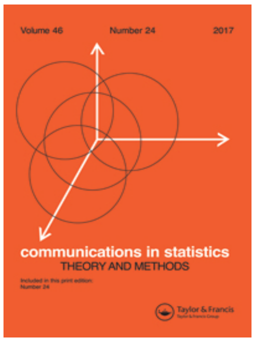

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

Efficiency of the generalized-difference-based weighted mixed almost unbiased two-parameter estimator in partially linear model

Fikri Akdeniz & Mahdi Roozbeh

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2017, VOL. 46, NO. 24, 12259–12280
https://doi.org/10.1080/03610926.2017.1295075

Taylor & Francis
Taylor & Francis Group

Check for updates

Efficiency of the generalized-difference-based weighted mixed almost unbiased two-parameter estimator in partially linear model

Fikri Akdeniz[a] and Mahdi Roozbeh[b]

[a]Department of Mathematics and Computer Science, Çağ University, Tarsus-Mersin, Turkey; [b]Faculty of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran

ABSTRACT
In this paper, a generalized difference-based estimator is introduced for the vector parameter β in partially linear model when the errors are correlated. A generalized-difference-based almost unbiased two-parameter estimator is defined for the vector parameter β. Under the linear stochastic constraint $r = R\beta + e$, we introduce a new generalized-difference-based weighted mixed almost unbiased two-parameter estimator. The perfor*mance of this new estimator over the generalized-difference-based estimator and generalized- dif*ference-based almost unbiased two-parameter estimator in terms of the MSEM criterion is investigated. The efficiency properties of the new estimator is illustrated by a simulation study. Finally, the performance of the new estimator is evaluated for a real dataset.

ARTICLE HISTORY
Received 24 April 2016
Accepted 7 February 2017

KEYWORDS
Difference-based estimator; Generalized-difference-based almost unbiased ridge estimator; Generalized-difference-based weighted mixed almost unbiased two-parameter ridge estimator; Partially linear model; Weighted mixed estimator.

MATHEMATICS SUBJECT CLASSIFICATION
Primary 62G08; Secondary 62J07

1. Introduction

Partially linear models have received considerable attention in statistics and econometrics. They have a wide range of applications. In these models, some of the relations are believed to be of certain parametric form, while others are not easily parameterized. Consider the partially linear model

$$y_i = x_i'\beta + f(u_i) + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{1}$$

where $x_i' = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is a vector of explanatory variables, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$ is an unknown p-dimensional parameter vector, the u_i are known and non random in some bounded domain $D \subset \Re$, $f(.)$ is an unknown smooth function, and ε_i's are independent and identically distributed random errors with $E(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma^2$, and are independent of (x_i, u_i). We shall call $f(u)$ the smooth part of the model and assume that it represents a smooth unparametrized functional relationship. The u's have bounded support, say the unit interval, and have been rearranged so that $u_1 \leq u_2 \leq \cdots \leq u_n$.

CONTACT Mahdi Roozbeh mahdi.roozbeh@semnan.ac.ir Department of Mathematics, Statistics, and Computer Sciences, Faculty of Mathematics, Semnan University, Semnan 35195-363, Iran.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lsta.

The goal is to estimate the unknown parameter vector β and non parametric function *f(u)* from the data $\{y_i, x_i, u_i\}$. In vector/matrix notation, the model (1) is written as

$$y = X\beta + f + \varepsilon \tag{2}$$

where $y = (y_1, \ldots, y_n)', \; X = [x_1, \ldots, x_n]', \; f = (f(u_1), \ldots, f(u_n))', \; \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$.

Partially linear regression models are more flexible than the standard linear regression models, since they combine both parametric and non parametric components when it is believed that the response variable y depends on variable X in a linear way but is non linearly related to other independent variable U.

In model (2), Yatchew's method does not require an estimator of the function *f*(.) and are often called difference-based estimation procedure, provided that *f*(.) is differentiable and the u_i's are closely spaced, it is possible to remove the effect of the function *f*(.) by differencing the data appropriately (Yatchew 2003).

In regression analysis, researchers often encounter the problem of multicollinearity. In case of multicollinearity, we know that the correlation matrix might have one or more small eigenvalues which causes the estimates of the regression coefficients to be large in absolute value. The least squares estimator performs poorly in the presence of multicollinearity. Multicollinearity is defined as the existence of nearly linear dependency among column vectors of the design matrix X in the linear model $y = X\beta + \varepsilon$. The existence of multicollinearity may lead to wide confidence intervals for the individual parameters or linear combination of the parameters and may produce estimates with wrong signs. Condition number is a measure of the presence of multicollinearity. If $X'X$ is ill conditioned with a large condition number, ridge regression estimator (Hoerl and Kennard 1970) can be used to estimate β.

To apply shrinkage estimators is well known as an efficient remedial measure in order to solve problems caused by multicollinearity. We assume that the condition number of the parametric component is large indicating that a biased estimation procedure is desirable. Its parametric part has the same structural form as the classical methods.

Akdeniz and Tabakan (2009) introduced a ridge estimator for the vector of parameters in a partially linear regression model when additional linear restrictions on the parameter vector are assumed to hold. A difference-based ridge regression estimator of regression parameters in the partial linear model is given in Tabakan and Akdeniz (2010). The difference-based estimation procedure is optimal in the sense that the estimator of the linear component is asymptotically efficient and asymptotically minimax rate optimal for the partial linear model. (Wang, Brown, and Cai 2011). Arumairajan and Wijekoon (2014) proposed stochastic restricted ordinary ridge estimator. Arashi et al. (2015) considered the estimation of the restricted ridge regression parameter in singular models. Arashi and Valizadeh (2015) proposed several estimators for estimating the biasing parameter in the study of partial linear models in the presence of multicollinearity. Generalized-difference-based estimator is introduced for the vector parameter β in the semiparametric regression model when the errors are correlated, by Akdeniz et al. (2015). Wu (2016) proposed a difference-based almost unbiased Liu estimator for the vector of β in partial linear model. Roozbeh (2015) obtained the necessary and sufficient conditions for the superiority of the shrinkage ridge type estimator over its counterpart in the semiparametric regression model when the errors are dependent and some non stochastic linear restrictions are imposed under a multicollinearity setting.

In this paper, a generalized-difference-based restricted estimator is introduced for the vector parameter β in partially linear model when the errors are correlated. A generalized-difference-based ridge estimator is defined for the vector parameter β. Under the linear stochastic constraint $r = R\beta + e$, we introduced a new generalized-difference-based weighted mixed almost unbiased two-parameter estimator.

2. Difference-based estimator

In this section, we use a difference-based technique to estimate the linear regression coefficient vector β. This technique has been used to remove the non parametric component in partially linear model by various authors (e.g., Yatchew 1997, 2003 and Klipple and Eubank 2007). Consider the following partially linear model

$$y = X\beta + f + \varepsilon \tag{3}$$

Yatchew (1997) suggested estimating β on the basis of the mth order differencing equation

$$\sum_{j=0}^{m} d_j y_{k-j} = \left(\sum_{j=0}^{m} d_j x_{k-j}\right)\beta + \left(\sum_{j=0}^{m} d_j f(u_{k-j})\right) + \left(\sum_{j=0}^{m} d_j \varepsilon_{k-j}\right), \quad k = m+1, \ldots, n \tag{4}$$

Now let $d = (d_0, d_1, \ldots, d_m)'$ be a $(m+1)$ vector, where m is the order of differencing and $d_0, d_1, \ldots, d_m$ are differencing weights that minimize $\min_{d_0,\ldots,d_m} \delta = \sum_{l=1}^{m} (\sum_{j=0}^{m-l} d_j d_{l+j})^2$ satisfying the conditions:

$$\sum_{j=0}^{m} d_j = 0 \quad \text{and} \quad \sum_{j=0}^{m} d_j^{\,2} = 1 \tag{5}$$

Let us define the $(n-m) \times n$ differencing matrix D whose elements satisfy Equation (5) as

$$D = \begin{bmatrix} d_0 & d_1 & d_2 & . & . & . & d_m & 0 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & d_0 & d_1 & d_2 & . & . & . & d_m & 0 & 0 & 0 & . & . & . & 0 \\ . & . & & & & & & & & & & & & & \\ . & . & . & & & & & & & & & & & & \\ . & . & . & . & & & & & & & & & & & \\ 0 & 0 & 0 & . & . & . & & d_0 & d_1 & d_2 & . & . & . & d_m & 0 \\ 0 & 0 & 0 & . & . & . & 0 & 0 & d_0 & d_1 & . & . & . & d_{m-1} & d_m \end{bmatrix}$$

The goal of this step is to eliminate the effect of the non parametric component $f(.)$. Applying the differencing matrix to model (3) permits direct estimation of the parametric effect. As a result of developments in Roozbeh, Arashi, and Niroumand (2011), it is known that the parameter vector β in Equation (3) can be estimated with parametric efficiency. We now show the difference-based estimators that can be used for this purpose. Since the data have been reordered $0 \leq u_1 \leq u_2 \leq \cdots \leq u_n \leq 1$ so that the values of the non parametric variable(s) are close, the application of the differencing matrix D in model (3) removes the non parametric effect in large samples (Yatchew 2003). If $f(.)$ is an unknown function that is the inferential object and has a bounded first derivative, then $Df(.)$ is close to 0, so that applying the differencing matrix we have

$$Dy = DX\beta + Df + D\varepsilon \tag{6}$$

which is approximately equal to $DX\beta + D\varepsilon$, or

$$\tilde{y} \cong \tilde{X}\beta + \tilde{\varepsilon} \tag{7}$$

where $\tilde{y} = Dy$, $\tilde{X} = DX$, and $\tilde{\varepsilon} = D\varepsilon$. The role of the constraints (5) is now evident (Yatchew 2003 p. 57, Klipple and Eubank 2007). Yatchew (2003) defines a simple differencing estimator of the parameter β in the semiparametric regression model. Thus, standard linear model considerations suggest estimating β by

$$\hat{\beta}_{\text{diff}} = \left[(DX)'(DX)\right]^{-1}(DX)'(Dy) \tag{8}$$

This estimator was first proposed in Yatchew (2003). Thus, differencing allows one to perform inferences on β as if there were no non parametric component $f(.)$ in model (7) (see Yatchew 1997).

3. Partially linear models with correlated errors

In this section, we consider the following partially linear model

$$y = X\beta + f + \varepsilon$$

with $E(\varepsilon) = 0$, and $E(\varepsilon\varepsilon') = \sigma^2 V$. So, $\tilde{\varepsilon} = D\varepsilon$ is a $(n-m)$vector of disturbances distributed with

$$E(\tilde{\varepsilon}) = 0 \quad \text{and } E(\tilde{\varepsilon}\tilde{\varepsilon}') = \sigma^2 DVD' = \sigma^2 V_D \tag{9}$$

where $V_D = DVD' \neq I_{n-m}$ is a known $(n-m) \times (n-m)$ symmetric positive definite (p.d.) matrix and $\sigma^2 > 0$ is an unknown parameter (see Roozbeh, Arashi, and Niroumand 2011). It is well known that adopting the linear model (7), the unbiased estimator of β is the following *generalized-difference-based estimator* (GDE) given by

$$\hat{\beta}_{\text{GDE}} = \left(\tilde{X}'V_D^{-1}\tilde{X}\right)^{-1}\tilde{X}'V_D^{-1}\tilde{y} \tag{10}$$

It is observed from Equation (10) that the properties of the GDE of β depends on the characteristics of the information matrix $G = \tilde{X}'V_D^{-1}\tilde{X}$.

If G:$p \times p$, $p << n - m$ matrix is ill-conditioned with a large condition number, then the $\hat{\beta}_{\text{GDE}}$ produces large sampling variances. Moreover, some regression coefficients may be statistically insignificant, and meaningful statistical inference becomes difficult for the researcher. We assume that the condition number of the G matrix is large indicating that a biased estimation procedure is desirable.

4. Weighted mixed regression and estimation of parameters

The use of prior information in linear regression analysis is well known to provide more efficient estimators of regression coefficients. The available prior information sometimes can be expressed in the form of exact, stochastic, or inequality restrictions.

We consider model (7):

$$\tilde{y} \cong \tilde{X}\beta + \tilde{\varepsilon}, \quad \tilde{\varepsilon} \sim (0, \sigma^2 DVD') = (0, \sigma^2 V_D) \tag{11}$$

When a set of stochastic linear constraints binding the regression coefficients in a linear regression model is available, Theil and Goldberger (1961) have proposed the method of mixed regression estimation. Their method typically assumes that the prior information in the form of stochastic linear constraints and sample information in the form of observations on the study variable and explanatory variables are equally important and therefore receive equal weights in the estimation procedure.

Totally, we do not have exact prior information such as $R\beta = r$, involving estimation of economic relations, industrial structures, production planning, etc. Therefore, stochastic uncertainty occurs in specifying linear programming due to economic and financial studies.

In addition to sample model (11), it is supposed that a set of stochastic linear constraints binding the regression coefficients is available in the form of independent prior information:

$$r = R\beta + e, \quad e \sim (0, \sigma^2 W) \tag{12}$$

where R is a $q \times p$ known matrix with rank $(R) = q$, and e is a $q \times 1$ vector of disturbances. W is assumed to be known and positive definite, the $q \times 1$ vector r can be interpreted as a random variable with expectation $E(r) = R\beta$. Therefore the restriction in Equation (12) does not hold exactly but in the mean and we assume r to be known, that is to be a realized value of the random vector, so that all expectations are conditional on r as, for example, $E(\hat{\beta}|r)$(Rao, Toutenburg, and Shalab 2008). In order to take the information in Equation (12) into account while constructing estimators $\hat{\beta}$ for β, we require that $E(R\hat{\beta}|r) = r$ (see Toutenburg et al. 2003).

In model (12), we have assumed the structure of the dispersion matrix of e, $E(ee') = \sigma^2 W$, that is, with the same factor of proportionality σ^2 as occurred in the sample model. Therefore, it may some times be more realistic to suppose that $E(ee') = W$. It is also assumed that the random vector ε is stochastically independent of e.

When the sample information given by Equation (11) and prior information is described by Equation (12) are to be assigned not necessarily equal weights on the basis of some extraneous considerations in the estimation of regression parameters, Schaffrin and Toutenburg (1990) have proposed the method of weighted mixed regression estimation. Following their technique, we obtain the generalized-difference-based weighted mixed estimator of β.

In order to incorporate the restrictions in Equation (12) in the estimation of parameters, we minimize

$$(\tilde{y} - \tilde{X}\beta)' V_D^{-1} (\tilde{y} - \tilde{X}\beta) + \omega (r - R\beta)' W^{-1} (r - R\beta) \tag{13}$$

with respect to β. This leads to the following solution for β:

$$\hat{\beta}_{\text{GDWME}}(\omega) = \hat{\beta}(\omega) = (G + \omega R' W^{-1} R)^{-1} \left(\tilde{X}' V_D^{-1} \tilde{y} + \omega R' W^{-1} r \right) \tag{14}$$

where ω is a non stochastic and non negative scalar weight with $0 \leq \omega \leq 1$ ($\omega = 0$ would lead to $\hat{\beta}_{\text{GDE}}$). It is seen that a value of ω between 0 and 1 specifies an estimator in which the prior information receives less weight in comparison with the sample information. On the other hand, a value of ω greater than 1 implies higher weight to the prior information which, of course, may be of little practical interest. Since

$$(G + \omega R' W^{-1} R)^{-1} = G^{-1} - \omega G^{-1} R' (W + \omega R G^{-1} R')^{-1} R G^{-1} \tag{15}$$

we have

$$\hat{\beta}_{\text{GDWME}}(\omega) = \hat{\beta}_{\text{GDE}} + \omega G^{-1} R' (W + \omega R G^{-1} R')^{-1} (r - R\hat{\beta}_{\text{GDE}}) \tag{16}$$

which is called *generalized-difference-based weighted mixed estimator* (GDWME).

If we substitute $\omega = 1$ in Equation (14) , we get

$$\hat{\beta}_{\text{GDME}} = (G + R' W^{-1} R)^{-1} \left(\tilde{X}' V_D^{-1} \tilde{y} + R' W^{-1} r \right) \tag{17}$$

which is the *generalized-difference-based mixed estimator* (GDME). The ordinary mixed estimator was proposed by Theil and Golberger (1961). This estimator gives equal weight to sample and prior information.

Generalized ridge estimator proposed by Hoerl and Kennard (1970) is defined as

$$\begin{aligned}\hat{\beta}_{\mathrm{GDRE}}(k) &= \left(\tilde{X}'V_D^{-1}\tilde{X} + kI\right)^{-1}\tilde{X}'V_D^{-1}\tilde{y}\\ &= (G+kI)^{-1}\tilde{X}'V_D^{-1}\tilde{y} = G_k^{-1}\tilde{X}'V_D^{-1}\tilde{y}\\ &= (I+kG^{-1})^{-1}\hat{\beta}_{\mathrm{GDE}} = T_k\hat{\beta}_{\mathrm{GDE}} \qquad (18)\end{aligned}$$

where $G_k = G + kI$, $T_k = G_k^{-1}G = GG_k^{-1}$.

Akdeniz and Erol (2003) discussed the almost unbiased ridge regression estimator (AURE), which is given as follows:

$$\hat{\beta}_{\mathrm{AURE}}(k) = \left(I - k^2S_k^{-2}\right)\hat{\beta}_{\mathrm{OLS}} \qquad (19)$$

where $S_k = X'X + kI$, $\hat{\beta}_{\mathrm{OLS}} = (X'X)^{-1}X'y$. Similarly, we define the *generalized-difference-based almost unbiased ridge estimator* (GDAURE) as follows:

$$\hat{\beta}_{\mathrm{GDAURE}}(k) = \left(I - k^2G_k^{-2}\right)\hat{\beta}_{\mathrm{GDE}}. \qquad (20)$$

Li and Yang (2010) proposed the stochastic mixed ridge estimator, Liu, Yang, and Wu (2013) introduced the weighted mixed almost unbiased ridge estimator based on the weighted mixed estimator. Liu et al. (2014) considered two kinds of weighted mixed almost unbiased estimators in a linear stochastic restricted regression model. Substituting $\hat{\beta}_{\mathrm{GDE}}$ with $\hat{\beta}_{\mathrm{GDAURE}}(k)$ in $\hat{\beta}_{\mathrm{GDWME}}(\omega)$, we describe a *generalized-difference-based weighted mixed almost unbiased ridge estimator* (GDWMAURE), as follows:

Since

$$(W + \omega RG^{-1}R')^{-1} = W^{-1} - \omega W^{-1}R(G + \omega R'W^{-1}R)^{-1}R'W^{-1} \qquad (21)$$

and

$$\begin{aligned}\omega G^{-1}R'(W + \omega RG^{-1}R')^{-1}r &= wG^{-1}R'\left[W^{-1} - \omega W^{-1}R(G + \omega R'G^{-1}R)^{-1}R'W^{-1}\right]r\\ &= \left[G^{-1} - \omega G^{-1}R'W^{-1}R(G + \omega R'W^{-1}R)^{-1}\right]\omega R'W^{-1}r\\ &= \left[G^{-1} - G^{-1}\{\,(G + \omega R'W^{-1}R) - G\,\}(G + \omega R'W^{-1}R)^{-1}\right]\\ &\quad \times \omega R'W^{-1}r\\ &= \left[G^{-1} - G^{-1}\left[I - G(G + \omega R'W^{-1}R)^{-1}\right]\right]\omega R'W^{-1}r\\ &= (G + \omega R'W^{-1}R)^{-1}\omega R'W^{-1}r \qquad (22)\end{aligned}$$

Using the equalities in Equations (20) and (22), we have

$$\begin{aligned}\hat{\beta}_{\mathrm{GDWMAURE}}(\omega, k) &= \hat{\beta}_{\mathrm{GDAURE}}(k) + \omega G^{-1}R'(W + \omega RG^{-1}R')^{-1}(r - R\hat{\beta}_{\mathrm{GDAURE}}(k))\\ &= \left(I - k^2G_k^{-2}\right)\hat{\beta}_{\mathrm{GDE}} + \omega G^{-1}R'(W + \omega RG^{-1}R')^{-1}\left(r - R\left(I - k^2G_k^{-2}\right)\hat{\beta}_{\mathrm{GDE}}\right)\\ &= (G + \omega R'W^{-1}R)^{-1}\left[\left(I - k^2G_k^{-2}\right)\tilde{X}'V_D^{-1}\tilde{y} + \omega R'W^{-1}r\right] \qquad (23)\end{aligned}$$

In fact, from the definition of $\hat{\beta}_{\mathrm{GDWMAURE}}(\omega, k)$, we can see that $\hat{\beta}_{\mathrm{GDWMAURE}}(\omega, k)$ is general estimator, and which includes the $\hat{\beta}_{\mathrm{GDE}}$, $\hat{\beta}_{\mathrm{GDME}}$, $\hat{\beta}_{\mathrm{GDAURE}}(k)$, and $\hat{\beta}_{\mathrm{GDWME}}(\omega)$ as special cases. Namely,

$$\text{if } k = 0, \omega = 0, \quad \text{then } \hat{\beta}_{\mathrm{GDWMAURE}}(0, 0) = \hat{\beta}_{\mathrm{GDE}} \qquad (24)$$

$$\text{if } k = 0 \quad \text{and } \omega = 1, \quad \text{then } \hat{\beta}_{\mathrm{GDWMAURE}}(\omega = 1, k = 0) = \hat{\beta}_{\mathrm{GDME}} \qquad (25)$$

if $\omega = 0$, then $\hat{\beta}_{\mathrm{GDWMAURE}}(\omega = 0, k) = G^{-1}(I - k^2G_k^{-2})\tilde{X}'V_D^{-1}\tilde{y}$, Observing that G^{-1} and $(I - k^2G_k^{-2})$ are commutative, we have

$$\hat{\beta}_{\text{AU}}(\omega = 0, k) == \left(I - k^2G_k^{-2}\right) G^{-1}\tilde{X}'V_D^{-1}\tilde{y} = \left(I - k^2G_k^{-2}\right) \hat{\beta}_{\text{GDE}} = \hat{\beta}_{\text{GDAURE}}(k) \tag{26}$$

$$\begin{aligned} \text{if } k = 0, \text{ then } \hat{\beta}_{\text{AU}}(\omega, k = 0) &= (G + \omega R'W^{-1}R)^{-1}\left[\tilde{X}'V_D^{-1}\tilde{y} + \omega R'W^{-1}r\right] \\ &= \hat{\beta}_{\text{GDWME}}(\omega), \end{aligned} \tag{27}$$

Following, Özkale and Kaçıranlar (2007) and Wu and Yang (2013), we may define the *generalized-difference-based almost unbiased two-parameter estimator* (GDAUTPE) in semi-parametric partial linear models as follows:

$$\hat{\beta}_{\text{GDAUTPE}}(k, d) = \left[I - k^2(1 - d)^2\left(\tilde{X}'V_D^{-1}\tilde{X} + kI\right)^{-2}\right]\hat{\beta}_{\text{GDE}} \tag{28}$$

The bias vector and MSE matrix of $\hat{\beta}_{\text{GDAUTPE}}(k, d)$ can be obtained as

$$b_1 = \text{Bias}(\hat{\beta}_{\text{GDAUTPE}}(k, d) = -k^2(1 - d)^2(G + kI)^{-2}\beta = -(I - G_{kd})^2\beta$$

and

$$\begin{aligned} \text{MSEM}(\hat{\beta}_{\text{GDAUTPE}}(k, d)) &= \text{Cov}(\hat{\beta}_{\text{GDAUTPE}}(k, d)) + b_1b_1{}' \\ &= \sigma^2G_{kd}G^{-1}G_{kd} + b_1b_1{}' \end{aligned} \tag{29}$$

where $G_{kd} = I - k^2(1 - d)^2G_K^{-2}$.

$\hat{\beta}_{\text{GDAUTPE}}(k, d)$ is a generalization of the *generalized-difference-based almost unbiased ridge estimator* (GDAURE) and the *generalized-difference-based almost unbiased Liu estimator* (GDAULE):

$$\text{when } d = 0, \quad \text{then } \hat{\beta}_{\text{GDAUTPE}}(k, d = 0) = \left(I - k^2G_K^{-2}\right)\hat{\beta}_{\text{GDE}} = \hat{\beta}_{\text{GDAURE}}(k) \tag{30}$$

$$\text{when } k = 1, \quad \text{then } \hat{\beta}_{\text{GDAUTPE}}(k = 1, d) = (I - (1 - d)^2(G + I)^{-2})\hat{\beta}_{\text{GDE}} = \hat{\beta}_{\text{GDAULE}}(d) \tag{31}$$

A modified two-parameter estimator is introduced for the vector of parameters in the linear regression model in Dorugade (2014).

5. Mean squared error matrix comparisons of estimators

In this section, we compare the underlying estimators. For the convenience of the following discussions, we give some lemmas here.

Lemma 1 (Farebrother 1976). *Let A be a positive definite matrix, namely $A > 0$, and let α be some vector, then $A - \alpha\alpha' \geq 0$ if and only if $\alpha'A^{-1}\alpha \leq 1$.*

Lemma 2. *Let $n \times n$ matrices $M > 0$ and $N \geq 0$, then $M > N$ if and only if $\lambda_{\max}(NM^{-1}) < 1$. (see Rao, Toutenburg, and Shalab 2008).*

Lemma 3 (Trenkler and Toutenburg 1990). *Let $\hat{\beta}_j = A_jy$, $j = 1, 2$ be two competing estimators of β. Suppose that $\Delta = \text{Cov}(\hat{\beta}_1) - \text{Cov}(\hat{\beta}_2) > 0$. Then*

$$\text{MSEM}(\hat{\beta}_1) - \text{MSEM}(\hat{\beta}_2) \geq 0$$

if and only if $b_2{}'(\Delta + b_1b_1{}')^{-1}b_2 \leq 1$, where b_j denotes bias vector of $\hat{\beta}_j$.

The mean squared error matrix (MSEM) of an estimator $\tilde{\beta}$ is defined as

$$\text{MSEM}(\tilde{\beta}) = E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)' = \text{Cov}(\tilde{\beta}) + \text{Bias}(\tilde{\beta})\text{Bias}(\tilde{\beta})'$$

where $\text{Cov}(\tilde{\beta})$ is the dispersion matrix and $\text{Bias}(\tilde{\beta}) = E(\tilde{\beta}) - \beta$ is the bias vector.

5.1. Proposed estimator: Generalized-difference-based weighted mixed almost unbiased two-parameter estimator

Using Equation (28) , we have generalized-difference-based weighted mixed almost unbiased two-parameter estimator (GDWMAUTPE):

$$\begin{aligned}\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d) = \hat{\beta}_{\text{GDAUTPE}}(k, d) + \omega G^{-1}R'(W + \omega RG^{-1}R')^{-1} \\ \times (r - R\hat{\beta}_{\text{GDAUTPE}}(k, d))\end{aligned} \tag{32}$$

$$B\left[\{ I - k^2(1-d)^2 G_k^{-2} \} \tilde{X}' V_D^{-1} \tilde{y} + \omega R' W^{-1} r\right] \tag{33}$$

where $B = (G + \omega R' W^{-1} R)^{-1}$.

$$\text{Bias}(\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d)) = -k^2(1-d)^2 B G_K^{-2} G\beta = b_2 \tag{34}$$

and

$$\text{MSEM}(\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d)) = \sigma^2 B(G_{kd} G G_{kd} + \omega^2 R' W^{-1} R)B + b_2 b_2' \tag{35}$$

5.2. MSEM comparison between $\hat{\beta}_{\text{GDAUTPE}}(k, d)$ and $\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d)$

Theorem 1. *Let* $k > 0$, $0 < d < 1$, *and* $0 < \omega < 1$. *When* $\lambda_{\max}(NM^{-1}) < 1$, $\hat{\beta}_{GD}(\omega, k, d) =: \tilde{\beta}_2$ *is superior to* $\hat{\beta}_{\text{GDAUTPE}}(k, d) =: \tilde{\beta}_1$ *in the MSEM sense, namely* $\Delta_1 \geq 0$ *if and only if* $b_2'(\Phi + b_1 b_1')^{-1} b_2 \leq 1$.

Proof. We consider the MSEM difference between $\hat{\beta}_{\text{GDAUTPE}}(k, d) =: \tilde{\beta}_1$ and $\hat{\beta}_{\text{GD}}(\omega, k, d) =: \tilde{\beta}_2$ as

$$\begin{aligned}\Delta_1 &= \text{MSEM}(\hat{\beta}_{\text{GDAUTPE}}(k, d)) - \text{MSEM}(\hat{\beta}_{\text{GD}}(\omega, k, d)) \\ &= \sigma^2 G_{kd} G^{-1} G_{kd} - \sigma^2 B G_{kd} G G_{kd} B' + B\omega^2 R' W^{-1} R B + b_1 b_1' - b_2 b_2' \\ &= \Phi + b_1 b_1' - b_2 b_2'\end{aligned} \tag{36}$$

Now consider

$$\begin{aligned}\Phi &= \text{Cov}(\tilde{\beta}_1) - \text{Cov}(\tilde{\beta}_2) = \sigma^2 G_{kd} G^{-1} G_{kd} - \sigma^2 \left[B G_{kd} G G_{kd} B + B\omega^2 R' W^{-1} R B\right] \\ &= \sigma^2 (M - N)\end{aligned} \tag{37}$$

where $M = G_{kd} G^{-1} G_{kd}$ and $N = [B G_{kd} G G_{kd} B + B\omega^2 R' W^{-1} R B]$.

Note that $M = G_{kd} G^{-1} G_{kd} > 0$, and $N = [B(G_{kd} G G_{kd} + \omega^2 R' W^{-1} R)B] > 0$. When $\lambda_{\max}\{B(G_{kd} G G_{kd} + \omega^2 R' W^{-1} R)B\ (G_{kd} G^{-1} G_{kd})^{-1}\} < 1$, we can get that $\Phi > 0$, by applying Lemma 2. Thus, from Equation (36) and applying Lemma 3, we have $\Delta_1 \geq 0$ if and only if $b_2'(\Phi + b_1 b_1')^{-1} b_2 \leq 1$. This result completes the proof. □

5.3. MSEM comparison of the GDE and GDWMAUTPE

First, a difference-based model (7) can be transformed to a canonical form by the orthogonal transformation. A symmetric matrix $G = \tilde{X}'V_D^{-1}\tilde{X}$ has an eigenvalue-eigenvector decomposition of the form $G = Q\Lambda Q'$, where Q is an orthogonal matrix such that

$$Q'GQ = Q'\tilde{X}'V_D^{-1}\tilde{X}Q = Q'\tilde{X}'V_D^{-1/2}V_D^{-1/2}\tilde{X}Q = \bar{X}'\bar{X} = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$$

where λ_i is the eigenvalue of G. Then, we get canonical form of model (7) as

$$V_D^{-1/2}\tilde{y} = V_D^{-1/2}\tilde{X}\beta + V_D^{-1/2}\tilde{\varepsilon}$$

and

$$y^* = X^*\beta + \varepsilon^* = X^*QQ'\beta + \varepsilon^* = \bar{X}\bar{\beta} + \varepsilon^* \tag{38}$$

where $y^* = V_D^{-1/2}\tilde{y}$, $\bar{X} = X^*Q = V_D^{-1/2}\tilde{X}Q$, $\varepsilon^* = V_D^{-1/2}\tilde{\varepsilon}$, $\bar{\beta} = Q'\beta$.

Finally, we compare MSEM values between the GDE and GDWMAUTPE. From Equations (10) and (35), the difference of MSEM values between the GDE and GDWMAUTPE can be computed by

$$\Delta_2 = \text{MSEM}(\hat{\beta}_{\text{GDE}}) - \text{MSEM}(\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d)). \tag{39}$$

Theorem 2. *The GDWMAUTPE is superior to the GDE according to the MSEM criterion, namely $\Delta_2 \geq 0$ if and only if*

$$b_2'\left[\sigma^2(G^{-1} - B(G_{kd}GG_{kd} + \omega^2R'W^{-1}R)B\right]^{-1}b_2 \leq 1.$$

Proof. Using Equations (10) and (35), we obtain

$$\Delta_2 = (\sigma^2G^{-1} - \sigma^2B) + (\sigma^2B - \text{MSEM}(\hat{\beta}(\omega))) + (\text{MSEM}(\hat{\beta}(\omega)) - \text{MSEM}(\tilde{\beta}_2)) - b_2b_2' \tag{40}$$

where

$$\begin{aligned}
\text{MSEM}(\hat{\beta}(\omega)) &= \sigma^2B(G + \omega^2R'W^{-1}R)B \\
\text{MSEM}(\hat{\beta}_{\text{GDE}}) &= \sigma^2G^{-1} \\
\text{MSEM}(\tilde{\beta}_2) &= \sigma^2B(G_{kd}GG_{kd} + \omega^2R'W^{-1}R)B.
\end{aligned}$$

(i) Consider the following parts of Equation (40)

$$\begin{aligned}
\sigma^2G^{-1} - \sigma^2B &= \sigma^2\{\, G^{-1} - (G + \omega R'W^{-1}R)^{-1} \,\} \\
&= \sigma^2\{\, G^{-1} - (G^{-1} - wG^{-1}R'(W + \omega RG^{-1}R')^{-1}RG^{-1} \,\} \\
&= \sigma^2\omega G^{-1}R'(W + \omega RG^{-1}R')^{-1}RG^{-1} > 0
\end{aligned}$$

(ii) Observing that $B > 0$, we have

$$\begin{aligned}
\sigma^2B - \text{MSEM}(\hat{\beta}(\omega)) &= \sigma^2B - \sigma^2B(G + \omega^2R'W^{-1}R)B \\
&= \sigma^2BB^{-1}B - \sigma^2B(G + \omega^2R'W^{-1}R)B \\
&= \sigma^2B(B^{-1} - G - \omega^2R'W^{-1}R)B \\
&= \sigma^2B(G + \omega R'W^{-1}R - G - \omega^2R'W^{-1}R)B \\
&= \sigma^2\omega(1 - \omega)B(R'W^{-1}R)B > 0
\end{aligned}$$

(iii)

$$\begin{aligned}\mathrm{MSEM}(\hat{\beta}(\omega)) - \mathrm{MSEM}(\tilde{\beta}_2) &= \sigma^2 B(G\omega^2 R'W^{-1}R)B - \sigma^2 B(G_{kd}GG_{kd} + \omega^2 R'W^{-1}R)B \\ &= \sigma^2 B(G - G_{kd}GG_{kd})B\end{aligned}$$

where $G_{kd} = I - k^2(1-d)^2 G_k^{-2} = I - T_{kd}$. Since $G = Q\Lambda Q'$, we can easily compute that

$$\begin{aligned}D &= Q'(G - G_{kd}GG_{kd})Q \\ &= \Lambda - \left[I - k^2(1-d)^2(\Lambda + kI)^{-2}\right]\Lambda\left[I - k^2(1-d)^2(\Lambda + kI)^{-2}\right] \\ &= T_{kd}\Lambda + \Lambda T_{kd} - T_{kd}\Lambda T_{kd}\end{aligned}$$

D is a diagonal matrix and ith element is a positive number:

$$\begin{aligned}d_i &= k^2(1-d)^2\frac{2\lambda_i}{(\lambda_i + k)^2} - k^4(1-d)^4\frac{\lambda_i}{(\lambda_i + k)^4} \\ &= \frac{k^2(1-d)^2\lambda_i}{(\lambda_i + k)^4}\left[(\sqrt{2}-1)k + \sqrt{2}\lambda_i + kd\right]\left[\sqrt{2}(\lambda_i + k) + k(1-d)\right] > 0, \\ &\qquad i = 1, 2, \ldots, p\end{aligned}$$

which means that $D > 0$. Then we have $\sigma^2 BDB > 0$. Applying Lemma 1, we can get that $\Delta_2 \geq 0$, if and only if $b_2'\Delta_2^{-1}b_2 \leq 1$. This completes the proof. □

6. Selection of biasing parameters *k, d*, and non stochastic weight ω

In this section, we give a method about to choose k, d, and w. First, consider the difference-based model (7) and its canonical form as

$$y^* = X^*\beta + \varepsilon^* = X^*QQ'\beta + \varepsilon^* = \bar{X}\bar{\beta} + \varepsilon^* \tag{41}$$

Note that, $\mathrm{MSEM}(\hat{\bar{\beta}}_{\mathrm{GD}}(\omega, k, d) = Q'\mathrm{MSEM}(\hat{\beta}_{\mathrm{GD}}(\omega, k, d))Q$. We see that

$$\Omega^* = \mathrm{MSEM}(\hat{\bar{\beta}}_{\mathrm{GD}}(\omega, k, d)) = Q'\left[\sigma^2 B(G_{kd}GG_{kd} + \omega^2 R'W^{-1}R)B' + b_2b_2'\right]Q$$

where

$$B = (G + \omega R'W^{-1}R)^{-1}, \qquad G_{kd} = I - k^2(1-d)^2(G + kI)^{-2},$$
$$b_2 = -k^2(1-d)^2 BG_k^{-2}G\beta, \quad G_k = G + kI$$

It is easy to compute that

$$\begin{aligned}\Omega^* &= \sigma^2(\Lambda + \omega\Psi)^{-1}\left\{\left[I - k^2(1-d)^2(\Lambda + kI)^{-2}\right]\Lambda\left[I - k^2(1-d)^2(\Lambda + kI)^{-2}\right]\right. \\ &\quad \left. + w^2\Psi\right\}(\Lambda + \omega\Psi)^{-1} + k^4(1-d)^4(\Lambda + \omega\Psi)^{-1} \\ &\quad \times (\Lambda + kI)^{-2}\Lambda\bar{\beta}\bar{\beta}'\Lambda(\Lambda + kI)^{-2}(\Lambda + \omega\Psi)^{-1}\end{aligned}$$

Ω^* is a diagonal matrix and ith diagonal element is

$$\eta_i^* = (\lambda_i + \omega\xi_i)^{-1}\left(\left[1 - \frac{k^2(1-d)^2}{(\lambda_i + k)^2}\right]^2 \sigma^2\lambda_i + \sigma^2\omega^2\xi_i + \frac{k^4(1-d)^4\lambda_i^2\bar{\beta}_i^2}{(\lambda_i + k)^4}\right)(\lambda_i + \omega\xi_i)^{-1}.$$

Thus, we obtain

$$\begin{aligned}\mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d)) = tr\Omega^* = \\ &= \sum_{i=1}^{p} \frac{\sigma^2\lambda_i\left[(\lambda_i+k)^2 - k^2(1-d)^2\right]^2 + \sigma^2\omega^2\xi_i(\lambda_i+k)^4 + k^4(1-d)^4\lambda_i^2\bar{\beta}_i^2}{(\lambda_i+k)^4(\lambda_i+\omega\xi_i)^2} \\ &= \sum_{i=1}^{p} \frac{\sigma^2\{(1-\upsilon_i^2)^2\lambda_i + \omega^2\xi_i\} + \upsilon_i^4\lambda_i^2\bar{\beta}_i^2}{(\lambda_i+\omega\xi_i)^2} \end{aligned} \quad (42)$$

where $\nu_i = \frac{k(1-d)}{\lambda_i+k}$.

Optimal values of k, d, and ω can be derived by minimizing Equation (42) . First with fixed values of d and w, it is easy to see that

$$\frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial \upsilon_i} = \frac{2\sigma^2(1-\upsilon_i^2)(-2\upsilon_i)\lambda_i + 4\upsilon_i^3\lambda_i^2\bar{\beta}_i^2}{(\lambda_i+\omega\xi_i)^2} \quad (43)$$

or

$$2\sigma^2\left(1-\upsilon_i^2\right)(-2\upsilon_i)\lambda_i + 4\upsilon_i^3\lambda_i^2\bar{\beta}_i^2 = 0 \quad (44)$$

For a fixed d and ω, note that $\frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial k} = \frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial \upsilon_i}\frac{\partial \upsilon_i}{\partial k} = 0$ implies $\frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial \upsilon_i} = 0$, since $\frac{\partial \upsilon_i}{\partial k} = \frac{(1-d)\lambda_i}{(\lambda_i+k)^2} \neq 0$. Simplifying Equation (44) using ν_i, we obtain

$$\lambda_i\upsilon_i^2\bar{\beta}_i^2 - \sigma^2 + \sigma^2\upsilon_i^2 = 0 \quad (45)$$

and

$$\upsilon_i = \sqrt{\frac{\sigma^2}{\lambda_i\bar{\beta}_i^2 + \sigma^2}} = \frac{k(1-d)}{\lambda_i + k}$$

Thus, the optimal choice of the parameter k is

$$k_{i\mathrm{opt}} = \frac{\lambda_i\sigma}{(1-d)\sqrt{\lambda_i\bar{\beta}_i^2 + \sigma^2} - \sigma} \quad (46)$$

After the unknown parameters σ^2 and $\bar{\beta}_i^2$ are replaced by their unbiased estimators, we get optimal estimator of k for a fixed d value as

$$\hat{k}_{i\mathrm{opt}} = \frac{\lambda_i\hat{\sigma}}{(1-d)\sqrt{\lambda_i\hat{\bar{\beta}}_i^2 + \hat{\sigma}^2} - \hat{\sigma}} \quad (47)$$

We can see that k is always positive, if we set a constraint on k-values in Equation (46) , so that it is positive, then the positiveness of the estimator in Equation (47) can be obtained. In this way, we have the following result: if

$$\hat{d} < 1 - \min\left[\frac{\hat{\sigma}}{\sqrt{\lambda_i\hat{\bar{\beta}}_i^2 + \hat{\sigma}^2}}\right] \quad (48)$$

is selected for all i, then $\hat{k}_{i\mathrm{opt}}$ in Equation (47) is always positive.

For a fixed k and w, we note that $\frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial d} = \frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial \upsilon_i}\frac{\partial \upsilon_i}{\partial d} = 0$ and $\frac{\partial \upsilon_i}{\partial d} = \frac{-k}{\lambda_i+k} \neq 0$, implies $\frac{\partial \mathrm{MSE}(\hat{\hat{\beta}}_{\mathrm{GD}}(\omega,k,d))}{\partial \upsilon_i} = 0$ or $\upsilon_i^2(\lambda_i\bar{\beta}_i^2 + \sigma^2) = \sigma^2$ and $\upsilon_i = \sqrt{\frac{\sigma^2}{\lambda_i\bar{\beta}_i^2+\sigma^2}} = \frac{k(1-d)}{\lambda_i+k}$.

Thus, we have

$$d_{i\text{opt}} = 1 - \frac{(\lambda_i + k)\sigma}{k} \frac{1}{\sqrt{\lambda_i \bar{\beta}_i^2 + \sigma^2}} \tag{49}$$

After the unknown parameters σ^2 and $\bar{\beta}_i^2$ are replaced by their unbiased estimators, we get the optimal estimator of d for a fixed k value as

$$\hat{d}_{i\text{opt}} = 1 - \frac{(\lambda_i + k)\hat{\sigma}}{k} \frac{1}{\sqrt{\lambda_i \hat{\bar{\beta}}_i^2 + \hat{\sigma}^2}} \tag{50}$$

Note that, $\hat{d}_{i\text{opt}}$ in Equation (50) is always less than 1, but it may be smaller than zero. $\hat{d}$ in Equation (48) is always less than 1 and greater than zero, that is $\hat{d} \in (0, 1)$.

Remark. We know that when $k = 1$, $\bar{\beta}(\omega, k, d)$ leads to the generalized-difference-based weighted mixed almost unbiased Liu estimator. Therefore, when $k = 1$, $\hat{d}_{i\text{opt}}$ in Equation (50) reduces to the estimate of d given in Akdeniz and Kaçıranlar (1995).

$$\hat{d}_{i\text{opt}} = 1 - \frac{(1 + \lambda_i)\hat{\sigma}}{\sqrt{\lambda_i \hat{\bar{\beta}}_i^2 + \hat{\sigma}^2}} \tag{51}$$

For a fixed k and d, note that

$$\frac{\partial \text{MSE}(\hat{\bar{\beta}}_{\text{GD}}(\omega, k, d))}{\partial \omega} = \frac{2\omega\xi_i\sigma^2(\omega\xi_i + \lambda_i) - 2\xi_i\sigma^2[(1 - v_i^2)^2\lambda_i + \omega^2\xi_i]}{(\lambda_i + \omega\xi_i)^3} = 0$$

then we have

$$\omega = (1 - v_i^2)^2 = \frac{(\lambda_i + kd)^2(\lambda_i + 2k - kd)^2}{(\lambda_i + k)^4}$$

and

$$\hat{\omega}_{\text{opt}} = \frac{(\lambda_i + \hat{k}\hat{d})^2(\lambda_i + 2\hat{k} - \hat{k}\hat{d})^2}{(\lambda_i + \hat{k})^4}, \quad 0 < \hat{\omega}_{\text{opt}} < 1 \tag{52}$$

7. Illustrative examples

In this section, we present some numerical examples to support our assertions. The process is categorized into two setups: the first part is devoted to the Monte Carlo simulation studies, and the second one is application of our proposed estimation method to the electricity consumption dataset collected in Germany.

7.1. Monte Carlo simulation studies

In this section, we continue the comparison of proposed estimators based on the scalar values of mean squared error matrix by some simulations and graphical results. Since, theoretically, these estimators are very difficult to compare, the Monte Carlo simulation studies have been conducted to compare the efficiency of the estimators. The scalar-valued mean squared error (SMSE) for any estimator $\tilde{\beta}$ is defined as

$$\text{SMSE}(\tilde{\beta}) = \text{tr}(\text{MSEM}(\tilde{\beta})) = tr(\text{Cov}(\tilde{\beta})) + \text{Bias}(\tilde{\beta})'\text{Bias}(\tilde{\beta})$$

To achieve different degrees of collinearity, following McDonald and Galarneau (1975) and Gibbons (1981), the explanatory variables were generated using the following device for $n = 30$ with 10,000 iteration from the following model:

$$x_{ij} = (1 - \gamma^2)^{\frac{1}{2}} z_{ij} + \gamma z_{ip}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p \tag{53}$$

where z_{ij} are independent standard normal pseudo-random numbers, and γ is specified so that the correlation between any two explanatory variables is given by γ^2. These variables are then standardized so that $X'X$ and $X'Y$ are in correlation forms. Four different sets of correlation corresponding to $\gamma = 0.75, 0.90, 0.95$, and 0.99 are considered. Then n observations for the dependent variable are determined by

$$y_i = \sum_{j=1}^{6} x_{ij}\beta_j + f(u_i) + \varepsilon_i, \quad i = 1, 2, \ldots, n \tag{54}$$

where

$$\beta = (3, 1, 3, -2, -5, 4)', \; f(u) = \exp\left\{\sin(u)\cos(3u) + \sqrt{u}\right\}$$

for $u \in [0, 3]$ and

$$\varepsilon \sim N_n(0, \sigma^2 V), \sigma^2 = 0.36, v_{ij} = \exp\{-\varphi|\mathrm{i} - \mathrm{j}|\}, \quad \varphi = 2, \quad i, \quad j = 1, \quad \infty \ldots, n$$

The parametric part of model (54), that is, β, is estimated by third-order differencing coefficients, $d_0 = 0.8582$, $d_1 = -0.3832$, $d_2 = -0.2809$, and $d_3 = -0.1942$, and then, the non parametric part is estimated by kernel methodology and cross-validation criteria.

Performance of estimations of the partially linear models greatly depends on the selection of smoothing parameter (Aydin 2014). Cross-validation directly estimates the model prediction error. It is the simplest and most widely used method for choosing the smoothing parameter for non parametric models. However, evaluating the leave one-out cross-validation tends to be time-consuming for a partially linear model with even a moderate sample size. It adds one more level of difficulty if we want to conduct intensive numerical simulation studies, such as using a bootstrap procedure to estimate the variance. Li, Zhang, and Wu (2011) proposed an algorithm to quickly compute cross-validation.

Optimal differencing weights do not have analytic expressions but may be calculated easily using an optimization routine. Hall, Kay, and Titterington (1990) presented weights to order $m = 10$. These contain some minor errors. Now, we define the $(n - 3) \times n$ differencing matrix as

$$D = \begin{pmatrix} d_0 & d_1 & d_2 & d_3 & 0 & 0 & \ldots & 0 \\ 0 & d_0 & d_1 & d_2 & d_3 & 0 & \ldots & 0 \\ \vdots & \ddots & & & & & & \vdots \\ 0 & 0 & \ldots & 0 & d_0 & d_1 & d_2 & d_3 \end{pmatrix}$$

For the restriction, we consider the following stochastic linear restrictions

$$r = R\beta + e, \quad R = \begin{pmatrix} 1 & 5 & -3 & -1 & -1 & 0 \\ -2 & -1 & 0 & -2 & 3 & 1 \\ 1 & 2 & 1 & 3 & -2 & 0 \\ 4 & -1 & 2 & 2 & 0 & -2 \\ 5 & 3 & 4 & -5 & 1 & 0 \end{pmatrix}$$

where $e \sim N_q(0, \sigma_e^2 W)$, $\sigma_e^2 = 0.0036$, $w_{ij} = (\frac{1}{n})^{|i-j|}$, $i, j = 1, \ldots, q$

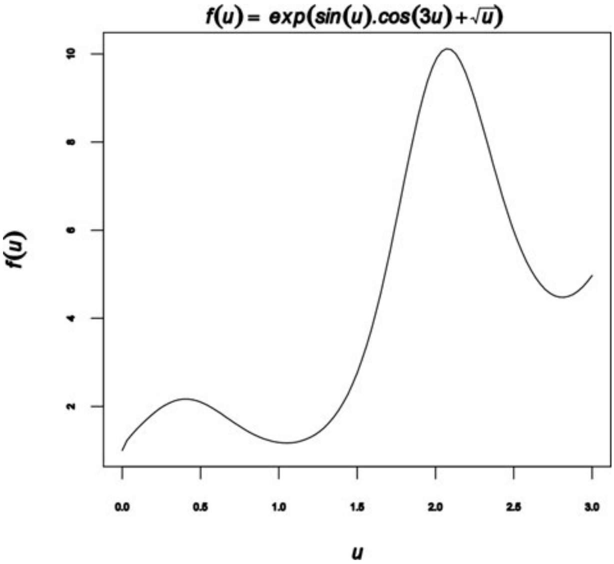

Figure 1. Non parametric function of model (54).

Monte Carlo simulation is performed with $M = 10^4$ replications, obtaining the estimators $\hat{\beta}_{(1)} = \hat{\beta}_{\text{GDE}}$, $\hat{\beta}_{(2)} = \hat{\beta}_{\text{GDWME}}(\hat{\omega}_{\text{opt}})$, $\hat{\beta}_{(3)} = \hat{\beta}_{\text{GDAURE}}(\hat{k}_{\text{opt}})$, $\hat{\beta}_{(4)} = \hat{\beta}_{\text{GDWMAURE}}(\hat{\omega}_{\text{opt}}, \hat{k}_{\text{opt}})$, $\hat{\beta}_{(5)} = \hat{\beta}_{\text{GDAUTPE}}(\hat{k}_{\text{opt}}, \hat{d}_{\text{opt}})$, and $\hat{\beta}_{(6)} = \hat{\beta}_{\text{GDWMAUTPE}}(\hat{\omega}_{\text{opt}}, \hat{k}_{\text{opt}}, \hat{d}_{\text{opt}})$ in the restricted partially linear model. The corresponding estimator of the non linear part for the ith method is obtained using kernel method as $\hat{f}_{(i)} = K(y - X\hat{\beta}_{(i)})$ for $i = 1, \ldots, 6$, and K is the smoother matrix.

The relative efficiencies of the above methods with respect to the first method are estimated as

$$\text{Eff}(\hat{\beta}_{(\text{i})}, \hat{f}_{(\text{i})}) = \frac{\frac{1}{\text{M}} \sum_{m=1}^{M} \|y^{(\text{m})} - X^{(\text{m})}\hat{\beta}_{(1)}^{(\text{m})} - \hat{f}_{(1)}^{(\text{m})}\|_2^2}{\frac{1}{\text{M}} \sum_{m=1}^{M} \|y^{(\text{m})} - X^{(\text{m})}\hat{\beta}_{(i)}^{(\text{m})} - \hat{f}_{(i)}^{(\text{m})}\|_2^2}, \quad i = 1, \ldots, 6$$

where $(X^{(\text{m})}, y^{(\text{m})})$ stands for the generated sample in the mth iteration, $\hat{\beta}_{(i)}^{(m)}$ and $\hat{f}_{(i)}^{(m)}$ are the estimators obtained in the mth iteration, and $\|v\|_2^2 = \sum_{i=1}^{q} v_i^2$ for $v = (v_1, \ldots, v_q)'$.

In Figure 1, the non parametric part of the model (54) is plotted. This function is difficult to be estimated and provides a good test case for the non parametric regression method. All computations were conducted using the statistical package R. In Tables 1–4, we computed

Table 1. Evaluation of parameters for proposed estimators with $\gamma = 0.75$.

Method Coefficients	GDE	GDWME	GDAURE	GDWMAURE	GDAUTPE	GDWMAUTPE
$\hat{\beta}_1$	2.1548	2.9986	2.1603	2.9309	3.0076	2.9402
$\hat{\beta}_2$	0.6391	1.0001	0.6550	1.0302	0.9741	1.0204
$\hat{\beta}_3$	2.1516	3.0007	2.1607	3.0372	3.0041	3.0299
$\hat{\beta}_4$	−1.5799	−2.0010	−1.5469	−2.0226	−2.0055	−2.0269
$\hat{\beta}_5$	−3.7948	−5.0011	−3.7458	−5.0095	−4.9756	−5.0195
$\hat{\beta}_6$	2.5220	3.9967	2.5811	3.8580	3.9897	3.8747
$\text{SM}\circ\text{SE}(\hat{\beta}_{(i)})$	50.3722	0.2438	18.5331	0.1272	17.7875	0.1201
$\hat{\text{mse}}(\hat{f}_{(i)}, f)$	71.8083	14.9874	30.4609	8.2367	27.0279	8.0640
$\text{Eff}(\hat{\beta}_{(i)}, \hat{f}_{(i)})$	1.0000	4.0229	2.1119	4.2516	2.3244	4.3590

Table 2. Evaluation of parameters for proposed estimators with $\gamma = 0.90$.

Method coefficients	GDE	GDWME	GDAURE	GDWMAURE	GDAUTPE	GDWMAUTPE
$\hat{\beta}_1$	2.6600	2.9855	2.6671	2.9288	3.1776	2.9397
$\hat{\beta}_2$	0.8546	1.0071	0.8649	1.0277	1.0484	1.0240
$\hat{\beta}_3$	2.6692	3.0046	2.6920	3.0369	3.1481	3.0304
$\hat{\beta}_4$	− 1.7046	− 2.0072	− 1.6707	− 2.0283	− 1.9794	− 2.0244
$\hat{\beta}_5$	− 4.2693	− 5.0023	− 4.2552	− 5.0187	− 4.9640	− 5.0146
$\hat{\beta}_6$	2.7645	3.9538	2.7945	3.8483	3.6171	3.8641
$\text{SM}\circ\text{SE}(\hat{\beta}_{(i)})$	24.9877	0.2043	11.0622	0.1151	10.7938	0.1042
$\text{m}\hat{\text{s}}\text{e}(\hat{f}_{(i)}, f)$	68.7767	16.5730	19.6194	9.0625	15.4960	8.8482
$\text{Eff}(\hat{\beta}_{(i)}, \hat{f}_{(i)})$	1.0000	4.0576	2.2212	4.3883	2.9329	4.5954

Table 3. Evaluation of parameters for proposed estimators with $\gamma = 0.95$.

Method coefficients	GDE	GDWME	GDAURE	GDWMAURE	GDAUTPE	GDWMAUTPE
$\hat{\beta}_1$	3.1384	2.9991	2.9300	2.9443	2.9163	2.9414
$\hat{\beta}_2$	1.1146	1.0038	1.0268	1.0248	1.0263	1.0235
$\hat{\beta}_3$	2.9666	2.9961	2.7683	3.0259	2.7534	3.0254
$\hat{\beta}_4$	− 1.9048	− 1.9956	− 1.8161	− 2.0173	− 1.8090	− 2.0206
$\hat{\beta}_5$	− 4.9334	− 4.9859	− 4.6444	− 5.0006	− 4.6102	− 5.0032
$\hat{\beta}_6$	3.7258	3.9869	3.4399	3.8717	3.4325	3.8664
$\text{SM}\circ\text{SE}(\hat{\beta}_{(i)})$	9.7697	0.1760	3.4386	0.0904	3.2596	0.0840
$\text{m}\hat{\text{s}}\text{e}(\hat{f}_{(i)}, f)$	53.4212	12.6956	28.0886	7.3261	28.4249	6.2016
$\text{Eff}(\hat{\beta}_{(i)}, \hat{f}_{(i)})$	1.0000	4.7585	2.6877	5.0156	2.6856	5.1868

Table 4. Evaluation of parameters for proposed estimators with $\gamma = 0.99$.

Method coefficients	GDE	GDWME	GDAURE	GDWMAURE	GDAUTPE	GDWMAUTPE
$\hat{\beta}_1$	1.1593	2.9891	1.1409	2.9286	2.7440	2.9616
$\hat{\beta}_2$	0.3943	1.0051	0.3861	1.0285	1.1517	1.0093
$\hat{\beta}_3$	1.3265	3.0079	1.2958	3.0393	3.1414	3.0190
$\hat{\beta}_4$	− 0.9883	− 2.0018	− 0.9572	− 2.0252	− 1.8914	− 2.0211
$\hat{\beta}_5$	− 2.4480	− 5.0017	− 2.3853	− 5.0147	− 4.9948	− 5.0192
$\hat{\beta}_6$	1.5533	3.9827	1.5933	3.8565	3.8582	3.9195
$\text{SM}\circ\text{SE}(\hat{\beta}_{(i)})$	279.9880	0.2644	73.7155	0.1441	71.1475	0.1258
$\text{m}\hat{\text{s}}\text{e}(\hat{f}_{(i)}, f)$	78.6052	14.1194	27.1817	7.9669	26.3492	7.7944
$\text{Eff}(\hat{\beta}_{(i)}, \hat{f}_{(i)})$	1.0000	4.4109	2.0800	4.6297	2.0854	5.2402

the proposed estimators at optimum values of parameters, $\hat{\omega}_{\text{opt}}, \hat{k}_{\text{opt}}, \hat{d}_{\text{opt}}$, respectively. We numerically estimated the SMSE's, efficiencies of proposed estimators relative to GDE, separately, and $\text{m}\hat{\text{s}}\text{e}(\hat{f}_{(\text{i})}, f) = \frac{1}{\text{nM}} \sum_{m=1}^{M} \| \hat{f}_{(\text{i})}^{(\text{m})} - f \|_2^2$ for all proposed estimators. The 3D diagrams as well as the 2D slices of SMSE versus parameters are plotted for proposed estimators in Figures 2 and 3. Since the results were similar across cases, to save space we reported only the results for $\gamma = 0.90$. Figure 4 shows the fitted function by kernel smoothing after estimating the linear part of the model by proposed estimators for $\gamma = 0.90$.

7.2. Application to electricity consumption dataset

To motivate the problem of estimation in the semiparametric partial linear model, we apply the electricity consumption, considered by Akdeniz Duran, Härdle, and Osipenko (2012). The variables are defined for 177 items as follows:

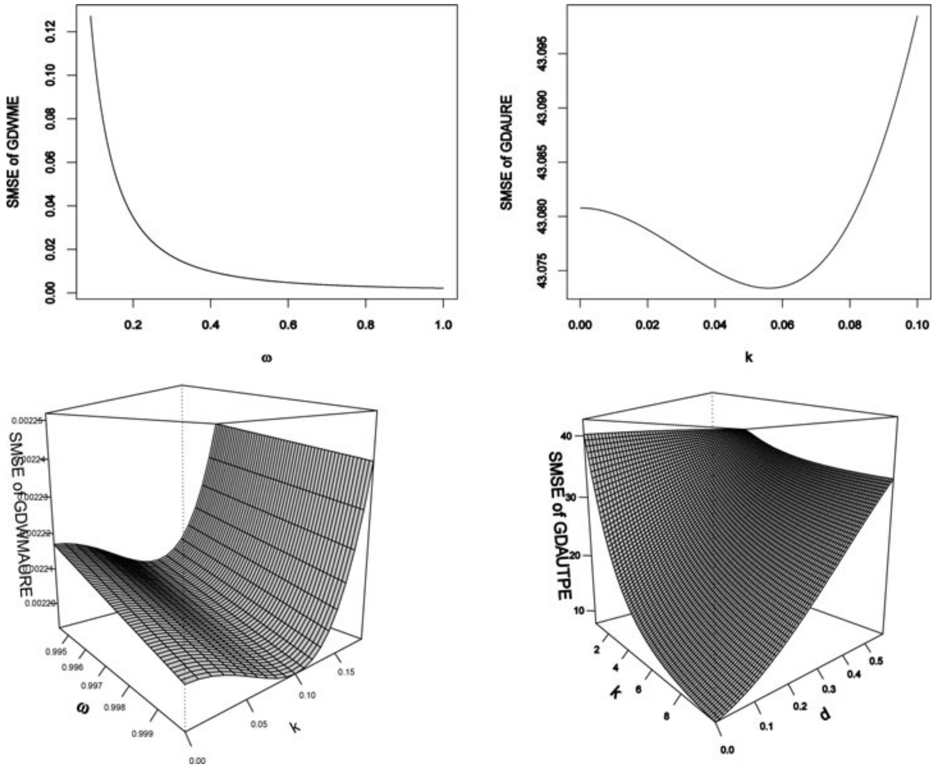

Figure 2. Diagram of SMSE versus parameters for $\gamma = 0.90$. Top left: SMSE($\hat{\beta}_{\text{GDWME}}(\omega)$); top right: SMSE($\hat{\beta}_{\text{GDAURE}}(k)$); bottom left: SMSE($\hat{\beta}_{\text{GDWMAURE}}(\omega, k)$); bottom right: SMSE($\hat{\beta}_{\text{GDAUTPE}}(k, d)$).

The dependent variable y is the log monthly electricity consumption per person (LEC) and the independent variables include log income per person (LI), log rate of electricity price to the gas price (LREG), and cumulated average temperature index (Temp) for the corresponding month taken as average of 20 German cities computed from the data of German weather service.

To detect the non parametric part of the model, by Yatchew (2003), the test statistic for the null hypothesis that the regression function has the parametric form, that is, $H_0 : f(u) = h(u; \beta)$ for a parametric function $h(.)$, against the non parametric alternative $f(u)$, when one uses optimal differencing weights, is

$$Z_0 = \sqrt{nm}\frac{\hat{\sigma}^2 - \hat{\sigma}^2_{\text{diff}}}{\hat{\sigma}^2_{\text{diff}}} \xrightarrow{D} N(0, 1) \tag{55}$$

where $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \mathrm{h}(\mathrm{u}; \hat{\beta})\right)^2$, $\hat{\sigma}^2_{\text{diff}} = \frac{\tilde{y}'(I-P)\tilde{y}}{\mathrm{tr}(D'(I-P)D)}$, $P = \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}$.

We consider Temp as a non parametric part (using a third-order differencing coefficients), because, it has the largest value of non parametric significance test statistics among those of other independent variables. The statistics of above test for all explanatory variables can be found in Table 5. We also use the added-variable plots to identify the parametric and non parametric components of the model (for more details, see Sheather 2009). Added-variable plots enable us to visually assess the effect of each predictor, having adjusted for the effects of

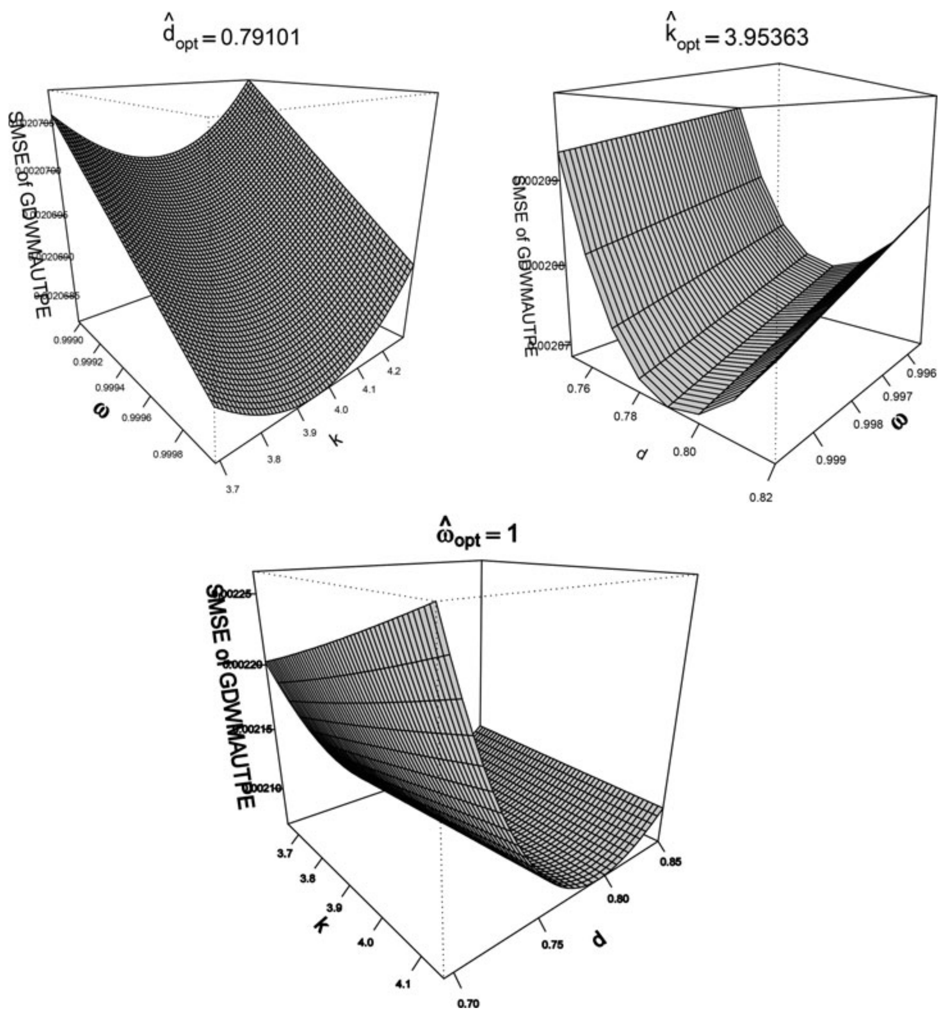

Figure 3. Diagram of SMSE($\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, d)$) versus parameters for $\gamma = 0.90$. Top left: SMSE($\hat{\beta}_{\text{GDWMAUTPE}}(\omega, k, \hat{d}_{\text{opt}})$); top right: SMSE($\hat{\beta}_{\text{GDWMAUTPE}}(\omega, \hat{k}_{\text{opt}}, d)$); bottom center: SMSE($\hat{\beta}_{\text{GDWMAUTPE}}(\hat{\omega}_{\text{opt}}, k, d)$).

the other predictors. By looking at added-variable plot (Figure 5), we consider Temp as a non parametric part and so, the specification of the semiparametric partial linear model is

$$(\text{LEC})_i = \sum_{j=1}^{11} \beta_j x_{ij} + \beta_{12}(\text{LI})_i + \beta_{13}(\text{LREG})_i + f(\text{Temp}_i) + \varepsilon_i \tag{56}$$

where $x_1, \ldots, x_{11}$ are dummy variables for the monthly effects. The ratio of largest eigenvalue to smallest eigenvalue for new design matrix in model (55) after applying differencing method

Table 5. Values of test statistics (5).

Variable	*Z0*
LI	1.99
LREG	2.06
Temp	2.95*

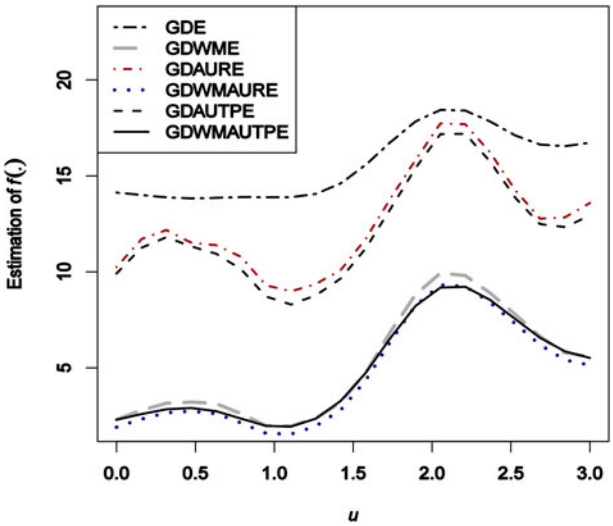

Figure 4. Estimation of the function under study by kernel approach for $\gamma = 0.90$..

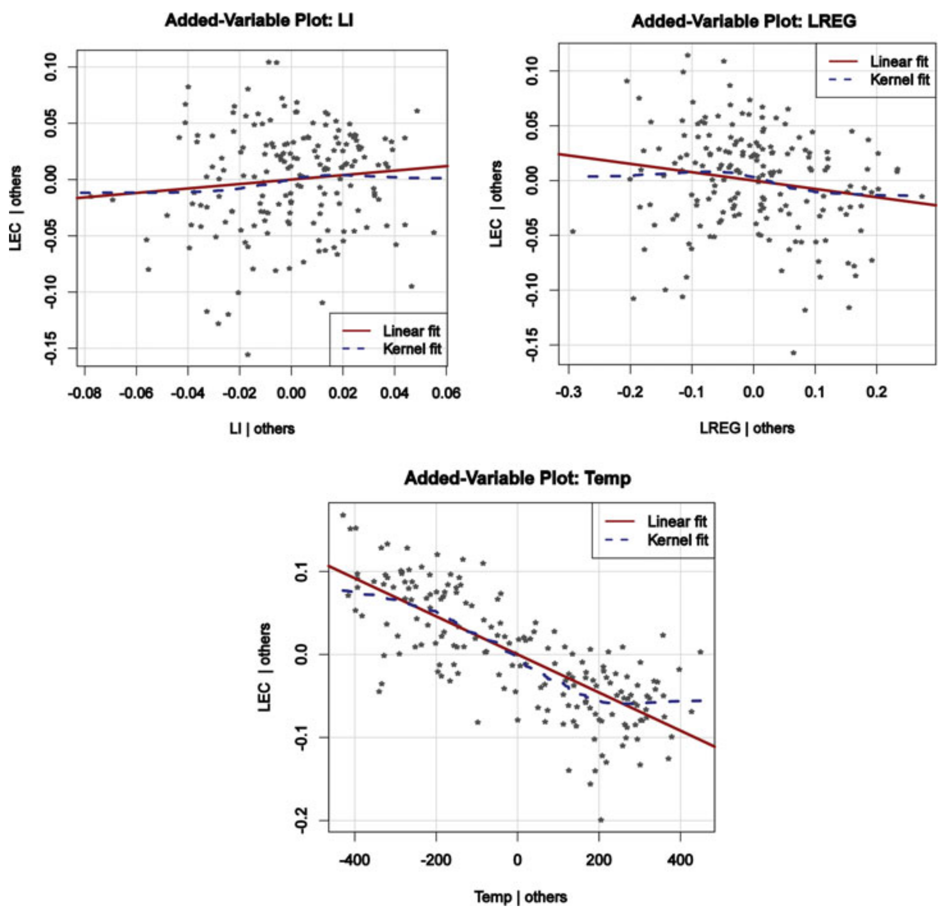

Figure 5. Added-variable plots of individual explanatory variables versus dependent variable, linear fit (red solid line) and kernel fit (blue dashed line).

Table 6. Evaluation of parameters for proposed estimators for real dataset.

Method variables	GDE	GDWME	GDAURE	GDWMAURE	GDAUTPE	GDWMAUTPE
*x*1	− 0.59549	− 0.18216	− 0.06522	− 0.10158	− 0.06620	0.10267
*x*2	− 0.16763	− 0.06374	− 0.10542	− 0.06659	− 0.10549	− 0.07666
*x*3	− 0.02016	− 0.00159	− 0.01829	− 0.00953	− 0.01829	− 0.01228
*x*4	− 0.00801	0.00551	− 0.00702	− 0.00369	− 0.00703	− 0.00325
*x*5	− 0.00619	0.00986	− 0.00537	0.00122	− 0.00537	− 0.00006
*x*6	− 0.01977	− 0.00256	− 0.01792	− 0.01030	− 0.01792	− 0.01271
*x*7	− 0.01178	− 0.00026	− 0.01034	− 0.00768	− 0.01034	− 0.00820
*x*8	− 0.02579	− 0.00631	− 0.02344	− 0.01293	− 0.02344	− 0.01660
*x*9	− 0.01444	0.00022	− 0.01297	− 0.00694	− 0.01298	− 0.00835
*x*10	− 0.01230	− 0.00063	− 0.01085	− 0.00908	− 0.01085	− 0.00781
*x*11	− 0.00878	0.00880	− 0.00766	0.00131	− 0.00766	− 0.00102
LI	− 0.00006	0.01563	0.00039	0.00744	0.00039	0.00645
LREG	− 0.00631	0.00763	− 0.00541	0.00122	− 0.00541	− 0.00088
RSS	0.39933	0.36578	0.34653	0.35402	0.34657	0.34291
*R*2	0.57175	0.60772	0.62837	0.62034	0.62833	0.63226

is approximately $\lambda_{13}/\lambda_1 = 220.3069$ and so, there exists a potential multicollinearity between the columns of design matrix.

After a primary evaluation of model (56), one might consider the stochastic restriction $rR\beta$, where

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad r = \begin{pmatrix} -0.02 \\ 0 \end{pmatrix}$$

We test the linear hypothesis $H_0 : r \simeq R\beta$ in the framework of our semiparametric partial linear model (56). The test statistic for H_0, given our observations, is

$$\chi^2_{\text{rank}(R)} = (R\hat{\beta}_{\text{diff}} - r)'\left(R\hat{\Sigma}_{\hat{\beta}_{\text{diff}}}R'\right)^{-1}(R\hat{\beta}_{\text{diff}} - r) = 0.05952$$

where $\hat{\Sigma}_{\hat{\beta}_{\text{diff}}} = (1 + \frac{1}{2m})\hat{\sigma}^2_{\text{diff}}(\tilde{X}'\tilde{X})^{-1}$ (see Yatchew [2]). The test statistic is not greater than upper α-quantile of χ^2 distribution. Thus, we conclude that the null hypothesis H_0 is not rejected.

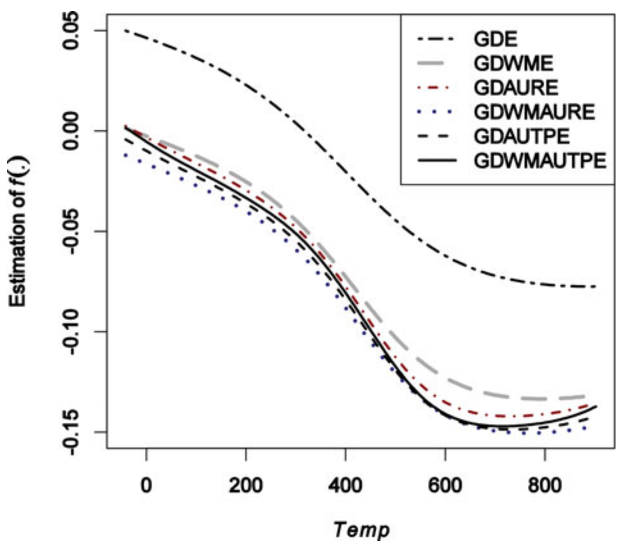

Figure 6. Estimations of non parametric part of model (56).

Table 6 shows a summery of the results. In this table, the RSS and R^2, respectively, are the residual sum of squares and coefficient of determination of the model, that is, RSS $= y - \hat{y}_2^2$, $\hat{y} = X\hat{\beta}_{(i)} + \hat{f}(u)$, $y =$ LEC and $R^2 = 1 -$ RSS$/S_{yy}$, which are calculated for each proposed estimators of β. For estimation of non parametric effect, at first we estimated the parametric effects by one of the proposed estimators and then, local polynomial approach was applied to fit LEC $- X\hat{\beta}_{(i)}$ on $u =$ Temp, where $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}$, LI, LREG) (Figure 6).

8. Conclusions

We considered the method of weighted mixed regression estimation to estimate the regression coefficients in generalized-difference-based semiparametric partially linear model. The generalized-difference-based weighted mixed almost unbiased ridge estimator, $\hat{\beta}_{\mathrm{GDWMAURE}}(\omega, k)$, is derived and its dominance over both the generalized-difference-based weighted mixed estimator, $\hat{\beta}(\omega)$, and the generalized-difference-based almost unbiased ridge estimator, $\hat{\beta}_{\mathrm{GDAURE}}(k)$, is studied under the criterion of mean squared error matrix. The generalized-difference-based weighted mixed almost unbiased two-parameter estimator, $\hat{\beta}_{\mathrm{GDWMAUTPE}}(w, k, d)$, in semiparametric partial linear models is proposed. After some theorems, the Monte Carlo simulation studies and a real data example have been conducted to compare the performance of the proposed estimators numerically. The results from the Monte Carlo simulations for $n = 30$, $P = 6$, and different γ are presented in Tables 1–4 and Figures 1–4. From these tables, it can be seen that the factor affecting the performance of the estimators is the degree of correlation (γ). It can be concluded that GDWMAUTPE is leading to be the best estimator among others for the parametric part of the model, since it offers smaller SMSE and mse values in all proposed estimators. Further GDE is the worst estimator for the parametric part in these examples. In general, the value of γ has positive effect on the performance of the proposed estimators with respect to GDE. In the real example study, a near dependency among the column of $X'X$ identified from $\lambda_{13}/\lambda_1 = 220.3069$, that is, the design matrix may be considered as being very ill-conditioned and we had to consider the ridge form of proposed estimators in our study. As it can be seen from Table 5 and Figure 5, the non linear relation between log monthly electricity consumption per person (LEC) and cumulated average temperature index (Temp) can be detected and so, the pure parametric model does not fit to the data and semiparametric partial linear model fits more significantly. Further, from Table 6 and Figure 6, it can be deduced that GDWMAUTPE is quite efficient in the sense that it has significant value of goodness of fit.

Acknowledgments

The authors are grateful to the anonymous reviewers and the associate editor for their valuable comments which helped to improve the quality of this work. The second author's research was supported in part by Research Councils of Semnan University, Iran.

References

Akdeniz, F., E. Akdeniz Duran, M. Roozbeh, and M. Arashi. 2015. Efficiency of the generalized difference-based Liu estimators in semiparametric regression models with correlated error. *Journal of Statistical Computation and Simulation* 85 (1):147–65.

Akdeniz, F., and H. Erol. 2003. Mean squared error comparisons of some biased estimators in linear regression. *Communications in Statistics Theory and Methods* 32 (12): 2391–415.

Akdeniz, F., and S. Kaçıranlar. 1995. On the almost unbiased generalized Liu estimator and unbiased estimation of the Bias and MSE. *Communications in Statistics-Theory and Methods* 24 (7):1789–97.

Akdeniz, F., and G. Tabakan. 2009. Restricted ridge estimators of the parameters in semiparametric regression model . *Communications in Statistics- Theory and Methods* 38 (11):1852–69.

Akdeniz Duran, E., W. K. Härdle, and M. Osipenko. 2012. Difference-based ridge and Liu type Estimators in semiparametric regression models. *Journal of Multivariate Analysis* 105:164–75.

Arashi, M., M. Janfada, and M. Norouzirad. 2015. Singular ridge regression with stochastic constraints. *Communications in Statistics-Theory and Methods* 44 (6):1281–92.

Arashi, M., and T. Valizadeh. 2015. Performance of Kibria's methods in partial linear model. *Statistical Papers* 56 (1):231–46.

Arumairajan, S., and P. Wijekoon. 2014. Improvement of ridge estimator when stochastic restrictions are avaiable in the linear regression model. *Journal of Statistical and Econometric Methods* 23 (1):35–48.

Aydin, D. 2014. Estimations of the partially linear models with smoothing spline based on different selection methods: A comparative study. *Pakistan Journal of Statistics* 30:35–56.

Dorugade, A. V. 2014. A modified two-parameter estimator in linear regression. *Statıstıcs ın Transıtıon New Series* 15 (1):23–36.

Farebrother, R. W. 1976. Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society: Series B* 38:248–50.

Gibbons, D. G. 1981. A simulation study of some ridge estimators. *Journal of the American Statistical Association* 76:131–9.

Hall, P., J. W. Kay, and D. M. Titterington. 1990. On estimat5ion of noise variance in two-Dimensional signal processing. *Advances in Applied Probability* 23:476–95.

Hoerl, A.E., and R. W. Kennard. 1970. Ridge regression: biased estimation for orthogonal problems. *Technometrics* 12:55–67.

Klipple, K., and R. L. Eubank. 2007. Difference-based variance estimators for partially linear models. *Festschrift in Honor of Distinguished Professor Mir Masoom Ali on the Occasion of his Retirement* May 18–19, p. 313–23.

Li, J., W. Zhang, and Z. Wu. 2011. Optimal zone for bandwidth selection in semiparametric Models. *Journal of Nonparametric Statistics* 23:701–17.

Li, Y., and H. Yang. 2010. A new stochastic mixed ridge estimator in linear regression. *Statistical Papers* 51 (2):315–23.

Liu, C., H. Yang, and J. Wu. 2013. On the weighted mixed almost unbiased ridge estimator in Stochastic restricted linear regression. *Journal of Applied Mathematics* Article ID:902715.

Liu, C., H. Jiang, X. Shi, and D. Liu. 2014. Two kinds of weighted biased estimators in Stochastic restricted regression model. *Journal of Applied Mathematics* Article ID:314875.

McDonald, M. C., and D. I. Galarneau. 1975. A Monte Carlo evaluation of ridge-type estimators. *Journal of the American Statistical Association* 70:407–16.

Özkale, M. R., and S. Kaçıranlar. 2007. Inference the restricted and unrestricted two-parameter estimators. *Communications in Statistics-Theory and Methods* 36:2707–25.

Rao, C. R., H. Toutenburg, and H. C. Shalab. 2008. *Linear models and generalizations: least squares and alternatives*. New York: Springer.

Roozbeh, M. 2015. Shrinkage ridge estimators in semiparametric regression models. *Journal of Multivariate Analysis* 136:56–74.

Roozbeh, M., M. Arashi, and H. A. Niroumand. 2011. Ridge regression methodology in partial linear models with correlated errors. *Journal of Statistical Computation and Simulation* l81 (4):517–28.

Schaffrin, B., and H. Toutenburg. 1990. Weighted mixed regression. *Zeitschrift fur Angewandte Mathematik und Mechanik* 70:735–8.

Sheather, S. J. 2009. *A modern approach to regression with R*. New York: Springer.

Tabakan, G., and F. Akdeniz. 2010. Difference-based ridge estimator of parameters in partial linear model. *Statistical Papers* 51:357–68.

Theil, H., and A. S. Goldberger. 1961. On pure and mixed statistical estimation in economics. *International Economic Review* 2:65–78.

Toutenburg, H., V. K. Srivastava, B. Schaffrin, and C. Heumann. 2003. Efficiency properties of weighted mixed regression estimation. *METRON- International Journal of Statistics* LXI (1):91–103.
Trenkler, G., and H. Toutenburg. 1990. Mean squared error matrix comparisons between biased estimators: an overview of recent results. *Statistical Papers* 31:165–79.
Wang, L., L. D. Brown, and T. T. Cai. 2011. A difference- based approach to semiparametric partial linear model. *Electronic Journal of Statistics* 5:619–41.
Wu, J. 2016. Performance of the difference-based almost unbiased Liu estimator in partial linear model. *Journal of Statistical Computation and Simulation* DOI:10.1080/00949655.2015.11336628.
Wu, J., and H. Yang. 2013. Efficiency of an almost unbiased two-parameter estimator in linear regression model. *Statistics* 47 (3):535–45.
Yatchew, A. 1997. An elementary estimator of the partial linear model. *Economics Letters* 57:135–43.
Yatchew, A. (2003). *Semiparametric Regression for the Applied Econometrican*. Cambridge: Cambridge University Press.