



Research article

Feasible robust Liu estimator to combat outliers and multicollinearity effects in restricted semiparametric regression model

W. B. Altukhaes^{1,2}, M. Roozbeh^{3,*} and N. A. Mohamed^{1,*}

¹ Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, Malaysia

² Department of Mathematics, College of Science and Humanities, Shaqra University, Sahqra 11961, Saudi Arabia

³ Department of Statistics, Faculty of Mathematics, Statistics and Computer Sciences, Semnan University, Semnan, Iran

* **Correspondence:** Email: mahdi.roozbeh@semnan.ac.ir, nuranisah_mohamed@um.edu.my.

Abstract: Regression analysis frequently encounters two issues: multicollinearity among the explanatory variables, and the existence of outliers in the data set. Multicollinearity in the semiparametric regression model causes the variance of the ordinary least-squares estimator to become inflated. Furthermore, the existence of multicollinearity may lead to wide confidence intervals for the individual parameters and even produce estimates with wrong signs. On the other hand, as is often known, the ordinary least-squares estimator is extremely sensitive to outliers, and it may be completely corrupted by the existence of even a single outlier in the data. Due to such drawbacks of the least-squares method, a robust Liu estimator based on the least trimmed squares (LTS) method for the regression parameters is introduced under some linear restrictions on the whole parameter space of the linear part in a semiparametric model. Considering that the covariance matrix of the error terms is usually unknown in practice, the feasible forms of the proposed estimators are substituted, and their asymptotic distributional properties are derived. Moreover, necessary and sufficient conditions for the superiority of the Liu type estimators over their counterparts for choosing the biasing Liu parameter d are extracted. The performance of the feasible type of robust Liu estimators is compared with the classical ones in constrained semiparametric regression models using extensive Monte-Carlo simulation experiments and a real data example.

Keywords: feasible generalized Liu estimator; least trimmed squares estimator; multicollinearity; outlier; semiparametric regression models

Mathematics Subject Classification: 62G08, 62G35, 62J05, 62J07

1. Introduction

The semiparametric regression model (SRM) is an appropriate tool to model a data set when the type of relationship between the dependent variable and some of the explanatory variables is linear parametric, but the link function of the dependent variable with other explanatory variables is not clear [1–3]. Consider the set of observations denoted as $(y_1, \mathbf{x}_1^\top, t_1), \dots, (y_n, \mathbf{x}_n^\top, t_n)$, which conform to the semiparametric regression model defined by

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (1.1)$$

where y_i is the value of the response variable for i th observation, $\mathbf{x}_i^\top = (x_{i1}, \dots, x_{ip})$ represents a vector of explanatory variables, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ denotes a vector of the unknown parameters, and the t_i 's are the observed points that match the boundaries of the domain $D \subset \mathbb{R}$ [4–6]. It is generally assumed that the unknown function $f(\cdot)$ is a smooth function, while the ε_i 's represent random errors that are considered to be independent of both (x_i, t_i) .

Since the semiparametric regression models combine both parametric and nonparametric components, the response variable depends on the explanatory variables in a linear form but has a nonlinear relationship with other explanatory variables, which are more flexible than standard linear regression models according to (1.1) [7–9]. There are different ways to estimate $\boldsymbol{\beta}$ and $f(\cdot)$. Some of the most important methods were introduced by [10,11].

The presence of nearly linear dependency among the columns of the design matrix $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$ is known as multicollinearity, and it is an issue that might arise in regression analysis. In this case, the matrix $\mathbf{S} = \mathbf{X}^\top \mathbf{X}$ contains one or more small eigenvalues, causing the regression coefficient estimations to be large in absolute value. The condition number is an effective measure for detecting the presence of multicollinearity. The matrix \mathbf{S} is ill-conditioned under multicollinearity because its condition number tends to an extremely large value. Multicollinearity makes the ordinary least-squares estimator (OLSE) perform badly. Also, multicollinearity in data may cause confidence intervals to be too large for either the individual parameters or their linear mixes, which may lead to inaccurate predictions. Applying shrinkage estimators is widely used as an effective solution to address the issues arising from multicollinearity [12–15]. In this study, the shrinkage estimator, suggested by Liu [16], is applied to solve the problem of multicollinearity. Liu [16] combined the Stein type estimator with the conventional ordinary ridge regression estimator to derive the Liu estimator, as described in [17,18]. Other alternative approaches to addressing the issue of multicollinearity can be found in [19–21].

Besides the multicollinearity problem, another typical issue that arises in regression analysis is the presence of outliers, which are observations that do not follow the pattern of the main bulk of the data. Outliers can cause problems like inflated sums of squares, estimate bias, p-value confusion, and more. To combat these problems, robust regression methods are used. The ordinary least-squares estimator is known to be extremely affected by outliers, so the least trimmed squares approach is used to estimate both components of SRM in this research.

The breakdown point of an estimator is the fundamental measurement that is used to evaluate its robustness. This breakdown point concept refers to the percentage of outlying observations (up to 50 percent) that can contaminate the estimation promiscuously. In computational geometry, the investigation of effective algorithms for robust estimation methods has been an important field of study. Several researchers have examined the robust least median of squares (LMS) method, which is the hyperplane that

minimizes the squared residual median [22]. Although the LMS estimator has been the subject of most publications on robust estimation in the field of linear models, Rousseeuw and Leroy [23] noted that LMS is not the optimal option due to its statistical features. They asserted that selecting the least trimmed squares is the better alternative option because both LTS and LMS have the same breakdown point, approximately 50%, but the objective function of LTS is smoother than LMS. Also, since LTS converges more quickly and is distributed asymptotically normally [22], it has superior statistical efficiency. For these reasons, LTS is a better choice as a starting point for two-step robust estimators such as the MM-estimator [24].

The main focus of this paper is to study a feasible generalized robust Liu estimator in a restricted semiparametric regression model. The organization of this article is as follows: Section 2 contains the classical estimator of a restricted semiparametric regression model based on the kernel method. After reviewing the concepts of Liu and least trimmed squares approaches in a semiparametric regression model, a new feasible robust Liu estimator in a restricted semiparametric regression model is suggested, and then its asymptotic bias and distributional covariance are derived in Section 3. Based on the obtained results, the feasible generalized robust Liu estimator is compared with the classical one in terms of the mean squared error. In Section 4, the efficiencies of the proposed estimators are assessed through Monte Carlo simulation experiments as well as with a real-world data example. Finally, some important findings are concluded in Section 5.

2. Feasible type of the classical estimators in RSRM

The estimators conform to certain restrictions in classical estimators. Let us examine the semiparametric regression model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + f(\mathbf{t}) + \boldsymbol{\varepsilon}, \quad (2.1)$$

where $\mathbf{y} = (y_1, \dots, y_n)^\top$, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$, $f(\mathbf{t}) = (f(t_1), \dots, f(t_n))^\top$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^\top$.

Generally, we assume that $\boldsymbol{\varepsilon}$ is a vector of disturbances that follows a distribution with $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $E(\boldsymbol{\varepsilon}^\top \boldsymbol{\varepsilon}) = \sigma^2 \mathbf{V}$, where σ^2 is an unknown parameter and \mathbf{V} is a symmetric and positive definite matrix. To estimate the linear part of model (2.1), we first remove the non-parametric effect by detrending. Given the assumption that $\boldsymbol{\beta}$ is known, a natural non-parametric estimator of $f(\cdot)$ is

$$\hat{f}(\mathbf{t}) = k(\mathbf{t})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}),$$

in which $k(\cdot)$ is a kernel function. Following [25], by substituting $\hat{f}(\mathbf{t})$ for $f(\mathbf{t})$ in Eq (2.1), the model may be reduced to

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad (2.2)$$

where $\tilde{\mathbf{y}} = (\mathbf{I}_n - \mathbf{K})\mathbf{y}$, $\tilde{\mathbf{X}} = (\mathbf{I}_n - \mathbf{K})\mathbf{X}$, and \mathbf{K} is the smoother matrix with (i, j) -th component $K_\omega(t_i, t_j)$ in which $K_\omega(\cdot)$ is a kernel function of order m with bandwidth parameter ω . We can simply use the following transformations to change model (2.2) into a standard regression model by multiplying $\mathbf{V}^{-1/2}$ on the both sides as follows:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}, \quad E(\tilde{\boldsymbol{\varepsilon}}) = \mathbf{0}, \quad E(\tilde{\boldsymbol{\varepsilon}}\tilde{\boldsymbol{\varepsilon}}^\top) = \sigma^2 \mathbf{I}_n,$$

where $\tilde{\mathbf{y}} = \mathbf{V}^{-1/2}\tilde{\mathbf{y}}$, $\tilde{\mathbf{X}} = \mathbf{V}^{-1/2}\tilde{\mathbf{X}}$, and $\tilde{\boldsymbol{\varepsilon}} = \mathbf{V}^{-1/2}\boldsymbol{\varepsilon}$. Now, the estimation of $\boldsymbol{\beta}$ is performed using the generalized least-squares estimator (GLSE), which is known to be the best linear unbiased estimator

$$\begin{aligned}
\hat{\beta}_{GLS} &= \operatorname{argmin}_{\beta} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta)^{\top} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta) \\
&= (\tilde{\mathbf{X}}^{\top} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^{\top} \tilde{\mathbf{y}} \\
&= \mathbf{C}^{-1} \tilde{\mathbf{X}}^{\top} \mathbf{V}^{-1} \tilde{\mathbf{y}},
\end{aligned} \tag{2.3}$$

where $\mathbf{C} = \tilde{\mathbf{X}}^{\top} \mathbf{V}^{-1} \tilde{\mathbf{X}}$.

In applications, the matrix \mathbf{V} is not known. Therefore, $\hat{\beta}_{GLS}$ in Eq (2.3) is not applicable because it is a function of covariance matrix (\mathbf{V}) which is not known. To solve this issue, we have to utilize a two-stage process and implement a feasible generalized least-squares estimator (FGLSE) by replacing the unknown parameter \mathbf{V} with a suitable estimator [26], $\mathbf{S} = \frac{1}{n-p} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\hat{\beta}_{LS})(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\hat{\beta}_{LS})^{\top}$, which is a consistent estimator, as follows:

$$\hat{\beta}_{FGLS} = \mathbf{C}_F^{-1} \tilde{\mathbf{X}}^{\top} \mathbf{S}^{-1} \tilde{\mathbf{y}}, \tag{2.4}$$

where $\mathbf{C}_F = \tilde{\mathbf{X}}^{\top} \mathbf{S}^{-1} \tilde{\mathbf{X}}$, and $\hat{\beta}_{LS}$ is the ordinary least-squares estimator $(\tilde{\mathbf{X}}^{\top} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^{\top} \tilde{\mathbf{y}}$. As demonstrated in Zellner [26], $\hat{\beta}_{FGLS} = \hat{\beta}_{GLS} + O_p(n^{-1})$, and consequently $\sqrt{n}(\hat{\beta}_{FGLS} - \beta)$ and $\sqrt{n}(\hat{\beta}_{GLS} - \beta)$ have the same normal asymptotical distribution, and so $\operatorname{Var}(\hat{\beta}_{FGLS}) = \mathbf{C}_F^{-1} + o(n^{-1})$, where $O_p(n^{-1})$ indicates an amount which is of order n^{-1} in probability and $o(n^{-1})$ indicates a quantity of a higher order of smallness than n^{-1} .

Interestingly, another method to handle the strong and extremely strong multicollinearity problems is to obtain the estimators under particular constraints on the unknown parameters, which may be exact or stochastic (see [27–29] for more details). By applying some constraints on the parameter space of the linear part, Durbin [30], Theil and Goldberger [31], and Theil [32] proposed the ordinary mixed estimator (OME) for the vector of the regression coefficient. Assume that we had prior knowledge regarding β in the sense of non-stochastic exact constraints [33–35], as follows:

$$\mathbf{R}\beta = \mathbf{r}$$

where \mathbf{R} is a known matrix $q \times p$ of prior information of rank $q < p$ and \mathbf{r} is a known $q \times 1$ vector. This restriction should come from an outside source (it might be determined, for example, by an outside source of information or an expert). Thus, when the regression parameters are restricted by a space of linear constraints non-stochastically represented by independent prior information, we provide the instruments necessary to compute the risk of estimators. Next, the performances of the new constrained estimators and the classical estimators may be compared under certain conditions. We show that the innovative constrained estimators outperform the classical ones in terms of least-risk functions, assuming linear restrictions. In these circumstances, certain non-sample information (a previous constraint on the parameters) may exist; they are often presented to the model as constraints. Compared to typical estimators, the restricted estimation performs better, and so, in this work, the restricted semiparametric regression model (RSRM) is fitted to the data set. The feasible generalized least-squares restricted estimator (FGLSRE) is derived by imposing a linear restriction as follows:

$$\begin{aligned}
\hat{\beta}_{FGLSR} &= \operatorname{argmin}_{\beta} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta)^{\top} \mathbf{S}^{-1} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta) \\
&\quad \text{s.t. } \mathbf{R}\beta = \mathbf{r} \\
&= \hat{\beta}_{FGLS} - \mathbf{C}_F^{-1} \mathbf{R}^{\top} (\mathbf{R} \mathbf{C}_F^{-1} \mathbf{R}^{\top})^{-1} (\mathbf{R} \hat{\beta}_{FGLS} - \mathbf{r}).
\end{aligned} \tag{2.5}$$

As it is known, the covariance matrix estimation of $\hat{\beta}_{FGLS}$ is equal to $\sigma^2 \mathbf{C}_F^{-1}$. So, the FGLSE and its

covariance matrix are significantly influenced by the features of matrix \mathbf{C}_F . The FGLS estimators gain susceptibility to various errors when \mathbf{C}_F is ill-conditioned. Also, some of the estimations of the regression coefficients, for instance, might have incorrect signs or be statistically insignificant; this could lead to unstable estimators, which are characterized by large confidence intervals for the specific parameters. Making valid statistical inferences becomes challenging in the presence of these errors, and so a biased estimation technique is introduced and utilized for RSRM under the multicollinearity problem.

3. Feasible robust Liu estimator in RSRM

Multicollinearity leads to $\mathbf{X}^T \mathbf{X}$ being ill-conditioned with a large condition number. When the signal to noise ratio $\boldsymbol{\beta}^T \boldsymbol{\beta} / \sigma^2$ is small and the condition number of $\mathbf{X}^T \mathbf{X}$ is large, the least-squares estimator is most severely affected by multicollinearity. In this situation, the high level of data noise is enlarged by $(\mathbf{X}^T \mathbf{X})^{-1}$, making the least-squares estimator highly unstable. To solve this drawback, Hoerl and Kennard [36] proposed the ridge estimator $\hat{\boldsymbol{\beta}}_k = (\mathbf{X}^T \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$ in the standard linear regression model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $E(\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}$, and it has become the most often used method for combating the multicollinearity problem that causes the least squares estimator to fail. Indeed, the ridge method solves the multicollinearity problem by adding a small non-stochastic constant k to the diagonal elements of $\mathbf{X}^T \mathbf{X}$ to decrease its condition number. In practical use, the biasing parameter k in ridge approach is typically rather modest. It is obvious that the condition number of $\mathbf{X}^T \mathbf{X} + k\mathbf{I}$ is a decreasing function of k . Thus, high values of k are needed to achieve small-scale control over the condition number of $\mathbf{X}^T \mathbf{X} + k\mathbf{I}$. Because of this, the small k selected in practice may not be sufficiently big to solve the severe multicollinearity problem of $\mathbf{X}^T \mathbf{X}$. So, the resultant ridge estimation may still be unstable since $\mathbf{X}^T \mathbf{X} + k\mathbf{I}$ has remained ill-conditioned. Furthermore, despite its practical effectiveness, the ridge estimator is a complicated function of k . Although the Stein-type estimator $\hat{\boldsymbol{\beta}}_c = c(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ is a linear function of c , the shrinkage of each element of $\hat{\boldsymbol{\beta}}_c$ is the same. To address these issues, Liu [12] proposed a new biased estimator $\hat{\boldsymbol{\beta}}_d = (\mathbf{X}^T \mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^T \mathbf{y} + d\hat{\boldsymbol{\beta}})$ by combining the advantages of the ridge and Stein-type estimators, which effectively solved the problem of ill-conditioning in standard regression model, where $0 < d < 1$ is a biasing parameter and $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$. It is obvious that when $d = 1$, $\hat{\boldsymbol{\beta}}_d = \hat{\boldsymbol{\beta}}$.

According to [12] the mean squared error (MSE) of the Liu estimator is obtained by

$$\text{MSE}(\hat{\boldsymbol{\beta}}_d) = \sigma^2 \sum_{j=1}^p \frac{(\lambda_j + d)^2}{\lambda_j (\lambda_j + 1)^2} + (d - 1)^2 \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2}, \quad (3.1)$$

where α_j^2 corresponds to the j th element of $\boldsymbol{\alpha} = \boldsymbol{\Gamma}^T \boldsymbol{\beta}$ and $\boldsymbol{\Gamma}$ is an orthogonal matrix such that $\mathbf{C}_F = \boldsymbol{\Gamma} \boldsymbol{\Lambda} \boldsymbol{\Gamma}^T$, in which $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_p)$ contains the eigenvalues of matrix \mathbf{C}_F . Consequently, the biasing parameter d is chosen by minimizing $\text{MSE}(\hat{\boldsymbol{\beta}}_d)$ as follows:

$$\hat{d} = 1 - \hat{\sigma}_{FGLS}^2 \left(\frac{\sum_{j=1}^p \frac{1}{\lambda_j (\lambda_j + 1)}}{\sum_{j=1}^p \frac{\hat{\alpha}_{jFGLS}^2}{(\lambda_j + 1)^2}} \right), \quad (3.2)$$

where $\hat{\sigma}_{FGLS}^2$ and $\hat{\alpha}_{jFGLS}^2$ are the unbiased estimators of σ^2 and α_j based on FGLSE, respectively, i.e., $\hat{\sigma}_{FGLS}^2 = \frac{1}{n-p} (\mathbf{y} - \tilde{\mathbf{X}} \hat{\boldsymbol{\beta}}_{FGLS})^T \mathbf{S}^{-1} (\mathbf{y} - \tilde{\mathbf{X}} \hat{\boldsymbol{\beta}}_{FGLS})$ and $\hat{\alpha}_{jFGLS} = \boldsymbol{\Gamma}^T \hat{\boldsymbol{\beta}}_{FGLS}$.

The feasible generalized least-squares Liu estimator (FGLSLE) can be extended by [37–39] as

follows:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{FGLSL}(d) &= (\widetilde{\mathbf{X}}^\top \mathbf{S}^{-1} \widetilde{\mathbf{X}} + \mathbf{I})^{-1} (\widetilde{\mathbf{X}}^\top \mathbf{S}^{-1} \widetilde{\mathbf{X}} + d\mathbf{I}) \widehat{\boldsymbol{\beta}}_{FGLS} \\ &= (\mathbf{C}_F + \mathbf{I})^{-1} (\mathbf{C} + d\mathbf{I}) \widehat{\boldsymbol{\beta}}_{FGLS} \\ &= \mathbf{F}_d \widehat{\boldsymbol{\beta}}_{FGLS}, \quad 0 \leq d \leq 1,\end{aligned}\tag{3.3}$$

where $\mathbf{F}_d = (\mathbf{C}_F + \mathbf{I})^{-1}(\mathbf{C}_F + d\mathbf{I})$.

Based on the fact that \mathbf{F}_d and \mathbf{C}_F^{-1} are commutative, the feasible generalized least-squares restricted Liu estimator (FGLSRLE) can be obtained for RSRM as follows [40–42]:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{FGLSRLE}(d) &= \underset{\boldsymbol{\beta}}{\operatorname{argmin}} (\widetilde{\mathbf{y}} - \widetilde{\mathbf{X}}\boldsymbol{\beta})^\top \mathbf{S}^{-1} (\widetilde{\mathbf{y}} - \widetilde{\mathbf{X}}\boldsymbol{\beta}) + (d\widehat{\boldsymbol{\beta}}_{FGLS} - \boldsymbol{\beta})^\top (d\widehat{\boldsymbol{\beta}}_{FGLS} - \boldsymbol{\beta}) \\ &\quad \text{s.t. } \mathbf{R}\boldsymbol{\beta} = \mathbf{r} \\ &= \mathbf{F}_d \widehat{\boldsymbol{\beta}}_{FGLS} - \mathbf{F}_d \mathbf{C}_F^{-1} \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F^{-1} \mathbf{R}^\top)^{-1} (\mathbf{R} \widehat{\boldsymbol{\beta}}_{FGLS} - \mathbf{r}).\end{aligned}\tag{3.4}$$

Lemma 3.1. If $\boldsymbol{\beta}$ satisfies the linear restriction $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$, then the properties (bias, covariance, and mean squared error) of the suggested estimator can be calculated directly as follows:

$$\operatorname{Bias}(\widehat{\boldsymbol{\beta}}_{FGLSRLE}(d)) = -(\mathbf{I} - \mathbf{F}_d)\boldsymbol{\beta} + o\left(n^{-\frac{1}{2}}\right),\tag{3.5}$$

$$\operatorname{Cov}(\widehat{\boldsymbol{\beta}}_{FGLSRLE}(d)) = \sigma^2 \mathbf{F}_d \mathbf{H} \mathbf{F}_d^\top + o(n^{-1}),\tag{3.6}$$

$$\operatorname{MSE}(\widehat{\boldsymbol{\beta}}_{FGLSRLE}(d)) = \sigma^2 \operatorname{tr}(\mathbf{F}_d \mathbf{H} \mathbf{F}_d^\top) + \boldsymbol{\beta}^\top (\mathbf{F}_d - \mathbf{I})^\top (\mathbf{F}_d - \mathbf{I}) \boldsymbol{\beta} + o(n^{-1}),\tag{3.7}$$

where $\mathbf{H} = \mathbf{C}_F^{-1} (\mathbf{I} - \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F^{-1} \mathbf{R}^\top)^{-1} \mathbf{R} \mathbf{C}_F^{-1})$.

Theorem 3.1. The MSE of FGLSRLE under the linear restriction $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ can be given by

$$\operatorname{MSE}(\widehat{\boldsymbol{\beta}}_{FGLSRLE}(d)) = \sigma^2 \sum_{j=1}^p \frac{(\lambda_j + d)^2}{(\lambda_j + 1)^2} m_{jj} + (d - 1)^2 \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} + o(n^{-1}),\tag{3.8}$$

where m_{jj} is the j th diagonal element of the matrix $\mathbf{M} = \boldsymbol{\Gamma}^\top \mathbf{H} \boldsymbol{\Gamma}$.

Proof. Using $(\mathbf{C}_F + \mathbf{I})^{-1} = \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + \mathbf{I})^{-1} \boldsymbol{\Gamma}^\top$ and $(\mathbf{C}_F + d\mathbf{I}) = \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + d\mathbf{I}) \boldsymbol{\Gamma}^\top$, we can write

$$\begin{aligned}\operatorname{tr}(\mathbf{F}_d \mathbf{H} \mathbf{F}_d^\top) &= \operatorname{tr}((\mathbf{C}_F + \mathbf{I})^{-1} (\mathbf{C}_F + d\mathbf{I}) \mathbf{H} (\mathbf{C}_F + d\mathbf{I}) (\mathbf{C}_F + \mathbf{I})^{-1}) \\ &= \operatorname{tr}(\boldsymbol{\Gamma}(\boldsymbol{\Lambda} + \mathbf{I})^{-1} \boldsymbol{\Gamma}^\top \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + d\mathbf{I}) \boldsymbol{\Gamma}^\top \mathbf{H} \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + d\mathbf{I}) \boldsymbol{\Gamma}^\top \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + \mathbf{I})^{-1} \boldsymbol{\Gamma}^\top) \\ &= \operatorname{tr}((\boldsymbol{\Lambda} + \mathbf{I})^{-2} (\boldsymbol{\Lambda} + d\mathbf{I})^2 \boldsymbol{\Gamma}^\top \mathbf{H} \boldsymbol{\Gamma}) \\ &= \sum_{j=1}^p \frac{(\lambda_j + d)^2}{(\lambda_j + 1)^2} m_{jj}.\end{aligned}$$

Also, from $(\mathbf{C}_F + \mathbf{I})^{-2} = \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + \mathbf{I})^{-2} \boldsymbol{\Gamma}^\top$, we have

$$\begin{aligned}\boldsymbol{\beta}^\top (\mathbf{F}_d - \mathbf{I})^\top (\mathbf{F}_d - \mathbf{I}) \boldsymbol{\beta} &= \boldsymbol{\alpha}^\top \boldsymbol{\Gamma}^\top ((\mathbf{C}_F + \mathbf{I})^{-1} (\mathbf{C}_F + d\mathbf{I}) - \mathbf{I})^\top ((\mathbf{C}_F + \mathbf{I})^{-1} (\mathbf{C}_F + d\mathbf{I}) - \mathbf{I}) \boldsymbol{\Gamma} \boldsymbol{\alpha} \\ &= \boldsymbol{\alpha}^\top \boldsymbol{\Gamma}^\top ((\mathbf{C}_F + d\mathbf{I}) - (\mathbf{C}_F + \mathbf{I})) (\mathbf{C}_F + \mathbf{I})^{-2} ((\mathbf{C}_F + d\mathbf{I}) - (\mathbf{C}_F + \mathbf{I})) \boldsymbol{\Gamma} \boldsymbol{\alpha} \\ &= (d - 1)^2 \boldsymbol{\alpha}^\top \boldsymbol{\Gamma}^\top \boldsymbol{\Gamma}(\boldsymbol{\Lambda} + \mathbf{I})^{-2} \boldsymbol{\Gamma}^\top \boldsymbol{\Gamma} \boldsymbol{\alpha} \\ &= (d - 1)^2 \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2}.\end{aligned}$$

So, the proof is completed. ■

As an important result of Theorem 3.1, the optimal value of the biasing parameter d can be obtained by differentiating the MSE function of GLSRLE as a function of d (same as $g(d)$) with respect to d , and solve it by setting the derivative equal to zero to extract the optimal value of d . Via direct calculation, we have

$$g'(d) = \frac{\partial \text{MSE}(\hat{\beta}_d)}{\partial d} = 2\sigma^2 \sum_{j=1}^p \frac{(\lambda_j + d)}{(\lambda_j + 1)^2} m_{jj} + 2(d - 1) \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} = 0.$$

So, it can be written

$$\begin{aligned} \left(\sigma^2 \sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2} + \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} \right) d &= \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} - \sigma^2 \sum_{j=1}^p \frac{\lambda_j m_{jj}}{(\lambda_j + 1)^2} \\ \Rightarrow d &= \frac{\sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} - \sigma^2 \sum_{j=1}^p \frac{\lambda_j m_{jj}}{(\lambda_j + 1)^2}}{\sigma^2 \sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2} + \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2}} \\ &= \frac{\sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2} \pm \sigma^2 \sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2} - \sigma^2 \sum_{j=1}^p \frac{\lambda_j m_{jj}}{(\lambda_j + 1)^2}}{\sigma^2 \sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2} + \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + 1)^2}} \\ &= 1 - \sigma^2 \frac{\sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2}}{\sum_{j=1}^p \frac{\alpha_j^2 + \sigma^2 m_{jj}}{(\lambda_j + 1)^2}}. \end{aligned}$$

According to the fact that $g''(d)$ is positive for all values of d , we can conclude the obtained extremum value of d minimizes the MSE function of GLSRLE. Now, for practical aspects, the following estimator can be used for the optimal d in applications:

$$\hat{d} = 1 - \hat{\sigma}_{FGLSR}^2 \left(\frac{\sum_{j=1}^p \frac{m_{jj}}{(\lambda_j + 1)^2}}{\sum_{j=1}^p \frac{\hat{\alpha}_{jFGLSR}^2 + \hat{\sigma}_{FGLSR}^2 m_{jj}}{(\lambda_j + 1)^2}} \right), \quad (3.9)$$

where $\hat{\sigma}_{FGLSR}^2$ and $\hat{\alpha}_{jFGLSR}^2$ are the unbiased estimators of σ^2 and α_j based on GLSRE, respectively, i.e., $\hat{\sigma}_{FGLSR}^2 = \frac{1}{n-(p+q)} (\mathbf{y} - \tilde{\mathbf{X}}\hat{\beta}_{FGLSR})^\top \mathbf{S}^{-1} (\mathbf{y} - \tilde{\mathbf{X}}\hat{\beta}_{FGLSR})$ and $\hat{\alpha}_{FGLSR} = \mathbf{\Gamma}^\top \hat{\beta}_{FGLSR}$, in which

$$\hat{\beta}_{FGLSR} = \hat{\beta}_{FGLS} - \mathbf{C}_F^{-1} \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F^{-1} \mathbf{R}^\top)^{-1} (\mathbf{R} \hat{\beta}_{FGLS} - \mathbf{r}). \quad (3.10)$$

As was mentioned earlier, outlier observations have the potential to significantly corrupt the least-squares estimators and all of the estimators based on it due to their significant impact on the objective function. The robust regression approach is a broad term that encompasses various estimating approaches. Least trimmed squares is a robust regression method introduced by [43]. LTS seeks to tackle this issue by minimizing the sum of the lowest h squared residuals following the removal of a specific percentage of extreme values. In this case, h serves as a threshold, and the proportion of the outlying data is represented by the ratio $\alpha = (n - h)/n$.

Typically, the value of h can be taken as $h = \lceil n(1 - \alpha) \rceil$, where $\lceil x \rceil$ stands for the ceiling of x . Some other authors suggest to take $h = \lfloor n/2 \rfloor + \lfloor (p + 1)/2 \rfloor$, $h = \lfloor n(1 - \alpha) \rfloor + \lfloor \alpha(p + 1) \rfloor$, or $h = \lfloor n(1 - \alpha) \rfloor + 1$ (see [44]). The LTS estimator is computed by solving the $\binom{n}{h}$ total least-squares fits combinations

of the index set $\{1, \dots, n\}$. Thus, for large values of sample size, finding the global minimum in the objective function of the LTS method takes time and space. To accelerate the process of finding the solution (LTS fit), we use an analogue of the FAST-LTS algorithm extended by [24].

Let z_i represent the indicator variable that signifies whether or not observation i is regarded as a normal observation. The optimization problem of a feasible robust estimator based on the LTS approach in RSRM can be developed as follows:

$$\begin{aligned} \min_{\beta, \mathbf{z}} \psi(\beta, \mathbf{z}) &= (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta)^\top \mathbf{S}^{-\frac{1}{2}} \mathbf{Z} \mathbf{S}^{-\frac{1}{2}} (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}\beta) \\ \text{s.t. } \mathbf{R}\beta &= \mathbf{r}, \\ \mathbf{e}^\top \mathbf{z} &= h, \\ z_i &\in \{0, 1\}, \quad i = 1, \dots, n. \end{aligned} \quad (3.11)$$

where \mathbf{Z} is the diagonal matrix with diagonal elements $\mathbf{z} = (z_1, \dots, z_n)^\top$, $\mathbf{e} = (1, \dots, 1)_{n \times 1}^\top$, and h is a positive integer. The resultant estimator is the feasible generalized least trimmed squares restricted estimator (FGLTSRE), which is provided by

$$\hat{\beta}_{FGLTSR}(\mathbf{z}) = \hat{\beta}_{FGLTS}(\mathbf{z}) - \mathbf{C}_F(\mathbf{z})^{-1} \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F(\mathbf{z})^{-1} \mathbf{R}^\top)^{-1} (\mathbf{R} \hat{\beta}_{FGLTS}(\mathbf{z}) - \mathbf{r}), \quad (3.12)$$

where $\mathbf{C}_F(\mathbf{z}) = \tilde{\mathbf{X}}^\top \mathbf{S}^{-1/2} \mathbf{Z} \mathbf{S}^{-1/2} \tilde{\mathbf{X}}$ and $\hat{\beta}_{FGLTS}(\mathbf{z}) = \mathbf{C}_F(\mathbf{z})^{-1} \tilde{\mathbf{X}}^\top \mathbf{S}^{-1/2} \mathbf{Z} \mathbf{S}^{-1/2} \tilde{\mathbf{y}}$.

Now, we aim to implement the robust estimators obtained previously via the Liu idea to extract the novel feasible robust Liu estimator that is resistant to the existence of multicollinearity and outliers in the data set. The feasible generalized least trimmed squares restricted Liu estimator (FGLTSRL) for RSRM using a two stages estimator for d and β can be extended as follows:

$$\hat{\sigma}_{FGLTSR}^2 = \frac{1}{n - (p + q)} (\mathbf{y} - \tilde{\mathbf{X}} \hat{\beta}_{FGLTSR}(\mathbf{z}))^\top \mathbf{S}^{-\frac{1}{2}} \mathbf{Z} \mathbf{S}^{-\frac{1}{2}} (\mathbf{y} - \tilde{\mathbf{X}} \hat{\beta}_{FGLTSR}(\mathbf{z}))$$

$$\hat{d}_{LTS} = 1 - \hat{\sigma}_{FGLTSR}^2 \left(\frac{\sum_{j=1}^p \frac{m_{jj}(\mathbf{z})}{(\lambda_j(\mathbf{z}) + 1)}}{\sum_{j=1}^p \frac{\hat{\alpha}_{jFGLTSR}^2(\mathbf{z}) + \hat{\sigma}_{FGLTSR}^2 m_{jj}(\mathbf{z})}{(\lambda_j(\mathbf{z}) + 1)^2}} \right)$$

$$\begin{aligned} \hat{\beta}_{FGLTSRL}(\hat{d}_{LTS}, \mathbf{z}) &= \\ \mathbf{F}_{\hat{d}_{LTS}}(\mathbf{z}) \hat{\beta}_{FGLTS}(\mathbf{z}) - \mathbf{F}_{\hat{d}_{LTS}}(\mathbf{z}) \mathbf{C}_F(\mathbf{z})^{-1} \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F(\mathbf{z})^{-1} \mathbf{R}^\top)^{-1} (\mathbf{R} \hat{\beta}_{FGLTS}(\mathbf{z}) - \mathbf{r}) \end{aligned} \quad (3.13)$$

where $\lambda_j(\mathbf{z})$ is the j th eigenvalue of matrix $\mathbf{C}_F(\mathbf{z}) = \mathbf{\Gamma}(\mathbf{z}) \mathbf{A}(\mathbf{z}) \mathbf{\Gamma}(\mathbf{z})^\top$, $m_{jj}(\mathbf{z})$ is the j th diagonal element of the matrix $\mathbf{M}(\mathbf{z}) = \mathbf{\Gamma}(\mathbf{z})^\top \mathbf{H}(\mathbf{z}) \mathbf{\Gamma}(\mathbf{z})$ in which $\mathbf{H}(\mathbf{z}) = \mathbf{C}(\mathbf{z})^{-1} (\mathbf{I} - \mathbf{R}^\top (\mathbf{C}(\mathbf{z})^{-1} \mathbf{R}^\top)^{-1} \mathbf{R} \mathbf{C}(\mathbf{z})^{-1})$, $\hat{\alpha}_{jFGLTSR}^2(\mathbf{z})$ is the j th element of $\hat{\alpha}_{FGLTSR}(\mathbf{z}) = \mathbf{\Gamma}^\top \hat{\beta}_{FGLTSR}(\mathbf{z})$, and $\mathbf{F}_{\hat{d}_{LTS}}(\mathbf{z}) = (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} (\mathbf{C}_F(\mathbf{z}) + \hat{d}_{LTS} \mathbf{I})$.

Theorem 3.2. The estimation of the MSE of the suggested estimator (5.2) under the linear restriction $\mathbf{R}\beta = \mathbf{r}$ can be given by

$$\begin{aligned} \text{MSE} \left(\hat{\beta}_{FGLTSRL}(\hat{d}_{LTS}, \mathbf{z}) \right) &= \hat{\sigma}_{FGLTSR}^2 \sum_{j=1}^p \frac{(\lambda_j(\mathbf{z}) + \hat{d}_{LTS})^2}{(\lambda_j(\mathbf{z}) + 1)^2} m_{jj}(\mathbf{z}) + (\hat{d}_{LTS} - \\ &1)^2 \sum_{j=1}^p \frac{\hat{\alpha}_{jFGLTSR}^2(\mathbf{z})}{(\lambda_j(\mathbf{z}) + 1)^2} + o(n^{-1}). \end{aligned} \quad (3.14)$$

Proof. The proof directly follows by mimicking the proof of Theorem 3.1. ■

Lemma 3.2. The covariance matrix of estimator $\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})$ is smaller compared to the covariance of estimator $\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})$ if and only if

$$0 < \tau < \min_{\mu_{ii} < 0} \left| \frac{1}{\mu_{ii}} \right|, \quad (3.15)$$

where $\tau = \frac{1}{1+d}$ and the μ_{ii} 's are the eigenvalues of matrix $\mathbf{H}(\mathbf{z})^{-1}\mathbf{N}(\mathbf{z})$, in which $\mathbf{N}(\mathbf{z}) = \mathbf{C}_F(\mathbf{z})\mathbf{H}(\mathbf{z}) + \mathbf{H}(\mathbf{z})\mathbf{C}_F(\mathbf{z})$.

Proof. The covariance matrix of the mentioned estimators can be written as

$$\begin{aligned} \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})\right) &= \sigma^2 \mathbf{F}_d(\mathbf{z})\mathbf{H}(\mathbf{z})\mathbf{F}_d(\mathbf{z})^\top + o(n^{-1}), \\ \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})\right) &= \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d = 1, \mathbf{z})\right) = \sigma^2 \mathbf{H}(\mathbf{z}) + o(n^{-1}). \end{aligned}$$

So, the difference $\Delta^* = \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})\right) - \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})\right)$ can be expressed as follows:

$$\begin{aligned} \Delta^* &= \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})\right) - \text{Cov}\left(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})\right) \\ &= \sigma^2 (\mathbf{H}(\mathbf{z}) - \mathbf{F}_d(\mathbf{z})\mathbf{H}(\mathbf{z})\mathbf{F}_d(\mathbf{z})^\top) \\ &= \sigma^2 \mathbf{F}_d(\mathbf{z}) \left(\mathbf{F}_d(\mathbf{z})^{-1} \mathbf{H}(\mathbf{z}) (\mathbf{F}_d(\mathbf{z})^\top)^{-1} - \mathbf{H}(\mathbf{z}) \right) \mathbf{F}_d(\mathbf{z})^\top \\ &= \sigma^2 (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \left((\mathbf{C}_F(\mathbf{z}) + \mathbf{I})\mathbf{H}(\mathbf{z})(\mathbf{C}_F(\mathbf{z}) + \mathbf{I}) - (\mathbf{C}_F(\mathbf{z}) + d\mathbf{I})\mathbf{H}(\mathbf{z})(\mathbf{C}_F(\mathbf{z}) + d\mathbf{I}) \right) (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \\ &= \sigma^2 (1 - d^2) (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \left(\mathbf{H}(\mathbf{z}) + \frac{1}{1+d} (\mathbf{C}_F(\mathbf{z})\mathbf{H}(\mathbf{z}) + \mathbf{H}(\mathbf{z})\mathbf{C}_F(\mathbf{z})) \right) (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \\ &= \sigma^2 (1 - d^2) \mathbf{U}(\mathbf{z}) (\mathbf{H}(\mathbf{z}) + \tau \mathbf{N}(\mathbf{z})) \mathbf{U}(\mathbf{z}), \end{aligned}$$

where $\mathbf{U}(\mathbf{z}) = (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1}$, $\tau = \frac{1}{1+d}$, and $\mathbf{N}(\mathbf{z}) = \mathbf{C}_F(\mathbf{z})\mathbf{H}(\mathbf{z}) + \mathbf{H}(\mathbf{z})\mathbf{C}_F(\mathbf{z})$ is a symmetric matrix. Since $\mathbf{H}(\mathbf{z}) = \mathbf{L}(\mathbf{z})^\top \mathbf{L}(\mathbf{z})$, in which

$$\mathbf{L}(\mathbf{z}) = \left(\mathbf{C}_F(\mathbf{z})^{-1} - \mathbf{C}_F(\mathbf{z})^{-1} \mathbf{R}^\top (\mathbf{R} \mathbf{C}_F^{-1} \mathbf{R}^\top)^{-1} \mathbf{R} \mathbf{C}_F(\mathbf{z})^{-1} \right)^{1/2},$$

and $\text{Rank}(\mathbf{L}(\mathbf{z})) = p - q < n$, then $\mathbf{H}(\mathbf{z})$ is a positive definite matrix. Therefore, a nonsingular matrix \mathbf{Q} exists such that $\mathbf{Q}^\top \mathbf{H}(\mathbf{z}) \mathbf{Q} = \mathbf{I}$ and $\mathbf{Q}^\top \mathbf{N}(\mathbf{z}) \mathbf{Q} = \mathbf{P}(\mathbf{z})$, where $\mathbf{P}(\mathbf{z})$ is a diagonal matrix and its diagonal elements are the roots of the polynomial equation $|\mathbf{H}(\mathbf{z})^{-1} \mathbf{N}(\mathbf{z}) - \mu \mathbf{I}| = 0$ (see Graybill [45], pp. 408; and Harville [46] pp. 563), and so we have

$$\begin{aligned} \Delta^* &= \sigma^2 (1 - d^2) \mathbf{U}(\mathbf{z}) (\mathbf{Q}^\top)^{-1} (\mathbf{Q}^\top \mathbf{H}(\mathbf{z}) \mathbf{Q} + \tau \mathbf{Q}^\top \mathbf{N}(\mathbf{z}) \mathbf{Q}) \mathbf{Q}^{-1} \mathbf{U}(\mathbf{z}) \\ &= \sigma^2 (1 - d^2) \mathbf{U}(\mathbf{z}) (\mathbf{Q}^\top)^{-1} (\mathbf{I} + \tau \mathbf{P}(\mathbf{z})) \mathbf{Q}^{-1} \mathbf{U}(\mathbf{z}), \end{aligned}$$

where $\mathbf{I} + \tau \mathbf{P}(\mathbf{z}) = \text{diag}(1 + \tau \mu_{11}, \dots, 1 + \tau \mu_{pp})$. And, since $\mathbf{N}(\mathbf{z}) = \mathbf{C}_F(\mathbf{z})\mathbf{H}(\mathbf{z}) + \mathbf{H}(\mathbf{z})\mathbf{C}_F(\mathbf{z}) \neq \mathbf{0}$, then at least one of diagonal elements of $\mathbf{P}(\mathbf{z})$ is nonzero. Assume $\mu_{ii} < 0$ for at least one i . Then, the positive definiteness of $\mathbf{I} + \tau \mathbf{P}(\mathbf{z})$ is ensured by

$$0 < \tau < \min_{\mu_{ii} < 0} \left| \frac{1}{\mu_{ii}} \right|.$$

As a result, for every $i = 1, \dots, p$, $1 + \tau \mu_{ii} > 0$, and $\mathbf{I} + \tau \mathbf{P}(\mathbf{z})$ is a positive definite matrix. Therefore, Δ^* is turned into a positive definite matrix. It is now clear that, if and only if (3.15) fulfills the criteria, $\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})$ has a smaller variance than $\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})$. ■

Next the necessary and sufficient condition is provided under which the FGLTSRLE in RSRM is preferable to the FGLTSRE in the sense of the mean squared error matrix (MSEM). The following lemma is required for the demonstration of the forthcoming theorem.

Lemma 3.3. (Farebrother [47]) Let \mathbf{A} be an $p \times p$ positive definite matrix, \mathbf{b} be an $(p \times 1)$ nonzero vector, and δ a positive scalar value. Then, $\delta \mathbf{A} - \mathbf{b} \mathbf{b}^\top$ is non-negative if and only if $\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{b} \leq \delta$.

Theorem 3.3. Let us be given the estimator $\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})$ under the linear regression model with true restrictions $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$. $\widehat{\boldsymbol{\beta}}_{FGLTSRL}(\widehat{d}_{LTS}, \mathbf{z})$ is MSEM superior to $\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})$ if and only if

$$\frac{1-d}{1+d} \boldsymbol{\beta}^\top \mathbf{G}(\mathbf{z})^{-1} \boldsymbol{\beta} \leq \sigma^2, \quad (3.16)$$

where $\mathbf{G}(\mathbf{z}) = \mathbf{H}(\mathbf{z}) + \tau \mathbf{N}(\mathbf{z})$.

Proof. We prove the necessary and sufficient conditions for the MSEM difference $\Delta = \text{MSEM}(\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})) - \text{MSEM}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z}))$, where

$$\begin{aligned} \text{MSEM}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})) &= \text{COV}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})) + \text{Bias}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})) \left(\text{Bias}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d, \mathbf{z})) \right)^\top \\ &= \sigma^2 \mathbf{F}_d(\mathbf{z}) \mathbf{H}(\mathbf{z}) \mathbf{F}_d(\mathbf{z})^\top + (\mathbf{I} - \mathbf{F}_d) \boldsymbol{\beta} \boldsymbol{\beta}^\top (\mathbf{I} - \mathbf{F}_d)^\top + o(n^{-1}), \\ \text{MSEM}(\widehat{\boldsymbol{\beta}}_{FGLTSR}(\mathbf{z})) &= \text{MSEM}(\widehat{\boldsymbol{\beta}}_{FGLTSRL}(d = 1, \mathbf{z})) = \sigma^2 \mathbf{H}(\mathbf{z}) + o(n^{-1}). \end{aligned}$$

According to the proof of Lemma 3.2., the difference Δ can be expressed as follows:

$$\begin{aligned} \Delta &= \sigma^2 \mathbf{F}_d(\mathbf{z}) \left(\mathbf{F}_d(\mathbf{z})^{-1} \mathbf{H}(\mathbf{z}) ((\mathbf{F}_d(\mathbf{z})^\top)^{-1} - \mathbf{H}(\mathbf{z})) \mathbf{F}_d(\mathbf{z})^\top \right. \\ &\quad \left. - (1-d)^2 (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \boldsymbol{\beta} \boldsymbol{\beta}^\top (\mathbf{C}_F(\mathbf{z}) + \mathbf{I})^{-1} \right) \\ &= \mathbf{U}(\mathbf{z}) \left\{ \sigma^2 (1-d^2) (\mathbf{H}(\mathbf{z}) + \tau \mathbf{N}(\mathbf{z})) - (1-d)^2 \boldsymbol{\beta} \boldsymbol{\beta}^\top \right\} \mathbf{U}(\mathbf{z}) \\ &= (1-d)^2 \mathbf{U}(\mathbf{z}) \left\{ \sigma^2 \frac{1+d}{1-d} (\mathbf{H}(\mathbf{z}) + \tau \mathbf{N}(\mathbf{z})) - \boldsymbol{\beta} \boldsymbol{\beta}^\top \right\} \mathbf{U}(\mathbf{z}) \\ &= (1-d)^2 \mathbf{U}(\mathbf{z}) \left(\sigma^2 \frac{1+d}{1-d} \mathbf{G}(\mathbf{z}) - \boldsymbol{\beta} \boldsymbol{\beta}^\top \right) \mathbf{U}(\mathbf{z}), \end{aligned}$$

where $\mathbf{G}(\mathbf{z}) = \mathbf{H}(\mathbf{z}) + \tau \mathbf{N}(\mathbf{z})$. Now, by using Lemma 3.3 and supposing that condition (3.15) is met, it is concluded that Δ is positive definite if and only if

$$\frac{1-d}{1+d} \boldsymbol{\beta}^\top \mathbf{G}(\mathbf{z})^{-1} \boldsymbol{\beta} \leq \sigma^2, \quad 0 < d < 1. \quad \blacksquare$$

Theorem 3.4. Let us be given the estimator $\widehat{\boldsymbol{\beta}}_{FGLSRL}(d)$ under the linear regression model with true restrictions $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$. $\widehat{\boldsymbol{\beta}}_{FGLSRL}(d)$ is MSEM superior to $\widehat{\boldsymbol{\beta}}_{FGLSR}$ if and only if

$$\frac{1-d}{1+d} \boldsymbol{\beta}^\top \mathbf{G}^{-1} \boldsymbol{\beta} \leq \sigma^2, \quad (3.17)$$

where $\mathbf{G} = \mathbf{H} + \tau \mathbf{N}$ and $\mathbf{N} = \mathbf{C}_F \mathbf{H} + \mathbf{H} \mathbf{C}_F$.

Proof. The desired result is simply obtained, similar to the proof of Theorem 3.4. ■

4. Illustrative experiments

To demonstrate the advantages of the improved techniques that have been proposed for the restricted semiparametric regression model in the presence of the multicollinearity and outlier problems simultaneously, we examine the theoretical findings using some numerical experiments in this section. We evaluate the performance of the proposed techniques in both a real-world data set and Monte Carlo simulation schemes.

4.1. The Monte Carlo simulation studies

We conduct a numerical analysis to evaluate the precision of our robust estimators for RSRM when dealing with contaminated data sets with outliers and multicollinearity. In each replication, the regressors are randomly generated using the following structure: Indeed, in order to reach the various levels of

multicollinearity, we used the approach proposed by [48,49], in which the explanatory variables were constructed using a device with 150 observations and 10^3 iterations, based on the model described below:

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

where z_{ij} are independent standard normal pseudo-random variables, and γ is chosen such that the correlation between any two explanatory variables is equal to γ^2 . These variables are subsequently normalized to ensure that $\mathbf{X}^T \mathbf{X}$ and $\mathbf{X}^T \mathbf{y}$ are in correlation forms. Four distinct sets of correlation values are investigated, specifically for $\gamma = 0.25, 0.50, 0.75,$ and 0.95 . For the dependent variable, n observations are then calculated by

$$y_i = \sum_{j=1}^5 x_{ij} \beta_j + f(t_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (4.1)$$

where

$$\begin{aligned} \boldsymbol{\beta} &= (-1, 4, 2, -5, -3)^T, \\ f(t) &= \exp\{\sin(t) \cos(t) + \sqrt{t}\}, t \in [0, 3], \\ \boldsymbol{\varepsilon}_{(n \times 1)} &= (\boldsymbol{\varepsilon}_1^T, \boldsymbol{\varepsilon}_2^T)^T, \end{aligned}$$

in which

$$\boldsymbol{\varepsilon}_1^T_{(h \times 1)} \sim N_h(\mathbf{0}, \sigma^2 \mathbf{V}), \sigma^2 = 1.64, [v_{ij}] = \exp\{-9|i - j|\}, h = [0.25n], [0.33n], [0.50n]$$

and

$$\boldsymbol{\varepsilon}_2^T_{((n-h) \times 1)} \sim i.i.d. \chi_1^2(15),$$

where $\chi_m^2(\delta)$ represents the m -degree of freedom non-central Chi-squared distribution with non-centrality parameter δ . The primary motivation behind selecting such structure for producing the error terms is to corrupt the data set and assess the resistance of the suggested techniques. In fact, we designated the last $n-h$ error terms as independent non-central Chi-squared distributed random variables and the first h error terms as dependent normal random variables. The non-centrality parameter leads to that the outliers lie on one side of the real regression model and bias the non-robust estimations. For the restriction, we consider the following stochastic linear restrictions:

$$\mathbf{R} = \begin{pmatrix} 1 & 5 & -3 & -1 & -1 \\ -2 & -1 & 0 & -2 & 3 \\ 1 & 2 & 1 & 3 & -2 \\ 4 & -1 & 2 & 2 & 0 \end{pmatrix}, \mathbf{r} = \mathbf{R}\boldsymbol{\beta}.$$

For estimating the nonparametric part of model (4.1), $f(\cdot)$, the weight proposed by [50] with the Gaussian kernel is used as follows:

$$W_\omega(t_j) = \frac{1}{n\omega} K\left(\frac{t_i - t_j}{\omega}\right) = \frac{1}{n\omega} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(t_i - t_j)^2}{2\omega^2}\right\}.$$

Also, the cross-validation (C.V.) approach is applied for obtaining the optimum value of bandwidth ω , which minimizes the C.V. criterion.

The non-parametric component of model (4.1) is presented in Figure 1. This wavy function is challenging to predict and offers a useful example for testing the proposed estimation techniques. All calculations were performed with R 4.3.1, the statistical software program. Tables 1–14 present a summary of the results. After iterating the process for all simulations, the minimum, maximum, mean, median, and standard deviation values of MSEs for the linear and non-linear estimators were reported in Tables 1 & 2, respectively, where

$$\text{mse}(\hat{f}_{(i)}, \mathbf{f}) = \frac{1}{nM} \sum_{m=1}^M \left\| \hat{f}_{(i)}^{(m)} - \mathbf{f} \right\|_2^2$$

$$\text{mse}(\hat{f}_{(i)}, \mathbf{f}) = \frac{1}{nM} \sum_{m=1}^M \left\| \hat{f}_{(i)}^{(m)} - \mathbf{f} \right\|_2^2, \hat{f}_{(i)} = \mathbf{K}(\mathbf{y} - \mathbf{X}\hat{\beta}_{(i)}^{(m)})$$

in which $\hat{\beta}_{(i)}^{(m)}$ and $\hat{f}_{(i)}^{(m)}$ are the i th estimators of the linear and non-linear parts ($i=1, \dots, 4$) obtained in the m th iteration for all of the four proposed approaches, and $\|\mathbf{v}\|_2^2 = \sum_{i=1}^q v_i^2$ for $\mathbf{v} = (v_1, \dots, v_p)^\top$. Also, PCDO is the percentage of the contaminated data with outliers ($\text{PCDO} = 100 \times \frac{n-h}{n} \%$).

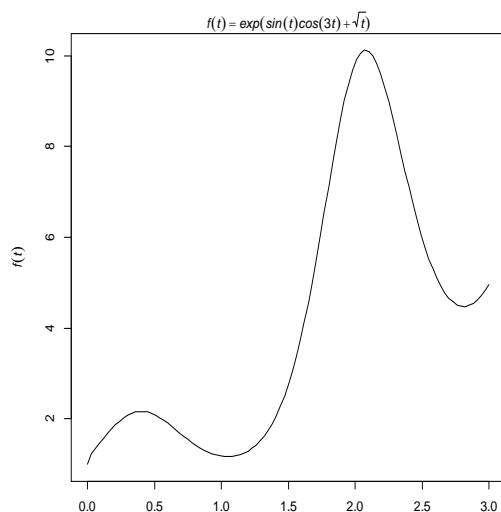


Figure 1. The nonlinear function of the simulated model.

Figure 2 shows the estimations of the non-linear part of model (4.1) using the proposed methods. In this figure, the nonparametric function is estimated by the kernel method after estimation of the linear part of model (4.1) by FGLSRE, FGLTSRE, FGLSRLE, and FGLTSRLE, respectively. To save space, the results have been only reported for $n = 150$ with PCDO=25%, 33%, and 50%, and $\gamma = 0.95$. From Figure 2, it is evident that the non-robust methods are completely corrupted by the outliers, especially for large values of PCDO.

Table 1. Mean squared error estimations of the proposed estimators for the linear part of the simulated data sets with n=150.

| | γ | 0.25 | 0.50 | 0.75 | 0.95 | 0.25 | 0.50 | 0.75 | 0.95 | 0.25 | 0.50 | 0.75 | 0.95 |
|----------|----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | | PCDO =25% | | | | PCDO =33% | | | | PCDO =50% | | | |
| FGLSRE | min | 1.03e+00 | 1.4501 | 1.9507 | 8.4008 | 1.5712 | 4.9800 | 4.4346 | 10.09635 | 2.7600 | 5.8985 | 8.4886 | 14.5415 |
| | max | 3.0344 | 3.7254 | 7.1833 | 38.9501 | 4.1060 | 5.6992 | 6.3436 | 23.1795 | 4.8609 | 12.5527 | 11.9615 | 39.6976 |
| | mean | 1.2606 | 2.2249 | 4.3570 | 16.5846 | 1.9013 | 3.9258 | 5.4962 | 19.9989 | 3.2705 | 9.3250 | 10.5542 | 26.3929 |
| | median | 1.1228 | 1.9902 | 4.1577 | 15.6974 | 1.8818 | 3.1400 | 5.2324 | 18.9837 | 3.1247 | 8.8504 | 10.2531 | 25.1407 |
| | S.D. | 1.3633 | 2.9277 | 2.5325 | 8.5307 | 2.2494 | 2.4276 | 3.7169 | 9.9713 | 6.9775 | 6.4574 | 5.7781 | 12.4557 |
| FGLTSRE | min | 7.00e-02 | 1.1141 | 1.1376 | 5.8009 | 0.0937 | 1.6500 | 2.3008 | 9.45506 | 0.0750 | 1.0046 | 2.7407 | 12.9800 |
| | max | 2.0108 | 3.1206 | 5.8441 | 17.4920 | 1.4185 | 4.0889 | 4.9014 | 18.2030 | 3.8996 | 4.8417 | 6.2896 | 28.6880 |
| | mean | 0.1579 | 1.9061 | 3.1640 | 13.6686 | 0.6468 | 3.1952 | 3.2871 | 12.0766 | 0.2389 | 2.2825 | 3.4733 | 15.1121 |
| | median | 0.0644 | 1.8409 | 3.0672 | 12.2794 | 0.4776 | 3.0808 | 3.1337 | 11.8957 | 0.1002 | 2.1237 | 3.2014 | 14.8865 |
| | S.D. | 0.6358 | 2.9805 | 2.2827 | 7.0157 | 1.1846 | 3.0649 | 3.4142 | 8.6359 | 1.3573 | 4.4232 | 3.6887 | 8.9685 |
| FGLSRLE | min | 4.10e-01 | 0.0491 | 0.0858 | 2.5456 | 1.0680 | 2.0301 | 2.7207 | 4.1117 | 2.5720 | 3.1352 | 3.2247 | 13.3250 |
| | max | 1.8972 | 2.8717 | 3.1734 | 7.6356 | 2.9482 | 4.8827 | 5.5675 | 12.1979 | 4.9784 | 7.1075 | 7.2553 | 24.1935 |
| | mean | 0.2437 | 0.6099 | 1.3926 | 3.8501 | 1.7648 | 2.9027 | 3.5436 | 6.2565 | 3.1045 | 4.3330 | 7.7663 | 17.4571 |
| | median | 0.1147 | 0.6019 | 0.1821 | 3.9388 | 0.6456 | 2.7531 | 3.2622 | 6.1038 | 3.0753 | 4.2515 | 7.2732 | 17.1550 |
| | S.D. | 0.3457 | 1.3459 | 2.5714 | 4.7190 | 2.4499 | 1.4511 | 2.7632 | 3.1329 | 6.8705 | 5.4671 | 3.7902 | 7.3125 |
| FGLTSRLE | min | 4.33e-03 | 0.0045 | 0.0037 | 0.0604 | 0.0345 | 0.0141 | 0.0109 | 0.6054 | 0.0017 | 0.1570 | 0.0048 | 1.1370 |
| | max | 1.9089 | 2.2738 | 1.1269 | 4.6501 | 0.9583 | 1.2217 | 1.3226 | 6.8717 | 1.6645 | 0.8509 | 0.8867 | 12.1068 |
| | mean | 0.1461 | 0.0992 | 0.1808 | 0.8070 | 0.9593 | 0.2282 | 0.3750 | 1.3672 | 0.2407 | 0.2941 | 0.4949 | 6.0419 |
| | median | 0.0581 | 0.0434 | 0.0745 | 0.3423 | 0.3887 | 0.0830 | 0.1660 | 0.9254 | 0.1183 | 0.3212 | 0.2139 | 5.8904 |
| | S.D. | 0.2205 | 1.1919 | 2.3091 | 3.1927 | 1.3874 | 1.2861 | 2.4600 | 2.9477 | 1.3493 | 1.9637 | 2.7041 | 3.9317 |

Table 2. Mean squared error estimations of the proposed estimators for the non-linear part of the simulated data sets with n=150.

| | γ | 0.25 | 0.50 | 0.75 | 0.95 | 0.25 | 0.50 | 0.75 | 0.95 | 0.25 | 0.50 | 0.75 | 0.95 |
|----------|----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | | PCDO =25% | | | | PCDO =33% | | | | PCDO =50% | | | |
| FGLSRE | min | 0.0374 | 0.3994 | 3.0404 | 5.0481 | 0.0449 | 0.8437 | 3.0321 | 6.0326 | 0.9591 | 1.0215 | 5.5270 | 10.11928 |
| | max | 6.2170 | 4.8515 | 5.4275 | 10.3723 | 2.7879 | 5.8039 | 4.4065 | 13.1375 | 9.6363 | 9.6595 | 7.6221 | 28.3078 |
| | mean | 0.6183 | 1.4321 | 3.4064 | 7.3234 | 0.3497 | 1.7602 | 3.5049 | 11.3692 | 1.6280 | 2.5842 | 5.5436 | 18.4228 |
| | median | 0.3508 | 1.2565 | 3.2481 | 7.2132 | 0.2345 | 1.3185 | 3.2925 | 11.2302 | 1.3582 | 2.3344 | 6.3158 | 16.2578 |
| | S.D. | 0.7293 | 1.9897 | 3.4365 | 3.3594 | 0.3291 | 2.6560 | 2.5909 | 4.4043 | 3.9471 | 5.6909 | 3.6291 | 9.4677 |
| FGLTSRE | min | 0.0287 | 0.4223 | 2.0463 | 4.0252 | 0.0343 | 0.6355 | 2.0375 | 3.0287 | 0.0215 | 0.6205 | 3.5196 | 7.0279 |
| | max | 4.0802 | 4.8972 | 4.5401 | 9.8426 | 2.5455 | 3.7042 | 2.6508 | 8.7396 | 6.3417 | 5.3711 | 5.4789 | 23.3797 |
| | mean | 0.4200 | 1.2698 | 2.2545 | 5.2060 | 0.2251 | 1.6775 | 2.3414 | 6.2519 | 0.5627 | 1.5171 | 3.4784 | 12.3705 |
| | median | 0.2451 | 1.1965 | 2.1852 | 4.9608 | 0.1736 | 1.2286 | 2.2140 | 6.1749 | 0.2998 | 1.2809 | 4.2729 | 11.2342 |
| | S.D. | 0.4864 | 2.7594 | 3.2296 | 3.1450 | 0.1909 | 2.4193 | 2.3521 | 4.2299 | 0.7026 | 1.6265 | 2.5539 | 7.3965 |
| FGLSRLE | min | 0.0281 | 0.0475 | 0.2469 | 1.0435 | 0.03421 | 0.5787 | 2.0245 | 2.9891 | 0.9282 | 0.9250 | 5.1206 | 8.0220 |
| | max | 5.9483 | 3.0696 | 3.9557 | 3.4799 | 2.5126 | 6.0783 | 2.5968 | 6.2644 | 9.0662 | 7.4303 | 6.9023 | 21.3871 |
| | mean | 0.5849 | 0.4537 | 0.9337 | 1.3566 | 0.3184 | 1.5949 | 2.5431 | 4.4030 | 1.6164 | 2.2959 | 5.1532 | 14.4144 |
| | median | 0.3256 | 0.2654 | 0.7635 | 1.1312 | 0.2174 | 1.3489 | 2.3237 | 4.2665 | 1.3497 | 2.0034 | 5.9194 | 14.2567 |
| | S.D. | 0.6946 | 0.5153 | 2.4658 | 2.3829 | 0.3070 | 2.6909 | 1.9268 | 4.4213 | 3.7313 | 2.7059 | 1.6403 | 6.4495 |
| FGLTSRLE | min | 0.0244 | 0.0450 | 0.0470 | 0.0259 | 0.0325 | 0.0369 | 0.0385 | 0.0340 | 0.0207 | 0.0253 | 0.0270 | 1.1271 |
| | max | 3.8884 | 3.0961 | 0.7341 | 1.0003 | 2.3036 | 3.9890 | 0.9417 | 1.6739 | 6.7319 | 1.9403 | 1.0525 | 5.0367 |
| | mean | 0.3968 | 0.2801 | 0.2675 | 0.7239 | 0.2173 | 0.4027 | 0.3713 | 0.5898 | 0.5475 | 0.5839 | 0.5652 | 3.3780 |
| | median | 0.2340 | 0.1993 | 0.1929 | 0.5678 | 0.1742 | 0.2428 | 0.2340 | 0.3023 | 0.3232 | 0.3432 | 0.3836 | 3.2335 |
| | S.D. | 0.4563 | 0.2756 | 2.2493 | 2.1681 | 0.1671 | 1.4512 | 1.3876 | 3.2692 | 0.6847 | 0.6430 | 1.5663 | 4.3896 |

Table 3. Evaluation of the parameters for the proposed methods with $\gamma = 0.25$ and PCDO = 25%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.0025 | -1.0022 | -1.0021 | -1.0019 |
| $\hat{\beta}_2$ | 3.9439 | 3.9519 | 3.9547 | 3.9586 |
| $\hat{\beta}_3$ | 1.8317 | 1.8558 | 1.8640 | 1.8757 |
| $\hat{\beta}_4$ | -4.8547 | -4.8754 | -4.8825 | -4.8926 |
| $\hat{\beta}_5$ | -2.9235 | -2.9344 | -2.9382 | -2.9435 |
| \hat{d} | 1.0000 | 1.0000 | 0.0002 | 0.4316 |
| $e^T z$ | 150.00 | 116.00 | 150.00 | 116.00 |

Table 4. Evaluation of the parameters for the proposed methods with $\gamma = 0.50$ and PCDO = 25%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.1020 | -1.0916 | -1.0014 | -1.0014 |
| $\hat{\beta}_2$ | 3.9059 | 3.9052 | 3.9693 | 3.9700 |
| $\hat{\beta}_3$ | 1.8176 | 1.8957 | 1.9079 | 1.9101 |
| $\hat{\beta}_4$ | -4.8057 | -4.8599 | -4.9204 | -4.9224 |
| $\hat{\beta}_5$ | -2.8398 | -2.8526 | -2.9581 | -2.9591 |
| \hat{d} | 1.000 | 1.000 | 0.0078 | 0.6775 |
| $e^T z$ | 150 | 131 | 150 | 131 |

Table 5. Evaluation of the parameters for the proposed methods with $\gamma = 0.70$ and PCDO = 25%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.7530 | -1.4123 | -1.1520 | -1.0019 |
| $\hat{\beta}_2$ | 3.1342 | 3.5500 | 3.8571 | 3.9585 |
| $\hat{\beta}_3$ | 1.1027 | 1.3500 | 1.2712 | 1.8755 |
| $\hat{\beta}_4$ | -4.2296 | -4.6705 | -4.7888 | -4.8925 |
| $\hat{\beta}_5$ | -2.1103 | -2.5318 | -2.7415 | -2.9434 |
| \hat{d} | 1.000 | 1.000 | 0.0054 | 0.6639 |
| $e^T z$ | 150 | 128 | 150 | 128 |

Table 6. Evaluation of the parameters for the proposed methods with $\gamma = 0.95$ and PCDO = 25%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -2.5080 | -2.0052 | -1.0036 | -1.0035 |
| $\hat{\beta}_2$ | 2.8241 | 3.0851 | 3.9207 | 3.9228 |
| $\hat{\beta}_3$ | 0.4724 | 1.1554 | 1.7622 | 1.7684 |
| $\hat{\beta}_4$ | -3.7103 | -4.1024 | -4.7946 | -4.8000 |
| $\hat{\beta}_5$ | -2.0002 | -2.0134 | -2.8919 | -2.8947 |
| \hat{d} | 1.000 | 1.000 | 0.6488 | 0.2912 |
| $e^T z$ | 150 | 122 | 122 | 108.392 |

Table 7. Evaluation of the parameters for the proposed methods with $\gamma = 0.25$ and PCDO = 33%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.0035 | -1.0034 | -1.0079 | -1.0059 |
| $\hat{\beta}_2$ | 3.9222 | 3.9244 | 3.8259 | 3.8707 |
| $\hat{\beta}_3$ | 1.7667 | 1.7732 | 1.4776 | 1.6122 |
| $\hat{\beta}_4$ | -4.7985 | -4.8042 | -4.5488 | -4.6651 |
| $\hat{\beta}_5$ | -2.8939 | -2.8969 | -2.7625 | -2.8237 |
| \hat{d} | 1.0000 | 1.0000 | 0.0067 | 0.6464 |
| $e^T z$ | 150 | 129 | 150 | 129 |

Table 8. Evaluation of the parameters for the proposed methods with $\gamma = 0.50$ and PCDO = 33%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.1029 | -1.1025 | -1.1023 | -1.0021 |
| $\hat{\beta}_2$ | 3.9065 | 3.9150 | 3.9199 | 3.9533 |
| $\hat{\beta}_3$ | 1.8096 | 1.8109 | 1.8298 | 1.8599 |
| $\hat{\beta}_4$ | -4.6355 | -4.7574 | -4.7703 | -4.8790 |
| $\hat{\beta}_5$ | -2.5134 | -2.7250 | -2.8017 | -2.9363 |
| \hat{d} | 1.0000 | 1.0000 | 0.0002 | 0.4261 |
| $e^T z$ | 150 | 130 | 150 | 130 |

Table 9. Evaluation of the parameters for the proposed methods with $\gamma = 0.75$ and PCDO = 33%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.9438 | -1.4133 | -1.6228 | -1.0026 |
| $\hat{\beta}_2$ | 2.9157 | 3.6282 | 3.5385 | 3.9429 |
| $\hat{\beta}_3$ | 1.0472 | 1.3847 | 1.3156 | 1.8287 |
| $\hat{\beta}_4$ | -3.7816 | -4.7140 | -4.6407 | -4.8521 |
| $\hat{\beta}_5$ | -2.0051 | -2.6021 | -2.5162 | -2.9221 |
| \hat{d} | 1.000 | 1.000 | 0.0001 | 0.4026 |
| $e^T z$ | 150 | 130 | 150 | 130 |

Table 10. Evaluation of the parameters for the proposed methods with $\gamma = 0.95$ and PCDO = 33%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -2.8086 | -2.0067 | -1.7142 | -1.0038 |
| $\hat{\beta}_2$ | 2.2198 | 3.0528 | 3.5466 | 3.9168 |
| $\hat{\beta}_3$ | 0.4293 | 1.2685 | 1.1197 | 1.7505 |
| $\hat{\beta}_4$ | -3.5071 | -4.0187 | -4.3579 | -4.7845 |
| $\hat{\beta}_5$ | -1.7406 | -2.1003 | -2.5726 | -2.8866 |
| \hat{d} | 1.000 | 1.000 | 0.0008 | 0.3654 |
| $e^T z$ | 150 | 118 | 150 | 118 |

Table 11. Evaluation of the parameters for the proposed methods with $\gamma = 0.25$ and PCDO = 50%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.8514 | -1.0015 | -1.8709 | -1.0010 |
| $\hat{\beta}_2$ | 3.9029 | 3.9678 | 3.8807 | 3.9790 |
| $\hat{\beta}_3$ | 1.6097 | 1.9034 | 1.7422 | 1.9371 |
| $\hat{\beta}_4$ | -4.1220 | -4.9166 | -4.2501 | -4.9457 |
| $\hat{\beta}_5$ | -2.4590 | -2.9561 | -2.4737 | -2.9714 |
| \hat{d} | 1.000 | 1.000 | 0.0013 | 0.0217 |
| $e^T z$ | 150 | 127 | 150 | 127 |

Table 12. Evaluation of the parameters for the proposed methods with $\gamma = 0.50$ and PCDO = 50%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.9411 | -1.3117 | -1.6436 | -1.0011 |
| $\hat{\beta}_2$ | 3.1654 | 3.4623 | 3.2789 | 3.9763 |
| $\hat{\beta}_3$ | 0.8961 | 1.5868 | 1.0367 | 1.9289 |
| $\hat{\beta}_4$ | -3.7103 | -4.7022 | -4.0454 | -4.9386 |
| $\hat{\beta}_5$ | -2.0528 | -2.3485 | -2.1512 | -2.9677 |
| \hat{d} | 1.000 | 1.000 | 0.0250 | 0.0203 |
| $e^T z$ | 150 | 123 | 150 | 123 |

Table 13. Evaluation of the parameters for the proposed methods with $\gamma = 0.75$ and PCDO = 50%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -1.0021 | -1.0023 | -1.0011 | -1.0012 |
| $\hat{\beta}_2$ | 3.9535 | 3.9485 | 3.9767 | 3.9726 |
| $\hat{\beta}_3$ | 1.8606 | 1.8456 | 1.9300 | 1.9178 |
| $\hat{\beta}_4$ | -4.8796 | -4.8667 | -4.9395 | -4.9290 |
| $\hat{\beta}_5$ | -2.9366 | -2.9298 | -2.9682 | -2.9626 |
| \hat{d} | 1.000 | 1.000 | 0.0018 | 0.0168 |
| $e^T z$ | 150 | 125 | 150 | 125 |

Table 14. Evaluation of the parameters for the proposed methods with $\gamma = 0.95$ and PCDO = 50%.

| Coefficients | Method | | | |
|-----------------|---------|---------|---------|----------|
| | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
| $\hat{\beta}_1$ | -3.1056 | -2.1161 | -2.6811 | -1.0016 |
| $\hat{\beta}_2$ | 3.0766 | 3.1668 | 3.0995 | 3.9690 |
| $\hat{\beta}_3$ | 0.3197 | 1.1703 | 1.0297 | 1.9072 |
| $\hat{\beta}_4$ | -2.6802 | -3.6548 | -3.1793 | -4.9165 |
| $\hat{\beta}_5$ | -1.1117 | -2.0683 | -1.9681 | -2.9567 |
| \hat{d} | 1.000 | 1.000 | 0.0003 | 0.0205 |
| $e^T z$ | 150 | 127 | 150 | 127 |

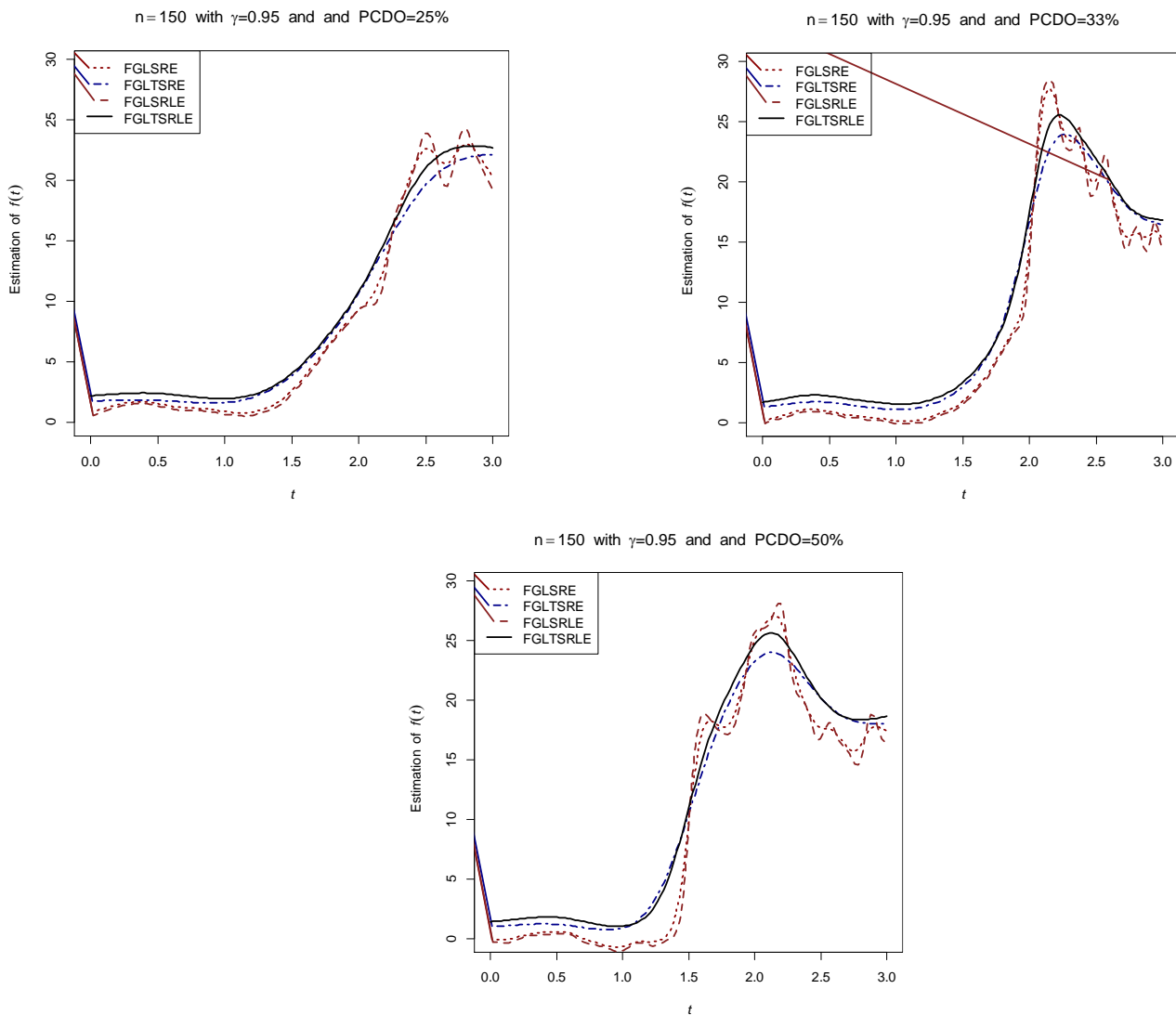


Figure 2. Estimation of the nonparametric function under study by kernel method for $n = 150$, $\gamma = 0.95$, PCDO = 25% (low), PCDO=33% (moderate), and PCDO=50% (high).

4.2. Real-world data analysis

We analyze the hedonic pricing of housing features in order to motivate the challenge of linearly restricted estimations in the semiparametric regression model. Lot size has a big impact on housing costs. Ho [51] fit this data set using semiparametric least squares. The information is based on 92 detached houses that were sold in the Ottawa region in 1987. Here is how the variables are defined: The sale price (SP) is the dependent variable, while the lot size (lot area = LT), square footage of housing (SFH), average neighborhood income (ANI), distance to the highway (DHW), garage availability (GAR), and fireplace (FP) are the independent variables. At first, the pure parametric model is fit as follows:

$$(SP)_i = \beta_0 + \beta_1(LT)_i + \beta_2(SFH)_i + \beta_3(FP)_i + \beta_4(DHW)_i + \beta_5(GAR)_i + \beta_6(ANI)_i + \varepsilon_i.$$

We use the added-variable charts to intuitively determine the parametric and nonparametric components of the model (see Sheather [52] for more details). Added-variable plots allow us to

examine each predictor's influence graphically after adjusting for the effects of the other explanatory variables. Based on the analysis of the added-variable plot (Figure 3), we identify ANI as a nonparametric component. Therefore, the SRM is specified accordingly:

$$(SP)_i = \beta_0 + \beta_1(LT)_i + \beta_2(SFH)_i + \beta_3(FP)_i + \beta_4(DHW)_i + \beta_5(GAR)_i + f(ANI)_i + \varepsilon_i. \quad (4.2)$$

The “mctest” package in R is used to detect multicollinearity in the design matrix, producing the following results. The Farrar-Glauber test and other pertinent tests for multicollinearity are provided. Overall Multicollinearity Diagnostics

| | MC Results | detection |
|------------------------|------------|-----------|
| Determinant X'X : | 0.005618 | 1 |
| Farrar Chi-Square: | 50.8378 | 1 |
| Red Indicator: | 0.2065 | 0 |
| Sum of Lambda Inverse: | 700.2104 | 1 |
| Theil's Method: | -0.7320 | 0 |
| Condition Number: | 200.4021 | 1 |

1 --> COLLINEARITY is detected by the test

0 --> COLLINEARITY is not detected by the test

The correlation graphs are displayed in Figure 4 for the real data set. It is evident from the output above and Figure 4 that the independent variables in the real data set under investigation exhibit substantial multicollinearity. So, to address the multicollinearity issue, the suggested estimating techniques must be used.

The restriction $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ may be identified as follows based on a basic investigation of the semiparametric regression model (4.2) using a robust Liu estimator:

$$\mathbf{R} = \begin{pmatrix} -1 & 0 & -1 & -1 & 1 \\ 1 & 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & -2 & 8 \end{pmatrix}, \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Now, the linear hypothesis $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ is examined in the framework of the restricted semiparametric regression model (4.2). The test statistic is computed as follows under $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$:

$$\chi_{rank(R)}^2 = (\mathbf{R}\hat{\boldsymbol{\beta}}_{FGLS} - \mathbf{r})^\top (\mathbf{R}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}} \mathbf{R}^\top)^{-1} (\mathbf{R}\hat{\boldsymbol{\beta}}_{FGLS} - \mathbf{r}) = 0.4781,$$

where $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}} = \hat{s}^2 (\tilde{\mathbf{X}}^\top \mathbf{S}^{-1} \tilde{\mathbf{X}})^{-1}$, in which $\hat{s}^2 = \frac{1}{n-p} (\mathbf{y} - \tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}_{FGLS})^\top \mathbf{S}^{-1} (\mathbf{y} - \tilde{\mathbf{X}}\hat{\boldsymbol{\beta}}_{FGLS})$. Consequently,

the restricted estimators are obtained. Table 15 shows a brief evaluation of the proposed estimators.

In this table, the values of \widehat{MSE} and R^2 are calculated, in which $R^2 = 1 - \frac{RSS}{S_{YY}}$ is the coefficient of

determination of the model, where $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ is the residual sum of squares and $\hat{y}_i = x_i \hat{\boldsymbol{\beta}} + \hat{f}(t_i)$. Compared to the other procedures, FGLTSRLE seems to be accurately effective based on the results that were obtained.

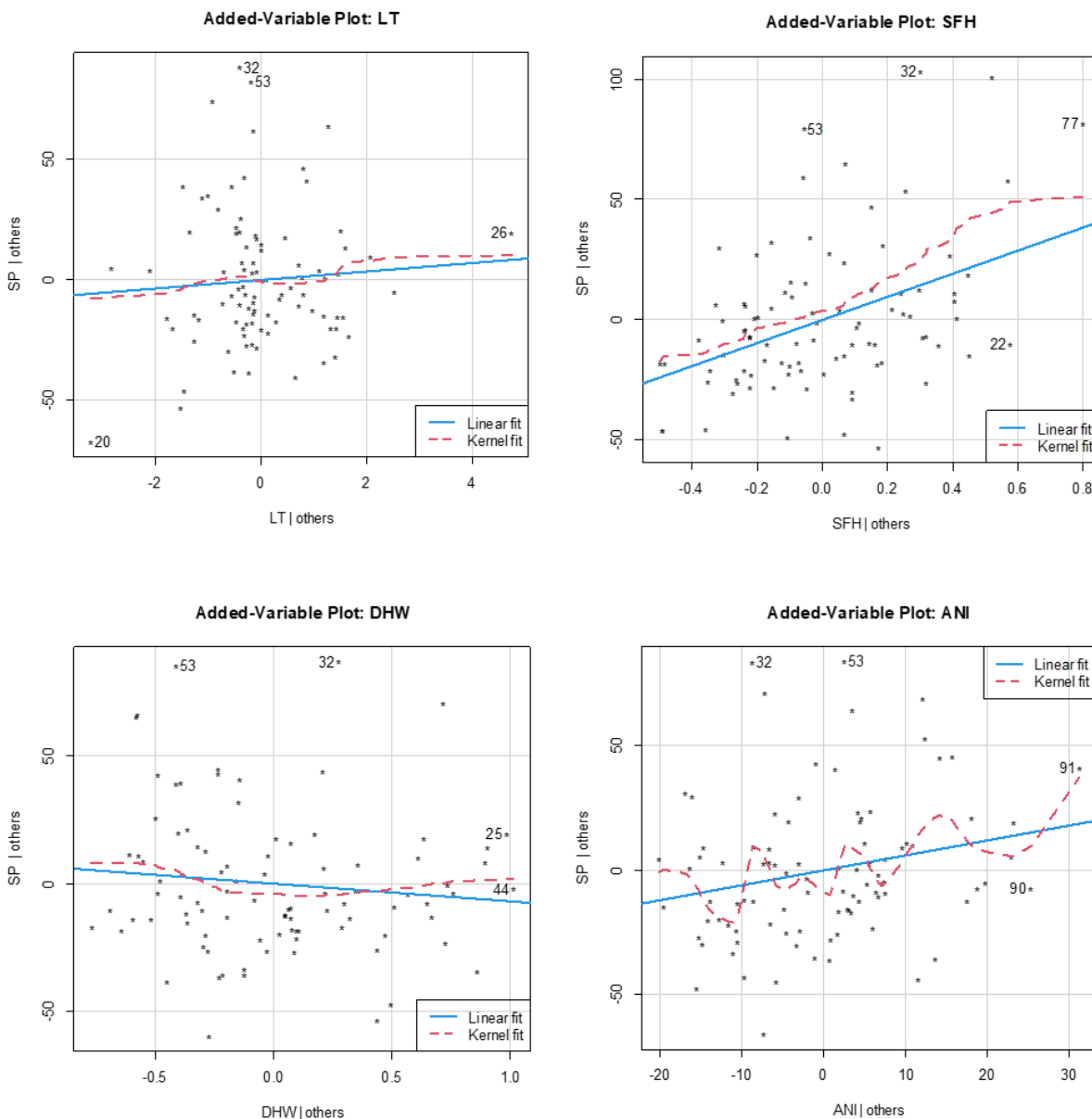


Figure 3. Added-variable plots of individual explanatory variables vs. dependent variable, linear fit (blue solid line), and kernel fit (red dashed line).

Table 15. Evaluation of parameters for proposed estimators for the real data set method.

| Coefficients | FGLSRE | FGLTSRE | FGLSRLE | FGLTSRLE |
|--------------|---------|---------|---------|---------------|
| LT | 0.7018 | 1.0509 | 0.8514 | 1.1235 |
| SFH | 46.7515 | 33.5686 | 38.9154 | 26.3747 |
| FP | 3.9311 | 2.5740 | 3.5210 | 1.9568 |
| DHW | -1.6147 | -0.7616 | -0.9952 | -0.4125 |
| GAR | 6.2476 | 4.3865 | 5.2015 | 2.9958 |
| $e^T z$ | 92.0000 | 86.0000 | 92.0000 | 86.0000 |
| \hat{d} | 1.0000 | 1.0000 | 0.1542 | 0.6741 |
| MSE | 926.80 | 456.41 | 809.59 | 335.17 |
| R^2 | 0.2346 | 0.6156 | 0.3354 | 0.7325 |

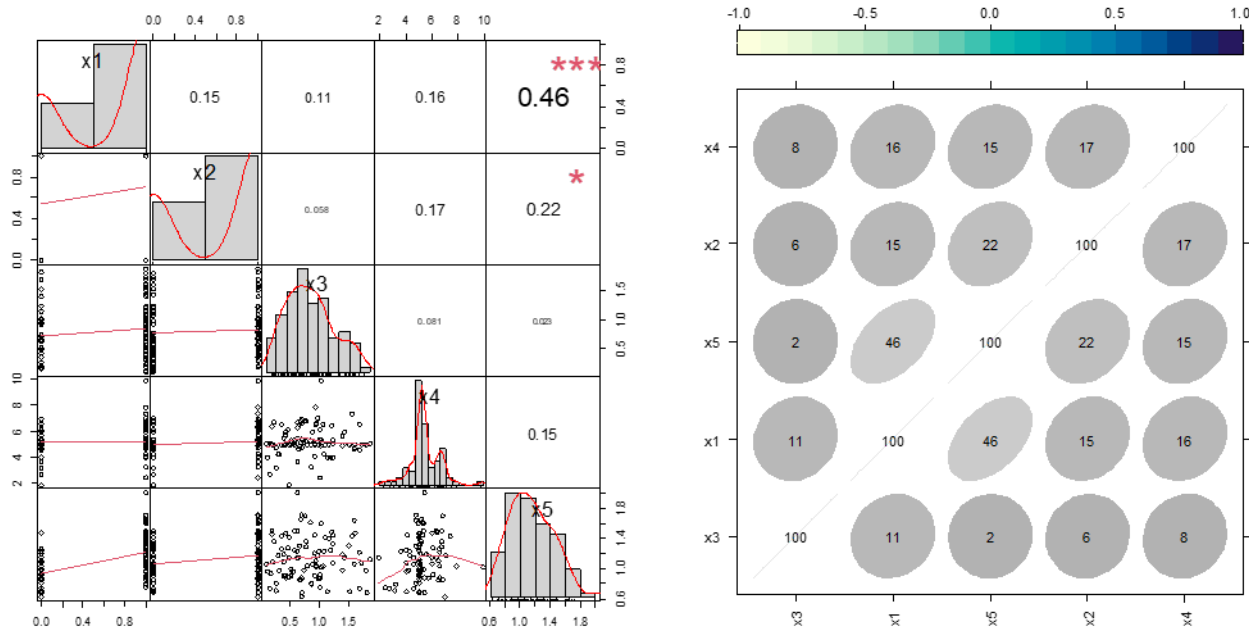


Figure 4. Visualization of the correlation plots for the explanatory variables of the real data set.

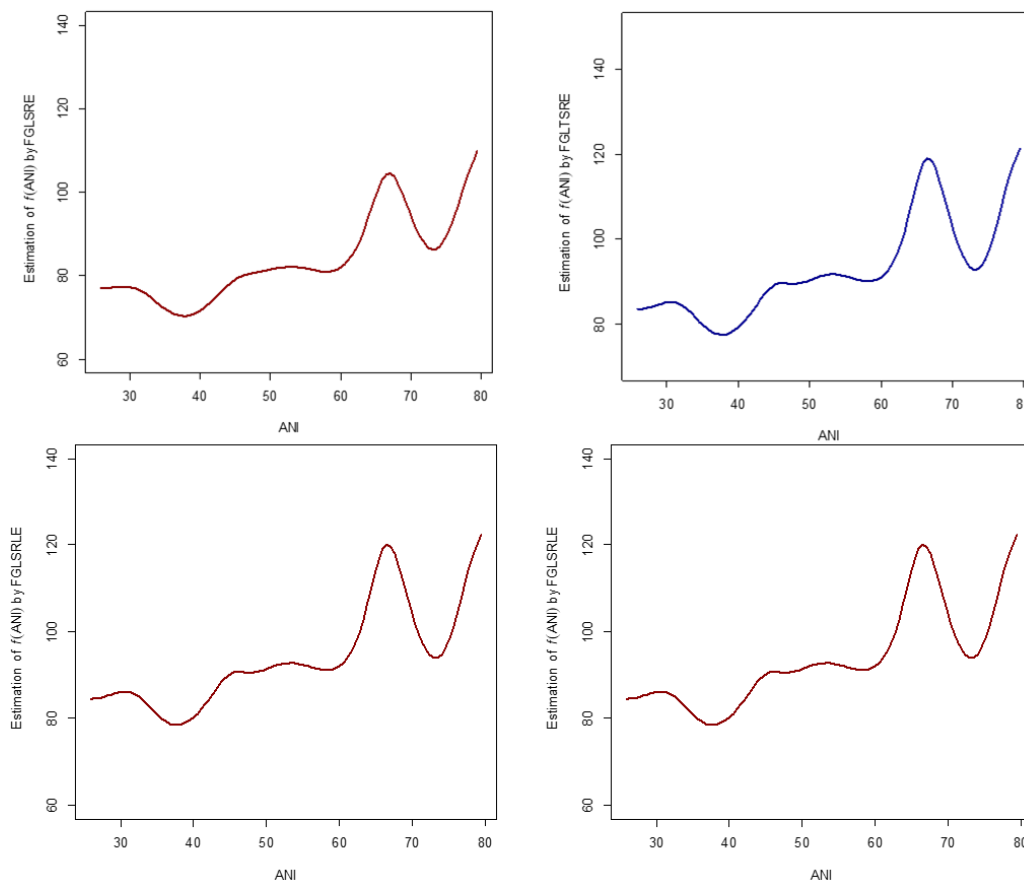


Figure 5. Estimations for the nonparametric part of model (4.2).

Following the proposed estimations of the linear component of the model (4.2), the estimations on the non-parametric function by kernel smoothing is shown in Figure 5. For estimation of the nonparametric effect, at first we estimated the parametric effects by one of the proposed methods, and then the kernel approach was applied to fit $SP_i - x_i^T \hat{\beta}$ on $ANI_i, i = 1, \dots, n$ for all proposed linear estimators, where $x_i^T = (LT_i, SFH_i, FP_i, DHW_i, GAR_i)$. Table 15 and Figure 5 demonstrate how the Liu type of robust and non-robust estimators outperform non-Liu forms in both parametric and nonparametric estimations due to the presence of multicollinearity in the design matrix. Furthermore, robust estimators outperform non-robust estimators in model prediction since the data set contains some outlier observations.

5. Conclusions

In this research, Liu and non-Liu types of the feasible generalized restricted robust estimator are suggested in a semiparametric regression model when some additional linear constraints held on the linear parameter space and the variance matrix of the error terms were unknown. We introduced robust Liu estimators in the presence of multicollinearity among column vectors of the design matrix of a semiparametric regression model and outliers in the data set. We also introduced some new estimators of d by minimizing the mean squared error of the proposed estimators. After extracting the MSEM superiority condition of a feasible generalized least trimmed squares restricted Liu estimator over a non-Liu type based on some theorems, comprehensive Monte-Carlo simulation experiments and a real data analysis were conducted to evaluate the effectiveness of the suggested estimators. The numerical experiments illustrated that the suggested methods can be effectively implemented to predict the dependent variable of restricted SRMs without being affected by the corruptive impact of multicollinearity or outlier issues. As a good topic for future research, it is proposed to derive the asymptotic distribution of the proposed estimator by the interested authors (see [53,54] for more details).

Author contributions

W. B. Altukhaes: Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing-original draft preparation, Writing-review and editing, Visualization, Funding acquisition; M. Roozbeh: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing-original draft preparation, Writing-review and editing, Visualization, Project administration; N. A. Mohamed: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Visualization, Supervision, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The first author would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work. The second author thank the Research Council of Semnan University for its support. The second and third authors would like to thank the Ministry of Higher Education Malaysia for their support in funding this research through the Fundamental Research Grant Scheme (Project No.: FP072-2023) awarded to Nur Anisah Mohamed and Mahdi Roozbeh.

The authors would like to thank the three anonymous reviewers and handling editor for their valuable comments and corrections to an earlier version of this paper, which significantly improved the quality of our work.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. P. Green, C. Jennison, A. Seheult, Analysis of field experiments by least squares smoothing, *J. Roy. Statist. Soc. Ser. B*, **47** (1985), 299–315. <https://doi.org/10.1111/j.2517-6161.1985.tb01358.x>
2. R.-F. Engle, C.-W.-J. Granger, J. Rice, A. Weiss, Semiparametric estimates of the relation between weather and electricity sales, *J. Am. Stat. Assoc.*, **81** (1986), 310–320. <https://doi.org/10.2307/2289218>
3. R.-L. Eubank, E.-L. Kambour, J.-T. Kim, K. Klipple, C.-S. Reese, M. Schimek, Estimation in partially linear models, *Comput. Stat. Data. Anal.*, **29** (1998), 27–34. [https://doi.org/10.1016/S0167-9473\(98\)00054-1](https://doi.org/10.1016/S0167-9473(98)00054-1)
4. P. Speckman, Kernel smoothing in partial linear models, *J. R. Stat. Soc. Ser. B*, **50** (1988), 413–436. <https://doi.org/10.1111/j.2517-6161.1988.tb01738.x>
5. R.-L. Eubank, *Nonparametric Regression and Spline Smoothing*, New York: Marcel Dekker, 1999. <https://doi.org/10.1201/9781482273144>
6. D. Ruppert, M.-P. Wand, R.-C. Carroll, *Semiparametric Regression*; Cambridge: Cambridge University Press, 2003. <https://doi.org/10.1017/CBO9780511755453>
7. W. Härdle, M. Müller, S. Sperlich, A. Werwatz, *Nonparametric and Semiparametric Models*, Berlin/Heidelberg: Springer, 2004. <https://doi.org/10.1007/978-3-642-17146-8>
8. Yatchew, *Semiparametric Regression for the Applied Econometrician*, Cambridge: Cambridge University Press, 2003. <https://doi.org/10.1017/CBO9780511615887>
9. F. Akdeniz, G. Tabakan, Restricted ridge estimators of the parameters in semiparametric regression model, *Commun. Stat. Theory Methods*, **38** (2009), 1852–1869. <https://doi.org/10.1080/03610920802470109>
10. D.-E. Akdeniz, W.-K. Hardle, M. Osipenko, Difference based ridge and Liu type estimators in semiparametric regression models, *J. Multivar. Anal.*, **105** (2012), 164–175. <https://doi.org/10.1016/j.jmva.2011.08.018>
11. M. Arashi, T. Valizadeh, Performance of Kibria’s methods in partial linear ridge regression model, *Stat. Pap.*, **56** (2015), 231–246. <https://doi.org/10.1007/s00362-014-0578-6>
12. F. Akdeniz, S. Kaçiranlar, On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE, *Commun. Stat. Theory Methods*, **24** (1995), 1789–1797. <https://doi.org/10.1080/03610929508831585>
13. M.-N. Akram, B.-M.-G. Kibria, M. Arashi, A.-F. Lukman, A new improved Liu estimator for the QSAR model with inverse Gaussian response, *Commun. Stat. Simul. Comput.*, **53** (2024), 1873–1888. <https://doi.org/10.1080/03610918.2022.2059088>

14. S. Kaçıranlar, N. Ozbay, E. Ozkan, H. Guler, Comparison of Liu and two parameter principal component estimator to combat multicollinearity, *Concurr. Comput. Pract. Exp.*, **34** (2022), e6737. <https://doi.org/10.1002/cpe.6737>
15. B. Kan, O. Alpu, B. Yazici, Robust ridge and robust Liu estimator for regression based on the LTS estimator, *J. Appl. Stat.*, **40** (2013), 644–655. <https://doi.org/10.1080/02664763.2012.750285>
16. K.-J. Liu, A new class of biased estimate in linear regression, *Commun. Stat. Theory Methods*, **22** (1993), 393–402. <https://doi.org/10.1080/03610929308831027>
17. M. Arashi, A.-F. Lukman, Z.-Y. Algamal, Liu regression after random forest for prediction and modeling in high dimension, *J. Chemom.*, **36** (2022), e3393. <https://doi.org/10.1002/cem.3393>
18. M. Arashi, M. Roozbeh, H.-A. Niroumand, A note on Stein-type shrinkage estimator in partial linear models, *Statistics*, **46** (2012), 673–685. <https://doi.org/10.1080/02331888.2011.553682>
19. M. Roozbeh, S. Babaie-Kafaki, M. Manavi, A heuristic algorithm to combat outliers and multicollinearity in regression model analysis, *Iran. J. Num. Anal. Opt.*, **12** (2022), 173–186. <https://doi.org/10.22067/IJNAO.2021.68160.1008>
20. M. Roozbeh, A. Rouhi, N.-A. Mohamed, F. Jahadi, Generalized support vector regression and symmetry functional regression approaches to model the high-dimensional data, *Symmetry*, **15** (2023), 1262. <https://doi.org/10.3390/sym15061262>
21. M. Roozbeh, S. Babaie-Kafaki, Z. Aminifard, A nonlinear mixed–integer programming approach for variable selection in linear regression model, *Commun. Stat. Simul. Comput.*, **11** (2023), 5434–5445. <https://doi.org/10.1080/03610918.2021.1990323>
22. P.-J. Rousseeuw, Least median of squares regression, *J. Am. Stat. Assoc.*, **79** (1984), 871–880. <https://doi.org/10.2307/2288718>
23. P.-J. Rousseeuw, A.-M. Leroy, *Robust Regression and Outlier Detection*, New York: John Wiley, 1987. <https://doi.org/10.1002/0471725382>
24. P.-J. Rousseeuw, K. van Driessen, Computing LTS regression for large data sets, *Data Min. Knowl. Discov.*, **12** (2006), 29–45. <https://doi.org/10.1007/s10618-005-0024-4>
25. M. Amini, M. Roozbeh, Optimal partial ridge estimation in restricted semiparametric regression models, *J. Multivar. Anal.*, **136** (2015), 26–40. <https://doi.org/10.1016/j.jmva.2015.01.005>
26. A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *J. Am. Stat. Assoc.*, **57** (1962), 348–368. <https://doi.org/10.2307/2281644>
27. M. Arashi, B.-M.-G. Kibria, T. Valizadeh, On ridge parameter estimators under stochastic subspace hypothesis, *J. Stat. Comput. Simul.*, **87** (2017), 966–983. <https://doi.org/10.1080/00949655.2016.1239104>
28. M.-H. Karbalaee, M. Arashi, S.-M.-M. Tabatabaey, Performance analysis of the preliminary test estimator with series of stochastic restrictions, *Commun. Stat. Theory Methods*, **47** (2018), 1–17. <https://doi.org/10.1080/03610926.2017.1300275>
29. M. Roozbeh, G. Hesamian, M.-G. Akbari, Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors, *J. Comput. Appl. Math.*, **378** (2020), 112940. <https://doi.org/10.1016/j.cam.2020.112940>
30. J. Durbin, A note on regression when there is extraneous information about one of the coefficients, *J. Am. Stat. Assoc.*, **48** (1990), 799–808. <https://doi.org/10.2307/2281073>
31. H. Theil, A.-S. Goldberger, On pure and mixed statistical estimation in economics, *Int. Econ. Rev.*, **2** (1961), 65–78. <https://doi.org/10.2307/2525589>

32. H. Theil, On the use of incomplete prior information in regression analysis, *J. Am. Stat. Assoc.*, **58** (1963), 401–411. <https://doi.org/10.2307/2283275>
33. R. Fallah, M. Arashi, S.-M.-M. Tabatabaey, On the ridge regression estimator with sub-space restriction, *Commun. Stat. Theory Methods*, **46** (2017), 11854–11865. <https://doi.org/10.1080/03610926.2017.1285928>
34. R. Fallah, M. Arashi, S.-M.-M. Tabatabaey, Shrinkage estimation in restricted elliptical regression model, *J. Iran. Stat. Soc.*, **17** (2018), 49–61. <https://doi.org/10.29252/jirss.17.1.49>
35. H. Toutenburg, *Prior Information in Linear Models*, New York: John Wiley, 1982. <https://doi.org/10.2307/2982032>
36. A.-E. Hoerl, R.-W. Kennard, Ridge regression: Biased estimation for non-orthogonal problems, *Technometrics*, **12** (1970), 69–82. <https://doi.org/10.2307/1271436>
37. M. Roozbeh, N.-A. Hamzah, Feasible robust estimator in restricted semiparametric regression models based on the LTS approach, *Commun. Stat. Simul. Comput.*, **46** (2017), 7332–7350. <https://doi.org/10.1080/03610918.2016.1236954>
38. F. Akdeniz, M. Roozbeh, Generalized difference-based weighted mixed almost unbiased ridge estimator in partially linear models, *Stat. Pap.* **60** (2019), 1717–1739. <https://doi.org/10.1007/s00362-017-0893-9>
39. F. Akdeniz, M. Roozbeh, E. Akdeniz, M.-N. Khan, Generalized difference-based weighted mixed almost unbiased liu estimator in semiparametric regression models, *Commun. Stat. Theory Methods*, **51** (2022), 4395–4416. <https://doi.org/10.1080/03610926.2020.1814340>
40. 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. <https://doi.org/10.1080/00949655.2010.519705>
41. K. Månsson, B.-M.-G. Kibria, G. Shukur, A restricted Liu estimator for binary regression models and its application to an applied demand system, *J. Appl. Stat.*, **43** (2016), 1119–1127. <https://doi.org/10.1080/02664763.2015.1092110>
42. K. Månsson, B.-M.-G. Kibria, Estimating the unrestricted and restricted Liu estimators for the Poisson regression model: Method and application, *Comput. Econ.*, **58** (2021), 311–326. <https://doi.org/10.1007/s10614-020-10028-y>
43. P.-J. Rousseeuw, Multivariate estimation with high breakdown point, *Math. Stat. Appl.*, **8** (1985), 283–297.
44. Alfons, C. Croux, S. Gelper, Sparse least trimmed squares regression for analyzing high-dimensional large data sets, *Ann. Appl. Stat.*, **7** (2013), 226–248. <https://doi.org/10.1214/12-AOAS575>
45. F.-A. Graybill, *Matrices with Applications in Statistics*, Wadsworth: Belmont, 1983.
46. D. Harville, *Matrix Algebra from a Statistician's Perspective*, New York: Springer Verlag, 1997. <https://doi.org/10.1007/b98818>
47. R. Farebrother, Further results on the mean square error of ridge regression, *J. Roy. Stat. Soc. Ser. B*, **38** (1976), 248–250. <https://doi.org/10.1111/j.2517-6161.1976.tb01588.x>
48. G.-C. McDonald, D.-I. Galarneau, A Monte Carlo evaluation of some ridge-type estimators, *J. Am. Stat. Assoc.*, **70** (1975), 407–416. <https://doi.org/10.2307/2285832>
49. D.-G. Gibbons, A simulation study of some ridge estimators, *J. Am. Stat. Assoc.*, **76** (1981), 131–139. <https://doi.org/10.2307/2287058>

50. M.-B. Priestley, M.-T. Chao, Non-Parametric Function Fitting, *J. R. Stat. Soc. Ser. B*, **34** (1972), 385–392. <https://doi.org/10.1111/j.2517-6161.1972.tb00916.x>
51. M. Ho, *Essays on the Housing Market*, PhD thesis, University of Toronto, 1995.
52. S.-J. Sheather, *A Modern Approach to Regression with R*, New York: Springer, 2009. <https://doi.org/10.1007/978-0-387-09608-7>
53. P. Cizek, Least trimmed squares in nonlinear regression under dependence, *J. Stat. Plann. Inference*, **136** (2005), 3967–3988. <https://doi.org/10.1016/j.jspi.2005.05.004>
54. J.-A. Visek, The least trimmed squares part III: Asymptotic normality, *Kybernetika*, **42** (2006), 203–224.



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)