power function. For the same measurement error, when the observation sample sizes increase, the estimated power function also increases. So our proposed procedure is feasible and easy to perform.

Second, we give the finite sample performance of the two-stage restricted estimator for nonparametric components with the restricted condition that $A\beta = 5$. For saving space, only the simulation results for estimators of f_1 are presented. We compare two estimators: one is the proposed two-stage restricted estimator for nonparametric component by (30) and the other is the benchmark estimator, which has the same form as the two-stage restricted estimator except that when we estimate f_1 we assume that f_2 is known. We use $h_1 = h_2 = n^{-1/5}$ to estimate $\hat{\beta}$, and the optimal bandwidth h chosen by the leave-one-out cross-validation method to estimate the nonparametric component f_1 . The performance of estimators is evaluated by the root averaged squared errors (RASEs):

$$\text{RASE}(\hat{f}_1) = \sqrt{n^{-1} \sum_{i=1}^n [\hat{f}_1(Z_{i1}) - f_1(Z_{i1})]^2}. \quad (34)$$

We take $n = 200, 300, 500$ and $\sigma_u^2 = 0.25, 0.4, 0.6$, respectively. For a given sample size, the sample means (SMs) and standard deviations (STDs) of RASEs are calculated. In each simulation, the repeated number is 1000. The simulation results are listed in Table 1.

From Table 1, it can be seen that, with the increase of the sample size n , the finite sample performance of the two estimators improves and the performance of the two-stage restricted estimator is close to that of the benchmark

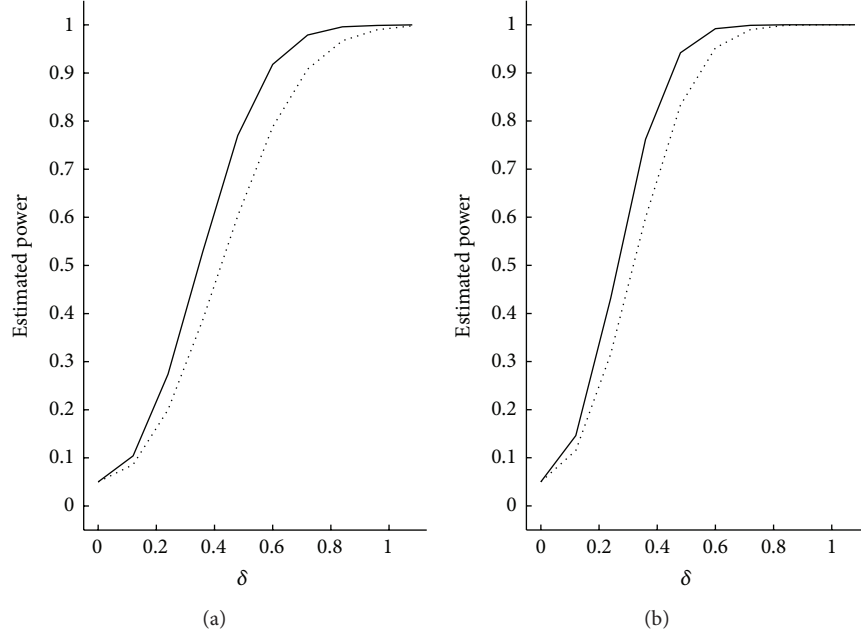


FIGURE 1: The plot of the estimated power function for sample sizes $n = 200$ and $n = 300$, where the solid line denotes $n = 300$ and the dotted line denotes $n = 200$. (a) shows the estimated power function curves with $\sigma_u^2 = 0.4$. (b) shows the estimated power function curves with $\sigma_u^2 = 0.25$.

TABLE 1: The finite sample performance of the estimators for f_1 under different sample sizes n and Σ_{uu} .

Σ_{uu}	$n = 200$	$n = 300$	$n = 500$
$\Sigma_{uu} = 0.25I_2$			
Two-stage			
SM	0.2249	0.1887	0.1530
STD	0.1152	0.0915	0.0552
Benchmark			
SM	0.2239	0.1762	0.1494
STD	0.0958	0.0764	0.0533
$\Sigma_{uu} = 0.4I_2$			
Two-stage			
SM	0.2787	0.2298	0.1923
STD	0.1245	0.1077	0.0930
Benchmark			
SM	0.2784	0.2284	0.1905
STD	0.1235	0.1043	0.0916
$\Sigma_{uu} = 0.6I_2$			
Two-stage			
SM	0.3253	0.2609	0.2162
STD	0.1728	0.1375	0.1038
Benchmark			
SM	0.3200	0.2602	0.2153
STD	0.1707	0.1344	0.1008

estimator. So the proposed two-stage restricted estimator of the nonparametric component has an oracle property.

5. Proofs

Firstly, some lemmas will be given.

Lemma 8. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d random vectors, where Y_i is scalar random variable. Further assume that $E|Y|^s < \infty$ and $\sup_x \int |y|^s p(x, y) dy < \infty$, where $p(x, y)$ denotes the joint density of (X, Y) . Let $K(\cdot)$ be a bounded positive function with a bounded support, satisfying a Lipschitz condition. Given that $n^{2\delta-1}h \rightarrow \infty$ for some $\delta < 1 - s^{-1}$, then

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^n \{K_h(X_i - x) Y_i - E[K_h(X_i - x) Y_i]\} \right| = O_p\left(\frac{\log(1/h)}{nh}\right)^{1/2}. \quad (35)$$

This is Lemma 7.1 of Fan and Huang [17].

Lemma 9. Suppose that assumptions (A1)–(A5) hold. Then the following asymptotic approximations hold uniformly over all the elements of the matrices:

$$\begin{aligned} S_d^c &= S_d - \frac{\mathbf{1}\mathbf{1}'}{n} + o_p\left(\frac{\mathbf{1}\mathbf{1}'}{n}\right), \\ (I - S_1^c S_2^c)^{-1} &= I + o_p\left(\frac{1}{n}\right), \end{aligned} \quad (36)$$

where $d = 1, 2$. Similarly, the second formula is true for $(I - S_2^c S_1^c)^{-1}$.

This lemma is Lemma 3.1 and Lemma 3.2 of Opsomer and Ruppert [1].

Lemma 10. Suppose that assumptions (A1)–(A5) hold. Then as $n \rightarrow \infty$, it holds that

$$\frac{1}{n} \mathbf{W}'(I - S)'(I - S) \mathbf{W} - \Sigma_{uu} = \frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \xrightarrow{P} \Gamma_{X|Z}, \quad (37)$$

where $\widetilde{\mathbf{W}} = (I - S)\mathbf{W}$.

The proof of this lemma can be found in the works of Wang et al. [10].

Lemma 11. Suppose that assumptions (A1)–(A5) hold. Then the corrected-profile least-squares estimator $\widehat{\beta}$ of β is asymptotically normal. Namely,

$$\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{\mathcal{D}} N(0, \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1}), \quad n \rightarrow \infty. \quad (38)$$

This lemma is just Theorem 1 of Liang et al. [11] and is same as Theorem 3.2 of Wang et al. [10].

Proof of Theorem 1. Let

$$\begin{aligned} J_n &= I - (\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A' [A(\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A']^{-1} A \\ &= I - \left(\frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \right)^{-1} A' \left[A \left(\frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \right)^{-1} A' \right]^{-1} A. \end{aligned} \quad (39)$$

By Lemma 10, we can obtain

$$J_n \xrightarrow{P} J = I - \Gamma_{X|Z}^{-1} A' (A \Gamma_{X|Z}^{-1} A')^{-1} A. \quad (40)$$

By the first equation of (18), we have

$$\begin{aligned} \widehat{\beta}_r - \beta &= \left\{ I - (\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A' \right. \\ &\quad \times \left. \left[A(\widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - n\Sigma_{uu})^{-1} A' \right]^{-1} A \right\} \\ &\quad \times (\widehat{\beta} - \beta) = J(\widehat{\beta} - \beta) + (J_n - J)(\widehat{\beta} - \beta). \end{aligned} \quad (41)$$

Using the results that $J_n - J = o_P(1)$ and $\widehat{\beta} - \beta = O_P(n^{-1/2})$, we get

$$(J_n - J)(\widehat{\beta} - \beta) = o_P(n^{-1/2}). \quad (42)$$

By the Slutsky theorem and Lemma 11, we can derive the result after some calculations. \square

Proof of Theorem 4. (1) Using Lemma 10, we know that

$$C_0 = \left(A \left(\frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \right)^{-1} A' \right)^{-1} \xrightarrow{P} (A \Gamma_{X|Z}^{-1} A')^{-1} = C. \quad (43)$$

Similarly, we can prove that

$$L_0 = A \widehat{\Gamma}_{X|Z}^{-1} \widehat{\Lambda} \widehat{\Gamma}_{X|Z}^{-1} A' \xrightarrow{P} A \Gamma_{X|Z}^{-1} \Lambda \Gamma_{X|Z}^{-1} A' = L. \quad (44)$$

Under the null hypothesis of testing problem (25) and by applying Lemma 11, we can prove that

$$\sqrt{n}(A\widehat{\beta} - \mathbf{b}) \xrightarrow{\mathcal{D}} N(0, L). \quad (45)$$

By the second equation of (18), we get

$$\frac{1}{\sqrt{n}} \widehat{\lambda} = \left[A \left(\frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \right)^{-1} A' \right]^{-1} \sqrt{n}(A\widehat{\beta} - \mathbf{b}). \quad (46)$$

Using (43), (31), and the Slutsky theorem, we derive that

$$\frac{1}{\sqrt{n}} \widehat{\lambda} \xrightarrow{\mathcal{D}} N(0, CLC'). \quad (47)$$

By (43)–(45), it is easy to derive that, under the null hypothesis of testing problem (25),

$$T_n = \frac{1}{n} \widehat{\lambda}' (C_0 L_0 C_0')^{-1} \widehat{\lambda} \xrightarrow{\mathcal{D}} \chi^2(k). \quad (48)$$

(2) Under the alternative hypothesis and again by applying Lemma 11, we have

$$\begin{aligned} \sqrt{n}(A\widehat{\beta} - \mathbf{b}) &= \sqrt{n}A(\widehat{\beta} - \beta) \\ &\quad + \sqrt{n}(A\beta - \mathbf{b}) \xrightarrow{\mathcal{D}} N(\sqrt{n}(A\beta - \mathbf{b}), L). \end{aligned} \quad (49)$$

Using the similar argument as in (47), we can prove that

$$\begin{aligned} \frac{1}{\sqrt{n}} \widehat{\lambda} &= \left[A \left(\frac{1}{n} \widetilde{\mathbf{W}}' \widetilde{\mathbf{W}} - \Sigma_{uu} \right)^{-1} A' \right]^{-1} \sqrt{n}(A\widehat{\beta} - \mathbf{b}) \xrightarrow{\mathcal{D}} \\ &N(\sqrt{n}(A\beta - \mathbf{b}), CLC'). \end{aligned} \quad (50)$$

Then under the alternative hypothesis, we have

$$T_n = \frac{1}{n} \widehat{\lambda}' (C_0 L_0 C_0')^{-1} \widehat{\lambda} \xrightarrow{\mathcal{D}} \chi^2(k, \zeta), \quad (51)$$

where $\chi^2(k, \zeta)$ denotes the asymptotic noncentral chi-square distribution with k degrees of freedom and the noncentral parameter is

$$\zeta = \lim_{n \rightarrow \infty} n(A\beta - \mathbf{b})' (CLC')^{-1} (A\beta - \mathbf{b}). \quad (52)$$

\square

Proof of Theorem 6. Using (31), by some simple calculations, we get

$$\begin{aligned} \widehat{f}_1^R(z_1) &= (1, 0) (\mathcal{X}'_1 \Omega \mathcal{X}_1)^{-1} \mathcal{X}'_1 \Omega (I - S_2^c S_1)^{-1} \\ &\quad \times (I - S_2^c) (\mathbf{Y} - \mathbf{W} \widehat{\beta}_r) \\ &= (1, 0) (\mathcal{X}'_1 \Omega \mathcal{X}_1)^{-1} \mathcal{X}'_1 \Omega (I - S_2^c S_1)^{-1} (I - S_2^c) \\ &\quad \times [(F_1 + F_2) + (\varepsilon - \mathbf{U}\beta) + \mathbf{W}(\beta - \widehat{\beta}_r)] \\ &= T_1 + T_2 + T_3, \end{aligned} \quad (53)$$

where

$$\begin{aligned}
 T_1 &= (1, 0) \left(\mathcal{X}'_1 \Omega \mathcal{X}_1 \right)^{-1} \mathcal{X}'_1 \Omega (I - S_2^c S_1^c)^{-1} \\
 &\quad \times (I - S_2^c) (F_1 + F_2), \\
 T_2 &= (1, 0) \left(\mathcal{X}'_1 \Omega \mathcal{X}_1 \right)^{-1} \mathcal{X}'_1 \Omega (I - S_2^c S_1^c)^{-1} \\
 &\quad \times (I - S_2^c) (\varepsilon - \mathbf{U} \beta), \\
 T_3 &= (1, 0) \left(\mathcal{X}'_1 \Omega \mathcal{X}_1 \right)^{-1} \mathcal{X}'_1 \Omega (I - S_2^c S_1^c)^{-1} \\
 &\quad \times (I - S_2^c) \mathbf{W} (\beta - \hat{\beta}_r).
 \end{aligned} \tag{54}$$

Referring to the proof of Theorem 4.1 in Opsomer and Ruppert [1] and using Lemma 9, we obtain that

$$\begin{aligned}
 (I - S_2^c S_1^c)^{-1} (I - S_2^c) (F_1 + F_2) &= F_1 + O_p(h_1^2 + h_2^2), \\
 (I - S_2^c S_1^c)^{-1} (I - S_2^c) (\varepsilon - \mathbf{U} \beta) &= (\varepsilon - \mathbf{U} \beta) (I + o_p(1)), \\
 (I - S_2^c S_1^c)^{-1} (I - S_2^c) \mathbf{W} &= (\mathbf{X} + \mathbf{U} - E(\mathbf{X} | Z_2)) (I + o_p(1)).
 \end{aligned} \tag{55}$$

For Z_{i1} in a small neighborhood of z_1 and $|Z_{i1} - z_1| < h$, for each $f_1(Z_{i1})$, we use the following Taylor expansion:

$$\begin{aligned}
 f_1(Z_{i1}) &= f_1(z_1) + f'_1(z_1)(Z_{i1} - z_1) + \frac{f''_1(z_1)}{2}(Z_{i1} - z_1)^2 \\
 &\quad + o(h^2).
 \end{aligned} \tag{56}$$

Then we derive that

$$\begin{aligned}
 F_1 &= \begin{pmatrix} f_1(Z_{11}) \\ f_1(Z_{21}) \\ \vdots \\ f_1(Z_{n1}) \end{pmatrix} = \mathcal{X}_1 \begin{pmatrix} f_1(z_1) \\ f'_1(z_1) \end{pmatrix} \\
 &\quad + \begin{pmatrix} \frac{f''_1(z_1)}{2}(Z_{11} - z_1)^2 + o(h^2) \\ \vdots \\ \frac{f''_1(z_1)}{2}(Z_{n1} - z_1)^2 + o(h^2) \end{pmatrix}.
 \end{aligned} \tag{57}$$

By Lemma 8 and the usual nonparametric regression results, we can easily get

$$T_1 = f_1(z_1) + \frac{f''_1(z_1)}{2} \mu_2 h^2 + o_p(h^2). \tag{58}$$

Using the result that

$$\left(\mathcal{X}'_1 \Omega \mathcal{X}_1 \right)^{-1} = \frac{1}{np_1(z_1)} H^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix}^{-1} H^{-1} (I + o_p(1)) \tag{59}$$

with $H = \text{diag}(1, h)$, after some simple calculations, we can get

$$\sqrt{nh} T_2 = \frac{1}{np_1(z_1)} \sqrt{nh} \sum_{i=1}^n K_h(Z_{i1} - z_1) (\varepsilon_i - U'_i \beta) + o_p(1). \tag{60}$$

Note that $\sqrt{nh} (1/n) \sum_{i=1}^n K_h(Z_{i1} - z_1) (\varepsilon_i - U'_i \beta)$ is asymptotically normal with mean 0 and variance

$$\begin{aligned}
 v_0 p_1(z_1) E \left[(\varepsilon_1 - U'_1 \beta)^2 | Z_1 = z_1 \right] \\
 = v_0 p_1(z_1) (\sigma^2(z_1) + \beta' \Sigma_{uu} \beta).
 \end{aligned} \tag{61}$$

By the Slutsky theorem, we can derive that

$$\sqrt{nh} T_2 \xrightarrow{\mathcal{D}} N \left(0, \frac{\sigma^2(z_1) + \beta' \Sigma_{uu} \beta}{p_1(z_1)} v_0 \right). \tag{62}$$

Using similar calculation to that of T_2 and the result derived from Theorem 1 that $\beta - \hat{\beta}_r = O_p(n^{-1/2})$, then we have

$$\begin{aligned}
 \sqrt{nh} T_3 &= \sqrt{nh} (p_1(z_1))^{-1} \frac{1}{n} \sum_{i=1}^n K_h(Z_{i1} - z_1) \\
 &\quad \times (X_i - E(X_i | Z_2)) O_p(n^{-1/2}) \\
 &\quad + o_p(1) = o_p(1).
 \end{aligned} \tag{63}$$

By using the Slutsky theorem again, we derive the desired result. \square

Acknowledgments

Wang's researches are supported by the NSF project (ZR2011AQ007) of Shangdong Province of China and NNSF projects (11201499 and 11301309) of China.

References

- [1] S. J. D. Opsomer and D. Ruppert, "Fitting a bivariate additive model by local polynomial regression," *Annals of Statistics*, vol. 25, no. 1, pp. 186–211, 1997.
- [2] Q. Li, "Efficient estimation of additive partially linear models," *International Economic Review*, vol. 41, no. 4, pp. 1073–1092, 2002.
- [3] J. Jiang, H. Zhou, X. Jiang, and J. Peng, "Generalized likelihood ratio tests for the structure of semiparametric additive models," *The Canadian Journal of Statistics*, vol. 35, no. 3, pp. 381–398, 2007.
- [4] W. A. Fuller, *Measurement Error Models*, John Wiley & Sons, New York, NY, USA, 1987.
- [5] C. L. Cheng and J. W. Van Ness, *Statistical Regression with Measurement Error*, Arnold, London, UK, 1999.
- [6] R. J. Carroll, D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu, *Measurement Error in Nonlinear Models*, Chapman and Hall, New York, NY, USA, 2nd edition, 2006.

- [7] H. Liang, W. Härdle, and R. J. Carroll, "Estimation in a semiparametric partially linear errors-in-variables model," *The Annals of Statistics*, vol. 27, no. 5, pp. 1519–1535, 1999.
- [8] J. You and G. Chen, "Estimation of a semiparametric varying-coefficient partially linear errors-in-variables model," *Journal of Multivariate Analysis*, vol. 97, no. 2, pp. 324–341, 2006.
- [9] L. Li and T. Greene, "Varying coefficients model with measurement error," *Biometrics*, vol. 64, no. 2, pp. 519–526, 2008.
- [10] X. Wang, F. Chen, and L. Lin, "Empirical likelihood inference for the parameter in additive partially linear EV models," *Communications in Statistics*, vol. 39, no. 19, pp. 3513–3524, 2010.
- [11] H. Liang, S. W. Thurston, D. Ruppert, T. Apanasovich, and R. Hauser, "Additive partial linear models with measurement errors," *Biometrika*, vol. 95, no. 3, pp. 667–678, 2008.
- [12] D. W. Jorgenson, *Econometric Modeling of Producer Behavior*, MIT Press, Cambridge, Mass, USA, 2000.
- [13] C. Wei and Q. Wang, "Statistical inference on restricted partially linear additive errors-in-variables models," *Test*, vol. 21, no. 4, pp. 757–774, 2012.
- [14] C. H. Wei and X. Z. Wu, "Profile Lagrange multiplier test for partially linear varying-coefficient regression models," *Journal of Systems Science and Mathematical Sciences*, vol. 28, no. 4, pp. 416–424, 2008.
- [15] W. Zhang, G. Li, and L. Xue, "Profile inference on partially linear varying-coefficient errors-in-variables models under restricted condition," *Computational Statistics & Data Analysis*, vol. 55, no. 11, pp. 3027–3040, 2011.
- [16] C. H. Wei and C. Liu, "Statistical inference on semi-parametric partial linear additive models," *Journal of Nonparametric Statistics*, vol. 24, no. 4, pp. 809–823, 2012.
- [17] J. Fan and T. Huang, "Profile likelihood inferences on semiparametric varying-coefficient partially linear models," *Bernoulli*, vol. 11, no. 6, pp. 1031–1057, 2005.

