Journal of Multivariate Analysis 136 (2015) 56–74

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Shrinkage ridge estimators in semiparametric regression models

Mahdi Roozbeh

*Department of Statistics, Faculty of Mathematics, Statistics and Computer Sciences, Semnan University, P.O. Box 35195-363, Semnan, Iran*

ARTICLE INFO

*Article history:*
Received 8 November 2013
Available online 19 January 2015

*AMS subject classifications:*
primary 62G08
secondary 62J05
62J07

*Keywords:*
Generalized restricted ridge estimator
Kernel smoothing
Linear restriction
Multicollinearity
Positive-rule shrinkage
Semiparametric regression model
Stein-type shrinkage

ABSTRACT

In this paper, ridge and non-ridge type shrinkage estimators and their positive parts are defined in the semiparametric regression model when the errors are dependent and some non-stochastic linear restrictions are imposed under a multicollinearity setting. The exact risk expressions in addition to biases are derived for the estimators under study and the region of optimality of each estimator is exactly determined. Also, necessary and sufficient conditions, for the superiority of the ridge type estimator over its counterpart, for selecting the ridge parameter $k$ are obtained. Lastly, a simulation study and real data analysis are performed to illustrate the efficiency of proposed estimators based on the minimum risk criterion. In this regard, kernel smoothing and modified cross-validation methods for estimating the non-parametric function are used.



## 1. Introduction

Let $(y_1, \boldsymbol{x}_1, t_1), \ldots, (y_n, \boldsymbol{x}_n, t_n)$ be observations that follow the semiparametric regression model (SRM)

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + f(t_i) + \epsilon_i, \quad i = 1, \ldots, n \tag{1.1}$$

where $\boldsymbol{x}_i' = (x_{i1}, x_{i2}, \ldots, x_{ip})$ is a vector of explanatory variables, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is an unknown $p$-dimensional parameter vector, the $t_i$'s are known and non-random in some bounded domain $D \in \mathbb{R}$, $f(t_i)$ is an unknown smooth function and $\epsilon_i$'s are independent and identically distributed random errors with mean 0, variance $\sigma^2$, which are independent of $(\boldsymbol{x}_i, t_i)$. Semiparametric regression models are more flexible than standard linear models since they have a parametric and a nonparametric component. They can be a suitable choice when one suspects that the response $y$ linearly depends on $x$, but that it is nonlinearly related to $t$.

Surveys regarding the estimation and application of the model (1.1) can be found in the monograph of Härdle et al. [13]. Bunea [10] suggested a consistent covariate selection technique in an SRM through penalized least squares criterion. He showed that the selected estimator of the linear part is asymptotically normal. For bandwidth selection in the context of kernel-based estimation in model (1.1), Li and Palta [23] and Li et al. [24] used cross-validation criteria for optimal bandwidth selection. Raheem et al. [30] considered absolute penalty and shrinkage estimators in PLMs where the vector of coefficients $\boldsymbol{\beta}$ in the linear part can be partitioned as $(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')$, $\boldsymbol{\beta}_1$ is the coefficient vector of the main effects, and $\boldsymbol{\beta}_2$ is the vector of the nuisance effects.

*E-mail addresses:* m.roozbeh.stat@gmail.com, mahdi.roozbeh@profs.semnan.ac.ir.

http://dx.doi.org/10.1016/j.jmva.2015.01.002

Now, consider a semiparametric regression model in the presence of multicollinearity. 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. For our purpose we only employ the ridge regression concept due to Horel and Kennard [17], to combat multicollinearity. There are a lot of works adopting ridge regression methodology to overcome the multicollinearity problem. To mention a few recent researches in full-parametric regression, see [32,31,29,1,12,28,14,15,19,21,22,2,20,27]. The main focus of this approach is to develop necessary tools for computing the risk function of regression coefficient in a semiparametric regression model based on the eigenvalues of design matrix. We are also seeking a new estimator for shrinkage parameter by making use of the existing ones in the literature. It will be shown that the new estimator performs better than all the others not only for the regression coefficient, but even for the non-parametric component as well.

The study is organized as follows: In Section 2, Stein-type shrinkage as well as its positive part are defined for the regression coefficient, while their biases and risks are driven in Section 3 and detailed analysis is incorporated to compare the performance of the proposed estimators for different values of the ridge parameter. In Section 4, the least/most values of the ridge parameter are identified for which the ridge estimators dominate each other. Section 5 contains the simulation studies and a real data example related to the hedonic prices of housing attributes to demonstrate the performance of the proposed estimators, numerically.

## 2. The proposed estimators

Consider the following semiparametric regression model

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{f}(t) + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{y} = (y_1, \dots, y_n)'$, $\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)'$ is an $n \times p$ matrix, $\boldsymbol{f}(t) = (f(t_1), \dots, f(t_n))'$ and $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)$. We assume that in general, $\boldsymbol{\epsilon}$ is a vector of disturbances, which is distributed as a multivariate normal, $N_n(\boldsymbol{0}, \sigma^2\boldsymbol{V})$, where $\boldsymbol{V}$ is a symmetric, positive definite known matrix and $\sigma^2$ is an unknown parameter.

In this paper we confine ourselves to the semiparametric kernel smoothing estimator of $\boldsymbol{\beta}$, which attains the usual parametric convergence rate $n^{1/2}$ without under smoothing the nonparametric component $f(\cdot)$ [34]. Assume that $(y_i, \boldsymbol{x}_i, t_i),\ i = 1, \dots, n$ satisfy model (1.1). Since $E(\varepsilon_i) = 0$, we have $f(t_i) = E(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})$ for $i = 1, \dots, n$. Hence, if we know $\boldsymbol{\beta}$, a natural nonparametric estimator of $f(\cdot)$ is

$$\hat{f}(t, \boldsymbol{\beta}) = \sum_{i=1}^{n} W_{ni}(t)(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}), \tag{2.2}$$

where the positive weight functions $W_{ni}(\cdot)$ satisfy three conditions below:

(i) $\max_{1\le i\le n} \sum_{j=1}^{n} W_{ni}(t_j) = O(1)$,
(ii) $\max_{1\le i,j\le n} W_{ni}(t_j) = O(n^{-2/3})$,
(iii) $\max_{1\le i\le n} \sum_{j=1}^{n} W_{ni}(t_j) I(|t_i - t_j| > c_n) = O(d_n)$,

where $I$ is the indicator function, $c_n$ satisfies $\limsup_{n\to\infty} nc_n^3 < \infty$, and $d_n$ satisfies $\lim sup_{n\to\infty} nd_n^3 < \infty$.

The above assumptions guarantee the existence of $\hat{f}(t, \boldsymbol{\beta})$ at the optimal convergence rate $n^{-4/5}$, in semiparametric regression models with probability one. See Müller [26] for more details.

To estimate $\boldsymbol{\beta}$, we use the generalized least squares estimator (GLSE) given by

$$\hat{\boldsymbol{\beta}}_G = argmin_{\boldsymbol{\beta}}\, SS(\boldsymbol{\beta}) = \boldsymbol{C}^{-1}\widetilde{\boldsymbol{X}}'\boldsymbol{V}^{-1}\widetilde{\boldsymbol{y}}, \quad \boldsymbol{C} = \widetilde{\boldsymbol{X}}'\boldsymbol{V}^{-1}\widetilde{\boldsymbol{X}}, \tag{2.3}$$

where $SS(\boldsymbol{\beta}) = (\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta})'\boldsymbol{V}^{-1}(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta})$, $\widetilde{\boldsymbol{y}} = (\widetilde{y}_1, \dots, \widetilde{y}_n)'$, $\widetilde{\boldsymbol{X}} = (\widetilde{\boldsymbol{x}}_1, \dots, \widetilde{\boldsymbol{x}}_n)'$, $\widetilde{y}_i = y_i - \sum_{j=1}^{n} W_{nj}(t_i)y_j$ and $\widetilde{\boldsymbol{x}}_i = \boldsymbol{x}_i - \sum_{j=1}^{n} W_{nj}(t_i)\boldsymbol{x}_j$ for $i = 1, \dots, n$.

In this section, we will discuss about a biased estimation technique under multicollinearity. Simultaneously, we assume that $\boldsymbol{\beta}$ satisfies a linear non stochastic constraint, i.e.,

$$\boldsymbol{H\beta} = \boldsymbol{h}, \tag{2.4}$$

where $\boldsymbol{H}$ is a $q \times p$ non zero matrix with rank $q < p$ and $\boldsymbol{h}$ is a $q \times 1$ vector. In this paper, we refer restricted semiparametric regression model (RSRM) to (2.1).

For the RSRM, one generally adopts the well-known generalized restricted estimator (GRE)

$$\hat{\boldsymbol{\beta}}_{GR} = \hat{\boldsymbol{\beta}}_G + \boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{HC}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{h} - \boldsymbol{H}\hat{\boldsymbol{\beta}}_G). \tag{2.5}$$

The GRE is widely applied as an unbiased estimator. In practice, the researchers often encounter the problem of multicollinearity. That is, $\boldsymbol{C}$ is always ill-conditioned due to linear relationship among the regressors of $\widetilde{\boldsymbol{X}}$ matrix. Therefore, the unknown coefficients, which are estimated by GLSE, are usually unstable and give misleading information. To overcome this problem, many studies on the general linear model without linear restriction have been made. In fact, the coefficient parameter $\boldsymbol{\beta}$ can be regarded as a vector in $p$ dimensions space. If there exists multicollinearity in $\boldsymbol{C}$, the $\hat{\boldsymbol{\beta}}_{GR}$ would be badly apart from the actual coefficient parameter in some directions of $p$ dimensions space. In order to overcome this problem, one of the best methods is to use additional conditions to restrict the coefficients. Following Swamy et al. [36], Swamy and Mehta [35] and Zhong and Yang [38], the restricted ridge estimator can be obtained by minimizing the sum of squared residuals with a spherical restriction and a linear restriction (2.4), i.e., the RSRM is transformed into an optimal problem with two restrictions:

$$\begin{aligned}&\min(\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{X}}\boldsymbol{\beta})'\boldsymbol{V}^{-1}(\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{X}}\boldsymbol{\beta})\\ &\quad\text{s.t.}\quad \boldsymbol{\beta}'\boldsymbol{\beta}\le\phi^2,\\ &\qquad\quad\ \boldsymbol{H}\boldsymbol{\beta}=\boldsymbol{h}.\end{aligned}$$

The resulting estimator is generalized restricted ridge estimator (GRRE), which is given by

$$\hat{\boldsymbol{\beta}}_{GR}(k)=\hat{\boldsymbol{\beta}}_{G}(k)+\boldsymbol{C}_k^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}_k^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{h}-\boldsymbol{H}\hat{\boldsymbol{\beta}}_{G}(k)),\tag{2.6}$$

where

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{G}(k)&=\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{G},\quad \boldsymbol{R}_k=(\boldsymbol{I}_p+k\boldsymbol{C}^{-1})^{-1}\\ &=\boldsymbol{C}_k^{-1}\widetilde{\boldsymbol{X}}'\boldsymbol{V}^{-1}\widetilde{\boldsymbol{y}},\quad \boldsymbol{C}_k=\boldsymbol{C}+k\boldsymbol{I}_p\end{aligned}\tag{2.7}$$

is the generalized unrestricted ridge estimator (GURE) and $k\ge 0$ is the ridge parameter.

From Saleh [31], the likelihood ratio criterion for testing the null hypothesis $\boldsymbol{H}\boldsymbol{\beta}=\boldsymbol{h}$, is given by

$$£_n=\frac{(\boldsymbol{H}\hat{\boldsymbol{\beta}}_{G}-\boldsymbol{h})'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\hat{\boldsymbol{\beta}}_{G}-\boldsymbol{h})}{qs^2},\tag{2.8}$$

where,

$$s^2=\frac{1}{n-p}(\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}}_{G})'\boldsymbol{V}^{-1}(\widetilde{\boldsymbol{y}}-\widetilde{\boldsymbol{X}}\hat{\boldsymbol{\beta}}_{G}),\tag{2.9}$$

is an unbiased estimator of $\sigma^2$.

Under the null hypothesis and normal theory, $£_n$ follows a central $F$-distribution with $(q, n-p)$ degrees of freedom, while, under the alternative, it follows the non-central $F$-distribution with $(q, n-p)$ degrees of freedom and non-centrality parameter $\frac{1}{2}\Delta^*$, where

$$\Delta^*=\frac{(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h})'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H})^{-1}(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h})}{\sigma^2}.$$

In many practical situations, along with the model one may suspect that $\boldsymbol{\beta}$ belongs to the sub-space defined by (2.4). In such situation one combines the estimate of $\boldsymbol{\beta}$ and the test-statistic to obtain improved estimators of $\boldsymbol{\beta}$. First, we consider the preliminary test generalized restricted ridge estimator (PTGRRE) defined by

$$\hat{\boldsymbol{\beta}}_{GR}^{PT}(k)=\hat{\boldsymbol{\beta}}_{GR}(k)+\left[1-I(£_n\le F_{q,n-p}(\alpha))\right]\left(\hat{\boldsymbol{\beta}}_{G}(k)-\hat{\boldsymbol{\beta}}_{GR}(k)\right),\tag{2.10}$$

where $F_{q,n-p}(\alpha)$ is the upper $\alpha$-level critical value $(0<\alpha<1)$ from the central $F$-distribution and $I(A)$ is the indicator function of the set $A$. This estimator has been considered by Saleh and Kibria [32]. The PTGRRE has the disadvantage that it depends on $\alpha$, the level of significance, and also it yields the extreme results, namely $\hat{\boldsymbol{\beta}}_{GR}(k)$ and $\hat{\boldsymbol{\beta}}_{G}(k)$ depending on the outcome of the test. Later, we will discuss in detail of the Stein-type generalized restricted ridge estimator (SGRRE) defined by

$$\hat{\boldsymbol{\beta}}_{GR}^{S}(k)=\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S}=\hat{\boldsymbol{\beta}}_{GR}(k)+(1-d£_n^{-1})\left(\hat{\boldsymbol{\beta}}_{G}(k)-\hat{\boldsymbol{\beta}}_{GR}(k)\right),\quad d=\frac{(q-2)(n-p)}{q(n-p+2)},\ q\ge 3,\tag{2.11}$$

where

$$\hat{\boldsymbol{\beta}}_{GR}^{S}=\hat{\boldsymbol{\beta}}_{GR}+(1-d£_n^{-1})(\hat{\boldsymbol{\beta}}_{G}-\hat{\boldsymbol{\beta}}_{GR}),\tag{2.12}$$

is the Stein-type generalized restricted estimator (SGRE).

The SGRRE has the disadvantage that it has strange behavior for small values of $£_n$. Also, the shrinkage factor $(1-d£_n^{-1})$ becomes negative for $£_n<d$. Hence, we consider the positive-rule Stein-type generalized restricted ridge estimator (PRSGRRE) defined by

$$\hat{\boldsymbol{\beta}}_{GR}^{S+}(k)=\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S+}=\hat{\boldsymbol{\beta}}_{GR}^{S}(k)-(1-d£_n^{-1})I(£_n\le d)\left(\hat{\boldsymbol{\beta}}_{G}(k)-\hat{\boldsymbol{\beta}}_{GR}(k)\right)\tag{2.13}$$

where

$$\hat{\boldsymbol{\beta}}_{GR}^{S+}=\hat{\boldsymbol{\beta}}_{GR}^{S}-(1-d£_n^{-1})I(£_n\le d)(\hat{\boldsymbol{\beta}}_{G}-\hat{\boldsymbol{\beta}}_{GR}),\tag{2.14}$$

is the positive-rule Stein-type generalized restricted estimator (PRSGRE). The main purpose of this study is to consider the performance of shrinkage ridge estimators $\hat{\boldsymbol{\beta}}_{GR}^{S}(k)$ and $\hat{\boldsymbol{\beta}}_{GR}^{S+}(k)$. Shrinkage estimators has been considered by Arashi and Tabatabaey [7], Arashi et al. [6], Arashi [4], Arashi et al. [8] and extended to monotone functional estimation in multidimensional models same as additive regression model, semi-parametric partially linear model and generalized linear model by Zhang et al. [37].

## 3. Characteristics of the estimators

In this section, we provide the expressions for the bias and the quadratic risk of the estimators $\hat{\boldsymbol{\beta}}^{S}_{GR}(k)$ and $\hat{\boldsymbol{\beta}}^{S+}_{GR}(k)$. For properties of GURE, GRRE, and PTGRRE we refer the reader to [17,33,32].

### 3.1. Biases of the estimators

In this subsection we present expressions for the biases of the SGRRE and PRSGRRE.

**Theorem 3.1.** *Biases of the SGRRE and PRSGRRE are given by*

$$\begin{aligned}
\boldsymbol{b}(\hat{\boldsymbol{\beta}}^{S}_{GR}(k)) &= -qd\boldsymbol{R}_k\delta E[\chi^{-2}_{q+2}(\Delta^*)] - k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}\\
\boldsymbol{b}(\hat{\boldsymbol{\beta}}^{S+}_{GR}(k)) &= \boldsymbol{R}_k\delta\left\{\frac{qd}{q+2}E\left[F^{-1}_{q+2,n-p}(\Delta^*)I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right]\right.\\
&\quad \left. - \frac{qd}{q+2}E[F^{-1}_{q+2,n-p}(\Delta^*)] - G_{q+2,n-p}(x', \Delta^*)\right\} - k\boldsymbol{C}_k^{-1}\boldsymbol{\beta},
\end{aligned}$$

*where*

$$\begin{aligned}
G_{q+2i,n-p}(x', \Delta^*) &= \sum_{r=0}^{\infty}\frac{e^{-\Delta^*/2}(\Delta^*/2)^r}{\Gamma(r+1)}I_{x'}\left[\frac{q+2i}{2}+r, \frac{n-p}{2}\right],\\
E[\chi^{-2}_{q+s}(\Delta^*)]^n &= \sum_{r=0}^{\infty}\frac{e^{-\Delta^*/2}(\Delta^*/2)^r}{\Gamma(r+1)}\frac{\Gamma\left(\frac{q+s}{2}+r-n\right)}{2^n\Gamma\left(\frac{q+s}{2}+r\right)},\\
&E\left[F^{-j}_{q+s,n-p}(\Delta^*)I\left(F_{q+s,n-p}(\Delta^*) < \frac{qd}{q+s}\right)\right]\\
&= \sum_{r=0}^{\infty}\frac{e^{-\Delta^*/2}(\Delta^*/2)^r}{\Gamma(r+1)}\left(\frac{q+s}{n-p}\right)^j\frac{B\left(\frac{q+s+2r-2j}{2}, \frac{n-p+2j}{2}\right)}{B\left(\frac{q+s+2r}{2}, \frac{n-p}{2}\right)} \times I_x\left(\frac{q+s+2r-2j}{2}, \frac{n-p+2j}{2}\right),\\
\delta &= \boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\boldsymbol{\beta}-\boldsymbol{h}),
\end{aligned}$$

*and* $x' = \frac{qF_\alpha}{n-p+qF_\alpha}$, $x = \frac{qd}{n-p+qd}$ *and* $I_{x'}$ *is the Pearson regularized incomplete beta function, i.e.,*

$$I_{x'}(a, b) = \frac{B_{x'}(a, b)}{B(a, b)} = \frac{\int_0^{x'} t^{a-1}(1-t)^{b-1}}{\int_0^1 t^{a-1}(1-t)^{b-1}}.$$

**Proof.** By making use of the materials in Judge and Bock [18] and Saleh [31], in a similar fashion as in Arashi et al. [5], we have

$$\begin{aligned}
\boldsymbol{b}\big(\hat{\boldsymbol{\beta}}^{S}_{GR}(k)\big) &= E\big(\hat{\boldsymbol{\beta}}_G(k) - \boldsymbol{\beta}\big) - qd\boldsymbol{R}_k\delta E\big(\chi^{-2}_{q+2}(\Delta^*)\big)\\
&= -qd\boldsymbol{R}_k\delta E\big(\chi^{-2}_{q+2}(\Delta^*)\big) - k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}.
\end{aligned}$$

Also,

$$\begin{aligned}
\boldsymbol{b}\big(\hat{\boldsymbol{\beta}}^{S+}_{GR}(k)\big) &= E\big(\hat{\boldsymbol{\beta}}^{S}_{GR}(k) - \boldsymbol{\beta}\big) - \boldsymbol{R}_k\delta G_{q+2,n-p}(x', \Delta^*) + \frac{qd}{q+2}\boldsymbol{R}_k\delta E\left[F^{-1}_{q+2,n-p}(\Delta^*)I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right]\\
&= \boldsymbol{R}_k\delta\left\{\frac{qd}{q+2}E\left[F^{-1}_{q+2,n-p}(\Delta^*)I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right]\right.\\
&\quad \left. - \frac{qd}{q+2}E\big(F^{-1}_{q+2,n-p}(\Delta^*)\big) - G_{q+2,n-p}(x', \Delta^*)\right\} - k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}.
\end{aligned}$$

### 3.2. Risks of the estimators

For the risk of the estimators, we consider $R(\hat{\boldsymbol{\beta}}; \boldsymbol{\beta}) = tr\big(\boldsymbol{M}(\hat{\boldsymbol{\beta}})\big)$ where

$$\boldsymbol{M}(\hat{\boldsymbol{\beta}}) = E\big[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\big],$$

is the mean-squared error matrix of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}}$ can be any estimator of $\boldsymbol{\beta}$.

**Theorem 3.2.** *Risks of SGRRE and PRSGRRE are given by (3.1) and (3.2) respectively:*

$$\begin{aligned} R\left(\hat{\boldsymbol{\beta}}_{GR}^{S}(k);\boldsymbol{\beta}\right) &= \sigma^2 tr\left(\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{R}_k\right) + k^2\boldsymbol{\beta}'\boldsymbol{C}_k^{-2}\boldsymbol{\beta} + 2qdk\boldsymbol{\delta}'\boldsymbol{R}_k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}E\left(\chi_{q+2}^{-2}(\Delta^*)\right) \\ &\quad - dq\sigma^2 tr\left(\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k\right)\Bigg\{(q-2)E\left(\chi_{q+2}^{-4}(\Delta^*)\right) \\ &\quad + \left[1 - \frac{(q+2)\boldsymbol{\delta}'\boldsymbol{R}_k^2\boldsymbol{\delta}}{2\sigma^2\Delta^* tr\left(\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k\right)}\right](2\Delta^*)E\left(\chi_{q+4}^{-4}(\Delta^*)\right)\Bigg\}, \end{aligned} \tag{3.1}$$

$$\begin{aligned} R\left(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k);\boldsymbol{\beta}\right) &= R\left(\hat{\boldsymbol{\beta}}_{GR}^{S}(k);\boldsymbol{\beta}\right) - \sigma^2 tr\left(\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k\right) \\ &\quad \times E\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*)\right)^2 I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ &\quad + \boldsymbol{\delta}'\boldsymbol{R}_k^2\boldsymbol{\delta}E\left[\left(1 - \frac{qd}{q+4}F_{q+4,n-p}^{-1}(\Delta^*)\right)^2 I\left(F_{q+4,n-p}(\Delta^*) \le \frac{qd}{q+4}\right)\right] \\ &\quad - 2\boldsymbol{\delta}'\boldsymbol{R}_k^2\boldsymbol{\delta}E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ &\quad - 2k\boldsymbol{\delta}'\boldsymbol{R}_k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right]. \end{aligned} \tag{3.2}$$

**Proof.** Note that $\boldsymbol{\Upsilon} = \boldsymbol{C}^{-1/2}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1/2}$ is a symmetric idempotent matrix of rank $q \le p$. Therefore, there exists an orthogonal matrix $\boldsymbol{Q}$, such that

$$\boldsymbol{Q}\boldsymbol{C}^{-1/2}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1/2}\boldsymbol{Q}' = \begin{bmatrix}\boldsymbol{I}_q & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix},$$

$$\boldsymbol{Q}\boldsymbol{C}^{-1/2}\boldsymbol{R}_k^2\boldsymbol{C}^{-1/2}\boldsymbol{Q}' = \begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix} = \boldsymbol{A},$$

where

$$\boldsymbol{A}_{11} = (\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k^2\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1},$$

$$\boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 = (\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h})'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k^2\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h}).$$

The matrices $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}$ are of order $q$ and $p-q$ respectively. Now, we define the random variable

$$\boldsymbol{w} = \boldsymbol{Q}\boldsymbol{C}^{1/2}\hat{\boldsymbol{\beta}}_G - \boldsymbol{Q}\boldsymbol{C}^{-1/2}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{h},$$

so,

$$\boldsymbol{w} \sim N_p(\boldsymbol{\eta}, \sigma^2\boldsymbol{I}_p), \qquad \boldsymbol{\eta} = \boldsymbol{Q}\boldsymbol{C}^{1/2}\boldsymbol{\beta} - \boldsymbol{Q}\boldsymbol{C}^{-1/2}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{h}.$$

Partitioning the vectors $\boldsymbol{w} = (\boldsymbol{w}_1', \boldsymbol{w}_2')'$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}_1', \boldsymbol{\eta}_2')'$, where are independent sub-vectors of order $q$ and $p-q$ respectively, $\boldsymbol{w}_1 \sim N_q(\boldsymbol{\eta}_1, \sigma^2\boldsymbol{I}_p)$ and $\boldsymbol{w}_2 \sim N_{p-q}(\boldsymbol{\eta}_2, \sigma^2\boldsymbol{I}_{p-q})$, we obtain

$$\begin{aligned} \hat{\boldsymbol{\beta}}_G - \boldsymbol{\beta} &= \boldsymbol{C}^{-1/2}\boldsymbol{Q}'(\boldsymbol{w} - \boldsymbol{\eta}), \\ \hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR} &= \boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\hat{\boldsymbol{\beta}}_G - \boldsymbol{h}) \\ &= \boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1/2}\boldsymbol{Q}'\boldsymbol{w}. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} (\hat{\boldsymbol{\beta}}_G - \boldsymbol{\beta})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR}) &= (\boldsymbol{w} - \boldsymbol{\eta})'\boldsymbol{Q}\boldsymbol{C}^{-1/2}\boldsymbol{R}_k^2\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1/2}\boldsymbol{Q}'\boldsymbol{w} \\ &= (\boldsymbol{w} - \boldsymbol{\eta})'\boldsymbol{A}\boldsymbol{Q}\boldsymbol{\Upsilon}\boldsymbol{Q}'\boldsymbol{w} = (\boldsymbol{w} - \boldsymbol{\eta})'\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_q & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{w} \\ &= (\boldsymbol{w}_1 - \boldsymbol{\eta}_1)'\boldsymbol{A}_{11}\boldsymbol{w}_1 + (\boldsymbol{w}_2 - \boldsymbol{\eta}_2)'\boldsymbol{A}_{21}\boldsymbol{w}_1. \end{aligned}$$

Similarly,

$$(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR}) = \boldsymbol{w}'\begin{bmatrix}\boldsymbol{I}_q & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_q & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\boldsymbol{w} = \boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1.$$

Further, we conclude that

$$\pounds_n = \frac{\boldsymbol{w}_1'\boldsymbol{w}_1}{qs^2} \sim F_{q,n-p,\Delta^*}, \qquad \Delta^* = \frac{(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h})(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}(\boldsymbol{H}\boldsymbol{\beta} - \boldsymbol{h})}{\sigma^2}.$$

Therefore, by making use of the results given in Judge and Bock [18] and (2.11) we get

$$\begin{aligned}
R(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}) &= E\Big[\big(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S} - \boldsymbol{\beta}\big)'\big(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S} - \boldsymbol{\beta}\big)\Big] \\
&= R\big(\hat{\boldsymbol{\beta}}_G(k); \boldsymbol{\beta}\big) - 2dE\Big[\pounds_n^{-1}\big(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_G - \boldsymbol{\beta}\big)'\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] + d^2E\Big[\pounds_n^{-2}(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&= R\big(\hat{\boldsymbol{\beta}}_G(k); \boldsymbol{\beta}\big) - 2dE\Big[\pounds_n^{-1}\big(\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1 - \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1 + \boldsymbol{w}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1 - \boldsymbol{\eta}_2'\boldsymbol{A}_{21}\boldsymbol{w}_1\big)\Big] \\
&\quad + d^2E\Big[\pounds_n^{-2}\big(\boldsymbol{w}_1'\boldsymbol{A}_{11}\boldsymbol{w}_1\big)\Big] + 2kd\boldsymbol{\delta}'\boldsymbol{R}_k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}E\big[\pounds_n^{-1}\boldsymbol{w}\big].
\end{aligned}$$

But,

$$\begin{aligned}
&E\left(\frac{qs^2}{\boldsymbol{w}'\boldsymbol{w}}\right) = E\left[\left(\frac{qs^2}{\boldsymbol{w}'\boldsymbol{w}}\right)\right] = E\left(\frac{q}{n-p}\frac{\chi^2_{n-p}}{\chi^2_q(\Delta^*)}\right) = qE\big[\chi_q^{-2}(\Delta^*)\big] \\
&E\left[\frac{q^2s^4}{(\boldsymbol{w}'\boldsymbol{w})^2}\right] = E\left[\frac{q^2s^4}{(\boldsymbol{w}'\boldsymbol{w})^2}\right] = \frac{q^2}{(n-p)}(n-p+2)E\big[\chi_q^{-4}(\Delta^*)\big].
\end{aligned}$$

Therefore, gathering all terms we can obtain

$$\begin{aligned}
R\big(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}\big) &= \sigma^2 tr\Big[\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{R}_k\Big] + 2qdk\boldsymbol{\delta}'\boldsymbol{R}_k(k)\boldsymbol{C}_k^{-1}\boldsymbol{\beta}E\big[\chi_{q+2}^{-2}(\Delta^*)\big] \\
&\quad - dqtr(\boldsymbol{A}_{11})\sigma^2\Big\{2E\big[\chi_{q+2}^{-2}(\Delta^*)\big] - (q-2)E\big[\chi_{q+2}^{-4}(\Delta^*)\big]\Big\} \\
&\quad + \boldsymbol{\eta}_1'\boldsymbol{A}_{11}\boldsymbol{\eta}_1 dq\Big\{(q-2)E\big[\chi_{q+4}^{-4}(\Delta^*)\big] + 2E[\chi_{q+2}^{-2}(\Delta^*)\big] - 2E\big[\chi_{q+4}^{-2}(\Delta^*)\big]\Big\}.
\end{aligned}$$

Now, we use the following identities to obtain the final form of the risk $\hat{\boldsymbol{\beta}}_{GR}^{S}(k)$ given by (3.1).

$$\begin{aligned}
&E\big[\chi_p^{-2}(\Delta^*)\big] - E\big[\chi_{p+2}^{-2}(\Delta^*)\big] = 2E\big[\chi_{p+2}^{-4}(\Delta^*)\big], \quad \forall p > 2, \\
&E\big[\chi_{p+2}^{-2}(\Delta^*)\big] - (p-2)E\big[\chi_{p+2}^{-4}(\Delta^*)\big] = \Delta^*E\big[\chi_{p+4}^{-4}(\Delta^*)\big].
\end{aligned}$$

By using (2.13), we have

$$\begin{aligned}
R\big(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k); \boldsymbol{\beta}\big) &= E\Big[\big(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k) - \boldsymbol{\beta}\big)'\big(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k) - \boldsymbol{\beta}\big)\Big] \\
&= R\big(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}\big) + E\Big[(1 - d\pounds_n^{-1})^2 I(\pounds_n \le F_\alpha)(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_{GR} - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&\quad - 2E\Big[\big(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S}(k) - \boldsymbol{\beta}\big)'(1 - d\pounds_n^{-1})I(\pounds_n \le d)\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_{GR} - \hat{\boldsymbol{\beta}}_{GR})\Big].
\end{aligned}$$

But, using (2.11)

$$\begin{aligned}
&E\Big[\big(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR}^{S} - \boldsymbol{\beta}\big)'(1 - d\pounds_n^{-1})I(\pounds_n \le F_\alpha)\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&\quad = E\bigg\{\Big[(\boldsymbol{R}_k\hat{\boldsymbol{\beta}}_{GR} - \boldsymbol{\beta}) + (1 - d\pounds_n^{-1})\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big]'(1 - d\pounds_n^{-1})I(\pounds_n \le d)\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\bigg\} \\
&\quad = E\bigg\{\Big[(1 - d\pounds_n^{-1})I(\pounds_n \le d)(\hat{\boldsymbol{\beta}}_{GR} - \boldsymbol{\beta})'\boldsymbol{R}_{\boldsymbol{D}}^2(k)(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] + \Big[(1 - d\pounds_n^{-1})^2 I(\pounds_n \le d)(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})'\boldsymbol{R}_{\boldsymbol{D}}^2(k)(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&\qquad + \Big[(1 - d\pounds_n^{-1})I(\pounds_n \le d)\boldsymbol{\beta}'(\boldsymbol{R}_k - \boldsymbol{I}_p)\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big]\bigg\}.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
R\big(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k); \boldsymbol{\beta}\big) &= R\big(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}\big) - E\Big[(1 - d\pounds_n^{-1})^2 I(\pounds_n \le F_\alpha)(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&\quad - 2E\Big[(1 - d\pounds_n^{-1})I(\pounds_n \le d)(\hat{\boldsymbol{\beta}}_{GR} - \boldsymbol{\beta})'\boldsymbol{R}_k^2(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big]
\end{aligned}$$

$$
\begin{aligned}
&- 2\boldsymbol{\beta}'(\boldsymbol{R}_k - \boldsymbol{I}_p)E\Big[(1 - d\pounds_n^{-1})I(\pounds_n \le d)\boldsymbol{R}_k(\hat{\boldsymbol{\beta}}_G - \hat{\boldsymbol{\beta}}_{GR})\Big] \\
&= R\big(\hat{\boldsymbol{\beta}}^S_{GR}(k); \boldsymbol{\beta}\big) - \sigma^2\Bigg\{tr\Big[\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k\Big] \\
&\times E\Bigg[\Big(1 - \frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)\Big)^2 I\Big(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\Big)\Bigg] \\
&+ \frac{\boldsymbol{\delta}'\boldsymbol{R}_k^2\boldsymbol{\delta}}{\sigma^2}E\Bigg[\Big(1 - \frac{qd}{q+4}F^{-1}_{q+4,n-p}(\Delta^*)\Big)^2 I\Big(F_{q+4,n-p}(\Delta^*) \le \frac{qd}{q+4}\Big)\Bigg]\Bigg\} \\
&- 2\boldsymbol{\delta}'\boldsymbol{R}^2_{\boldsymbol{D}}(k)\boldsymbol{\delta}E\Bigg[\Big(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*) - 1\Big)I\Big(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\Big)\Bigg] \\
&- 2k\boldsymbol{\delta}'\boldsymbol{R}_k\boldsymbol{C}_k^{-1}\boldsymbol{\beta}E\Bigg[\Big(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*) - 1\Big)I\Big(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\Big)\Bigg]. \quad \square
\end{aligned}
$$

## 4. Performance of the estimators

In this section, we compare the underlying estimators with the usual ones under normality assumption. It is clear that for the positive-semi definite matrix $\boldsymbol{C}$ there exists an orthogonal matrix $\boldsymbol{\Gamma}$ such that $\boldsymbol{C} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}'$ and $\boldsymbol{\Lambda} = \boldsymbol{\Gamma}'\boldsymbol{C}\boldsymbol{\Gamma} =$ Diag $(\lambda_1, \ldots, \lambda_p)$ where

$$\lambda_1 > \lambda_2 > \cdots > \lambda_p \ge 0, \tag{4.1}$$

are the eigenvalues of $\boldsymbol{C}$ (see [3]). It is easy to see that the eigenvalues of $\boldsymbol{R}_k = (\boldsymbol{I}_p + k\boldsymbol{C}^{-1})^{-1}$ and $\boldsymbol{C}_k = \boldsymbol{C} + k\boldsymbol{I}_p$ are $\frac{\lambda_i}{\lambda_i+k}$ and $\lambda_i + k$, $i = 1, \ldots, p$, respectively. We note that the eigenvectors of $\boldsymbol{C}$, $\boldsymbol{R}_k$ and $\boldsymbol{C}_k$ are all the same. With this background we get the following identities:

$$
\begin{aligned}
\boldsymbol{\beta}'\boldsymbol{C}_k^{-2}\boldsymbol{\beta} &= \boldsymbol{\beta}'\boldsymbol{\Gamma}(\boldsymbol{\Lambda} + kI_p)^{-2}\boldsymbol{\Gamma}'\boldsymbol{\beta} = \boldsymbol{\alpha}'(\boldsymbol{\Lambda} + k\boldsymbol{I}_p)^{-2}\boldsymbol{\alpha} \\
&= \sum_{i=1}^{p}\frac{\alpha_i^2}{(\lambda_i + k)^2}, \quad \boldsymbol{\alpha} = \boldsymbol{\Gamma}'\boldsymbol{\beta},
\end{aligned} \tag{4.2}
$$

$$tr(\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{R}_k) = tr(\boldsymbol{C}^{-1}\boldsymbol{R}_k^2) = \sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i + k)^2}, \tag{4.3}$$

$$
\begin{aligned}
tr\big[\boldsymbol{R}_k\boldsymbol{C}^{-1}\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{R}_k\big] &= tr\big[\boldsymbol{\Gamma}(\boldsymbol{\Lambda} + kI_p)^{-1}\boldsymbol{\Gamma}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{\Gamma}(\boldsymbol{\Lambda} + k\boldsymbol{I}_p)^{-1}\boldsymbol{\Gamma}'\big] \\
&= tr\big[\boldsymbol{\Gamma}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{\Gamma}(\boldsymbol{\Lambda} + kI_p)^{-2}\big] \\
&= \sum_{i=1}^{p}\frac{h^*_{ii}}{(\lambda_i + k)^2},
\end{aligned} \tag{4.4}
$$

where $h^*_{ii} \ge 0$ is the $i$th diagonal element of $\boldsymbol{\Gamma}'\boldsymbol{H}'(\boldsymbol{H}\boldsymbol{C}^{-1}\boldsymbol{H}')^{-1}\boldsymbol{H}\boldsymbol{\Gamma} = \boldsymbol{H}^*$. Let us now compare the performance of the ridge estimator with its usual counterpart. Comparison results concerning GURE, GRRE and PTGRRE are well-known in the literature. Thus we focus on SGRRE and PRSGRRE in the sequel.

**Theorem 4.1.** *The SGRRE and PRSGRRE are superior to other proposed estimators in semiparametric regression models under following conditions:*

i. *There always exist a positive $k \in (0, k_1^*)$ and $(0, k_2^*)$ such that SGRRE has smaller risk value than SGRE under $H_0 : \boldsymbol{H\beta} = \boldsymbol{h}$ and $H_A : \boldsymbol{H\beta} \ne \boldsymbol{h}$, respectively, where*

$$k_1^* = \frac{\sigma^2 \min\limits_{1\le i\le p}\left\{\lambda_i - dh^*_{ii}\right\}}{\max\limits_{1\le i\le p}\left\{\alpha_i^2\lambda_i\right\}}, \tag{4.5}$$

$$k_2^* = \frac{\min\limits_{1\le i\le p}\left\{\sigma^2\left\{\lambda_i - qdh^*_{ii}\left[(q-2)E\Big(\chi^{-4}_{q+2}(\Delta^*)\Big) + \Big(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^*h^*_{ii}}\Big)(2\Delta^*)E\Big(\chi^{-4}_{q+4}(\Delta^*)\Big)\right]\right\} + qd\alpha_i\lambda_i^2\delta_iE\Big(\chi^{-2}_{q+2}(\Delta^*)\Big)\right\}}{\max\limits_{1\le i\le p}\left\{\alpha_i\lambda_i\left[\alpha_i - qd\delta_iE\Big(\chi^{-2}_{q+2}(\Delta^*)\Big)\right]\right\}}. \tag{4.6}$$

ii. *A sufficient condition for SGRRE to have risk value less than or equal to GURE is that there exists a value of $k \in (0, k_3^*)$, where $k_3^*$ is given by*

$$k_3^* = \frac{\sigma^2 \min\limits_{1\le i\le p}\left\{h_{ii}^*\left[(q-2)E\left(\chi_{q+2}^{-4}(\Delta^*)\right) + \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)(2\Delta^*)E\left(\chi_{q+4}^{-4}(\Delta^*)\right)\right]\right\}}{\max\limits_{1\le i\le p}\left\{2\alpha_i\lambda_i\delta_i E\left(\chi_{q+2}^{-2}(\Delta^*)\right)\right\}}. \tag{4.7}$$

iii. *A sufficient condition for SGRRE to have smaller risk value than GRRE is that $k \in (0, k_4^*)$ where $k_4^*$ is given by*

$$k_4^* = \frac{\max\limits_{1\le i\le p}\left\{\sigma^2 h_{ii}^*\left\{1 - dq\left[(q-2)E(\chi_{q+2}^{-4}(\Delta^*)) + \left(1 - \frac{(q-2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)(2\Delta^*)E(\chi_{q+4}^{-4}(\Delta^*))\right]\right\} - \lambda_i^2\delta_i^2\right\}}{\min\limits_{1\le i\le p}\left\{2\alpha_i\lambda_i\delta_i\left[1 - qdE(\chi_{q+2}^{-2}(\Delta^*))\right]\right\}}. \tag{4.8}$$

iv. *A sufficient condition for SGRRE to have smaller risk value than PTGRRE is that $k \in (0, k_5^*)$ where $k_5$ is given by*

$$\begin{aligned} k_5^* = {} & \max_{1\le i\le p}\left\{\sigma^2 h_{ii}^*\left\{G_{q+2,n-p}(x', \Delta^*) - dq\left[(q-2)E\left(\chi_{q+2}^{-4}(\Delta^*)\right) + \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)(2\Delta^*)E\left(\chi_{q+4}^{-4}(\Delta^*)\right)\right]\right\}\right. \\ & \left. - \lambda_i^2\delta_i^2\left[2G_{q+2,n-p}(x', \Delta^*) - G_{q+4,n-p}(x', \Delta^*)\right]\right\} \Big/ \\ & \min_{1\le i\le p}\left\{2\alpha_i\lambda_i\delta_i\left[G_{q+2,n-p}(x', \Delta^*) - qdE\left(\chi_{q+2}^{-2}(\Delta^*)\right)\right]\right\}. \end{aligned} \tag{4.9}$$

v. *There always exist a positive $k \in (0, k_6^*)$ and $(0, k_7^*)$ such that PRSGRRE has smaller risk value than PRSGRE under $H_0 : \boldsymbol{H\beta} = \boldsymbol{h}$ and $H_A : \boldsymbol{H\beta} \neq \boldsymbol{h}$, respectively, where*

$$k_6^* = \frac{\sigma^2 \min\limits_{1\le i\le p}\left\{\lambda_i - dh_{ii}^* - h_{ii}^* E\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(0)\right)^2 I\left(F_{q+2,n-p}(0) \le \frac{qd}{q+2}\right)\right]\right\}}{\max\limits_{1\le i\le p}\left\{\alpha_i^2\lambda_i\right\}}, \tag{4.10}$$

$$k_7^* = \frac{f_1(\alpha, \Delta^*)}{f_2(\alpha, \Delta^*)}, \tag{4.11}$$

*with*

$$\begin{aligned} f_1(\alpha, \Delta^*) = {} & \min_{1\le i\le p}\left\{\sigma^2\left\{\lambda_i - dqh_{ii}^*\left[(q-2)E\left(\chi_{q+2}^{-4}(\Delta^*)\right) + \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)\right](2\Delta^*)E\left(\chi_{q+4}^{-4}(\Delta^*)\right)\right.\right. \\ & - h_{ii}^* E\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*)\right)^2 I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ & \left. + \frac{\lambda_i^2\delta_i^2}{\sigma^2}E\left[\left(1 - \frac{qd}{q+4}F_{q+4,n-p}^{-1}(\Delta^*) - 1\right)^2 I\left(F_{q+4,n-p}(\Delta^*) \le \frac{qd}{q+4}\right)\right]\right\} \\ & + (\alpha_i - 2\delta_i)\lambda_i^2\delta_i E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ & \left. + dq\alpha_i\lambda_i^2\delta_i E\left(\chi_{q+2}^{-2}(\Delta^*)\right)\right\} \end{aligned}$$

*and*

$$f_2(\alpha, \Delta^*) = \max_{1\le i\le p}\alpha_i\lambda_i\left\{\alpha_i + \delta_i E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] - qd\delta_i E\left(\chi_{q+2}^{-2}(\Delta^*)\right)\right\}.$$

vi. *PRSGRRE has smaller risk value than SGRRE for all positive $k$.*

vii. *A sufficient condition for PRSGRRE to have risk value less than or equal to GURE is that $k \in (0, k_3^*)$ where $k_3^*$ is given by (4.7).*

viii. *A sufficient condition for PRSGRRE to have smaller risk value than GRRE is that $k \in (0, k_8^*)$ where $k_8^*$ is given by (4.12).*

$$k_8^* = \frac{f_3(\alpha, \Delta^*)}{f_4(\alpha, \Delta^*)}, \tag{4.12}$$

*with*

$$f_3(\alpha,\Delta^*) = \max_{1\le i\le p}\Bigg\{\sigma^2 h_{ii}\Bigg\{1 - dq\Bigg[(q-2)E\big(\chi^{-4}_{q+2}(\Delta^*)\big) + \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h^*_{ii}}\right)(2\Delta^*)E\big(\chi^{-4}_{q+4}(\Delta^*)\big)\Bigg]\Bigg\} - \lambda_i^2\delta_i^2$$
$$-\sigma^2\Bigg\{h^*_{ii}E\left[\left(1-\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)\right)^2 I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]$$
$$+\frac{\lambda_i^2\delta_i^2}{\sigma^2}E\left[\left(1-\frac{qd}{q+4}F^{-1}_{q+4,n-p}(\Delta^*)\right)^2 I\left(F_{q+4,n-p}(\Delta^*)\le\frac{qd}{q+4}\right)\right]\Bigg\}$$
$$-2\lambda_i^2\delta_i^2E\left[\left(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)-1\right)I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]\Bigg\}$$

*and*

$$f_4(\alpha,\Delta^*) = \min_{1\le i\le p}\Bigg\{2\alpha_i\lambda_i\delta_i\Bigg\{1 - qdE\big(\chi^{-2}_{q+2}(\Delta^*)\big)$$
$$+E\left[\left(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)-1\right)I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]\Bigg\}\Bigg\}.$$

ix. *A sufficient condition for PRSGRRE to have smaller risk value than PTGRRE is that $k \in (0, k_9^*)$ where $k_9^*$ is given by (4.13).*

$$k_9^* = \frac{f_5(\alpha,\Delta^*)}{f_6(\alpha,\Delta^*)}, \tag{4.13}$$

*with*

$$f_5(\alpha,\Delta^*) = \max_{1\le i\le p}\Bigg\{\sigma^2 h^*_{ii}\Bigg\{G_{q+2,n-p}(x',\Delta^*) - dq\Bigg[(q-2)E(\chi^{-4}_{q+2}(\Delta^*))$$
$$+\left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h^*_{ii}}\right)(2\Delta^*)E(\chi^{-4}_{q+4}(\Delta^*))\Bigg]\Bigg\} - \lambda_i^2\delta_i^2\left[2G_{q+2,n-p}(x',\Delta^*) - G_{q+4,n-p}(x',\Delta^*)\right]$$
$$-\sigma^2\Bigg\{h^*_{ii}E\left[\left(1-\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)\right)^2 I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]$$
$$+\frac{\lambda_i^2\delta_i^2}{\sigma^2}E\left[\left(1-\frac{qd}{q+4}F^{-1}_{q+4,n-p}(\Delta^*)\right)^2 I\left(F_{q+4,n-p}(\Delta^*)\le\frac{qd}{q+4}\right)\right]\Bigg\}$$
$$-2\lambda_i^2\delta_i^2E\left[\left(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)-1\right)I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]\Bigg\}$$

*and*

$$f_6(\alpha,\Delta^*) = \min_{1\le i\le p}\Bigg\{2\alpha_i\lambda_i\delta_i\Bigg\{G_{q+2,n-p}(x',\Delta^*) - qdE(\chi^{-2}_{q+2}(\Delta^*))$$
$$+E\left[\left(\frac{qd}{q+2}F^{-1}_{q+2,n-p}(\Delta^*)-1\right)I\left(F_{q+2,n-p}(\Delta^*)\le\frac{qd}{q+2}\right)\right]\Bigg\}\Bigg\}.$$

**Proof.** We only prove the first and last items. The proofs of the others are similar and just require a straight forward calculation.

Proof of (i): Using Eqs. (4.2)–(4.4) under $H_0 : \boldsymbol{H\beta} \ne \boldsymbol{h}$, the risk expression is given by

$$R\big(\hat{\boldsymbol{\beta}}^S_{GR}(k);\boldsymbol{\beta}\big) = \sum_{i=1}^p \frac{1}{(\lambda_i+k)^2}\left[\sigma^2(\lambda_i - dh^*_{ii}) + k^2\alpha_i^2\right]. \tag{4.14}$$

Differentiating with respect to $k$, we find

$$\frac{\partial R\big(\hat{\boldsymbol{\beta}}^S_{GR}(k);\boldsymbol{\beta}\big)}{\partial k} = 2\sum_{i=1}^p \frac{1}{(\lambda_i+k)^3}\left[k\alpha_i^2\lambda_i - \sigma^2(\lambda_i - dh^*_{ii})\right]. \tag{4.15}$$

Thus, a sufficient condition for (4.15) to be negative is that $0 < k < k_1^*$ where

$$k_1^* = \frac{\sigma^2 \min\limits_{1\le i\le p}\left\{\lambda_i - dh_{ii}^*\right\}}{\max\limits_{1\le i\le p}\left\{\alpha_i^2\lambda_i\right\}}.$$

Now, under $H_A : \boldsymbol{H}\boldsymbol{\beta} \neq \boldsymbol{h}$, the risk expression is given by

$$\begin{aligned} R\big(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}\big) &= \sum_{i=1}^{p} \frac{1}{(\lambda_i + k)^2}\Bigg\{\sigma^2\Big[\lambda_i - qdh_{ii}^*(q-2)E\big(\chi_{q+2}^{-4}(\Delta^*)\big) \\ &\quad + 2qdh_{ii}^*\Delta^*\left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)E\big(\chi_{q+4}^{-4}(\Delta^*)\big)\Big] + 2qdk\alpha_i\lambda_i\delta_i E\big(\chi_{q+2}^{-2}(\Delta^*)\big) + k^2\alpha_i^2\Bigg\}, \end{aligned} \tag{4.16}$$

which $\delta_i$ is the $i$th element of $\boldsymbol{\delta}$. Differentiating with respect to $k$, gives

$$\begin{aligned} \frac{\partial R\big(\hat{\boldsymbol{\beta}}_{GR}^{S}(k); \boldsymbol{\beta}\big)}{\partial k} &= 2\sum_{i=1}^{p} \frac{1}{(\lambda_i + k)^3}\Bigg\{k\lambda_i\alpha_i\Big[\alpha_i - qd\delta_i E\big(\chi_{q+2}^{-2}(\Delta^*)\big)\Big] - \sigma^2\Bigg\{\lambda_i - dqh_{ii}^*\Big[(q-2)E\big(\chi_{q+2}^{-4}(\Delta^*)\big) \\ &\quad + \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)(2\Delta^*)E\big(\chi_{q+4}^{-4}(\Delta^*)\big)\Big]\Bigg\} - qd\alpha_i\lambda_i^2\delta_i E\big(\chi_{q+2}^{-2}(\Delta^*)\big)\Bigg\}. \end{aligned} \tag{4.17}$$

Hence, a sufficient condition for (4.17) to be negative is that $0 < k < k_2^*$. Thus, $\hat{\boldsymbol{\beta}}_{GR}^{S}(k)$ has risk value less than the risk of $\hat{\boldsymbol{\beta}}^{S}$ for $0 < k < k_2^*$.

Proof of (ix): By making use of risk difference we obtain

$$\begin{aligned} R\big(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k); \boldsymbol{\beta}\big) - R\big(\hat{\boldsymbol{\beta}}_{GR}^{PT}(k); \boldsymbol{\beta}\big) &= \sum_{i=1}^{p} \frac{1}{(\lambda_i + k)^2}\Bigg\{\sigma^2 h_{ii}^*\Bigg\{G_{q+2,n-p}(x', \Delta^*) - dq\Big[(q-2)E\big(\chi_{q+2}^{-4}(\Delta^*)\big) \\ &+ \left(1 - \frac{(q+2)\lambda_i^2\delta_i^2}{2\sigma^2\Delta^* h_{ii}^*}\right)(2\Delta^*)E\big(\chi_{q+4}^{-4}(\Delta^*)\big)\Big]\Bigg\} - \lambda_i^2\delta_i^2\Big[2G_{q+2,n-p}(x', \Delta^*) - G_{q+4,n-p}(x', \Delta^*)\Big] \\ &- \sigma^2\Bigg\{h_{ii}^* E\left[\left(1 - \frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*)\right)^2 I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ &+ \frac{\lambda_i^2\delta_i^2}{\sigma^2}E\left[\left(1 - \frac{qd}{q+4}F_{q+4,n-p}^{-1}(\Delta^*)\right)^2 I\left(F_{q+4,n-p}(\Delta^*) \le \frac{qd}{q+4}\right)\right]\Bigg\} \\ &- 2\lambda_i^2\delta_i^2 E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right] \\ &- 2k\alpha_i\lambda_i\delta_i\Bigg\{G_{q+2,n-p}(x', \Delta^*) - qdE(\chi_{q+2}^{-2}(\Delta^*)) + E\left[\left(\frac{qd}{q+2}F_{q+2,n-p}^{-1}(\Delta^*) - 1\right) I\left(F_{q+2,n-p}(\Delta^*) \le \frac{qd}{q+2}\right)\right]\Bigg\}\Bigg\}. \end{aligned}$$

This difference is non positive ($\le 0$), whenever

$$k_9^* = \frac{f_5(\alpha, \Delta^*)}{f_6(\alpha, \Delta^*)}. \quad \square$$

## 5. Numerical results

In this section we proceed with some numerical computations as proofs of our assertions. The process is categorized into two setups as follows.

### 5.1. The simulation studies

In this section, we examine the risk function performance of the proposed estimators comparatively. To achieve different degrees of collinearity, following McDonald and Galarneau [25] and Gibbons [11] the explanatory were generated using the following device for $n = 1000$ from the following model:

$$x_{ij} = (1-\gamma^2)^{\frac{1}{2}} z_{ij} + \gamma z_{ip}, \quad i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, p \tag{5.1}$$

where $z_{ij}$ are independent standard normal pseudo-random numbers, and $\gamma$ is specified so that the correlation between any two explanatory variables is given by $\gamma^2$. These variables are then standardized so that $\boldsymbol{X}'\boldsymbol{X}$ and $\boldsymbol{X}'\boldsymbol{Y}$ are in correlation

**Table 1**
Evaluation of SGRRE at different $k$ values in model (5.2) with $\gamma = 0.80$.

| $k$ coefficients | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | $k_{opt} = 0.299$ | 0.30 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −1.007411 | −1.009063 | −1.010710 | −1.012353 | −1.013991 | −1.015626 | −0.993752 | −1.017255 |
| $\hat{\beta}_2$ | −0.900560 | −0.901983 | −0.903402 | −0.904818 | −0.906231 | −0.907640 | −0.995390 | −0.909045 |
| $\hat{\beta}_3$ | 1.985613 | 1.985895 | 1.986177 | 1.986459 | 1.986741 | 1.987024 | 2.004879 | 1.987306 |
| $\hat{\beta}_4$ | 3.024487 | 3.025410 | 3.026332 | 3.027252 | 3.028171 | 3.029088 | 3.004253 | 3.030004 |
| $\hat{\beta}_5$ | −5.073518 | −5.075917 | −5.078312 | −5.080702 | −5.083089 | −5.085472 | −5.002103 | −5.087850 |
| $\hat{\beta}_6$ | 3.999300 | 4.002245 | 4.005182 | 4.008111 | 4.011032 | 4.013946 | 3.996077 | 4.016851 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}^S_{GR}(k)]$ | 0.000000 | 0.000035 | 0.000142 | 0.000319 | 0.000568 | 0.000887 | 0.001268 | 0.001277 |
| $\hat{R}[\hat{\boldsymbol{\beta}}^S_{GR}(k); \boldsymbol{\beta}]$ | 0.274546 | 0.274159 | 0.273845 | 0.273602 | 0.273429 | 0.273328 | 0.273296 | 0.273296 |
| $\hat{\Delta}$ | 0.000000 | 0.000386 | 0.000700 | 0.000944 | 0.001116 | 0.001218 | 0.001249 | 0.001249 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.027114 | 0.027007 | 0.026916 | 0.026843 | 0.026787 | 0.026747 | 0.026725 | 0.026725 |

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

$$y_i = \sum_{j=1}^{6} x_{ji}\beta_j + f(t_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{5.2}$$

where

$$\boldsymbol{\beta} = (-1, -1, 2, 3, -5, 4)',$$
$$f(t) = \frac{1}{5}\Big[\phi\big(t; -7, 1.44\big) + \phi\big(t; -3.5, 1\big) + \phi\big(t; 0, 0.64\big) + \phi\big(t; 3.5, 0.36\big) + \phi\big(t; 7, 0.16\big)\Big],$$

which is a mixture of normal densities for $t \in [-9, 9]$ and $\phi(x; \mu, \sigma^2)$ is a normal density function with mean $\mu$ and variance $\sigma^2$. The main reason of selecting such structure for nonlinear part is to check the efficiency of nonparametric estimations for wavy function. This function is difficult to be estimated and provides a good test case for the nonparametric regression method. It is plotted in Fig. 1. Moreover, $\boldsymbol{\epsilon} \sim N_n(\boldsymbol{0}, \sigma^2\boldsymbol{V})$ for which the elements of $\boldsymbol{V}$ are $v_{ij} = (\frac{1}{n})^{|i-j|}$ and $\sigma^2 = 4.00$. To compare the performance of the proposed restricted estimators, we consider the parametric restriction $\boldsymbol{H\beta} = \boldsymbol{0}$, where

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

For the weight function $W_{ni}(t_j)$, we use

$$W_{ni}(t_j) = \frac{1}{nh_n}K\left(\frac{t_i - tj}{h_n}\right) = \frac{1}{nh_n} \cdot \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(t_i - t_j)^2}{2h_n^2}\right\}, \quad h_n = 0.01,$$

which is Priestley and Chao's weight with the Gaussian kernel. We also apply the cross-validation (C.V.) method to select the optimal bandwidth $h_n$, which minimizes the following C.V. function

$$C.V.(h_n) = \frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{\boldsymbol{y}}^{(-i)} - \widetilde{\boldsymbol{X}}^{(-i)}\hat{\boldsymbol{\beta}}^{(-i)}\right)^2,$$

where $\hat{\boldsymbol{\beta}}^{(-i)}$ obtain by replacing $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{y}}$ with $\widetilde{\boldsymbol{X}}^{(-i)} = \left(\tilde{x}_{jk}^{(-i)}\right)$, $1 \le k \le n$, $1 \le j \le p$, $\boldsymbol{y}^{(-i)} = \left(\tilde{y}_1^{(-i)}, \ldots, \tilde{y}_n^{(-i)}\right)$, $\tilde{x}_{sk}^{(-i)} = x_{sk} - \sum_{j\neq i}^{n} W_{nj}(t_i)x_{sj}$, $\tilde{y}_k^{(-i)} = y_k - \sum_{j\neq i}^{n} W_{nj}(t_i)y_j$. Here $\boldsymbol{y}^{(-i)}$ is the predicted value of $\boldsymbol{y} = (y_1, \ldots, y_n)$ at $\boldsymbol{x}_i = (x_{1i}, x_{2i}, \ldots, x_{pi})$ with $y_i$ and $x_i$ left out of the estimation of the $\boldsymbol{\beta}$.

All computations were conducted using the statistical package R 3.0.0. The ratio of largest eigenvalue to smallest eigenvalue of design matrix in model (5.2) is approximately $\lambda_5/\lambda_1 = 672.11$, 1549.80 and 14495.28 for $\gamma = 0.80$, 0.90 and 0.99, respectively. In Tables 1–6, we computed the proposed estimators, respectively. We numerically estimated the $QB(\cdot)$ (quadratic bias), $R(\cdot)$, $\Delta = R(\hat{\boldsymbol{\beta}}^{S\text{ or }S+}_{GR}; \boldsymbol{\beta}) - R[\hat{\boldsymbol{\beta}}^{S\text{ or }S+}_{GR}(k); \boldsymbol{\beta}]$ and $mse[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})] = \frac{1}{n}\sum_{i=1}^{n}[\hat{f}(t_i) - f(t_i)]^2$ for different values of $k$ and $\gamma$. We found the best values of $k$ ($k_{opt}$) by plotting $\Delta$ versus $k$ for each proposed estimators (Fig. 2). In this figure, in the left part, the risk of SGRRE (solid line) and PRSGDRRE (dash line) and in the right part, the $\Delta$ for SGRRE (solid line) and PRSGDRRE (dash line) versus ridge parameter $k$ are plotted for different values of $\gamma$. In the continuation, Fig. 3 shows the fitted function by kernel smoothing after estimation of the linear part of the model by $\hat{\boldsymbol{\beta}}^{S}_{GR}(k_{opt})$ and $\hat{\boldsymbol{\beta}}^{S+}_{GR}(k_{opt})$, that is, $\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}^{S\text{ or }S+}_{GR}(k_{opt})$, respectively, for $\gamma = 0.80$, 0.90 and 0.99. The minimum of C.V. approximately occurred at $h_n = 0.8210$ for the model (5.2) with $n = 1000$. The diagram of C.V. versus $h_n$ is also plotted in Fig. 4 for more clarification by using the modified cross-validation method of Bowman et al. [9] in kerdiest package. Three automatic bandwidth selection methods

**Table 2**
Evaluation of PRSGRRE at different $k$ values in model (5.2) with $\gamma = 0.80$.

| $k$ coefficients | 0 | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | $k_{opt} = 0.2635$ | 0.30 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −1.004043 | −1.002080 | −1.000122 | −0.998168 | −0.996220 | −0.994276 | −0.993752 | −0.992337 |
| $\hat{\beta}_2$ | −1.004043 | −1.002393 | −1.000747 | −0.999104 | −0.997466 | −0.995831 | −0.995390 | −0.994200 |
| $\hat{\beta}_3$ | 2.008086 | 2.007478 | 2.006869 | 2.006261 | 2.005652 | 2.005044 | 2.004879 | 2.004435 |
| $\hat{\beta}_4$ | 3.012129 | 3.010631 | 3.009135 | 3.007640 | 3.006147 | 3.004655 | 3.004253 | 3.003166 |
| $\hat{\beta}_5$ | −5.020214 | −5.016768 | −5.013326 | −5.009888 | −5.006456 | −5.003028 | −5.002103 | −4.999604 |
| $\hat{\beta}_6$ | 4.016171 | 4.012340 | 4.008517 | 4.004703 | 4.000897 | 3.997101 | 3.996077 | 3.993312 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}^{S+}_{GR}(k)]$ | 0.000000 | 0.000035 | 0.000142 | 0.000318 | 0.000564 | 0.000878 | 0.000975 | 0.001260 |
| $\hat{R}[\hat{\boldsymbol{\beta}}^{S+}_{GR}(k); \boldsymbol{\beta}]$ | 0.242314 | 0.241977 | 0.241712 | 0.241518 | 0.241395 | 0.241342 | 0.241340 | 0.241359 |
| $\hat{\Delta}$ | 0.000000 | 0.000336 | 0.000601 | 0.000795 | 0.000919 | 0.000971 | 0.000973 | 0.000954 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.027894 | 0.027931 | 0.027969 | 0.028006 | 0.028045 | 0.028083 | 0.028094 | 0.028123 |

**Table 3**
Evaluation of SGRRE at different $k$ values in model (5.2) with $\gamma = 0.90$.

| $k$ coefficients | 0 | 0.10 | 0.20 | 0.30 | $k_{opt} = 0.396$ | 0.40 | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −0.969980 | −0.968469 | −0.966971 | −0.965485 | −0.964070 | −0.964012 | −0.962550 | −0.961101 |
| $\hat{\beta}_2$ | −0.968565 | −0.967294 | −0.966033 | −0.964783 | −0.963591 | −0.963541 | −0.962310 | −0.961087 |
| $\hat{\beta}_3$ | 2.025682 | 2.024862 | 2.024040 | 2.023219 | 2.022431 | 2.022398 | 2.021576 | 2.020755 |
| $\hat{\beta}_4$ | 3.009494 | 3.008089 | 3.006686 | 3.005284 | 3.003940 | 3.003884 | 3.002487 | 3.001091 |
| $\hat{\beta}_5$ | −5.061974 | −5.058700 | −5.055436 | −5.052184 | −5.049071 | −5.048942 | −5.045711 | −5.042491 |
| $\hat{\beta}_6$ | 3.992947 | 3.989439 | 3.985956 | 3.982498 | 3.979202 | 3.979066 | 3.975658 | 3.972274 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}^{S}_{GR}(k)]$ | 0.000000 | 0.000423 | 0.001679 | 0.003745 | 0.006472 | 0.006602 | 0.010228 | 0.014604 |
| $\hat{R}[\hat{\boldsymbol{\beta}}^{S}_{GR}(k); \boldsymbol{\beta}]$ | 0.540776 | 0.537865 | 0.535830 | 0.534651 | 0.534309 | 0.534311 | 0.534793 | 0.536081 |
| $\hat{\Delta}$ | 0.000000 | 0.002910 | 0.004945 | 0.006124 | 0.006466 | 0.006464 | 0.005982 | 0.004695 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.090390 | 0.090073 | 0.089766 | 0.089467 | 0.089189 | 0.089177 | 0.088897 | 0.088625 |

**Table 4**
Evaluation of PRSGRRE at different $k$ values in model (5.2) with $\gamma = 0.90$.

| $k$ coefficients | 0 | 0.10 | 0.20 | 0.30 | $k_{opt} = 0.3485$ | 0.40 | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −0.969980 | −0.963358 | −0.956797 | −0.950295 | −0.947163 | −0.943853 | −0.937468 | −0.931141 |
| $\hat{\beta}_2$ | −0.968565 | −0.963038 | −0.957559 | −0.952127 | −0.949509 | −0.946740 | −0.941400 | −0.936105 |
| $\hat{\beta}_3$ | 2.025682 | 2.023431 | 2.021176 | 2.018918 | 2.017822 | 2.016658 | 2.014396 | 2.012131 |
| $\hat{\beta}_4$ | 3.009494 | 3.005066 | 3.000644 | 2.996229 | 2.994090 | 2.991821 | 2.987420 | 2.983026 |
| $\hat{\beta}_5$ | −5.061974 | −5.050552 | −5.039177 | −5.027850 | −5.022373 | −5.016569 | −5.005336 | −4.994149 |
| $\hat{\beta}_6$ | 3.992947 | 3.978997 | 3.965168 | 3.951459 | 3.944853 | 3.937868 | 3.924394 | 3.911034 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}^{S+}_{GR}(k)]$ | 0.000000 | 0.000421 | 0.001661 | 0.003688 | 0.004944 | 0.006468 | 0.009970 | 0.014164 |
| $\hat{R}[\hat{\boldsymbol{\beta}}^{S+}_{GR}(k); \boldsymbol{\beta}]$ | 0.476736 | 0.474222 | 0.472578 | 0.471787 | 0.471705 | 0.471830 | 0.472692 | 0.474355 |
| $\hat{\Delta}$ | 0.000000 | 0.002514 | 0.004158 | 0.004949 | 0.005031 | 0.004905 | 0.004044 | 0.002381 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.090390 | 0.090555 | 0.090723 | 0.090894 | 0.090978 | 0.091068 | 0.091245 | 0.091424 |

**Table 5**
Evaluation of SGRRE at different $k$ values in model (5.2) with $\gamma = 0.99$.

| $k$ coefficients | 0 | 0.10 | 0.20 | 0.30 | 0.40 | $k_{opt} = 0.422$ | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −1.430442 | −1.366404 | −1.308985 | −1.257172 | −1.210152 | −1.200383 | −1.167265 | −1.127966 |
| $\hat{\beta}_2$ | −1.267638 | −1.205877 | −1.150665 | −1.100998 | −1.056067 | −1.046750 | −1.015215 | −0.977902 |
| $\hat{\beta}_3$ | 1.487994 | 1.500567 | 1.510120 | 1.517217 | 1.522304 | 1.523191 | 1.525738 | 1.527803 |
| $\hat{\beta}_4$ | 3.046166 | 3.024189 | 3.001347 | 2.977979 | 2.954340 | 2.949123 | 2.930622 | 2.906971 |
| $\hat{\beta}_5$ | −4.945296 | −4.876890 | −4.811160 | −4.747965 | −4.687173 | −4.674108 | −4.628655 | −4.572293 |
| $\hat{\beta}_6$ | 5.025435 | 4.854127 | 4.700977 | 4.563172 | 4.438455 | 4.412585 | 4.324997 | 4.221296 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}^{S}_{GR}(k)]$ | 0.000000 | 0.080730 | 0.274931 | 0.532892 | 0.824584 | 0.891303 | 1.131696 | 1.443021 |
| $\hat{R}[\hat{\boldsymbol{\beta}}^{S}_{GR}(k); \boldsymbol{\beta}]$ | 5.748832 | 5.391173 | 5.179867 | 5.082957 | 5.076463 | 5.085263 | 5.142092 | 5.265685 |
| $\hat{\Delta}$ | 0.000000 | 0.357659 | 0.568965 | 0.665874 | 0.672369 | 0.663569 | 0.606740 | 0.483147 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.040411 | 0.040889 | 0.041339 | 0.04176 | 0.042156 | 0.042239 | 0.042526 | 0.042871 |

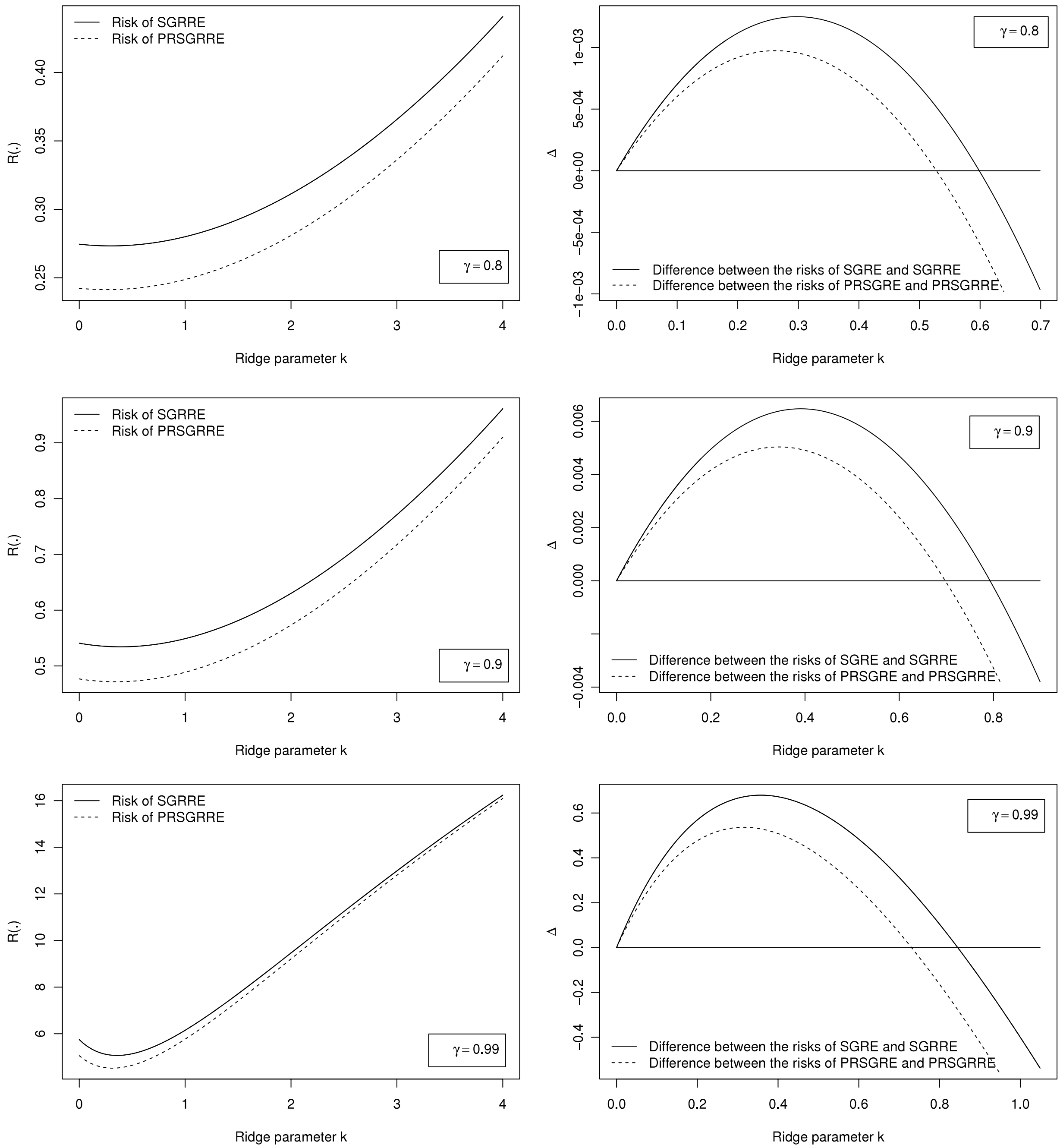


**Fig. 1.** The diagrams of $\Delta$ and $R(\cdot)$ versus $k$ for different values of $\gamma$.

for nonparametric kernel distribution function estimation are implemented by this package: the plug-in of Altman and Leger, the plug-in of Polanski and Baker, and the modified cross-validation of Bowman, Hall and Prvan. The C.V. values at different values of $h_n$ are showed in Table 7.

The numerical calculations based on Theorem 4.1 in the simulation study showed that:

- At most $k$ for dominance of SGRRE over SGRE are 0.598, 0.792 and 0.844 for $\gamma = 0.80,\ 0.90$ and 0.99, respectively. So, it can be found that ridge parameter $k$ is an increasing function of $\gamma$ for dominance of SGRRE over SGRE in the sense of risk function.
- At most $k$ for dominance of PRSGRRE over PRSGRE are 0.527, 0.697 and 0.730 for $\gamma = 0.80,\ 0.90$ and 0.99, respectively. So, it can be found that ridge parameter $k$ is an increasing function of $\gamma$ for dominance of PRSGRRE over PRSGRE in the sense of risk function.
- The PRSGRRE is better than SGRRE for all values of $k$ and $\gamma$ in the sense of risk function (Fig. 2).

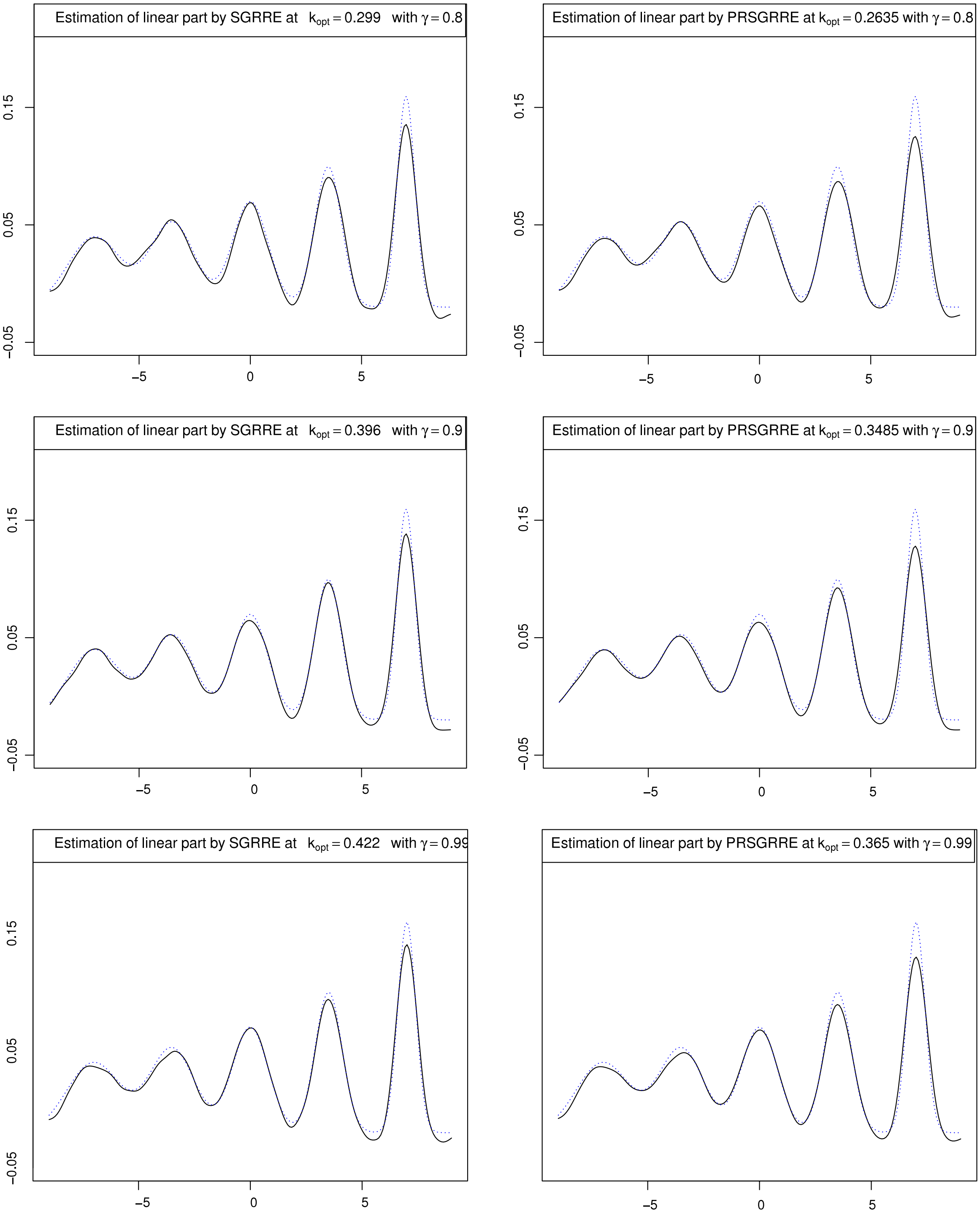


**Fig. 2.** Estimation of the function under study (mixtures of normal densities) by kernel approach for $n = 1000$. Solid black lines are the estimates and blue dashed lines are the true functions.

### *5.2. Real data example*

To motivate the problem of linearly constrained estimation in the semiparametric regression model, we consider the hedonic prices of housing attributes. Housing prices are very much affected by lot size. The semiparametric regression model that follows was estimated by Ho [16] using semiparametric least squares. The data consist of 92 detached homes in the Ottawa area that were sold during 1987. The variables are defined as follows: The dependent variable $y$ is sale price (SP),

**Table 6**
Evaluation of PRSGRRE at different $k$ values in model (5.2) with $\gamma = 0.99$.

| $k$ coefficients | 0 | 0.10 | 0.20 | 0.30 | $k_{opt} = 0.365$ | 0.40 | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| $\hat{\beta}_1$ | −1.430442 | −1.340968 | −1.260708 | −1.188254 | −1.144802 | −1.122480 | −1.062466 | −1.007459 |
| $\hat{\beta}_2$ | −1.267638 | −1.179763 | −1.101084 | −1.030198 | −0.987757 | −0.965978 | −0.907508 | −0.854034 |
| $\hat{\beta}_3$ | 1.487994 | 1.496773 | 1.501729 | 1.503622 | 1.503499 | 1.503050 | 1.500489 | 1.496318 |
| $\hat{\beta}_4$ | 3.046166 | 3.007328 | 2.967702 | 2.927731 | 2.901722 | 2.887740 | 2.847976 | 2.808619 |
| $\hat{\beta}_5$ | −4.945296 | −4.834679 | −4.728454 | −4.626386 | −4.562164 | −4.528251 | −4.433838 | −4.342949 |
| $\hat{\beta}_6$ | 5.025435 | 4.786284 | 4.572376 | 4.379811 | 4.264579 | 4.205458 | 4.046779 | 3.901690 |
| $\hat{QB}[\hat{\boldsymbol{\beta}}_{GR}^{S+}(k)]$ | 0.000000 | 0.078374 | 0.259474 | 0.489560 | 0.650382 | 0.738286 | 0.988618 | 1.231199 |
| $\hat{R}[\hat{\boldsymbol{\beta}}_{GR}^{S+}(k); \boldsymbol{\beta}]$ | 5.060371 | 4.752241 | 4.584068 | 4.525083 | 4.534121 | 4.552216 | 4.647883 | 4.798485 |
| $\hat{\Delta}$ | 0.000000 | 0.308130 | 0.476303 | 0.535288 | 0.526250 | 0.508155 | 0.412488 | 0.261886 |
| $\hat{mse}[\hat{f}(\boldsymbol{t}), f(\boldsymbol{t})]$ | 0.040411 | 0.041246 | 0.042091 | 0.042936 | 0.043482 | 0.043774 | 0.044601 | 0.045414 |

**Table 7**
The values of C.V. at different $h_n$ values.

| $h_n$ | C.V. |
|---|---|
| 0.0900 | 3.0095 |
| 0.4555 | 3.0093 |
| 0.8210* | 3.0092 |
| 1.1865 | 3.0093 |
| 1.5520 | 3.0097 |
| 1.9175 | 3.0106 |
| 2.2830 | 3.0120 |
| 2.6485 | 3.0140 |
| 3.0140 | 3.0168 |
| 3.3795 | 3.0204 |
| 3.7451 | 3.0249 |
| 4.1106 | 3.0304 |
| 4.4761 | 3.0371 |
| 4.8416 | 3.0448 |

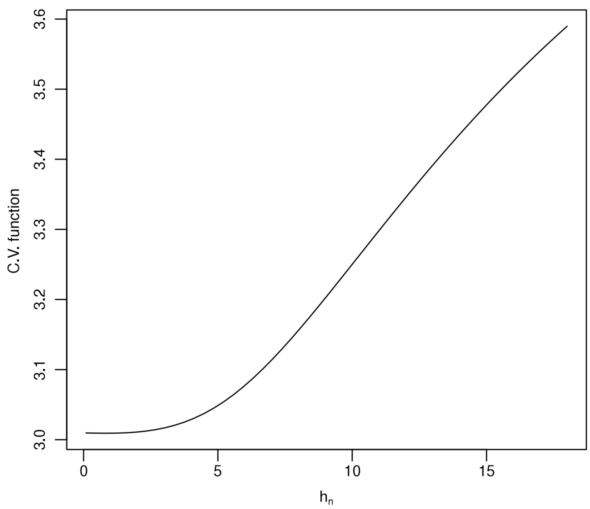


**Fig. 3.** The diagram of C.V. versus $h_n$.

the independents variable include lot size (lot area = LT), square footage of housing (SFH), average neighborhood income (ANI), distance to highway (DHW), presence of garage (GAR), fireplace (FP). We first consider the pure parametric model:

$$(SP)_i = \beta_0 + \beta_1(LT)_i + \beta_2(SFH)_i + \beta_3(FP)_i + \beta_4(DHW)_i + \beta_5(GAR)_i + \beta_6(ANI)_i + \epsilon_i. \tag{5.3}$$

In order to detecting the correlation between variables, we can take a look at correlation matrix given by Table 8. As it can be investigated, there exists a potential multicollinearity between SFH & FP and DHW & ANI. Also, the eigenvalues of $\boldsymbol{X}'\boldsymbol{X}$ are as $\lambda_7 = 2.385563e+05$, $\lambda_6 = 2.302869e+02$, $\lambda_5 = 2.390021e+01$, $\lambda_4 = 1.882493e+01$, $\lambda_3 = 1.561710e+01$, $\lambda_2 = 6.658370$ and $\lambda_1 = 1.682933$. It is easy to see that the condition number is approximately equal to $1.4175e+005$. So, the design matrix $\boldsymbol{X}$ is morbidity badly.

**Table 8**
Correlation matrix.

| Variable | SP | LT | SFH | FP | DHW | GAR | ANI |
|---|---|---|---|---|---|---|---|
| SP | 1.0000 | 0.14591 | 0.4774 | 0.3367 | −0.1001 | 0.2995 | 0.3415 |
| LT | 0.1459 | 1.00000 | 0.1544 | 0.1595 | 0.0814 | 0.1699 | 0.1334 |
| SFH | 0.4774 | 0.15445 | 1.0000 | 0.4633 | 0.0229 | 0.2230 | 0.2779 |
| FP | 0.3367 | 0.15953 | 0.4633 | 1.0000 | 0.1085 | 0.1456 | 0.3814 |
| DHW | −0.1001 | 0.08145 | 0.0229 | 0.1085 | 1.0000 | 0.0579 | −0.1096 |
| GAR | 0.2995 | 0.16991 | 0.2230 | 0.1456 | 0.0579 | 1.0000 | 0.0270 |
| ANI | 0.3415 | 0.13344 | 0.2779 | 0.3814 | −0.1096 | 0.0270 | 1.0000 |

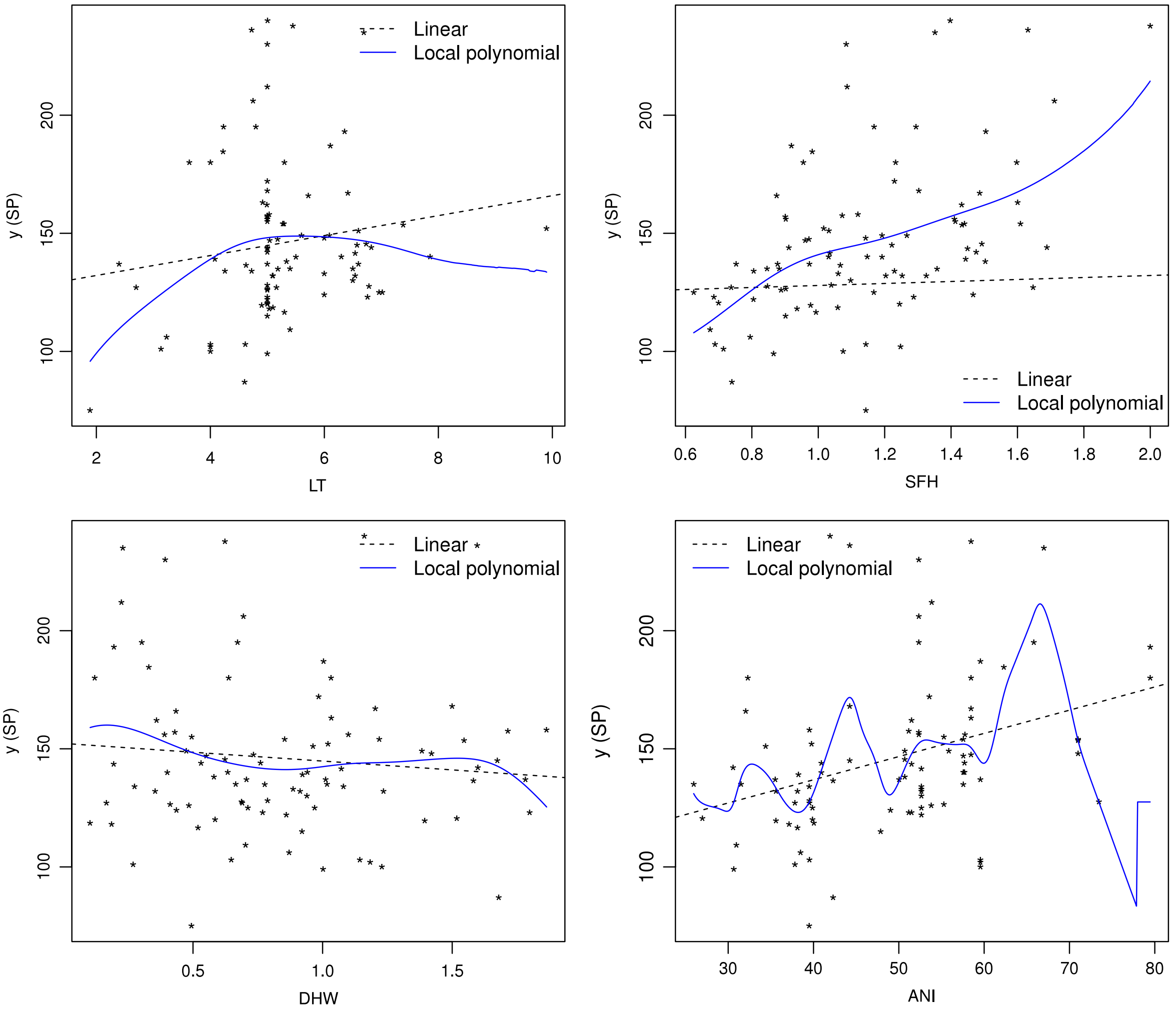


**Fig. 4.** Plots of individual explanatory variables versus dependent variable, linear fit (black dash line) and local polynomial fit (blue solid line).

An appropriate approach is to replace the pure parametric model with semiparametric model. Akdeniz and Tabakan [1] proposed a semiparametric model for this data by taking the lot area as the nonparametric component.

We plotted the dependent variable (SP) versus explanatory variables (except for the binary ones) to suspect the type of relation (linear or non linear) between them in Fig. 5. According to this figure, we consider the average neighborhood income (ANI) as a non-parametric part. So, the specification of the semiparametric model is

$$(SP)_i = \beta_0 + \beta_1(LT)_i + \beta_2(SFH)_i + \beta_3(FP)_i + \beta_4(DHW)_i + \beta_5(GAR)_i + f(ANI)_i + \epsilon_i. \tag{5.4}$$

To compare the performance of the proposed restricted estimators, we consider the parametric restriction $\boldsymbol{H\beta} = \mathbf{0}$, where

$$\boldsymbol{H} = \begin{pmatrix} -1 & 0 & -1 & -1 & 1 \\ 1 & 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & -2 & 8 \end{pmatrix}.$$

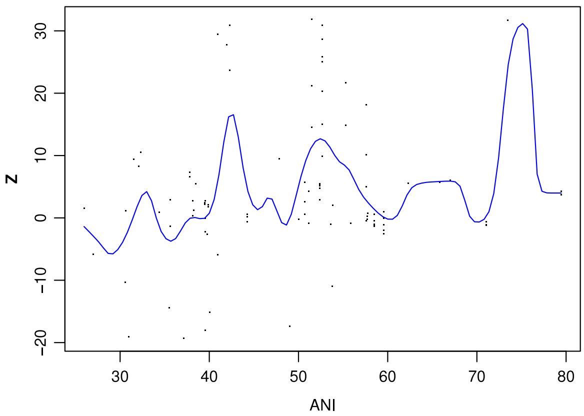


**Fig. 5.** Estimation of nonlinear effect of ANI on dependent variable by kernel fit.

**Table 9**
Fitting of parametric and semiparametric models to housing prices data.

| Variable | Parametric estimates | | Semiparametric estimates | |
|---|---|---|---|---|
| | Coef | SE | Coef | SE |
| Intercept | 62.6695 | 20.6901 | – | – |
| LT | 0.8146 | 2.6487 | 9.4009 | 2.252 |
| SFH | 39.2694 | 11.9940 | 73.4560 | 11.147 |
| FP | 6.0551 | 7.6806 | 5.3394 | 8.147 |
| DHW | −8.2151 | 6.7718 | −2.2271 | 7.282 |
| GAR | 14.4684 | 6.3840 | 12.9860 | 7.197 |
| ANI | 0.5600 | 0.2829 | – | – |
| $s^2$ | 798.0625 | | 149.7332 | |
| $RSS$ | 33860.06 | | 88236.41 | |
| $R^2$ | 0.3329 | | 0.8675 | |

**Table 10**
Evaluation of SGRRE at different $k$ values for housing prices data.

| $k$ variable | 0.0 | 0.10 | 0.20 | $k_{opt}=0.216$ | 0.30 | 0.40 | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| Intercept | – | – | – | – | – | – | – | – |
| LT | 9.0436 | 9.1394 | 9.2355 | 9.3316 | 9.4279 | 9.2508 | 9.5242 | 9.6204 |
| SFH | 84.0027 | 84.1060 | 84.1951 | 84.2709 | 84.3342 | 84.2081 | 84.3855 | 84.4256 |
| FP | 0.4605 | 0.0362 | 0.3764 | 0.7779 | 1.1687 | 0.4413 | 1.5492 | 1.9198 |
| DHW | −4.0864 | −4.3823 | −4.6735 | −4.9600 | −5.2419 | −4.7197 | −5.5195 | −5.7927 |
| GAR | 4.9240 | 4.7341 | 4.5486 | 4.3675 | 4.1905 | 4.5194 | 4.0174 | 3.8482 |
| ANI | – | – | – | – | – | – | – | – |
| $\hat{R}\left(\hat{\boldsymbol{\beta}}^S_{GR}(k);\boldsymbol{\beta}\right)$ | 257.8926 | 254.0920 | 252.6262 | 252.5986 | 253.3512 | 256.1286 | 260.8252 | 267.3135 |
| $\hat{\Delta}$ | 0.000000 | 3.800539 | 5.266399 | 5.293957 | 4.541358 | 1.763961 | −2.932665 | −9.420913 |
| $\widehat{mse}\left(\hat{f}(\cdot),f(\cdot)\right)$ | 2241.811 | 2241.08 | 2240.502 | 2240.428 | 2240.121 | 2240.039 | 2240.025 | 2240.183 |

We test the linear hypothesis $H_0:\boldsymbol{H\beta}=\mathbf{0}$ in the framework of our semiparametric model (2.1). The test statistic for $H_0$, given our observations, is

$$\chi^2_{rank(\boldsymbol{H})}\simeq(\boldsymbol{H}\hat{\boldsymbol{\beta}}_G-\boldsymbol{h})'\left(\boldsymbol{H}\hat{\Sigma}_{\hat{\boldsymbol{\beta}}}\boldsymbol{H}'\right)^{-1}(\boldsymbol{H}\hat{\boldsymbol{\beta}}_G-\boldsymbol{h})=0.4781,$$

where $\hat{\Sigma}_{\hat{\boldsymbol{\beta}}}=\hat{\sigma}^2\left(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}}\right)^{-1}$. Thus we conclude that the null hypothesis $H_0$ is not rejected.

Table 9 summarizes the results. The "parametric estimates" refers to a model in which ANI enters. In the "semiparametric estimates", we have used the kernel regression procedure with optimal bandwidth $h_n=3.38$ for estimating $f(ANI)$. For estimating non parametric effect, first we estimated the parametric effects and then, kernel approach applied to fit $\boldsymbol{Z}_i=SP_i-\boldsymbol{x}_i\hat{\boldsymbol{\beta}}_G$ on $ANI_i$ for $i=1,\dots,n$ , where $\boldsymbol{x}_i=(LT_i,SFH_i,FP_i,DHW_i,GAR_i)$ (Fig. 6).

The ratio of largest eigenvalue to smallest eigenvalue for new design matrix in model (5.4) is approximately $\lambda_5/\lambda_1=427.9926$ and so, there exists a potential multicollinearity between the columns of design matrix. Now, in order to overcome the multicollinearity for better performance of the estimators, we used the proposed estimators for model (5.4). The SGRRE and PRSGRRE for different values of ridge parameter are given in Tables 10 and 11, respectively. As it can be seen, $\hat{\boldsymbol{\beta}}^{S+}_{GR}(k_{opt})$

**Table 11**
Evaluation of PRSGRRE at different $k$ values for housing prices data.

| $k$ variable | 0.0 | 0.10 | $k_{opt} = 0.1955$ | 0.20 | 0.30 | 0.40 | 0.50 | 0.60 |
|---|---|---|---|---|---|---|---|---|
| Intercept | – | – | – | – | – | – | – | – |
| LT | 9.2261 | 9.3554 | 9.4760 | 9.4816 | 9.6049 | 9.7254 | 9.8432 | 9.9584 |
| SFH | 78.6146 | 77.7054 | 76.8600 | 76.8207 | 75.9594 | 75.1206 | 74.3034 | 73.5069 |
| FP | 2.9530 | 3.2292 | 3.4835 | 3.4952 | 3.7516 | 3.9987 | 4.2370 | 4.4667 |
| DHW | −3.1365 | −3.0335 | −2.9375 | −2.9331 | −2.8350 | −2.7393 | −2.6458 | −2.5544 |
| GAR | 9.0426 | 9.1124 | 9.1762 | 9.1791 | 9.2429 | 9.3038 | 9.3622 | 9.4179 |
| ANI | – | – | – | – | – | – | – | – |
| $\hat{R}\left(\hat{\boldsymbol{\beta}}_{GR}^{S+}(k); \boldsymbol{\beta}\right)$ | 234.0246 | 230.6884 | 229.6706 | 230.8286 | 232.6824 | 234.0243 | 239.1255 | 246.0051 |
| $\hat{\Delta}$ | 0.000000 | 3.336275 | 4.353975 | 3.196063 | 1.342234 | 0.000319 | −5.100843 | −11.98048 |
| $\widehat{mse}\left(\hat{f}(\cdot), f(\cdot)\right)$ | 2225.762 | 2233.685 | 2240.480 | 2240.838 | 2248.805 | 2256.785 | 2264.771 | 2272.758 |

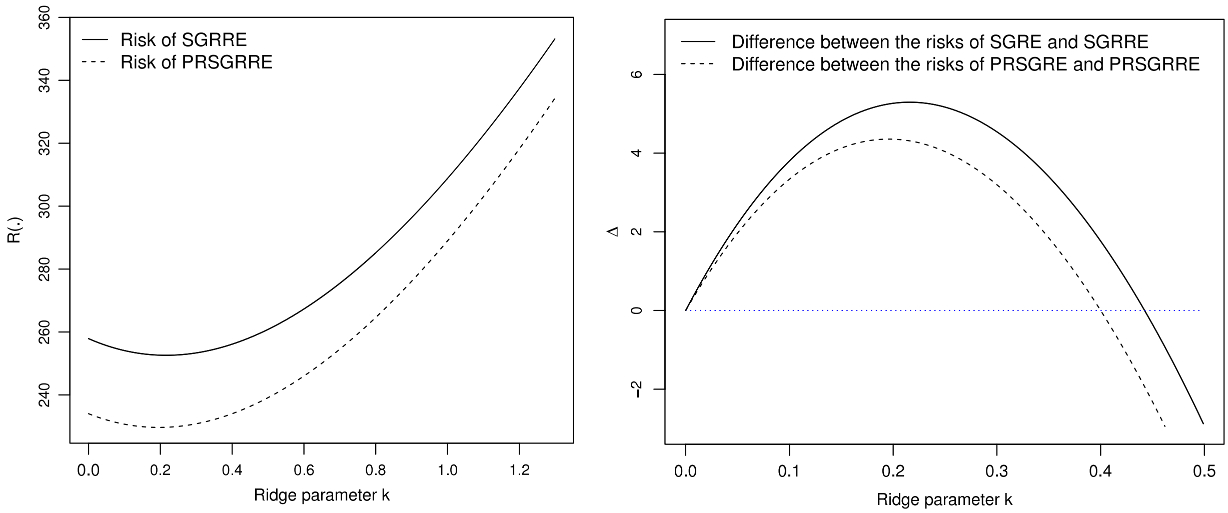


**Fig. 6.** The diagrams of $\Delta$ and $R(\cdot)$ versus $k$ for housing prices data.

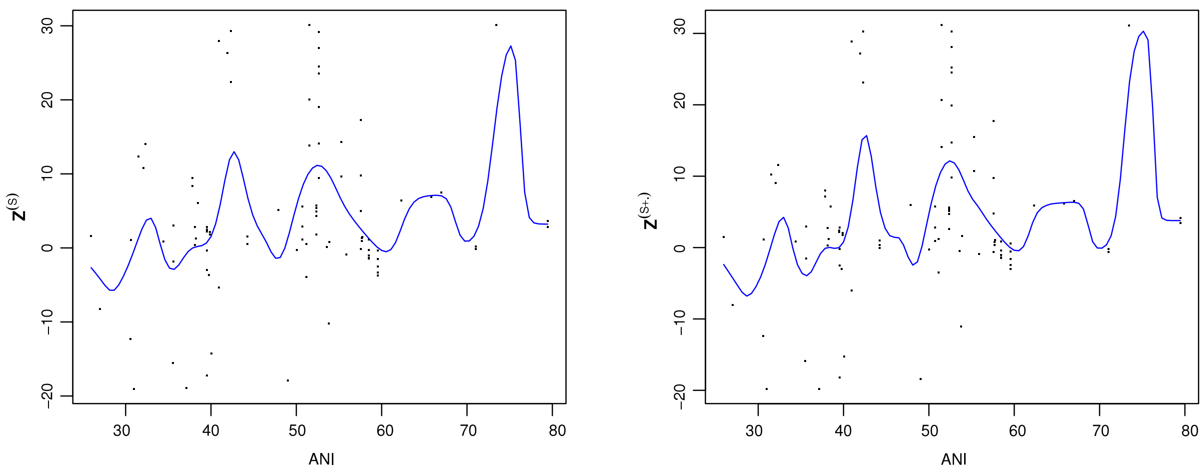


**Fig. 7.** Estimation of $f(ANI)$ by kernel regression after removing the linear part by proposed estimators in housing prices data.

is the best estimator for linear part of the semiparametric regression model in the sense of risk. In Fig. 7, the $\Delta$ and $R(\cdot)$ of SGRRE (solid line) and PRSGDRRE (dash line) versus ridge parameter $k$ are plotted. Finally, we estimated the non parametric effect ($f(ANI)$) after estimation of the linear part by $\hat{\boldsymbol{\beta}}_{GR}^{S \text{ or } S+}(k_{opt})$ in Fig. 7, i.e., we used kernel fit to regress $\boldsymbol{Z}^{(S) \text{ or } (S+)} = SP - \boldsymbol{X}\hat{\boldsymbol{\beta}}_{GR}^{S \text{ or } S+}(k_{opt})$ on $ANI$.

Based on Theorem 4.1 in the real example study, it can be said that:

- At most $k$ for dominance of SGRRE over SGRE is 0.442.
- At most $k$ for dominance of PRSGRRE over PRSGRE is 0.4.
- The PRSGRRE is better than SGRRE for all values of $k$ in the sense of risk function (Fig. 7).

## References

[1] F. Akdeniz, G. Tabakan, Restricted ridge estimators of the parameters in semiparametric regression model, Comm. Statist. Theory Methods 38 (2009) 1852–1869.
[2] E. Akdeniz Duran, W.K. Härdle, M. Osipenko, Difference based ridge and Liu type estimators in semiparametric regression models, J. Multivariate Anal. 105 (2012) 164–175.
[3] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley, New York, 1984.
[4] M. Arashi, Preliminary test and Stein estimators in simultaneous linear equations, Linear Algebra Appl. 436 (5) (2012) 1195–1211.
[5] M. Arashi, B.M. Golam Kibria, M. Norouzirad, S. Nadarajah, Improved preliminary test and Stein-rule Liu estimators for the ill-conditioned elliptical linear regression model, J. Multivariate Anal. 124 (2014) 53–74.
[6] M. Arashi, A.K. Md. Ehsanes Saleh, S.M.M. Tabatabaey, Estimation of parameters of parallelism model with elliptically distributed errors, Metrika 71 (2010) 79–100.
[7] M. Arashi, S.M.M. Tabatabaey, Improved variance estimation under sub-space restriction, J. Multivariate Anal. 100 (2009) 1752–1760.
[8] M. Arashi, S.M.M. Tabatabaey, H. Soleimani, Simple regression in view of elliptical models, Linear Algebra Appl. 437 (2012) 1675–1691.
[9] A.W. Bowman, P. Hall, T. Prvan, Cross-validation for the smoothing of distribution functions, Biometrika 85 (1998) 799–808.
[10] F. Bunea, Consistent covariate selection and post model selection inference in semiparametric regression, Ann. Statist. 32 (2004) 898–927.
[11] D.G. Gibbons, A simulation study of some ridge estimators, J. Amer. Statist. Assoc. 76 (1981) 131–139.
[12] M. Gisela, B.M.G. Kibria, One some ridge regression estimators: an empirical comparisons, Comm. Statist. Simulation Comput. 38 (3) (2009) 621–630.
[13] W. Härdle, H. Liang, J. Gao, Partially Linear Models, Physika Verlag, Heidelberg, 2000.
[14] M. Hassanzadeh Bashtian, M. Arashi, S.M.M. Tabatabaey, Ridge estimation under the stochastic restriction, Comm. Statist. Theory Methods 40 (2011) 3711–3747.
[15] M. Hassanzadeh Bashtian, M. Arashi, S.M.M. Tabatabaey, Using improved estimation strategies to combat multicollinearity, J. Stat. Comput. Simul. 81 (12) (2011) 1773–1797.
[16] M. Ho, Essays on the Housing Market (Ph.D. dissertation), University of Toronto, 1995, unpublished.
[17] A.E. Horel, R.W. Kennard, Ridge regression: biased estimation for non-orthogonal problems, Technometrics 12 (1970) 55–67.
[18] W. Judge, M.E. Bock, The Statistical Implication of Pre-test and Stein-rule Estimators in Econometrics, North-Holland, New York, 1978.
[19] S. Kaciranlar, S. Sakalliogu, M.R. Ozkale, H. Guler, More on the restricted ridge regression estimation, J. Stat. Comput. Simul. 81 (2011) 1433–1448.
[20] B.M.G. Kibria, Some Liu and Ridge type estimators and their properties under the ill-conditioned Gaussian linear regression model, J. Stat. Comput. Simul. 82 (2012) 1–17.
[21] B.M.G. Kibria, A.K.Md.E. Saleh, Preliminary test ridge regression estimators with student's t errors and conflicting test-statistics, Metrika 59 (2004) 105–124.
[22] B.M.G. Kibria, A.K.Md.E. Saleh, Improving the estimators of the parameters of a probit regression model: a ridge regression approach, J. Statist. Plann. Inference 142 (6) (2011) 1421–1435.
[23] J. Li, M. Palta, Bandwidth selection through cross-validation for semi-parametric varying-coefficient partially linear models, J. Stat. Comput. Simul. 79 (2009) 1277–1286.
[24] J. Li, W. Zhang, Z. Wu, Optimal zone for bandwidth selection in semiparametric models, J. Nonparametr. Stat. 23 (2011) 701–717.
[25] G.C. McDonald, D.I. Galarneau, A Monte Carlo evaluation of some ridge-type estimators, J. Amer. Statist. Assoc. 70 (1975) 407–416.
[26] M. Müller, Semiparametric Extensions to Generalized Linear Models, Habilitationsschrift, 2000.
[27] S. Najarian, M. Arashi, B.M.G. Kibria, A simulation study on some restricted ridge regression estimators, Comm. Statist. Simulation Comput. 42 (2013) 871–879.
[28] M.R. Ozkale, Comment on Ridge estimation to the restricted linear model, Comm. Statist. Theory Methods 38 (2009) 1094–1097.
[29] M.R. Ozkale, S. Kaciranlar, The restricted and unrestricted two parameter estimators, Comm. Statist. Theory Methods 36 (2007) 2707–2727.
[30] S.M.E. Raheem, S.E. Ahmed, K.A. Doksum, Absolute penalty and shrinkage estimation in partially linear models, Comput. Statist. Data Anal. 56 (2012) 874–891.
[31] A.K.Md.E. Saleh, Theory of Preliminary Test and Stein-type Estimation with Applications, John Wiley, New York, 2006.
[32] A.K.Md.E. Saleh, B.M.G. Kibria, Performances of some new preliminary test ridge regression estimators and their properties, Comm. Statist. A 22 (1993) 2747–2764.
[33] N. Sarkar, A new estimator combining the ridge regression and the restricted least squares method of estimation, Comm. Statist. A 21 (1992) 1987–2000.
[34] P. Speckman, Kernel smoothing in partial linear models, J. R. Stat. Soc. Ser. B 50 (1988) 413–436.
[35] P.A.V.B. Swamy, J.S. Mehta, A note on minimum average risk estimator for coefficients in linear models, Comm. Statist. Theory Methods 6 (1977) 1161–1186.
[36] P.A.V.B. Swamy, J.S. Mehta, P.N. Rappoport, Two methods of evaluating Hoerl and Kennard's ridge regression, Comm. Statist. 12 (1978) 1133–1155.
[37] C. Zhang, J. Li, J. Menga, On Steins lemma, dependent covariates and functional monotonicity in multi-dimensional modeling, J. Multivariate Anal. 99 (2008) 2285–2303.
[38] Z. Zhong, Hu Yang, Ridge estimation to the restricted linear model, Comm. Statist. Theory Methods 36 (11) (2007) 2099–2115.