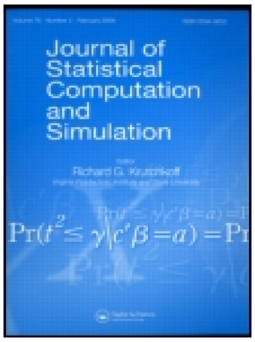

Taylor & Francis
Taylor & Francis Group

Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: https://www.tandfonline.com/loi/gscs20

Efficiency of the generalized difference-based Liu estimators in semiparametric regression models with correlated errors

Fikri Akdeniz, Esra Akdeniz Duran, Mahdi Roozbeh & Mohammad Arashi

To cite this article: Fikri Akdeniz, Esra Akdeniz Duran, Mahdi Roozbeh & Mohammad Arashi (2015) Efficiency of the generalized difference-based Liu estimators in semiparametric regression models with correlated errors, Journal of Statistical Computation and Simulation, 85:1, 147-165, DOI: 10.1080/00949655.2013.806924

To link to this article: https://doi.org/10.1080/00949655.2013.806924

Published online: 11 Jun 2013.

Journal of Statistical Computation and Simulation, 2015
Vol. 85, No. 1, 147–165, http://dx.doi.org/10.1080/00949655.2013.806924

Taylor & Francis
Taylor & Francis Group

Efficiency of the generalized difference-based Liu estimators in semiparametric regression models with correlated errors

Fikri Akdeniz[a]*, Esra Akdeniz Duran[b], Mahdi Roozbeh[c] and Mohammad Arashi[d]

[a]*Department of Mathematics and Computer Science, Çağ University, Tarsus, Turkey;* [b]*Department of Statistics, İstanbul Medeniyet University, İstanbul, Turkey;* [c]*Faculty of Mathematics, Statistics and Computer Sciences, Semnan University, Semnan, Iran;* [d]*Department of Statistics, School of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran*

(*Received 31 October 2012; final version received 16 May 2013*)

In this paper, a generalized difference-based estimator is introduced for the vector parameter β in the semiparametric regression model when the errors are correlated. A generalized difference-based Liu estimator is defined for the vector parameter β in the semiparametric regression model. Under the linear nonstochastic constraint $R\beta = r$, the generalized restricted difference-based Liu estimator is given. The risk function for the $\hat{\beta}_{\mathrm{GRD}}(\eta)$ associated with weighted balanced loss function is presented. The performance of the proposed estimators is evaluated by a simulated data set.

Keywords: balanced loss function; difference-based estimator; generalized Liu estimator; generalized difference-based restricted Liu estimator; semiparametric regression model

AMS Subject Classification: Primary: 62G08; Secondary: 62J07

1. Introduction

Semiparametric regression models have received considerable attention in statistics and econometrics. In these models, some of the relations are believed to be of certain parametric form while others are not easily parameterized. Consider the semiparametric regression 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 nonrandom in some bounded domain $D \subset \Re$, $f(\cdot)$ is an unknown smooth function, and ε's are independent and identically distributed random errors with mean 0 and variance σ^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_i have bounded support, say the unit interval, and have been arranged so that $u_1 \leq u_2 \leq \cdots \leq u_n$. The goal is to estimate the unknown parameter vector β and nonparametric function $f(u)$ from the data $\{y_i, X_i, u_i\}$. This will be done through a

*Corresponding author. Email: fikriakdeniz@gmail.com

difference-based estimation. In the vector/matrix notation, the model (1) is written as

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

where $y = (y_1, \ldots, y_n)'$, $X = [X_1, \ldots, X_n]'$, $f = (f(u_1), \ldots, f(u_n))'$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$.

Semiparametric linear regression models are more flexible than the standard linear regression models since they combine both parametric and nonparametric components when it is believed that the response variable y depends on variable X in a linear way but is nonlinearly related to other independent variable U. All the existing approaches for the semiparametric regression model are based on different nonparametric regression procedures. There have been several approaches to estimate β and f. Among the most important approaches are given by several researchers.[1–7]

In model (2), Yatchew [8] concentrates on the estimation of the linear component and used differencing to eliminate bias induced by the presence of the nonparametric component. Yatchew's method does not require an estimator of the function f and are often called the difference-based estimation procedure. Provided that $f(\cdot)$ is differentiable and the u's are closely spaced, it is possible to remove the effect of the function f by differencing the data appropriately.[9]

In regression analysis, researchers often encounter the problem of multicollinearity. In the case of multicollinearity we know that the correlation matrix might have one or more small eigenvalues which cause 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 [10] or Liu estimator [11] 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. For the purposes of this paper, we will employ the shrinkage estimator that was proposed by Liu [11] to combat multicollinearity. Liu [11] combined the Stein [12] estimator with the ordinary ridge regression estimator to obtain what we call the Liu (**Li**near **U**nified) estimator (see [13,14]). 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.

In this paper, a generalized restricted difference-based estimator is introduced for the vector parameter β in the semi parametric regression model when the errors are correlated. A generalized difference-based Liu estimator is defined for the vector parameter β in the semiparametric regression model. Under the linear nonstochastic constraint $R\beta = r$, the generalized restricted difference-based Liu estimator is given.

We also examine the risk performance of the estimators under study when the weighted balanced loss function (BLF) is used.

In order to compare the performance of the different estimators, a simulation study is conducted where the risk function is used as performance criteria and factors, including the degree of correlation, the sample size, the variance of the dependent variable and the number of explanatory variables are varied.

The paper is organized as follows. In Section 2, the model and difference-based estimator is defined. The generalized difference-based Liu estimator and restricted generalized difference-based Liu estimator of β are introduced in Section 3. Section 4 gives the risk function for the $\hat{\beta}_{GRD}(\eta)$ associated with weighted BLF. Comparison results are given in Section 5. The performance of the new estimator is evaluated by a simulated data set in Section 6.

2. The model and 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 nonparametric component in the semiparametric regression model by various authors (e.g. [8,9,15,16]). Consider the following semiparametric regression model

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

where f is an unknown smooth function and has a bounded first derivative.

Yatchew [8] suggested estimating β on the basis of the mth order differencing equation

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

where $d_0, d_1, \dots, d_m$ are differencing weights.

2.1. *How does the approximation work?*

Suppose u_i are equally spaced on the unit interval and $f' \leq L$. By the mean value theorem, for some $u_i^* \in [u_{i-1}, u_i]$ we have

$$f(u_i) - f(u_{i-1}) = f'(u_i^*)(u_i - u_{i-1}) \leq \frac{L}{n}$$

Note that with $m = 1$ from (4) we have

$$\begin{aligned} y_i - y_{i-1} &= (x_i - x_{i-1})\beta + f(u_i) - f(u_{i-1}) + \varepsilon_i - \varepsilon_{i-1} \\ &= (x_i - x_{i-1})\beta + O\left(\frac{1}{n}\right) + \varepsilon_i - \varepsilon_{i-1} \\ &\approx (x_i - x_{i-1})\beta + \varepsilon_i - \varepsilon_{i-1}. \end{aligned}$$

We then estimate the linear regression coefficient β by the ordinary least-squares estimator based on the differences. Then we obtain the least-squares estimate $\hat{\beta}_{\text{diff}} = \frac{\sum (y_i - y_{i-1})(x_i - x_{i-1})}{\sum (x_i - x_{i-1})^2}$.

Now let $d = (d_0, d_1, \dots, d_m)'$ be a $(m+1)$-vector, where m is the order of differencing and $d_0, d_1, \dots, d_m$ are differencing weights minimizing $\min_{d_0,\dots,d_m} \delta = \sum_{l=1}^{m} \left(\sum_{j=0}^{m-l} d_j d_{l+j}\right)^2$ 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 to have first and last rows $[d', 0'_{n-m-1}]$, $[0'_{n-m-1}, d']$ respectively, with the ith row $[0_i, d', 0'_{n-m-i-1}]$, $i = 2, \dots, (n-m-1)$, where 0_r indicates an r-vector or all zero elements

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

Applying the differencing matrix to model (3) permits direct estimation of the parametric effect. As a result of developments in Speckman.[4] it is known that the parameter vector β in (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 ordered so that the values of the nonparametric variable(s) are close, the application of the differencing matrix D in model (3) removes the nonparametric effect in large samples. 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 by applying the differencing matrix we have

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

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

or

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

where $\tilde{y} = Dy$, $\tilde{X} = DX$ and $\tilde{\varepsilon} = D\varepsilon$. So that the role of the constraints (5) is now evident.[9, p. 57,15] Yatchew [9] defines a simple differencing estimator of the parameter β in the semiparametric regression model. Thus, standard linear models considerations suggest estimating β by

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

This estimator was first proposed in [8]. Thus, differencing allows one to perform inferences on β as if there were no nonparametric component f in the model (3) (see [8,17]). The modified estimator of σ^2, defined as

$$\hat{\sigma}^2 = \frac{\tilde{y}'(I-P)\tilde{y}}{\text{tr}(D'(I-P)D)}, \tag{9}$$

where $\text{tr}(\cdot)$ is the trace function for a square matrix and P is the projection matrix and defined as

$$P = \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'. \tag{10}$$

(see [9]).

3. Semiparametric regression models with correlated errors

In this section we consider the following semiparametric 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

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

where $V_D = \text{DVD}' \neq I_{n-m}$ is a known $(n-m) \times (n-m)$ symmetric positive-definite (p.d.) matrix and $\sigma^2 > 0$ is an unknown parameter (see Roozbeh et al. [17]). It is well known that adopting the linear model (7), the unbiased estimator of β is the following generalized difference-based estimator given by

$$\hat{\beta}_{GD} = (\tilde{X}' V_D^{-1} \tilde{X})^{-1} \tilde{X}' V_D^{-1} \tilde{y}. \tag{11}$$

and the modified estimator of σ^2, defined as

$$\begin{aligned} \hat{\sigma}^2_{GD} &= \frac{(\tilde{y} - \tilde{X}\hat{\beta}_{GD})' V_D^{-1} (\tilde{y} - \tilde{X}\hat{\beta}_{GD})}{\mathrm{tr}[D'(I - \tilde{P})D]} \\ &= \frac{\tilde{y}' V_D^{-1/2} (I - \tilde{P}) V_D^{-1/2} \tilde{y}}{\mathrm{tr}[D'(I - \tilde{P})D]}, \end{aligned} \tag{12}$$

where $\tilde{P}$ is the projection matrix and defined as $\tilde{P} = V_D^{-1/2} \tilde{X} (\tilde{X}' V_D^{-1} \tilde{X})^{-1} \tilde{X}' V_D^{-1/2}$. It is observed from Equation (11) that the properties of the generalized difference-based estimator of β depend on the characteristics of the information matrix $\tilde{X}' V_D^{-1} \tilde{X} = G$.

If $G : p \times p, p << n - m$ matrix is ill-conditioned with a large condition number, then the $\hat{\beta}_{GD}$ 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.

3.1. *Generalized difference-based restricted Liu estimator*

In this section, we will discuss a biased estimation technique when the matrix G appears to be ill-conditioned. In the literature, there are various biased estimation methods to combat the multicollinearity problem, such as ridge regression estimator $\hat{\beta}(k) = (Z'Z + kI)^{-1} Z'y$, $k > 0$, [10] and the Liu estimator $\hat{\beta}_\eta = (Z'Z + I)^{-1}(Z'Z + \eta I)\hat{\beta}_{\mathrm{OLS}}, 0 < \eta < 1$.[11,13] As a remedy, following [11] we suggest to use the following estimator, namely, generalized difference-based Liu estimator:

$$\hat{\beta}_{GD}(\eta) = (\tilde{X}' V_D^{-1} \tilde{X} + I)^{-1} (\tilde{X}' V_D^{-1} \tilde{y} + \eta \hat{\beta}_{GD}). \tag{13}$$

Applying a penalizing function of the squared norm $\| \eta\hat{\beta}_{GD} - \beta \|^2$ to the generalized least-squares objective $(\tilde{y} - \tilde{X}\beta)' V_D^{-1} (\tilde{y} - \tilde{X}\beta)$ for the vector of regression coefficients yields a conditional objective:

$$F(\beta) = \arg\min_{\beta} [(\tilde{y} - \tilde{X}\beta)' V_D^{-1} (\tilde{y} - \tilde{X}\beta) + \| \eta\hat{\beta}_{GD} - \beta \|^2]. \tag{14}$$

The first-order condition of the objective (14) minimized by vector β is

$$\frac{\partial F}{\partial \beta} = 0. \tag{15}$$

From condition (15) of minimizing (14) by the vector β, we obtain the generalized difference-based Liu estimator given by Equation (13).

Now, we consider the linear nonstochastic constraint

$$R\beta = r. \tag{16}$$

For a given $q \times p$ matrix R with rank $q < p$ and a given $q \times 1$ known vector r. We will call an estimator $\beta^*(y)$ for β a restricted estimator with respect to $R\beta = r$, if it satisfies $R\beta^*(y) = r$ for all

$n \times 1$ vectors y (see [18]). Subject to the linear rectriction (16), the generalized difference-based restricted estimator is given by

$$\hat{\beta}_{GRD} = \hat{\beta}_{GD} + G^{-1}R'(RG^{-1}R')^{-1}(r - R\hat{\beta}_{GD}). \tag{17}$$

where $G = \tilde{X}'V_D^{-1}\tilde{X}$ is the information matrix.We can obtain generalized difference-based restricted Liu estimator to improve the generalized difference-based estimator by minimizing the sum of squared residuals with a restriction $R\beta = r$. Now, we can state the following theorem.

THEOREM 1 *The estimate of β in model* (7) *obtained by using Equation* (14)*, satisfying condition* (16)*, is equal to*

$$\hat{\beta}_{GRD}(\eta) = \hat{\beta}_{GD}(\eta) + (G+I)^{-1}R'[R(G+I)^{-1}R']^{-1}(r - R\hat{\beta}_{GD}(\eta)). \tag{18}$$

Proof We construct the following Lagrange function:

$$L(\beta, \lambda) = (\tilde{y} - \tilde{X}\beta)'V_D^{-1}(\tilde{y} - \tilde{X}\beta) + \| \eta\hat{\beta}_{GD} - \beta \|^2 - 2\lambda'(R\beta - r). \tag{19}$$

By differencing function L with respect to β and λ, we obtain

$$\frac{1}{2}\frac{\partial L}{\partial \beta} = -\tilde{X}'V_D^{-1}\tilde{y} + \tilde{X}'V_D^{-1}\tilde{X}\beta - \eta\hat{\beta}_{GD} + \beta - R'\lambda = 0, \tag{20}$$

$$\frac{1}{2}\frac{\partial L}{\partial \lambda} = R\beta - r = 0. \tag{21}$$

Solving Equation (20) with respect to β, we obtain

$$\beta = \hat{\beta}_{GD}(\eta) + (G+I)^{-1}R'\lambda. \tag{22}$$

If we substitute β into Equation (21), we have

$$R\hat{\beta}_{GD}(\eta) + R(G+I)^{-1}R'\lambda = r. \tag{23}$$

Solving Equation (23) with respect to λ, we obtain the optimal value of λ as

$$\hat{\lambda} = [R(G+I)^{-1}R']^{-1}(r - R\hat{\beta}_{GD}(\eta)). \tag{24}$$

Finally, the estimate of β may be written as

$$\hat{\beta}_{GRD}(\eta) = \hat{\beta}_{GD}(\eta) + (G+I)^{-1}R'[R(G+I)^{-1}R']^{-1}(r - R\hat{\beta}_{GD}(\eta)). \tag{25}$$

It is easy to see that $\hat{\beta}_{GRD}(\eta)$ satisfies $R\hat{\beta}_{GRD}(\eta) = r$. Thus, $\hat{\beta}_{GRD}(\eta)$ is the *generalized difference-based restricted Liu estimator* of β in model (7). ∎

THEOREM 2 *If β satisfies the linear restriction* (16)*, then*

$$\text{bias}(\hat{\beta}_{GRD}(\eta)) = E(\hat{\beta}_{GRD}(\eta)) - \beta = -(1-\eta)M\beta. \tag{26}$$

where $C = G + I$ and $M = C^{-1} - C^{-1}R'(RC^{-1}R')^{-1}RC^{-1}$.

Proof The bias of the generalized difference-based Liu estimator is calculated as follows:

$$E(\hat{\beta}_{GD}(\eta)) = E[(G+I)^{-1}(G+\eta I)\hat{\beta}_{GD}]$$
$$= [I-(1-\eta)C^{-1}]\beta. \tag{27}$$

and

$$\text{bias}(\hat{\beta}_{GD}(\eta)) = -(1-\eta)C^{-1}\beta.$$

Then the bias of $\hat{\beta}_{GD}$ is obtained by letting $\eta = 1$ in Equation (27) as follows $E(\hat{\beta}_{GD} - \beta) = 0$.

Let $\beta_0 = R'(RR')^{-1}r$, it can be clearly seen that β_0 satisfies the linear restriction $R\beta = r$. So we have

$$\begin{aligned}\hat{\beta}_{GRD}(\eta) &= \hat{\beta}_{GD}(\eta) + C^{-1}R'(RC^{-1}R')^{-1}(r - R\hat{\beta}_{GD}(\eta)) \\ &= C^{-1}C\hat{\beta}_{GD}(\eta) + C^{-1}R'(RC^{-1}R')^{-1}r - C^{-1}R'(RC^{-1}R')^{-1}RC^{-1}C\hat{\beta}_{GD}(\eta) \\ &= MC\hat{\beta}_{GD}(\eta) + \beta_0 - MC\beta_0,\end{aligned} \tag{28}$$

From (27) and (28) we have

$$\begin{aligned}E(\hat{\beta}_{GRD}(\eta)) &= MC[I-(1-\eta)C^{-1}]\beta + \beta_0 - MC\beta_0 \\ &= MC(\beta-\beta_0) + \beta_0 - (1-\eta)M\beta.\end{aligned} \tag{29}$$

Since the restrictions $R\beta = r$ are assumed to be true and $R\beta_0 = r$, we can easily obtain

$$\begin{aligned}MC(\beta-\beta_0) &= (I - R^{g_3}R)(\beta-\beta_0), \\ &= (\beta-\beta_0) - R^{g_3}(R\beta - R\beta_0) = \beta - \beta_0,\end{aligned} \tag{30}$$

where $R^{g_3} = R^- = C^{-1}R'(RC^{-1}R')^{-1}$ is a normalized generalized inverse of R (see Pringle and Rayner [19, p. 15]). Thus, we obtain

$$\begin{aligned}E(\hat{\beta}_{GRD}(\eta)) &= \beta - \beta_0 - (1-\eta)M\beta + \beta_0 \\ &= \beta - (1-\eta)M\beta\end{aligned} \tag{31}$$

and

$$\text{bias}(\hat{\beta}_{GRD}(\eta)) = -(1-\eta)M\beta. \tag{32}$$

Since M is a nonzero matrix, $\hat{\beta}_{GRD}(\eta)$ can be an unbiased estimator if and only if $\eta = 1$.

Thus, the proof is completed. ■

3.2. *The risk function for the* $\hat{\beta}_{GRD}(\eta)$ *associated with weighted BLF*

Roozbeh et al. [17] calculated the risk function for the generalized difference-based restricted ridge estimator under the weighted BLF. In this section, we calculate the risk function for the proposed estimator, $\hat{\beta}_{GRD}(\eta)$. To derive the risk function of the estimator, it is necessary to specify the loss function under study. Considering the goodness-of-fit and precision of estimation together, Zellner [20] has considered the following loss function:

$$L(\beta^*, \beta) = w(\tilde{X}\beta^* - \tilde{y})'(\tilde{X}\beta^* - \tilde{y}) + (1-w)(\tilde{X}\beta^* - \tilde{X}\beta)'(\tilde{X}\beta^* - \tilde{X}\beta), \tag{33}$$

where w is a nonstochastic weight such that $0 \leq w \leq 1$. Now, we consider the following weighted BLF:

$$L(\hat{\beta}_{GRD},\beta) = w(\tilde{X}\hat{\beta}_{GRD} - \tilde{y})'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD} - \tilde{y}) + (1-w)(\tilde{X}\hat{\beta}_{GRD} - \tilde{X}\beta)'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD} - \tilde{X}\beta) \tag{34}$$

$\hat{\beta}_{GRD}$ is the *generalized difference-based restricted estimator* of β and V_D^{-1} is the weight matrix. The risk function associated with the BLF given by Equation (34) is as follows:

$$R(\hat{\beta}_{GRD},\beta) = E(L(\hat{\beta}_{GRD},\beta)). \tag{35}$$

Substituting $\tilde{y} \approx \tilde{X}\beta + \tilde{\varepsilon}$ in Equation (35), the risk function of the estimator under study can bewritten as

$$\begin{aligned} R(\hat{\beta}_{GRD},\beta) = E(L(\hat{\beta}_{GRD},\beta) &= wE\{(\tilde{X}\hat{\beta}_{GRD} - \tilde{y})'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD} - \tilde{y}) \\ &\quad + (1-w)(\tilde{X}\hat{\beta}_{GRD} - \tilde{X}\beta)'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD} - \tilde{X}\beta)\} \\ &= E\{(\hat{\beta}_{GRD} - \beta)'\tilde{X}'V_D^{-1}\tilde{X}(\hat{\beta}_{GRD} - \beta)\} + wE[\tilde{\varepsilon}'V_D^{-1}\tilde{\varepsilon}] \\ &\quad - 2wE[(\hat{\beta}_{GRD} - \beta)\tilde{X}'V_D^{-1}\tilde{\varepsilon}]. \end{aligned} \tag{36}$$

(i) Since $E(\tilde{\varepsilon}) = 0$ and $E(\tilde{\varepsilon}\tilde{\varepsilon}') = \sigma^2 V_D$, then we have

$$E(\tilde{\varepsilon}'V_D^{-1}\tilde{\varepsilon}) = \operatorname{tr}(\tilde{\varepsilon}\tilde{\varepsilon}'V_D^{-1}) = \sigma^2 \operatorname{tr}(V_D V_D^{-1}) = \sigma^2 \operatorname{tr}(I_{n-m}) = \sigma^2(n-m). \tag{37}$$

(ii) Since $\hat{\beta}_{GD} = (\tilde{X}'V_D^{-1}\tilde{X})^{-1}(\tilde{X}'V_D^{-1}\tilde{y})$ and $\hat{\beta}_{GRD} = \hat{\beta}_{GD} + G^{-1}R'(RG^{-1}R')^{-1}(r - R\hat{\beta}_{GD})$, from Theorem 2, we have

$$\begin{aligned} \hat{\beta}_{GRD} &= M_0 G\hat{\beta}_{GD} + \beta_0 - M_0 G\beta_0 \\ &= M_0 G(G^{-1}\tilde{X}'V_D^{-1}\tilde{y}) + \beta_0 - M_0 G\beta_0 \\ &= \beta - M_0\tilde{X}'V_D^{-1}\tilde{\varepsilon}, \end{aligned}$$

where $G = \tilde{X}'V_D^{-1}\tilde{X}$. Then we have

$$\hat{\beta}_{GRD} - \beta = M_0\tilde{X}'V_D^{-1}\tilde{\varepsilon}. \tag{38}$$

Thus, we obtain

$$\begin{aligned} E\{(\hat{\beta}_{\text{GRD}} - \beta)'\tilde{X}'V_D^{-1}\tilde{\varepsilon}\} = E[\tilde{\varepsilon}'V_D^{-1}\tilde{X}M_0\tilde{X}'V_D^{-1}\tilde{\varepsilon}] &= E\{\operatorname{tr}[\tilde{\varepsilon}\tilde{\varepsilon}'V_D^{-1}\tilde{X}M_0\tilde{X}'V_D^{-1}]\} \\ &= \sigma^2 \operatorname{tr}(M_0 G). \end{aligned}$$

(iii)

$$\begin{aligned} E\{(\hat{\beta}_{\text{GRD}} - \beta)'\tilde{X}'V_D^{-1}\tilde{X}(\hat{\beta}_{\text{GRD}} - \beta)\} &= \operatorname{tr}\{E[\tilde{\varepsilon}'V_D^{-1}\tilde{X}M_0 G M_0\tilde{X}'V_D^{-1}\tilde{\varepsilon}] \\ &= \sigma^2 \operatorname{tr}(M_0 G), \end{aligned}$$

where $M_0 G M_0 = M_0$. Finally, we have

$$R(\hat{\beta}_{GRD},\beta) = w\sigma^2(n-m) + \sigma^2\operatorname{tr}(M_0 G) - 2w\sigma^2\operatorname{tr}(M_0 G).$$

Since

$$M_0 G = (G^{-1} - G^{-1}R'(RG^{-1}R')^{-1}RG^{-1})G = I - R^- R,$$

and

$$\text{tr}(M_0G) = \text{tr}(I - R^-R) = \text{tr}(I_p) - \text{tr}(RR^-) = \text{tr}(I_p) - \text{tr}(I_q) = p - q,$$

then we obtain

$$R(\hat{\beta}_{GRD}, \beta) = w\sigma^2(n-m) + \sigma^2(p-q) - 2w\sigma^2(p-q). \tag{39}$$

Now, we consider the following weighted BLF:

$$\begin{aligned} L(\hat{\beta}_{GRD}(\eta), \beta) &= w(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{y})'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{y}) \\ &\quad + (1-w)(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{X}\beta)'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{X}\beta), \end{aligned} \tag{40}$$

where $\hat{\beta}_{GRD}(\eta)$ is the *generalized difference-based restricted Liu estimator* of β and V_D^{-1} is the weight matrix. The risk function associated with the BLF given by Equation (40) is as follows:

$$R(\hat{\beta}_{GRD}(\eta), \beta) = E(L(\hat{\beta}_{GRD}(\eta), \beta)). \tag{41}$$

Substituting $\tilde{y} \approx \tilde{X}\beta + \tilde{\varepsilon}$ in Equation (41), the risk function of the estimator under study can be written as

$$\begin{aligned} R(\hat{\beta}_{GRD}(\eta), \beta) &= wE\{(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{y})'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{y}) \\ &\quad + (1-w)(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{X}\beta)'V_D^{-1}(\tilde{X}\hat{\beta}_{GRD}(\eta) - \tilde{X}\beta)\} \\ &= E\{(\hat{\beta}_{GRD}(\eta) - \beta)'\tilde{X}'V_D^{-1}\tilde{X}(\hat{\beta}_{GRD}(\eta) - \beta)\} \\ &\quad + wE[\tilde{\varepsilon}'V_D^{-1}\tilde{\varepsilon}] - 2wE[(\hat{\beta}_{GRD}(\eta) - \beta)\tilde{X}'V_D^{-1}\tilde{\varepsilon}] \end{aligned} \tag{42}$$

(iv) Since $\hat{\beta}_{GD}(\eta) = (\tilde{X}'V_D^{-1}\tilde{X} + I)^{-1}(\tilde{X}'V_D^{-1}\tilde{y} + \eta\hat{\beta}_{GD})$, from Theorem 2, we have

$$\begin{aligned} \hat{\beta}_{GRD}(\eta) &= MC\hat{\beta}_{GD}(\eta) + \beta_0 - MC\beta_0 \\ &= MCC^{-1}(G + \eta I)\hat{\beta}_{GD} + \beta_0 - MC\beta_0 \\ &= M(G + \eta I)G^{-1}\tilde{X}'V_D^{-1}\tilde{y} + \beta_0 - MC\beta_0 \\ &= \beta - (1-\eta)M\beta + M(G + \eta I)G^{-1}\tilde{X}'V_D^{-1}\tilde{\varepsilon}, \end{aligned}$$

where $C = G + I$. Then we obtain,

$$\begin{aligned} E\{(\hat{\beta}_{GRD}(\eta) - \beta)'\tilde{X}'V_D^{-1}\tilde{\varepsilon}\} &= E[\{M(G + \eta I)G^{-1}\tilde{X}'V_D^{-1}\tilde{\varepsilon} - ((1-\eta)M\beta)\}'\tilde{X}'V_D^{-1}\tilde{\varepsilon}] \\ &= \sigma^2\,\text{tr}[M(G + \eta I)]. \end{aligned} \tag{43}$$

Since $MC = I - R^-R$ and $\text{tr}(I - R^-R) = p - q$, thus we have $\text{tr}[M(G + \eta I)] = \text{tr}[MC - (1-\eta)M] = \text{tr}(\text{MC}) - (1-\eta)\text{tr}\,M = (p-q) - (1-\eta)\text{tr}\,M$ and $E\{(\hat{\beta}_{GRD}(\eta) - \beta)'\tilde{X}'V_D^{-1}\tilde{\varepsilon}\} = \sigma^2(p-q) - \sigma^2(1-\eta)\text{tr}\,M = \sigma^2(p-q) - \sigma^2(1-\eta)\text{tr}\,M$.

(v)

$$
\begin{aligned}
&E\{(\hat{\beta}_{GRD}(\eta)-\beta)'\tilde{X}'V_D^{-1}\tilde{X}(\hat{\beta}_{GRD}(\eta)-\beta)\}\\
&=\sigma^2\,\mathrm{tr}[M(G+\eta I)G^{-1}(G+\eta I)MG]+(1-\eta)^2\{\beta' MGM\beta\}\\
&=\sigma^2\,\mathrm{tr}[(G+\eta I)G^{-1}(G+\eta I)MGM]+(1-\eta)^2\{\beta' MGM\beta\}\\
&=\sigma^2\,\mathrm{tr}[(G+\eta I)G^{-1}(G+\eta I)(M-M^2)]+(1-\eta)^2\{\beta' MGM\beta\}.
\end{aligned}
\tag{44}
$$

Finally, combining (37), (43) and (44) we have

$$
\begin{aligned}
R(\hat{\beta}_{GRD}(\eta),\beta)&=w\sigma^2(n-m)-2w\sigma^2(p-q)-2w(1-\eta)\sigma^2\mathrm{tr}\,M\\
&\quad+\sigma^2\,\mathrm{tr}[(G+\eta I)G^{-1}(G+\eta I)(M-M^2)]+(1-\eta)^2\{\beta' MGM\beta\}.
\end{aligned}
\tag{45}
$$

4. Comparison results

In this section, we provide the necessary and sufficient conditions for which the estimator $\hat{\beta}_{GRD}(\eta)$ performs better than $\hat{\beta}_{GRD}$ in the sense that $R(\hat{\beta}_{GRD}(\eta),\beta)\le R(\hat{\beta}_{GRD},\beta)$ From Equations (35) and (45), the difference $\tilde{\Delta}=R(\hat{\beta}_{GRD},\beta)-R(\hat{\beta}_{GRD}(\eta),\beta)$ is given by

$$
\begin{aligned}
\tilde{\Delta}&=[]\{w\sigma^2(n-m)+\sigma^2(p-q)-2w\sigma^2(p-q)\}-\{w\sigma^2(n-m)-2w\sigma^2(p-q)\\
&\quad-2w(1-\eta)\sigma^2\mathrm{tr}M+\sigma^2\mathrm{tr}[(G+\eta I)G^{-1}(G+\eta I)(M-M^2)]+(1-\eta)^2\,\mathrm{tr}(\beta' MGM\beta)\}\\
&=\sigma^2(p-q)+2w(1-\eta)\sigma^2\,\mathrm{tr}\,M-\sigma^2\,\mathrm{tr}[(G+\eta I)G^{-1}(G+\eta I)(M-M^2)]\\
&\quad-(1-\eta)^2\,\mathrm{tr}(\beta' MGM\beta)\}
\end{aligned}
\tag{46}
$$

when $R\beta=r$. Since, $MCM=M$ and

$$
[(G+\eta I)G^{-1}(G+\eta I)(M-M^2)]=(C-(1-2\eta)I+\eta^2G^{-1})(M-M^2),
$$

we have

$$
\begin{aligned}
&\mathrm{tr}(C-(1-2\eta)I+\eta^2G^{-1})(M-M^2)\\
&=\mathrm{tr}\{CM-(1-2\eta)M-CM^2+(1-2\eta)M^2+\eta^2G^{-1}(M-M^2)\}\\
&=(p-q)-(1-2\eta)\mathrm{tr}(M)-\mathrm{tr}(MCM)+(1-2\eta)\mathrm{tr}(M^2)+\eta^2\mathrm{tr}(G^{-1}(M-M^2)).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\tilde{\Delta}&=2\sigma^2(1-w)(1-\eta)\,\mathrm{tr}\,M-2\sigma^2(1-\eta)\,\mathrm{tr}\,M^2\\
&\quad+\sigma^2(1-\eta^2)\mathrm{tr}(G^{-1}(M-M^2))-(1-\eta)^2\beta' MGM\beta,
\end{aligned}
\tag{47}
$$

$$
\frac{\mathrm{d}}{\mathrm{d}\eta}\tilde{\Delta}=-2\sigma^2(1-w)\,\mathrm{tr}\,M+2\sigma^2\,\mathrm{tr}\,M^2-2\sigma^2\eta\,\mathrm{tr}[G^{-1}(M-M^2)]+2(1-\eta)\beta' MGM\beta=0
$$

$$
\begin{aligned}
\Rightarrow \eta_{\mathrm{opt}}&=\frac{-\sigma^2(1-w)\,\mathrm{tr}\,M+\sigma^2\,\mathrm{tr}\,M^2+\beta' MGM\beta}{\beta' MGM\beta+\sigma^2\,\mathrm{tr}[G^{-1}(M-M^2)]}\\
&=\frac{-\sigma^2\,\mathrm{tr}(M-M^2)+\sigma^2 w\,\mathrm{tr}M+\beta'(M-M^2)\beta}{\beta'(M-M^2)\beta+\sigma^2\,\mathrm{tr}[G^{-1}(M-M^2)]},
\end{aligned}
\tag{48}
$$

$$
\frac{\mathrm{d}^2}{\mathrm{d}\eta^2}\tilde{\Delta}=-2\sigma^2\,\mathrm{tr}[G^{-1}(M-M^2)]-2\beta' MGM\beta.
\tag{49}
$$

So, if $-\beta'(M-M^2)\beta < \sigma^2\,\mathrm{tr}[G^{-1}(M-M^2)]$, we can conclude that η_{opt} maximizes the $\tilde{\Delta}$ and is the best η.

5. Simulation study

In this section, we examine the risk function performance of the proposed estimators. Our sampling experiment consists of different combinations of η and w, i.e.

$$\eta = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\} \quad \text{and} \quad w = \{0.1, 0.3, 0.5, 0.7, 0.9\}. \quad \text{and}$$

$$w = \{0.1, 0.3, 0.5, 0.7, 0.9\}.$$

To achieve different degrees of collinearity, following McDonald and Galarneau [21] and Gibbons [22] the explanatory were generated using the following device for $n = 1000$:

$$x_{ij} = (1-\gamma^2)^{1/2} z_{ij} + \gamma z_{ip}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, p, \tag{50}$$

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. Three different set of correlation corresponding to $\gamma = 0.80, 0.90$ and 0.99 are considered. Then n observations for the dependent variable are determined by

$$y_i = \sum_{j=1}^{6} x_{ji}\beta_j + f(t_i) + \varepsilon_i, \quad i = 1, 2, \ldots, n, \tag{51}$$

where $\beta = (3, 1, 3, 2, -5, 4)$ and $\varepsilon \sim N(0, \sigma^2 V)$ for which the elements of V are

$$v_{ij} = \left(\frac{1}{n}\right)^{|i-j}, \quad \sigma^2 = 6 \quad \text{and}$$

$$f(t_i) = \frac{1}{6} \sum_{j=1, j\neq 4,6} \varphi\left(t_i; j, \left[\frac{j+2}{10}\right]^j\right),$$

which is mixture of normals for $t_i = 10i/n$ and $\phi(x; \mu, \sigma^2)$ is a normal density function with mean μ and variance σ^2.

In model (51) the parametric effect, β, is estimated by a differencing procedure. Optimal differencing weights do not have analytic expressions, but may be calculated easily using an optimization routine. Hall et al. [23] present weights to order $m = 10$. These contain some minor errors. Optimal difference sequences for $1 \leq m \leq 10$ can be found in [9]. For the simulation study, we used order $m = 4$. For orders higher than 4, the estimators did not give any better results, therefore we did not include them in the simulation study for restrictive purposes. The fourth-order optimal differencing weights, for example $d_0 = 0.8873, d_1 = -0.3099, d_2 = 0.2464$, $d_3 = -0.1901$ and $d_4 = -0.1409$ in which $m = 4$ (see [9, p. 61]). Now, we define the $(n-4) \times n$ differencing matrix as

$$D = \begin{pmatrix} d_0 & d_1 & d_2 & d_3 & d_4 & 0 & 0 & \cdots & 0 \\ 0 & d_0 & d_1 & d_2 & d_3 & d_4 & 0 & \cdots & 0 \\ \vdots & \ddots & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & d_0 & d_1 & d_2 & d_3 & d_4 \end{pmatrix}$$

Table 1. Evaluation of parameters and risk functions for different values of η and w ($n = 1000$ and $\gamma = 0.8$).

η	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\hat{\beta}_1$	3.0342	3.0314	3.0285	3.0257	3.0229	3.0201	3.0173	3.0117	3.01	3.0090	3.0062
$\hat{\beta}_2$	0.9824	0.9839	0.9853	0.9868	0.9882	0.9896	0.9911	0.9939	0.9939	0.9953	0.9967
$\hat{\beta}_3$	2.9795	2.9812	2.9829	2.9845	2.9862	2.9879	2.9896	2.9929	2.9929	2.9926	2.9962
$\hat{\beta}_4$	2.0074	2.0068	2.0062	2.0056	2.0049	2.0043	2.0037	2.0025	2.0025	2.0019	2.0013
$\hat{\beta}_5$	−4.9994	−4.9995	−4.9995	−4.9995	−4.9996	−4.9996	−4.9997	−4.9998	−4.9998	−4.9998	−4.9999
$\hat{\beta}_6$	4.0643	4.0589	4.0536	4.0483	4.0430	4.0377	4.0325	4.0221	4.0221	4.0169	4.0117
$R(w = 0.1)$	444.000	443.977	443.959	443.945	443.937	443.933	443.934	443.940	443.950	443.965	443.984
$R(w = 0.3)$	473.600	473.575	473.555	473.540	473.530	473.525	473.524	473.528	473.537	473.550	473.568
$R(w = 0.5)$	503.200	503.173	503.152	503.135	503.124	503.117	503.114	503.117	503.124	503.135	503.151
$R(w = 0.7)$	532.800	532.772	532.748	532.730	532.717	532.708	532.704	532.705	532.710	532.720	532.735
$R(w = 0.9)$	562.400	562.370	562.345	562.325	562.310	562.300	562.294	562.293	562.297	562.306	562.318
$\tilde{\Delta}(w = 0.1)$	−0.0153	0.0075	0.0256	0.0388	0.0472	0.0509	0.0499	0.0443	0.0340	0.0192	0.0000
$\tilde{\Delta}(w = 0.3)$	−0.0318	−0.0071	0.0125	0.0274	0.0375	0.0428	0.0434	0.0394	0.0308	0.0176	0.0000
$\tilde{\Delta}(w = 0.5)$	−0.0482	−0.0219	−0.0005	0.0160	0.0277	0.0347	0.0370	0.0346	0.0276	0.0160	0.0000
$\tilde{\Delta}(w = 0.7)$	−0.0646	−0.0366	−0.0135	0.0046	0.0180	0.0266	0.0305	0.0298	0.0244	0.0144	0.0000
$\tilde{\Delta}(w = 0.9)$	−0.0810	−0.0513	−0.0266	−0.0067	0.0083	0.0185	0.0241	0.0250	0.0212	0.0128	0.0000
$\hat{m}$	0.0378	0.0370	0.0362	0.0355	0.0349	0.0342	0.0337	0.0331	0.0326	0.0322	0.0315

Table 2. Evaluation of parameters and risk functions for different values of η and w ($n = 1000$ and $\gamma = 0.9$).

η	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\hat{\beta}_1$	3.0070	3.0036	3.0001	2.9967	2.9933	2.9899	2.9865	2.9831	2.9797	2.9764	2.9730
$\hat{\beta}_2$	0.9963	0.9981	0.9999	1.0016	1.0034	1.0051	1.0069	1.0086	1.0103	1.0120	1.0137
$\hat{\beta}_3$	2.9957	2.9978	2.9999	3.0019	3.0039	3.0060	3.0080	3.0100	3.0120	3.0140	3.0160
$\hat{\beta}_4$	2.0015	2.0007	2.0000	1.9992	1.9985	1.9978	1.9970	1.9963	1.9955	1.9948	1.9941
$\hat{\beta}_5$	−4.9998	−4.9999	−4.9999	−5.0000	−5.0001	−5.0001	−5.0002	−5.0002	−5.0003	−5.0003	−5.0004
$\hat{\beta}_6$	4.0132	4.0067	4.0003	3.9938	3.9874	3.9810	3.9747	3.9683	3.9620	3.9557	3.9495
$R(w = 0.1)$	444.000	443.971	443.949	443.933	443.922	443.918	443.919	443.926	443.939	443.957	443.980
$R(w = 0.3)$	473.600	473.569	473.545	473.526	473.514	473.507	473.507	473.512	473.522	473.539	473.560
$R(w = 0.5)$	503.200	503.167	503.141	503.120	503.106	503.097	503.094	503.097	503.106	503.120	503.140
$R(w = 0.7)$	532.800	532.765	532.736	532.714	532.698	532.687	532.682	532.683	532.690	532.702	532.720
$R(w = 0.9)$	562.400	562.363	562.332	562.308	562.289	562.277	562.270	562.269	562.274	562.284	562.300
$\tilde{\Delta}(w = 0.1)$	−0.0191	0.0092	0.0315	0.0477	0.0581	0.0626	0.0614	0.0544	0.0418	0.0236	0.0000
$\tilde{\Delta}(w = 0.3)$	−0.0393	−0.0088	0.0154	0.0338	0.0462	0.0527	0.0535	0.0485	0.0379	0.0217	0.0000
$\tilde{\Delta}(w = 0.5)$	−0.0594	−0.0269	−0.0005	0.0198	0.0343	0.0428	0.0456	0.0426	0.0340	0.0197	0.0000
$\tilde{\Delta}(w = 0.7)$	−0.0796	−0.0450	−0.0165	0.0059	0.0223	0.0329	0.0377	0.0367	0.0300	0.0178	0.0000
$\tilde{\Delta}(w = 0.9)$	−0.0998	−0.0631	−0.0325	−0.0080	0.0104	0.0230	0.0298	0.0308	0.0261	0.0158	0.0000
$\hat{m}$	0.06706	0.06705	0.06704	0.06703	0.06701	0.03700	0.06699	0.06698	0.06697	0.06696	0.06695

Table 3. Evaluation of parameters and risk functions for different values of η and w ($n = 1000$ and $\gamma = 0.99$).

η	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\hat{\beta}_1$	3.0347	3.0295	3.0244	3.0193	3.0143	3.0092	3.0042	2.9993	2.9943	2.9894	2.9845
$\hat{\beta}_2$	0.9822	0.9848	0.9874	0.9900	0.9926	0.9952	0.9978	1.0003	1.0028	1.0053	1.0078
$\hat{\beta}_3$	2.9792	2.9823	2.9853	2.9884	2.9914	2.9944	2.9974	3.0004	3.0033	3.0063	3.0092
$\hat{\beta}_4$	2.0075	2.0064	2.0053	2.0042	2.0031	2.0020	2.0009	1.9998	1.9987	1.9977	1.9966
$\hat{\beta}_5$	−4.9994	−4.9995	−4.9996	−4.9996	−4.9997	−4.9998	−4.9999	−5.0000	−5.0000	−5.0001	−5.0002
$\hat{\beta}_6$	4.0651	4.0554	4.0458	4.0363	4.0268	4.0173	4.0080	3.9986	3.9894	3.9802	3.9710
$R(w = 0.1)$	444.000	443.958	443.926	443.902	443.887	443.881	443.883	443.893	443.911	443.937	443.971
$R(w = 0.3)$	473.600	473.555	473.520	473.493	473.475	473.466	473.465	473.472	473.487	473.511	473.542
$R(w = 0.5)$	503.200	503.152	503.113	503.084	503.063	503.051	503.047	503.051	503.064	503.085	503.113
$R(w = 0.7)$	532.800	532.749	532.707	532.675	532.651	532.636	532.629	532.631	532.640	532.658	532.684
$R(w = 0.9)$	562.400	562.346	562.301	562.266	562.239	562.221	562.211	562.210	562.217	562.232	562.255
$\tilde{\Delta}(w = 0.1)$	−0.0283	0.0131	0.0456	0.0692	0.0841	0.0906	0.0886	0.0784	0.0601	0.0339	0.0000
$\tilde{\Delta}(w = 0.3)$	−0.0574	−0.0128	0.0226	0.0492	0.0671	0.0764	0.0774	0.0700	0.0546	0.0312	0.0000
$\tilde{\Delta}(w = 0.5)$	−0.0865	−0.0389	−0.0003	0.0292	0.0500	0.0623	0.0661	0.0616	0.0490	0.0284	0.0000
$\tilde{\Delta}(w = 0.7)$	−0.1155	−0.0649	−0.0233	0.0092	0.0330	0.0481	0.0549	0.0532	0.0435	0.0256	0.0000
$\tilde{\Delta}(w = 0.9)$	−0.1446	−0.0909	−0.0463	−0.0108	0.0159	0.0340	0.0436	0.0448	0.0379	0.0229	0.0000
$\widehat{\text{MSE}}(\hat{f}(u), f(u))$	0.09180	0.08918	0.08664	0.08416	0.08176	0.07942	0.07715	0.07494	0.07279	0.07071	0.06861

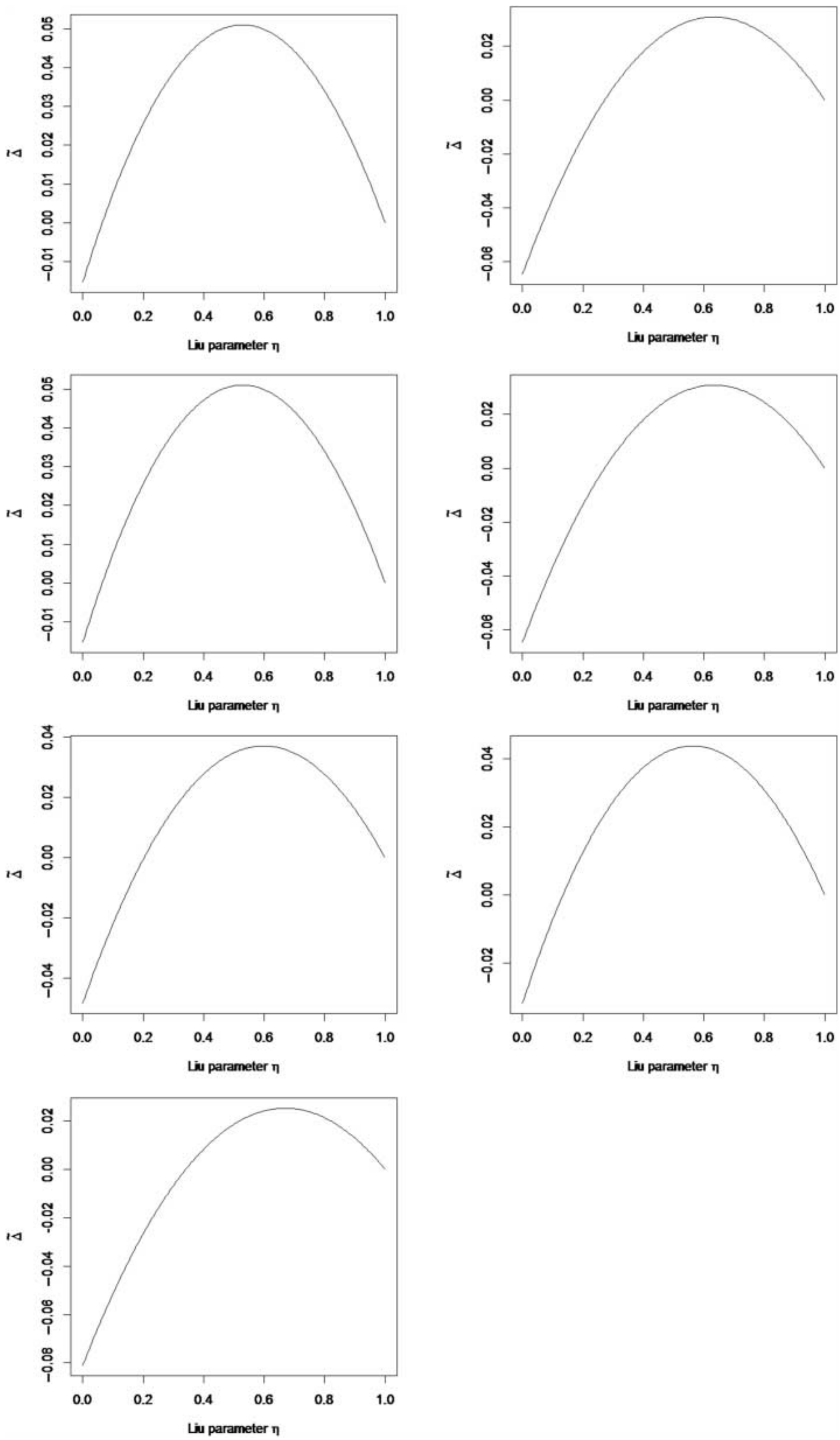

Figure 1. The diagram of $\tilde{\Delta}$ versus η for different values w and $\gamma = 0.8$. Top left: $w = 0.1$; top right: $w = 0.3$; middle left: $w = 0.5$; middle right: $w = 0.7$ and bottom left: $w = 0.9$.

For the linear restriction, the R matrix is given by

$$R = \begin{pmatrix} 1 & 5 & -3 & -1 & -1 & 0 \\ -2 & -1 & 0 & -2 & 3 & 1 \\ 1 & 2 & 1 & 3 & -2 & 0 \\ 4 & -1 & 2 & 2 & 0 & -2 \\ 5 & 3 & 4 & -5 & 1 & 0 \end{pmatrix}$$

and r is a vector, i.e. $r = [00000]'$.

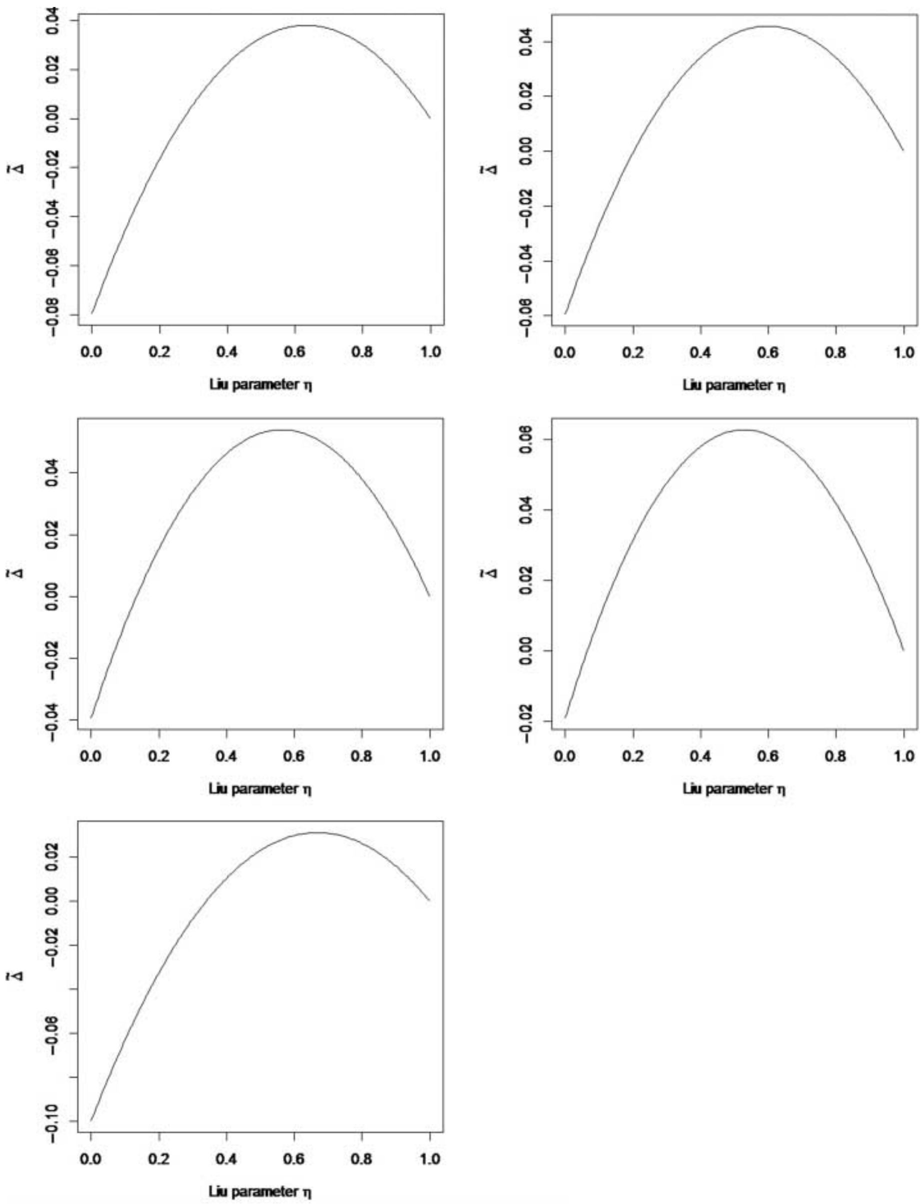

Figure 2. The diagram of $\tilde{\Delta}$ versus η for different values w and $\gamma = 0.9$. Top left: $w = 0.1$; top right: $w = 0.3$; middle left: $w = 0.5$; middle right: $w = 0.7$ and bottom left: $w = 0.9$.

All computations were conducted using the R statistical system. The matrix G has condition numbers 44.4, 78.49, and 787.82 for $\gamma = 0.8$, $\gamma = 0.9$ and $\gamma = 0.99$, respectively, which implies the existence of multicollinearity in the data set. In Tables 1–3 we computed the generalized difference-based restricted Liu estimators of parameters and risk functions values. We numerically calculated the $R(\hat{\beta}_{GRD}, \beta)$, $\tilde{\Delta}$ and

$$\widehat{\mathrm{MSE}}(\hat{f}(u), f(u)) = \frac{1}{n}\sum_{i=1}^{n}[\hat{f}(u_i) - f(u_i)]^2$$

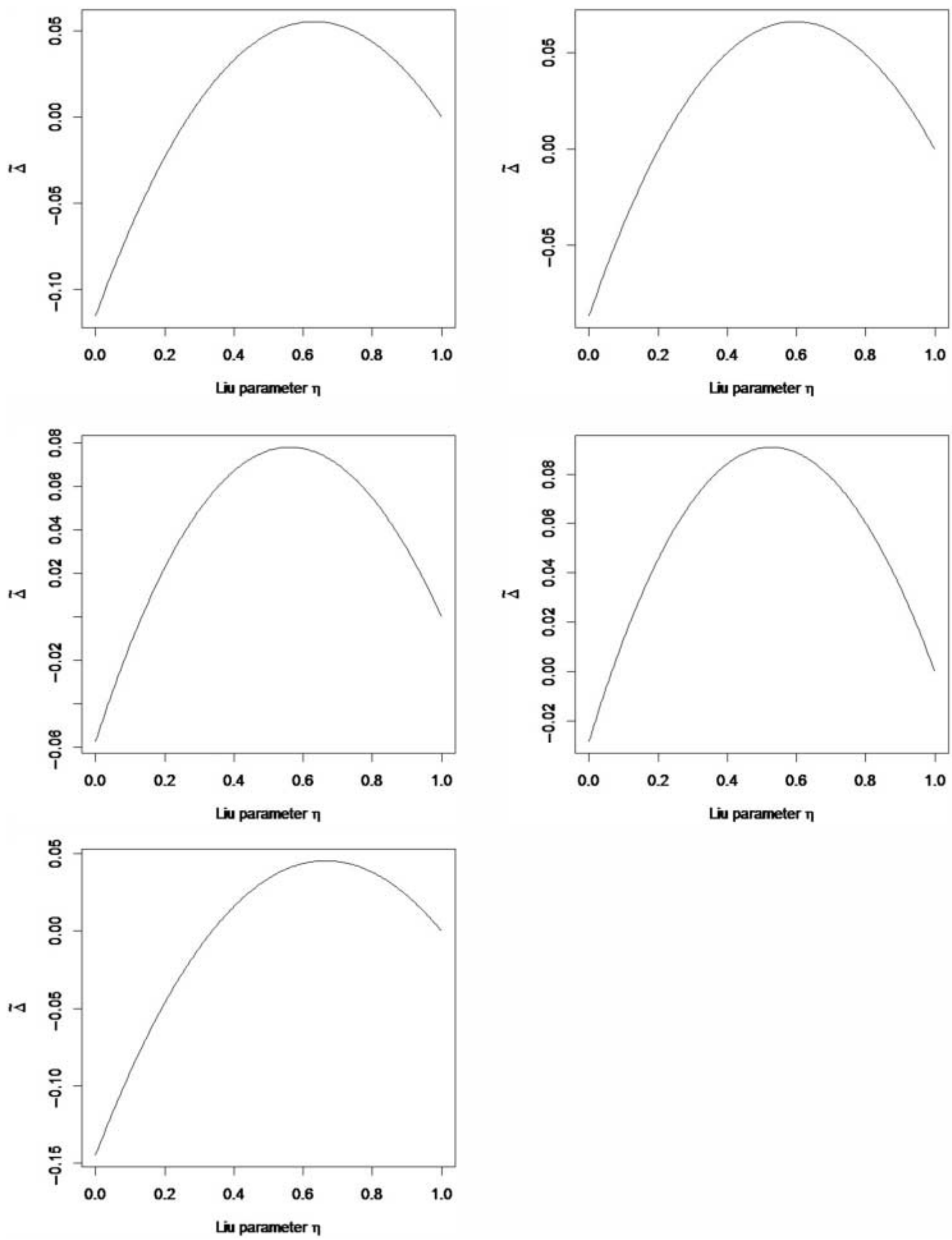

Figure 3. The diagram of $\tilde{\Delta}$ versus η for different values w and $\gamma = 0.99$. Top left: $w = 0.1$; top right: $w = 0.3$; middle left: $w = 0.5$; middle right: $w = 0.7$ and bottom left: $w = 0.9$.

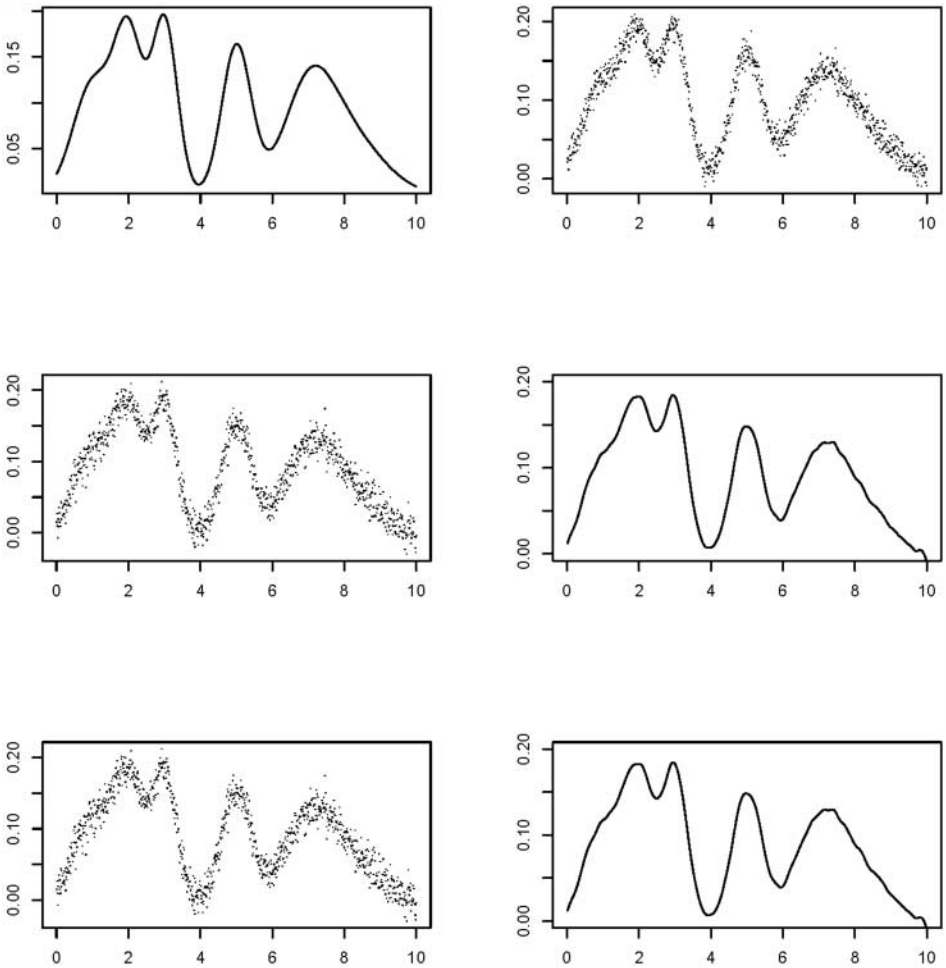

Figure 4. Estimation of the function under study by local linear regression for $n = 1000$. The function (top left), the data (top right) and the residuals obtained after estimation of the linear part of the model by $\hat{\beta}_{GRD}$, that is, $y - X\hat{\beta}_{GRD}$ (middle left), the fitted function for $\eta = 1$ (middle right), the residuals obtained after estimation of the linear part of the model by $\hat{\beta}_{GRD}(0.5)$ (bottom left) and the fitted function for $\eta = 0.5$ (bottom right).

for different η and w values when $n = 1000$. In Figures 1–3 we plotted the $\tilde{\Delta}$ versus Liu parameter η for different values of w and γ. Our methods were applied to several simulated data sets. Because the results were similar across cases, to save space, we reported here only the results for $n = 1000$, $\eta = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ and $w = \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

In Figure 4, we plotted the nonparametric part of the model in the top left plot. This function is difficult to estimate and provides a good test for nonparametric regression methods. The function is spatially inhomogeneous which means that its smoothness (second derivatives) varies over u. The top right plot shows $n = 1000$ data points after removing the linear part, i.e. $y - X\beta$. The middle left and right plot shows the residuals which are obtained after estimation of the linear part of the model by $\hat{\beta}_{GRD}$ and the fitted function, respectively. The bottom left and right plots are the middle part when $\hat{\beta}_{GRD}$ is replaced with $\hat{\beta}_{GRD}(0.5)$ when $\eta = 0.9$.

6. Conclusions

In this paper, we proposed the generalized difference-based Liu estimator, $\hat{\beta}_{GD}(\eta)$, in a semiparametric regression model when the errors were dependent and some additional linear restrictions were held on the whole parameter space β. In the presence of multicollinearity in a semiparametric regression model, we introduced the generalized difference-based restricted Liu estimator, $\hat{\beta}_{GRD}(\eta)$.

The risk functions of the proposed estimators under the weighted BLF were obtained.

We continued the comparison study by some simulation strategies and graphical results. The experiment was taken for different values of Liu parameter η and weight coefficient w in the weighted BLF.

According to Tables 1–3 and Figures 1–3, it can be realized that for all the combinations of η and w, $\hat{\beta}_{GRD}(\eta)$ is better than $\hat{\beta}_{GRD}$ if $\eta > a$, so that a is an increasing function of w.

References

[1] Green P, Jennison C, Seheult A. Analysis of field experiments by least squares smoothing. J Roy Statist Soc Ser B.1985;47:299–315.
[2] Engle RF, Granger CWJ, Rice J, Weiss A. Semiparametric estimates of the relation between weather and electricity sales. J Amer Statist Assoc. 1986;81:310–320.
[3] Eubank RL, Kambour EL, Kim JT, Klipple K, Reese CS, Schimek M. Estimation in partially linear models. Comput Statist Data Anal. 1988;29:27–34.
[4] Speckman P. Kernel somoothing in partial linear models. J Roy Statist Soc Ser B. 1988;50(3):413–436.
[5] Eubank RL. Nonparametric regression and spline smoothing. New York: Marcel Dekker; 1999.
[6] Ruppert D, Wand MP, Carroll RC. Semiparametric regression. Cambridge: Cambridge University Press; 2003.
[7] Härdle W, Müller M, Sperlich S, Werwatz A. Nonparametric and semiparmetric models. Berlin: Springer-Verlag; 2004.
[8] Yatchew A. An elementary estimator of the partial linear model. Econ Lett.1997;57:135–143.
[9] Yatchew A. Semiparametric regression for the applied econometrican. Cambridge: Cambridge University Press; 2003.
[10] Hoerl AE, Kennard RW. Ridge regression: biased estimation for orthogonal problems. Technometrics.1970;12: 55–67.
[11] Liu KJ. A new class of biased estimate in linear regression. Commun Statist Theory Methods. 1993;22:393–402.
[12] Stein C. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. Proc Third Berkeley Symp Math Statist Prob.1956;1:197–206.
[13] Akdeniz F, Kaçıranlar S. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Commun Statist-Theory Methods.1995;24(7):1789–1797.
[14] Gruber MHJ. Improving efficiency by shrinkage: the James–Stein and ridge regression estimators. New York: Marcell Dekker, Inc.; 1998.
[15] Klipple K, Eubank RL. Difference-based variance estimators for partially linear models. Festschrift in honor of Distinguished Professor Mir Masoom Ali on the occasion of his retirement; 2007 May 18–19, 2007. p. 313–323.
[16] Brown LD, Levine M. Variance estimation in nonparametric regression via the difference sequence method. Ann Statist. 2007;35:2219–2232.
[17] Roozbeh M, Arashi M, Niroumand HA. Ridge regression methodology in partial linear models with correlated errors. J Statist Comput Simul. 2011;81(4):517–528.
[18] Gross J. Restricted ridge estimation. Statist Probab Lett. 2003;65:57–64.
[19] Pringle PM, Rayner AA. Generalized inverse of matrices and its applications to statistics. New York: Hafner Publishing Co.; 1971.
[20] Zellner A. Bayesian and non-Bayesian estimation using balanced loss functions. In: Gupta SS, Berger JO, editors. Statistical decision theory and related topics, Vol. V. New York: Springer; 1994. p. 377–390.
[21] McDonald GC, Galarneau DI. A Monte Carlo evaluation of some Ridge-type estimators. J Amer Statist Assoc. 1975;70:407–416.
[22] Gibbons DG. A simulation study of some ridge estimators. J Amer Statist Assoc. 1981;76:131–139.
[23] Hall P, Kay JW, Titterington DM. On estimation of noise variance in two-dimentional signal processing. Adv Appl Probab. 1990;23:476–495.