Robust Liu Estimator Used to Combat Some Challenges in Partially Linear Regression Model by Improving LTS Algorithm Using Semidefinite Programming

Waleed B. Altukhaes [1,2], Mahdi Roozbeh [3,*] and Nur A. Mohamed [1,*]

1 Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, Kuala Lumpur 50603, Malaysia; waltokhaes@su.edu.sa
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 3513119111, Iran
* Correspondence: mahdi.roozbeh@semnan.ac.ir (M.R.); nuranisah_mohamed@um.edu.my (N.A.M.)

Abstract: Outliers are a common problem in applied statistics, together with multicollinearity. In this paper, robust Liu estimators are introduced into a partially linear model to combat the presence of multicollinearity and outlier challenges when the error terms are not independent and some linear constraints are assumed to hold in the parameter space. The Liu estimator is used to address the multicollinearity, while robust methods are used to handle the outlier problem. In the literature on the Liu methodology, obtaining the best value for the biased parameter plays an important role in model prediction and is still an unsolved problem. In this regard, some robust estimators of the biased parameter are proposed based on the least trimmed squares (LTS) technique and its extensions using a semidefinite programming approach. Based on a set of observations with a sample size of n, and the integer trimming parameter $h \leq n$, the LTS estimator computes the hyperplane that minimizes the sum of the lowest h squared residuals. Even though the LTS estimator is statistically more effective than the widely used least median squares (LMS) estimate, it is less complicated computationally than LMS. It is shown that the proposed robust extended Liu estimators perform better than classical estimators. As part of our proposal, using Monte Carlo simulation schemes and a real data example, the performance of robust Liu estimators is compared with that of classical ones in restricted partially linear models.

Keywords: generalized Liu estimator; least trimmed squares estimator; linear restriction; multicollinearity; outlier; partially linear regression models

MSC: 62G08; 62G35; 62J05; 62J07

1. Introduction

When the nature of the relationship between the response variable and some of the explanatory variables is unclear but the link function of the mean of the dependent variable is expected to have a linear parametric relationship to certain other explanatory variables, partially linear models (PLMs) are the suitable models to apply for predicting or modeling the data set. Consider the set of observations denoted as $(y_1, x_1^\top, t_1), \ldots, (y_n, x_n^\top, t_n)$, which conform to the partially linear model defined by

$$y_i = x_i^\top \beta + f(t_i) + \varepsilon_i, \qquad i = 1, 2, \ldots, n, \tag{1}$$

where y_i is the value of the response variable of the ith observation, $x_i^\top = (x_{i1}, \ldots, x_{ip})$ represents a vector of the explanatory variables, $\beta = (\beta_1, \ldots, \beta_p)^\top$ denotes a vector of the unknown parameters, and the observed points that match the boundaries of the domain

$D \subset \mathbb{R}$ are denoted by t_i. It is generally assumed that the unknown function $f(.)$ is a smooth function, while ε_i represents random errors that are considered to be independent of (x_i, t_i).

PLMs that combine both parametric and nonparametric components are more flexible than a conventional multiple regression model in cases where the assumption of a linear relationship between the dependent variable and some of its predictors (x) is made, while its relationship to the other explanatory variables (t) has an unknown non-linear shape [1,2]. There have been several approaches to estimating the parameters in a PLM. Among the most important approaches are those given by several researchers [1–11].

The presence of nearly linear dependency among the columns of the design matrix $X = (x_1, \ldots, x_n)^\top$ is known as multicollinearity, and it is a common issue that might arise in a regression analysis. In this case, the matrix $S = X^\top X$ contains one or more small eigenvalues, causing the regression coefficient estimations to be large in their absolute value. The condition number is an effective measure for recognizing the presence of multicollinearity. The matrix 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, the existence of multicollinearity in the data may cause the confidence intervals to be too large for either 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. In this study, the shrinkage estimator suggested by Liu [12] is applied to solve the problem of multicollinearity. Liu [12] combined the Stein estimator with the conventional ordinary ridge regression estimator to derive the Liu estimator, as described in [13–15]. Other alternative approaches to addressing the issue of multicollinearity can be found in the research papers [16–19].

Besides the multicollinearity problem, another typical issue that arises in regression analyses is the existence 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 the PLM used 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 an 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 [20]. Although the LMS estimator has been the subject of most publications on robust estimation in the field of linear models, Rousseeuw and Leroy [21] have noted that the LMS is not the optimal option due to its statistical features. They asserted that selecting the least trimmed squares is the better option because both the LTS and LMS have the same breakdown point, approximately 50%, but the LTS offers some advantages in comparison to the LMS. Compared to the LMS, the objective function of the LTS is smoother. Also, since the LTS converges more quickly and is normally distributed asymptotically [20], it has superior statistical efficiency. For these reasons, the LTS is a better choice as a starting point for two-step robust estimators such as the MM estimator [22].

The main innovation of this paper is that it proposes novel robust Liu estimators for the parameters of a restricted PLM and robust estimations for the biasing parameter of Liu estimations based on the LTS approach and its extensions to solve multicollinearity and outlier problems simultaneously. These proposed estimators are based on an improved LTS produced using a semidefinite programming approach. The organization of this article is as follows: Section 2 contains the classical estimator of a restricted partially linear model based on the kernel method. After reviewing the concepts of Liu and the least trimmed squares approach to a restricted partially linear model in Sections 3 and 4, respectively, our

new robust Liu estimators for a restricted partially linear model are suggested based on the semidefinite programming in Section 5, and their asymptotic biases and distributional covariances are then derived. In Section 6, the efficiencies of the proposed estimators are assessed through vast Monte Carlo simulation experiments and a real-world data example. Lastly, some important findings are concluded in Section 7.

2. The Classical Estimators under Restriction

The estimators that conform to certain restrictions are called classical estimators. Let us examine the following partially linear model:

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

where $y = (y_1, \ldots, y_n)^\top$, $X = (x_1, \ldots, x_n)^\top$, $f(t) = (f(t_1), \ldots, f(t_n))^\top$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\top$.

Generally, we assume that ϵ is a vector of the disturbances that follows the distribution of $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon^\top) = \sigma^2 V$, where σ^2 is an unknown parameter and V is a known matrix that is symmetric and positive definite.

To estimate the linear part of the model (2), we first remove the non-parametric effect by detrending. Given the assumption that β is known, a natural non-parametric estimator of $f(.)$ is

$$\widehat{f}(t) = k(t)(y - X\beta),$$

where $k(.)$ is a kernel function. Following [4], by substituting $\widehat{f}(t)$ for $f(t)$ in Equation (2), the model may be reduced to

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

where $\widetilde{y} = (I_n - K)y$, $\widetilde{X} = (I_n - K)X$, and K is the smoother matrix with the (i, j)-th component $K_\omega(t_i, t_j)$, in which $K_\omega(\cdot)$ is a kernel function of order m with the bandwidth parameter ω. Now, the estimation of β is performed using the generalized least-squares estimator (GLSE), which is known to be the best unbiased linear estimator

$$\begin{aligned}
\widehat{\beta}_{GLS} &= argmin_\beta (\widetilde{y} - \widetilde{X}\beta)^\top V^{-1}(\widetilde{y} - \widetilde{X}\beta) \\
&= C^{-1}\widetilde{X}^\top V^{-1}\widetilde{y},
\end{aligned} \tag{4}$$

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

Interestingly, another suitable method for handling strong and extremely strong multicollinearity problems is to obtain the estimators under particular restrictions on unknown parameters, which may be exact or stochastic (see [23–29] for more details). Assume that we had prior knowledge regarding β in the sense of its non-stochastic exact constraints, as follows:

$$R\beta = r$$

where R is a known matrix $q \times p$ of prior information of rank $q < p$ and 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 group of linear constraints non-stochastically represented by independent prior information, we provide the instruments necessary to compute the risk of the estimators. Next, the performances of the new constrained estimators and classical estimators may be compared under certain conditions. We show that our innovative constrained estimators outperform the classical ones in terms of the least risk functions, assuming linear restrictions. In these circumstances, certain non-sample information (a previous constraint on the parameters) may exist; it is often presented to the model as constraints. Compared to typical estimators, the restricted estimator performs better, and so, in this research work, the restricted partially linear model (RPLM) is fitted to the data set. The selection of the complete row's rank assumption is based on convenience and is backed by the fact that every consistent linear equation may be transformed into an equivalent

equation with a coefficient matrix that has a full row rank. The generalized least-squares restricted estimator (GLSRE) is derived by imposing a linear restriction, as follows:

$$\widehat{\boldsymbol{\beta}}_{GLSR} = argmin_{\beta} \left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta}\right)^{\top} \boldsymbol{V}^{-1} \left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta}\right)$$
$$\text{s.t. } \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r} \tag{5}$$
$$= \widehat{\boldsymbol{\beta}}_{GLS} - \boldsymbol{C}^{-1}\boldsymbol{R}^{\top} \left(\boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\right)^{-1} \left(\boldsymbol{R}\widehat{\boldsymbol{\beta}}_{GLS} - \boldsymbol{r}\right).$$

As such, it is known that the covariance matrix of $\widehat{\boldsymbol{\beta}}_{GLS}$ is equal to $\sigma^2 \boldsymbol{C}^{-1}$. Thus, the GLSE and its covariance matrix are significantly influenced by the features of matrix $\boldsymbol{C}$. The GLS estimators become susceptible to various errors when $\boldsymbol{C}$ 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 the generation of large confidence intervals for 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 an RPLM with a multicollinearity problem.

3. Restricted Liu Estimator in a Partially Linear Model

As was mentioned, multicollinearity leads to $\boldsymbol{X}^{\top}\boldsymbol{X}$ being ill conditioned, with a large condition number. When the condition number of $\boldsymbol{X}^{\top}\boldsymbol{X}$ is large, the least-squares estimator is more severely affected by multicollinearity. In this case, the high level of data noise is magnified by $\left(\boldsymbol{X}^{\top}\boldsymbol{X}\right)^{-1}$, making the least-squares estimator highly unreliable. To combat this drawback of the least-squares estimator, Hoerl and Kennard [30] proposed adding a ridge estimator $\widehat{\boldsymbol{\beta}}_k = \left(\boldsymbol{X}^{\top}\boldsymbol{X} + k\boldsymbol{I}\right)^{-1} \boldsymbol{X}^{\top}\boldsymbol{y}$ to the standard linear regression model $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, with $E(\boldsymbol{\varepsilon}) = 0$ and $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\top}) = \sigma^2\boldsymbol{I}$, and it has become the most often used method for solving the multicollinearity problem that causes the least-squares estimator to fail. Indeed, the ridge method solves the multicollinearity problem by adding a small constant k to the diagonal of $\boldsymbol{X}^{\top}\boldsymbol{X}$ to decrease its condition number. In practical use, the shrinkage parameter k in the ridge approach is often rather modest. It is obvious that the condition number of $\boldsymbol{X}^{\top}\boldsymbol{X} + k\boldsymbol{I}$ is a decreasing function of k. Thus, high values of k are needed to achieve small-scale control over the condition number of $\boldsymbol{X}^{\top}\boldsymbol{X} + k\boldsymbol{I}$. Because of this, the small k selected in practice may not be big enough to solve the severe ill-conditioning problem of $\boldsymbol{X}^{\top}\boldsymbol{X}$. As such, the resultant ridge estimation may still be unstable since $\boldsymbol{X}^{\top}\boldsymbol{X} + k\boldsymbol{I}$ has remained ill conditioned. Furthermore, despite its practical effectiveness, the ridge estimator is a complicated function of k. Although the Stein-type estimator, $\widehat{\boldsymbol{\beta}}_c = c\left(\boldsymbol{X}^{\top}\boldsymbol{X}\right)^{-1} \boldsymbol{X}^{\top}\boldsymbol{y}$, is a linear function of c, the shrinkage of each element of $\widehat{\boldsymbol{\beta}}_c$ is the same. To address these issues, Liu [12] proposed a new biased estimator $\widehat{\boldsymbol{\beta}}_d = \left(\boldsymbol{X}^{\top}\boldsymbol{X} + \boldsymbol{I}\right)^{-1} \left(\boldsymbol{X}^{\top}\boldsymbol{y} + d\widehat{\boldsymbol{\beta}}\right)$ by combining the advantages of the ridge and Stein-type estimators, which effectively solved the problem of ill conditioning in the standard regression model, where $0 < d < 1$ is the biasing parameter and $\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^{\top}\boldsymbol{X}\right)^{-1} \boldsymbol{X}^{\top}\boldsymbol{y}$. It is obvious that when $d = 1$, $\widehat{\boldsymbol{\beta}}_d = \widehat{\boldsymbol{\beta}}$.

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

$$\text{MSE}\left(\widehat{\boldsymbol{\beta}}_d\right) = \sigma^2 \sum_{j=1}^{p} \frac{\left(\lambda_j + d\right)^2}{\lambda_j \left(\lambda_j + 1\right)^2} + (d-1)^2 \sum_{j=1}^{p} \frac{\alpha_j^2}{\left(\lambda_j + 1\right)^2} \tag{6}$$

where α_j^2 corresponds to the jth element of $\boldsymbol{\alpha} = \boldsymbol{\Gamma}^{\top}\boldsymbol{\beta}$ and $\boldsymbol{\Gamma}$ is an orthogonal matrix such that $\boldsymbol{C} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^{\top}$, in which $\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda_p)$ contains the eigenvalues of matrix $\boldsymbol{C}$. Consequently, the biasing parameter d is chosen by minimizing the MSE of $\left(\widehat{\boldsymbol{\beta}}_d\right)$ as follows:

$$\widehat{d} = 1 - \widehat{\sigma}_{GLS}^2 \left(\frac{\sum_{j=1}^{p} \frac{1}{\lambda_j(\lambda_j+1)}}{\sum_{j=1}^{p} \frac{\widehat{\alpha}_{jGLS}^2}{\left(\lambda_j+1\right)^2}} \right) \tag{7}$$

where $\widehat{\sigma}^2$ and $\widehat{\alpha}_{jGLS}^2$ are the unbiased estimators of σ^2 and α_j based on the GLSE, respectively, i.e., $\widehat{\sigma}_{GLS}^2 = \frac{1}{n-p}\left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{GLS}\right)^{\top} \boldsymbol{V}^{-1}\left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{GLS}\right)$ and $\widehat{\boldsymbol{\alpha}}_{GLS} = \boldsymbol{\Gamma}^{\top}\widehat{\boldsymbol{\beta}}_{GLS}$.

The generalized least-squares Liu estimator (GLSLE) investigated by [31] is defined as follows:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{GLSL}(d) &= \left(\widetilde{\boldsymbol{X}}^{\top}\boldsymbol{V}^{-1}\widetilde{\boldsymbol{X}} + \boldsymbol{I}\right)^{-1}\left(\widetilde{\boldsymbol{X}}^{\top}\boldsymbol{V}^{-1}\widetilde{\boldsymbol{X}} + d\boldsymbol{I}\right)\widehat{\boldsymbol{\beta}}_{GLS} \\
&= (\boldsymbol{C} + \boldsymbol{I})^{-1}(\boldsymbol{C} + d\boldsymbol{I})\widehat{\boldsymbol{\beta}}_{GLS} \\
&= \boldsymbol{F}_d\widehat{\boldsymbol{\beta}}_{GLS}, \; 0 \le d \le 1,
\end{aligned}
\tag{8}
$$

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

Based on the fact that $\boldsymbol{F}_d$ and $\boldsymbol{C}^{-1}$ are commutative, the generalized least-squares restricted Liu estimator (GLSRLE) can be defined as follows for an RPLM [32–34]:

$$
\widehat{\boldsymbol{\beta}}_{GLSRL}(d) = \boldsymbol{F}_d\widehat{\boldsymbol{\beta}}_{GLS} - \boldsymbol{F}_d\boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\left(\boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\right)^{-1}\left(\boldsymbol{R}\widehat{\boldsymbol{\beta}}_{GLS} - \boldsymbol{r}\right).
\tag{9}
$$

Lemma 1. *If $\boldsymbol{\beta}$ satisfies the linear restriction $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$, then the bias vector, covariance matrix, and mean squared error functions of proposed estimator can be evaluated by direct calculations, as follows:*

$$
\mathrm{Bias}\left(\widehat{\boldsymbol{\beta}}_{GLSRL}(d)\right) = -(\boldsymbol{I} - \boldsymbol{F}_d)\boldsymbol{\beta},
\tag{10}
$$

$$
\mathrm{Cov}\left(\widehat{\boldsymbol{\beta}}_{GLSRL}(d)\right) = \sigma^2 tr\left(\boldsymbol{F}_d\boldsymbol{H}\boldsymbol{F}_d^{\top}\right),
\tag{11}
$$

$$
\mathrm{MSE}\left(\widehat{\boldsymbol{\beta}}_{GLSRL}(d)\right) = \sigma^2 tr\left(\boldsymbol{F}_d\boldsymbol{H}\boldsymbol{F}_d^{\top}\right) + \boldsymbol{\beta}^{\top}(\boldsymbol{F}_d - \boldsymbol{I})^{\top}(\boldsymbol{F}_d - \boldsymbol{I})\boldsymbol{\beta},
\tag{12}
$$

where $\boldsymbol{H} = \boldsymbol{C}^{-1}\left(\boldsymbol{I} - \boldsymbol{R}^{\top}\left(\boldsymbol{R}\boldsymbol{C}^{-1}\boldsymbol{R}^{\top}\right)^{-1}\boldsymbol{R}\boldsymbol{C}^{-1}\right).$

Theorem 1. *The mean squared error of the GLSRLE under the linear restriction $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{r}$ can be given by*

$$
MSE\left(\widehat{\boldsymbol{\beta}}_{GLSRL}(d)\right) = \sigma^2\sum_{j=1}^{p}\frac{(\lambda_j + d)^2}{(\lambda_j + 1)^2}m_{jj} + (d-1)^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j + 1)^2},
\tag{13}
$$

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

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

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

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

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

So, the proof is completed. □

As an important result of Theorem 1, the optimal value of the biasing parameter d can be obtained by differentiating the mean squared error function of the GLSRLE with respect to d as follows:

$$\hat{d} = 1 - \hat{\sigma}^2_{GLSR} \left(\frac{\sum_{j=1}^{p} \frac{m_{jj}}{(\lambda_j + 1)}}{\sum_{j=1}^{p} \frac{\hat{\alpha}^2_{jGLSR} + \hat{\sigma}^2_{GLSR} m_{jj}}{(\lambda_j + 1)^2}} \right), \tag{14}$$

where $\hat{\sigma}^2_{GLSR}$ and $\hat{\alpha}^2_{jGLSR}$ are the unbiased estimators of σ^2 and α_j based on the GLSRE, respectively, i.e., $\hat{\sigma}^2_{GLSR} = \frac{1}{n-(p-q)} \left(\widetilde{y} - \widetilde{X} \hat{\beta}_{GLSR} \right)^{\top} V^{-1} \left(\widetilde{y} - \widetilde{X} \hat{\beta}_{GLSR} \right)$ and $\hat{\alpha}_{GLSR} = \Gamma^{\top} \hat{\beta}_{GLSR}$, for which

$$\hat{\beta}_{GLSR} = \hat{\beta}_{GLS} - C^{-1} R^{\top} \left(R C^{-1} R^{\top} \right)^{-1} \left(R \hat{\beta}_{GLS} - r \right). \tag{15}$$

4. Extension of the Least Trimmed Squares (LTS) Estimator Using Semidefinite Programming in an RPLM

As is known, outliers have the potential to significantly corrupt the least-squares estimator and all of the estimators based on it due to their significant impact on the objective function. A robust regression approach is a broad term that encompasses various estimating approaches. The least trimmed squares is a robust regression method introduced by [35]. The LTS seeks to combat 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 are represented by the ratio $\alpha = (n - h)/n$. Typically, the value of h can be taken as $h = [[n(1 - \alpha)]]$, where $[[x]]$ stands for the ceiling of x. Some other authors suggest taking $h = [n/2] + [(p + 1)/2]$, $h = [n(1 - \alpha)] + [\alpha(p + 1)]$ or $h = [n(1 - \alpha)] + 1$ [36]. The LTS estimator is computed by solving the $\binom{n}{h}$ total least-squares fit combinations of the index set $\{1, \ldots, n\}$. Thus, for large sample sizes, 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 analog of the FAST-LTS algorithm extended by Rousseeuw and van Driessen [22].

Let z_i represent the indicator variable that signifies whether or not observation i is regarded as a normal observation. The objective function of the LTS in an RPLM will be examined as follows:

$$min_{\beta,z} \psi(\beta, z) = \left(\widetilde{y} - \widetilde{X} \beta \right)^{\top} V^{-\frac{1}{2}} Z V^{-\frac{1}{2}} \left(\widetilde{y} - \widetilde{X} \beta \right)$$
$$s, t, \quad R\beta = r,$$
$$e^{\top} z = h,$$
$$z_i \in \{0.1\}. \ i = 1 \ldots n. \tag{16}$$

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

$$\hat{\beta}_{GLTSR}(z) = \hat{\beta}_{GLTS}(z) - C(z)^{-1} R^{\top} \left(R C(z)^{-1} R^{\top} \right)^{-1} \left(R \hat{\beta}_{GLTS}(z) - r \right), \tag{17}$$

where $C(z) = \widetilde{X}^{\top} V^{-1/2} Z V^{-1/2} \widetilde{X}$ and $\hat{\beta}_{GLTS}(z) = C(z)^{-1} \widetilde{X}^{\top} V^{-1/2} Z V^{-1/2} \widetilde{y}$.

According to the research conducted in [37,38], we may define a relaxation problem in the RPLM called a relaxed least trimmed squares (RLTS) problem as follows:

$$min_{\beta,z^*}\psi(\beta,z^*) = \left(\widetilde{y} - \widetilde{X}\beta\right)^\top V^{-\frac{1}{2}}Z^*V^{-\frac{1}{2}}\left(\widetilde{y} - \widetilde{X}\beta\right)$$

$$s,t, \quad R\beta = r,$$
$$e^\top z^* = h,$$
$$0 \le z_i^* \le 1. \quad i = 1.\dots.n, \tag{18}$$

where Z^* is the diagonal matrix with diagonal elements $z^* = \left(z_1^*.\dots.,z_n^*\right)^\top$. The resultant estimator of the above optimization problem is the generalized relaxed least trimmed squares-restricted estimator (GRLTSRE), which is given as follows:

$$\widehat{\beta}_{GRLTSR}(z^*) = \widehat{\beta}_{GRLTS}(z^*) - C(z^*)^{-1}R^\top\left(RC(z^*)^{-1}R^\top\right)^{-1}\left(R\widehat{\beta}_{GRLTS}(z^*) - r\right) \tag{19}$$

where $C(z^*) = \widetilde{X}^\top V^{-1/2}Z^*V^{-1/2}\widetilde{X}$ and $\widehat{\beta}_{GRLTS}(z^*) = C(z^*)^{-1}\widetilde{X}^\top V^{-1/2}Z^*V^{-1/2}\widetilde{y}$.

Here, we propose an extension of the RLTS problem in RSRM, called ERLTS, based on the optimization of the following objective function:

$$min_{\beta,z^{**}}\psi(\beta,z^{**}) = \left(\widetilde{y} - \widetilde{X}\beta\right)^\top V^{-\frac{1}{2}}Z^{**}V^{-\frac{1}{2}}\left(\widetilde{y} - \widetilde{X}\beta\right)$$

$$s,t, \quad R\beta = r,$$
$$h_1 \le e^\top z^{**} \le h_2,$$
$$0 \le z_i^{**} \le 1. \quad i = 1,\dots.n. \tag{20}$$

where Z^{**} is the diagonal matrix with diagonal elements $z^{**} = \left(z_1^{**}.\dots.,z_n^{**}\right)^\top$ and the positive integers h_1 and h_2 are such that $h_1 \le h \le h_2$. The generalized extended relaxed least trimmed squares-restricted estimator (GERLTSRE) is the solution to this optimization problem, obtained by semidefinite programming as follows:

$$\widehat{\beta}_{GERLTSR}(z^{**}) = \widehat{\beta}_{GERLTS}(z^{**}) - C(z^{**})^{-1}R^\top\left(RC(z^{**})^{-1}R^\top\right)^{-1}\left(R\widehat{\beta}_{GERLTS}(z^{**}) - r\right), \tag{21}$$

where $C(z^{**}) = \widetilde{X}^\top V^{-1/2}Z^{**}V^{-1/2}\widetilde{X}$ and $\widehat{\beta}_{GERLTS}(z^{**}) = C(z^{**})^{-1}\widetilde{X}^\top V^{-1/2}Z^{**}V^{-1/2}\widetilde{y}$.

5. Extended LTS Liu Estimator in an RPLM

In this section, we try to implement the three types of LTS estimators introduced in the previous section, based on Liu's idea, to extract novel robust Liu estimators that are resistant to the existence of multicollinearity and outliers in the data set. The robust form of the Liu estimator based on the classical LTS estimator, considered by Kan et al. [39], can be extended as follows:

$$\widehat{\beta}_{LTSL}(d,z) = \left(X^\top ZX + I\right)^{-1}\left(X^\top ZX + dI\right)\widehat{\beta}_{LTS}(z). \tag{22}$$

Now, we adopt and utilize this estimator within the proposed robust estimators previously defined for the RPLM to obtain two-stage estimators for the parameters, as follows:

1. Robust estimators for d and β based on the LTS approach to the RPLM (and subsequently the generalized least trimmed squares-restricted Liu, GLTSRL, method):

$$\widehat{\sigma}_{GLTSR}^2 = \frac{1}{n-(p-q)}\left(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GLTSR}(z)\right)^\top V^{-1/2}ZV^{-1/2}\left(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GLTSR}(z)\right)$$

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

$$\widehat{\beta}_{GLTSRL}(\hat{d}_{LTS}.z) =$$
$$F_{\hat{d}_{LTS}}(z)\,\widehat{\beta}_{GLTS}(z) - F_{\hat{d}_{LTS}}(z)C(z)^{-1}R^\top\big(RC(z)^{-1}R^\top\big)^{-1}\big(R\widehat{\beta}_{GLTS}(z) - r\big) \tag{23}$$

where $\lambda_j(z)$ is the jth eigenvalue of matrix $C(z) = \Gamma(z)\Lambda(z)\Gamma(z)^\top$ and $m_{jj}(z)$ is the jth diagonal element of the matrix $M(z) = \Gamma(z)^\top H(z)\Gamma(z)$, in which $H(z) = C(z)^{-1}\big(I - R^\top(C(z)^{-1}R^\top)^{-1}RC(z)^{-1}\big)$, $\widehat{\alpha}^2_{jGLTSR}(z)$ is the jth element of $\widehat{\alpha}_{GLTSR}(z) = \Gamma^\top\widehat{\beta}_{GLTSR}(z)$ and $F_{\hat{d}_{LTS}}(z) = (C(z) + I)^{-1}(C(z) + \hat{d}_{LTS}I)$.

2. Robust estimators for d and β based on the RLTS approach to the RSRM (and subsequently the generalized relaxed least trimmed squares-restricted Liu, GRLTSRL, method):

$$\widehat{\sigma}^2_{GRLTSR} = \frac{1}{n - (p - q)}\big(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GRLTSR}(z^*)\big)^\top V^{-1/2}Z^*V^{-1/2}\big(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GRLTSR}(z^*)\big)$$

$$\hat{d}_{RLTS} = 1 - \widehat{\sigma}^2_{GRLTSR}\left(\frac{\sum_{j=1}^p \frac{m_{jj}(z^*)}{(\lambda_j(z^*)+1)}}{\sum_{j=1}^p \frac{\widehat{\alpha}^2_{jRGLTSR}(z^*)+\widehat{\sigma}^2_{GRLTSR}m_{jj}(z^*)}{(\lambda_j(z^*)+1)^2}}\right)$$

$$\widehat{\beta}_{GRLTSRL}(\hat{d}_{RLTS}, z^*) =$$
$$F_{\hat{d}_{RLTS}}(z^*)\,\widehat{\beta}_{GRLTS}(z^*) - F_{\hat{d}_{RLTS}}(z^*)C(z^*)^{-1}R^\top\big(RC(z^*)^{-1}R^\top\big)^{-1}\big(R\widehat{\beta}_{GRLTS}(z^*) - r\big) \tag{24}$$

where $\lambda_j(z^*)$ is the jth eigenvalue of matrix $C(z^*) = \Gamma(z^*)\Lambda(z^*)\Gamma(z^*)^\top$ and $m_{jj}(z^*)$ is the jth diagonal element of the matrix $M(z^*) = \Gamma(z^*)^\top H(z^*)\Gamma(z^*)$, in which $H(z^*) = C(z^*)^{-1}\big(I - R^\top(C(z^*)^{-1}R^\top)^{-1}RC(z^*)^{-1}\big)$, $\widehat{\alpha}^2_{jGRLTSR}(z^*)$ is the jth element of $\widehat{\alpha}_{GRLTSR}(z^*) = \Gamma^\top\widehat{\beta}_{GRLTSR}(z^*)$ and $F_{\hat{d}_{RLTS}}(z^*) = (C(z^*) + I)^{-1}(C(z^*) + \hat{d}_{RLTS}I)$.

3. Robust estimators for d and β based on the ERLTS approach to the RSRM (and subsequently the generalized extended relaxed least trimmed squares-restricted Liu, GERLTSRL, method):

$$\widehat{\sigma}^2_{GERLTSR} = \frac{1}{n - (p - q)}\big(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GERLTSR}(z^{**})\big)^\top V^{-1/2}Z^{**}V^{-1/2}\big(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GERLTSR}(z^{**})\big)$$

$$\hat{d}_{ERLTS} = 1 - \widehat{\sigma}^2_{GERLTSR}\left(\frac{\sum_{j=1}^p \frac{m_{jj}(z^{**})}{(\lambda_j(z^{**})+1)}}{\sum_{j=1}^p \frac{\widehat{\alpha}^2_{jGERLTSR}(z^{**})+\widehat{\sigma}^2_{GERLTSR}m_{jj}(z^{**})}{(\lambda_j(z^{**})+1)^2}}\right)$$

$$\widehat{\beta}_{GLTSRL}(\hat{d}_{ERLTS}, z^{**}) =$$
$$F_{\hat{d}_{ERLTS}}(z^{**})\,\widehat{\beta}_{GERLTS}(z^{**}) - F_{\hat{d}_{ERLTS}}(z^{**})C(z^{**})^{-1}R^\top\big(RC(z^{**})^{-1}R^\top\big)^{-1}\big(R\widehat{\beta}_{GERLTS}(z^{**}) - r\big) \tag{25}$$

where $\lambda_j(z^{**})$ is the jth eigenvalue of matrix $C(z^{**}) = \Gamma(z^{**})\Lambda(z^{**})\Gamma(z^{**})^\top$ and $m_{jj}(z^{**})$ is the jth diagonal element of the matrix $M(z^{**}) = \Gamma(z^{**})^\top H(z^{**})\Gamma(z^{**})$, in which $H(z^{**}) = C(z^{**})^{-1}\big(I - R^\top(C(z^{**})^{-1}R^\top)^{-1}RC(z^{**})^{-1}\big)$, $\widehat{\alpha}^2_{jGERLTSR}(z^{**})$ is the jth element of $\widehat{\alpha}_{GERLTSR}(z^{**}) = \Gamma^\top\widehat{\beta}_{GERLTSR}(z^{**})$ and $F_{\hat{d}_{ERLTS}}(z^{**}) = (C(z^{**}) + I)^{-1}(C(z^{**}) + \hat{d}_{ERLTS}I)$.

Theorem 2. *The mean squared error of the proposed estimators (23)–(25) under the linear restriction $R\beta = r$ can be given by*

$$\widehat{\text{MSE}}\left(\widehat{\boldsymbol{\beta}}_{GLTSRL}(\hat{d}_{LTS}.\boldsymbol{z})\right) = \widehat{\sigma}^2_{GLTSR}\sum\nolimits_{j=1}^{p} \frac{\left(\lambda_j(\boldsymbol{z}) + \hat{d}_{LTS}\right)^2}{\left(\lambda_j(\boldsymbol{z}) + 1\right)^2} m_{jj}(\boldsymbol{z}) + \left(\hat{d}_{LTS} - 1\right)^2 \sum\nolimits_{j=1}^{p} \frac{\widehat{\alpha}^2_{jGLTSR}(\boldsymbol{z})}{\left(\lambda_{j(\boldsymbol{z})} + 1\right)^2}, \tag{26}$$

$$\widehat{\text{MSE}}\left(\widehat{\boldsymbol{\beta}}_{GRLTSRL}(\hat{d}_{RLTS}.\boldsymbol{z}^*)\right) = \widehat{\sigma}^2_{GRLTSR}\sum\nolimits_{j=1}^{p} \frac{\left(\lambda_j(\boldsymbol{z}^*) + \hat{d}_{RLTS}\right)^2}{\left(\lambda_j(\boldsymbol{z}^*) + 1\right)^2} m_{jj}(\boldsymbol{z}^*) + \left(\hat{d}_{RLTS} - 1\right)^2 \sum\nolimits_{j=1}^{p} \frac{\widehat{\alpha}^2_{jGRLTSR}(\boldsymbol{z}^*)}{\left(\lambda_j(\boldsymbol{z}^*) + 1\right)^2}, \tag{27}$$

$$\widehat{\text{MSE}}\left(\widehat{\boldsymbol{\beta}}_{GERLTSRL}(\hat{d}_{LTS}.\boldsymbol{z}^{**})\right) = \widehat{\sigma}^2_{GERLTSR}\sum\nolimits_{j=1}^{p} \frac{\left(\lambda_j(\boldsymbol{z}^{**}) + \hat{d}_{ERLTS}\right)^2}{\left(\lambda_j(\boldsymbol{z}^{**}) + 1\right)^2} m_{jj}(\boldsymbol{z}^{**}) + (\hat{d}_{ERLTS} -$$
$$1)^2 \sum\nolimits_{j=1}^{p} \frac{\widehat{\alpha}^2_{jGERLTSR}(\boldsymbol{z}^{**})}{\left(\lambda_j(\boldsymbol{z}^{**}) + 1\right)^2}. \tag{28}$$

Proof. The proof directly follows by mimicking the proof of Theorem 1. □

Remark 1. *According to [40,41], because the variance matrix* $\boldsymbol{V}$ *is typically unknown in practice, the suggested estimators are non-applicable because they rely on the unknown variance matrix of the error terms. To address this issue, we must replace the unknown* $\boldsymbol{V}$ *with a consistent estimator of it using feasible two-stage estimators, as follows:*

$$\widehat{\boldsymbol{V}} = \frac{1}{n - (p - q)} \left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{ERLTSR}(\boldsymbol{z}^{**})\right)^\top \left(\widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}_{ERLTSR}(\boldsymbol{z}^{**})\right), \tag{29}$$

where

$$\widehat{\boldsymbol{\beta}}_{ERLTSR}(\boldsymbol{z}^{**}) = \widehat{\boldsymbol{\beta}}_{ERLTS}(\boldsymbol{z}^{**}) - \left(\widetilde{\boldsymbol{X}}^\top \boldsymbol{Z}^{**}\widetilde{\boldsymbol{X}}\right)^{-1}\boldsymbol{R}^\top \left(\boldsymbol{R}(\widetilde{\boldsymbol{X}}^\top \boldsymbol{Z}^{**}\widetilde{\boldsymbol{X}})^{-1}\boldsymbol{R}^\top\right)^{-1}(\boldsymbol{R}\widehat{\boldsymbol{\beta}}_{ERLTS}(\boldsymbol{z}^{**}) - \boldsymbol{r}),$$

in which $\widehat{\boldsymbol{\beta}}_{ERLTS}(\boldsymbol{z}^{**}) = \left(\widetilde{\boldsymbol{X}}^\top \boldsymbol{Z}^{**}\widetilde{\boldsymbol{X}}\right)^{-1}\widetilde{\boldsymbol{X}}^\top \boldsymbol{Z}^{**}\widetilde{\boldsymbol{y}}$ *is an ordinary ERLTS estimator of the parameter* $\boldsymbol{\beta}$.

6. Illustrative Experiments

To illustrate the advantages of the improved techniques that have been proposed for a restricted partially linear model in the presence of simultaneous multicollinearity and outlier problems, we continue our examination with some numerical experiments in this section. We evaluate the performance of the proposed techniques on both a real-world data set and some Monte Carlo simulation schemes.

6.1. The Monte Carlo Simulation Schemes

We conducted a numerical analysis to evaluate the precision of our robust estimators for an RSRM when dealing with contaminated data sets with outliers and multicollinearity. In each replication, the regressors were randomly generated using the following structure. Indeed, in order to reach different levels of multicollinearity, according to [42,43], explanatory variables were constructed using a device over 150 observations and 10^3 iterations, based on the model described below:

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

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 $\boldsymbol{X}^\top \boldsymbol{X}$ and $\boldsymbol{X}^\top \boldsymbol{y}$ are in correlating forms. Four distinct sets of correlation values are investigated, specifically $\gamma = 0.25,\ 0.50, 0.75$ and 0.95. For the dependent variable, n observations are then calculated by

$$y_i = \sum\nolimits_{j=1}^{5} x_{ij}\beta_j + f(t_i) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{30}$$

where

$$\boldsymbol{\beta} = (-1, 4, 2, -5, -3)^\top,$$

$$f(t) = exp\left\{\sin(t)\cos(t) + \sqrt{t}\right\}, \; t \in [0,\, 3],$$

$$\boldsymbol{\varepsilon}_{(n \times 1)} = \left(\boldsymbol{\varepsilon}_1^\top, \boldsymbol{\varepsilon}_2^\top\right)^\top,$$

in which

$$\boldsymbol{\varepsilon}_1^\top{}_{(h \times 1)} \sim N_h\left(\boldsymbol{0},\, \sigma^2 \boldsymbol{V}\right),\; \sigma^2 = 1.64,\; [v_{ij}] = exp\{-9|i-j|\},$$

$$h = [0.25n],\; [0.33n],\; [0.50n]$$

and

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

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

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

For estimating the nonparametric part of the model (16), $f(.)$, the weight proposed by Priestley and Chao [44] 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 to obtain the optimum value of the bandwidth ω, which minimizes the C.V. criterion.

The non-parametric component of the model (30) is presented in Figure 1. This wave 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 the mean squared errors of the linear and non-linear estimators are reported in Tables 1 and 2, respectively, where

$$\hat{\text{mse}}\left(\hat{\boldsymbol{\beta}}_{(i)}^{(m)}, \boldsymbol{\beta}\right) = \frac{1}{M} \Sigma_{m=1}^M \left\|\hat{\boldsymbol{\beta}}_{(i)}^{(m)} - \boldsymbol{\beta}\right\|_2^2,$$

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

in which $\hat{\beta}_{(i)}^{(m)}$ and $\hat{f}_{(i)}^{(m)}$ are the ith estimators of the linear and non-linear parts ($i = 1$, ..., 8) obtained in the mth iteration of the model for all of the proposed approaches and $\|\boldsymbol{v}\|_2^2 = \Sigma_{i=1}^q v_i^2$ for $\boldsymbol{v} = (v_1, \ldots, v_p)^\top$. Moreover, PCDO is the percentage of the data contaminated with outliers (PCDO = $100 \times \frac{n-h}{n}\%$).
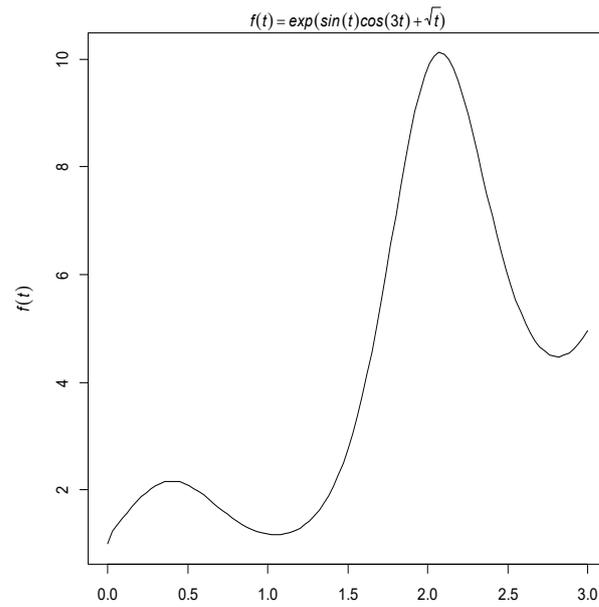
Figure 1. The non-linear function of the simulated model.

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	0.5	0.75	0.95	0.25	0.5	0.75	0.95	0.25	0.5	0.75	0.95
		PCDO = 25%				**PCDO = 33%**				**PCDO = 50%**			
GLSRE	min	1.0300	1.4501	1.9507	8.4008	1.5712	4.9800	4.4346	10.09635	2.76	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	4.357	16.5846	1.9013	3.9258	5.4962	19.9989	3.2705	9.325	10.5542	26.3929
	median	1.1228	1.9902	4.1577	15.6974	1.8818	3.14	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
GLTSRE	min	7.00×10^{-2}	1.1141	1.1376	5.8009	0.0937	1.65	2.3008	9.45506	0.075	1.0046	2.7407	12.98
	max	2.0108	3.1206	5.8441	17.492	1.4185	4.0889	4.9014	18.203	3.8996	4.8417	6.2896	28.688
	mean	0.1579	1.9061	3.164	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
GRLTSRE	min	1.68×10^{-2}	1.0801	1.0411	5.9207	0.0924	1.4001	2.4216	9.2008	0.0478	1.0009	2.2008	12.975
	max	2.4459	2.3196	5.78	17.0799	1.0422	4.0354	5.0202	19.7405	2.3377	3.901	4.9601	29.8601
	mean	0.1909	1.6353	3.2186	13.8984	0.9112	3.237	3.383	12.5147	0.2552	2.313	3.5289	15.3241
	median	0.0814	1.5541	3.089	12.3911	0.3803	3.1006	3.1606	11.9148	0.1238	2.142	3.2449	14.9558
	S.D.	0.6859	2.7016	2.3204	7.3535	1.352	2.3558	3.5929	8.3451	1.347	4.532	3.7338	8.2607
GERLTSRE	min	0.168	1.0601	1.0009	5.3707	0.0468	1.6021	2.2019	9.5476	0.0409	1.0537	2.8286	11.9307
	max	0.0675	2.2113	5.6578	17.7935	1.3463	4.0402	4.3671	22.228	2.7772	3.4363	5.7513	27.8879
	mean	0.652	1.601	3.2984	13.6677	**0.6243**	3.2104	3.2984	12.1435	0.2313	2.2695	3.4561	15.1132
	median	1.21×10^{-2}	1.5381	3.0651	11.9635	0.335	3.0759	3.1266	11.4454	0.105	2.1204	3.201	14.0838
	S.D.	1.8293	2.6617	2.2433	7.9852	1.0163	2.2913	3.4852	9.8424	1.3169	4.3761	3.6389	8.8768
GLSRLE	min	4.10×10^{-1}	0.0491	0.0858	2.5456	1.068	2.0301	2.7207	4.1117	2.572	3.1352	3.2247	13.325
	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	4.333	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	17.155
	S.D.	0.3457	1.3459	2.5714	4.719	2.4499	1.4511	2.7632	3.1329	6.8705	5.4671	3.7902	7.3125
GLTSRLE	min	4.33×10^{-3}	0.0045	0.0037	0.0604	0.0345	0.0141	0.0109	0.6054	0.0017	0.157	0.0048	1.137
	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	0.807	0.9593	0.2282	0.375	1.3672	0.2407	0.2941	0.4949	6.0419
	median	**0.0581**	0.0434	0.0745	0.3423	0.3887	0.083	0.166	0.9254	0.1183	0.3212	0.2139	5.8904
	S.D.	0.2205	1.1919	2.3091	3.1927	1.3874	1.2861	2.46	2.9477	1.3493	1.9637	2.7041	3.9317
GRLTSRLE	min	4.68×10^{-3}	0.0038	0.0057	0.0564	0.0143	0.0181	0.0079	0.5208	0.0178	0.0908	0.0028	1.4756
	max	2.0459	2.3196	0.78	5.0979	0.9422	2.0354	1.0202	6.7405	1.3377	0.851	0.7601	12.4909
	mean	0.19	0.13	0.23	0.9417	0.8516	0.2361	0.3832	1.4642	0.2553	0.3128	0.5285	7.3356
	median	0.0797	0.0585	0.1052	0.4345	**0.2379**	0.1001	0.1633	0.997	0.1238	0.142	0.2439	6.0558
	S.D.	0.2856	1.205	2.3272	3.3762	1.3412	1.3557	2.5913	3.2963	1.3472	1.1317	2.7331	4.2272
GERLTSRLE	min	4.20×10^{-3}	0.0034	0.0028	0.0155	0.01	0.0042	0.0067	0.605	0.0005	0.0309	0.0035	1.0607
	max	1.7169	2.3475	0.7972	4.4159	0.9508	1.2151	1.1933	7.082	1.4948	0.7753	0.726	10.3495
	mean	0.1513	**0.0918**	**0.1475**	**0.7416**	0.7909	**0.177**	**0.3214**	**1.264**	**0.2204**	**0.2847**	**0.4853**	**5.9774**
	median	0.0583	**0.0405**	**0.063**	**0.3076**	0.244	**0.0769**	**0.1461**	**0.8669**	**0.1021**	**0.124**	**0.2089**	**4.9675**
	S.D.	0.2326	1.1759	2.2753	3.2737	1.2888	1.3181	2.5431	3.219	1.3065	1.3905	2.6622	3.898

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%**			
GLSRE	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	9.6595	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
GLTSRE	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
GRLTSRE	min	0.0240	0.3670	2.0519	4.0360	0.0332	0.5453	2.0242	3.0242	0.0306	0.6308	3.5312	7.0221
	max	5.4595	4.1415	4.9134	9.7839	2.5251	5.1927	2.5555	8.0322	5.5357	4.1633	5.4292	23.3308
	mean	0.4834	1.3173	2.2997	5.2344	0.2503	1.7619	2.4152	6.3061	0.6057	1.5726	3.5296	12.4126
	median	0.2703	1.2140	2.2041	4.9715	0.1908	1.2534	2.2389	6.1972	0.3364	1.3213	4.3018	11.2465
	S.D.	0.5791	2.8124	3.2748	3.1874	0.1944	2.5513	2.4929	4.3212	0.6974	1.6619	2.6000	7.4440
GERLTSRE	min	0.0334	0.3976	2.0415	4.0255	0.0471	0.6224	2.0297	3.0262	0.8227	0.6209	3.5229	7.0202
	max	3.8208	4.9542	4.6089	9.7576	2.5110	4.4394	2.6821	8.9765	4.5546	3.5910	5.1143	23.1862
	mean	0.4353	1.2627	2.2505	5.2062	0.2374	1.5857	2.3452	6.2557	0.5574	1.5074	3.4712	12.3650
	median	0.2510	1.1923	2.1858	4.9607	0.1768	1.2293	2.2099	5.1798	0.3178	1.2860	4.2724	11.2312
	S.D.	0.4995	2.7534	3.2167	3.1454	0.1903	2.4407	2.3909	4.2392	0.6332	2.5743	2.5192	7.3882
GLSRLE	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
GLTSRLE	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
GRLTSRLE	min	0.0243	0.0387	0.0477	0.0360	0.0332	0.0248	0.0242	0.0242	0.0306	0.0308	0.0312	1.1252
	max	5.4595	2.7415	0.9134	0.9839	2.5251	3.1927	0.5555	1.5322	5.5357	1.1633	1.0092	5.3308
	mean	0.4815	0.3250	0.3085	0.7399	0.2446	0.4634	0.4155	0.5997	0.6058	0.5724	0.5293	3.4111
	median	0.2697	0.2193	0.2117	0.5791	0.1790	0.2594	0.2417	0.2933	0.3370	0.3213	0.3018	3.2465
	S.D.	0.5784	0.3178	2.2802	2.1918	0.1959	1.5513	1.4915	3.3142	0.6976	0.6615	0.5996	4.4402
GERLTSRLE	min	0.0319	0.0412	0.0420	0.0260	0.0333	0.0316	0.0305	0.0340	0.0199	0.0214	0.0226	1.1241
	max	3.6001	2.7898	0.8330	0.9889	2.2619	3.6611	0.9160	1.1491	4.8652	1.0579	0.9669	4.9801
	mean	0.4030	**0.2760**	**0.2616**	0.7361	**0.2094**	**0.4001**	**0.2877**	**0.5106**	**0.5357**	**0.5301**	**0.4937**	**3.3124**
	median	0.2375	**0.1943**	**0.1927**	0.5737	**0.1663**	**0.2142**	**0.2022**	**0.2926**	**0.3028**	**0.3061**	**0.2925**	**3.1817**
	S.D.	0.4617	0.2738	2.2417	2.1835	0.1577	2.4814	1.4375	3.2898	0.6099	0.5975	0.5406	4.3932

Table 3. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.25$ and PCDO = 25%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.0025	−1.0022	−1.0020	−1.0023	−1.0021	−1.0019	−1.0019	−1.0019
$\hat{\beta}_2$	3.9439	3.9519	3.9561	3.9489	3.9547	3.9586	3.9577	3.9581
$\hat{\beta}_3$	1.8317	1.8558	1.8682	1.8466	1.8640	1.8757	1.8732	1.8744
$\hat{\beta}_4$	−4.8547	−4.8754	−4.8861	−4.8675	−4.8825	−4.8926	−4.8905	−4.8916
$\hat{\beta}_5$	−2.9235	−2.9344	−2.9401	−2.9303	−2.9382	−2.9435	−2.9423	−2.9429
$\hat{d}$	1.0000	1.0000	1.0000	1.0000	0.0002	0.4316	0.1179	0.3106
$e^{\top}z$	150.00	116.00	108.3896	109.8749	150.00	116.00	108.3896	109.8749

Table 4. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.50$ and PCDO = 25%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.1020	−1.0916	−1.0817	−1.0816	−1.0014	−1.0014	−1.0014	−1.0013
$\hat{\beta}_2$	3.9059	3.9052	3.9220	3.9244	3.9693	3.9700	3.9683	3.9714
$\hat{\beta}_3$	1.8176	1.8957	1.8859	1.8932	1.9079	1.9101	1.9048	1.9141
$\hat{\beta}_4$	−4.8057	−4.8599	−4.8814	−4.8877	−4.9204	−4.9224	−4.9178	−4.9258
$\hat{\beta}_5$	−2.8398	−2.8526	−2.8481	−2.8514	−2.9581	−2.9591	−2.9567	−2.9610
$\hat{d}$	1.000	1.000	1.000	1.000	0.0078	0.6775	0.3618	0.5881
$e^{\top}z$	150	131	108.3894	110.1718	150	131	108.3894	110.1718

Table 5. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.75$ and PCDO = 25%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.7530	−1.4123	−1.4125	−1.4124	−1.1520	−1.0019	−1.0020	−1.0018
$\hat{\beta}_2$	3.1342	3.5500	3.5456	3.5475	3.8571	3.9585	3.9555	3.9600
$\hat{\beta}_3$	1.1027	1.3500	1.3369	1.3425	1.2712	1.8755	1.8665	1.8799
$\hat{\beta}_4$	−4.2296	−4.6705	−4.6592	−4.6640	−4.7888	−4.8925	−4.8847	−4.8963
$\hat{\beta}_5$	−2.1103	−2.5318	−2.5259	−2.5284	−2.7415	−2.9434	−2.9393	−2.9454
$\hat{d}$	1.000	1.000	1.000	1.000	0.0054	0.6639	0.3307	0.5648
$e^{\top}z$	150	128	108.392	110.461	150	128	108.392	110.461

Table 6. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.95$ and PCDO = 25%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−2.5080	−2.0052	−2.0053	−2.0061	−1.2354	−1.0036	−1.0035	−1.0036
$\hat{\beta}_2$	2.8241	3.0851	3.0845	3.0656	3.8306	3.9207	3.9228	3.9203
$\hat{\beta}_3$	0.4724	1.1554	1.1534	1.1967	1.3618	1.7622	1.7684	1.7610
$\hat{\beta}_4$	−3.7103	−4.1024	−4.1006	−4.1517	−4.7447	−4.7946	−4.8000	−4.7936
$\hat{\beta}_5$	−2.0002	−2.0134	−2.0024	−2.0167	−2.7917	−2.8919	−2.8947	−2.8914
$\hat{d}$	1.000	1.000	1.000	1.000	0.0052	0.6488	0.2912	0.5223
$e^{\top}z$	150	122	108.392	109.5919	150	122	108.392	109.5919

Table 7. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.25$ and PCDO = 33%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.0035	−1.0034	−1.0038	−1.0033	−1.0079	−1.0059	−1.0055	−1.0051
$\hat{\beta}_2$	3.9222	3.9244	3.9168	3.9265	3.8259	3.8707	3.8789	3.8881
$\hat{\beta}_3$	1.7667	1.7732	1.7504	1.7795	1.4776	1.6122	1.6366	1.6643
$\hat{\beta}_4$	−4.7985	−4.8042	−4.7844	−4.8095	−4.5488	−4.6651	−4.6862	−4.7101
$\hat{\beta}_5$	−2.8939	−2.8969	−2.8865	−2.8998	−2.7625	−2.8237	−2.8348	−2.8474
$\hat{d}$	1.0000	1.0000	1.0000	1.0000	0.0067	0.6464	0.2873	0.5200
$e^{\top}z$	150	129	108.392	109.981	150	129	108.392	109.981

Table 8. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.50$ and PCDO = 33%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.1029	−1.1025	−1.1022	−1.1026	−1.1023	−1.0021	−1.0021	−1.0021
$\hat{\beta}_2$	3.9065	3.9150	3.9212	3.9123	3.9199	3.9533	3.9533	3.9539
$\hat{\beta}_3$	1.8096	1.8109	1.8237	1.8108	1.8298	1.8599	1.8599	1.8618
$\hat{\beta}_4$	−4.6355	−4.7574	−4.7737	−4.7504	−4.7703	−4.8790	−4.8790	−4.8806
$\hat{\beta}_5$	−2.5134	−2.7250	−2.7335	−2.7213	−2.8017	−2.9363	−2.9363	−2.9372
$\hat{d}$	1.0000	1.0000	1.0000	1.0000	0.0002	0.4261	0.1148	0.3031
$e^\top z$	150	130	108.3896	109.8711	150	130	108.3896	109.8711

Table 9. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.75$ and PCDO = 33%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.9438	−1.4133	−1.4127	−1.4135	−1.6228	−1.0026	−1.0026	−1.0026
$\hat{\beta}_2$	2.9157	3.6282	3.6400	3.6226	3.5385	3.9429	3.9429	3.9433
$\hat{\beta}_3$	1.0472	1.3847	1.3200	1.3677	1.3156	1.8287	1.8286	1.8298
$\hat{\beta}_4$	−3.7816	−4.7140	−4.7446	−4.7994	−4.6407	−4.8521	−4.8520	−4.8530
$\hat{\beta}_5$	−2.0051	−2.6021	−2.6182	−2.6944	−2.5162	−2.9221	−2.9221	−2.9226
$\hat{d}$	1.000	1.000	1.000	1.000	0.0001	0.4026	0.0932	0.2683
$e^\top z$	150	130	108.3896	109.8729	150	130	108.3896	109.8729

Table 10. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.95$ and PCDO = 33%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−2.8086	−2.0067	−2.0043	−2.0079	−1.7142	−1.0038	−1.0038	−1.0038
$\hat{\beta}_2$	2.2198	3.0528	3.0049	3.0262	3.5466	3.9168	3.9158	3.9171
$\hat{\beta}_3$	0.4293	1.2685	1.2648	1.2186	1.1197	1.7505	1.7474	1.7512
$\hat{\beta}_4$	−3.5071	−4.0187	−4.1537	−4.0497	−4.3579	−4.7845	−4.7819	−4.7851
$\hat{\beta}_5$	−1.7406	−2.1003	−2.1104	−2.1630	−2.5726	−2.8866	−2.8852	−2.8869
$\hat{d}$	1.000	1.000	1.000	1.000	0.0008	0.3654	0.0722	0.2141
$e^\top z$	150	118	108.3896	109.8748	150	118	108.3896	109.8748

Table 11. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.25$ and PCDO = 50%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.8514	−1.0015	−1.0010	−1.0016	−1.8709	−1.0010	−1.0010	−1.0010
$\hat{\beta}_2$	3.9029	3.9678	3.9783	3.9647	3.8807	3.9790	3.9783	3.9781
$\hat{\beta}_3$	1.6097	1.9034	1.9350	1.8940	1.7422	1.9371	1.9350	1.9343
$\hat{\beta}_4$	−4.1220	−4.9166	−4.9439	−4.9085	−4.2501	−4.9457	−4.9439	−4.9432
$\hat{\beta}_5$	−2.4590	−2.9561	−2.9705	−2.9518	−2.4737	−2.9714	−2.9705	−2.9701
$\hat{d}$	1.000	1.000	1.000	1.000	0.0013	0.0217	0.0009	0.0121
$e^\top z$	150	127	108.3898	109.5876	150	127	108.3898	109.5876

Table 12. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.50$ and PCDO = 50%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.9411	−1.3117	−1.3111	−1.3119	−1.6436	−1.0011	−1.0011	−1.0010
$\hat{\beta}_2$	3.1654	3.4623	3.4761	3.4590	3.2789	3.9763	3.9761	3.9758
$\hat{\beta}_3$	0.8961	1.5868	1.6282	1.6771	1.0367	1.9289	1.9283	1.9274
$\hat{\beta}_4$	−3.7103	−4.7022	−4.7380	−4.7939	−4.0454	−4.9386	−4.9380	−4.9373
$\hat{\beta}_5$	−2.0528	−2.3485	−2.4674	−2.4642	−2.1512	−2.9677	−2.9674	−2.9670
$\hat{d}$	1.000	1.000	1.000	1.000	0.0250	0.0203	0.0011	0.0122
$e^\top z$	150	123	108.3898	109.5875	150	123	108.3898	109.5875

Table 13. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.75$ and PCDO = 50%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−1.0021	−1.0023	−1.0012	−1.0026	−1.0011	−1.0012	−1.0012	−1.0013
$\hat{\beta}_2$	3.9535	3.9485	3.9731	3.9435	3.9767	3.9726	3.9732	3.9721
$\hat{\beta}_3$	1.8606	1.8456	1.9194	1.8306	1.9300	1.9178	1.9195	1.9164
$\hat{\beta}_4$	−4.8796	−4.8667	−4.9304	−4.8537	−4.9395	−4.9290	−4.9304	−4.9278
$\hat{\beta}_5$	−2.9366	−2.9298	−2.9634	−2.9230	−2.9682	−2.9626	−2.9634	−2.9620
$\hat{d}$	1.000	1.000	1.000	1.000	0.0018	0.0168	0.0008	0.0089
$e^\top z$	150	125	108.3898	109.5875	150	125	108.3898	109.5875

Table 14. Evaluation of the parameters of the proposed estimation method when $\gamma = 0.95$ and PCDO = 50%.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
$\hat{\beta}_1$	−3.1056	−2.1161	−2.1115	−2.1171	−2.6811	−1.0016	−1.0015	−1.0014
$\hat{\beta}_2$	3.0766	3.1668	3.1671	3.1436	3.0995	3.9690	3.9676	3.9692
$\hat{\beta}_3$	0.3197	1.1703	1.1701	1.1708	1.0297	1.9072	1.9028	1.9085
$\hat{\beta}_4$	−2.6802	−3.6548	−3.9149	−3.8948	−3.1793	−4.9165	−4.9161	−4.9168
$\hat{\beta}_5$	−1.1117	−2.0683	−2.0552	−2.0867	−1.9681	−2.9567	−2.9558	−2.9577
$\hat{d}$	1.000	1.000	1.000	1.000	0.0003	0.0205	0.0022	0.0122
$e^\top z$	150	127	108.3898	109.5876	150	127	108.3898	109.5876

Figure 2 shows the estimations of the non-linear part of model (30) using the proposed methods. In this figure, the nonparametric function is estimated via the kernel method after the estimation of the linear part of the model (30) using all of the eight proposed methods. To save space, the results have been only reported for $n = 150$ with a 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.

The results from the Monte Carlo simulations for $n = 150$, $p = 5$ and $\gamma = 0.25$, 0.50, 0.75 and 0.95 are presented in Tables 1–14 and Figure 2. From these tables, it can be seen that the factors affecting the performance of the estimators are the degree of correlation (γ) between the explanatory variables and the percentage of the data contaminated with outliers (PCDO). From these tables and Figure 2, it is clearly concluded that if the levels of PCDO and multicollinearity increased, then the mean squared error estimations for both the linear and non-linear parts of the GLSRE would be highly increased. Also, increasing the level of the PCDO increases the mean squared error estimations for both the linear and non-linear parts of non-robust estimators, while increasing the level of multicollinearity increases the mean squared error estimations for both the linear and non-linear parts of non-Liu estimators. From Tables 1–14 and Figure 2, it can be seen that the GERLTSRLE often performs better than the other methods since it offers the smallest mean squared error estimations (bold values) for both the linear and non-linear parts of the models in most of the simulated schemes.
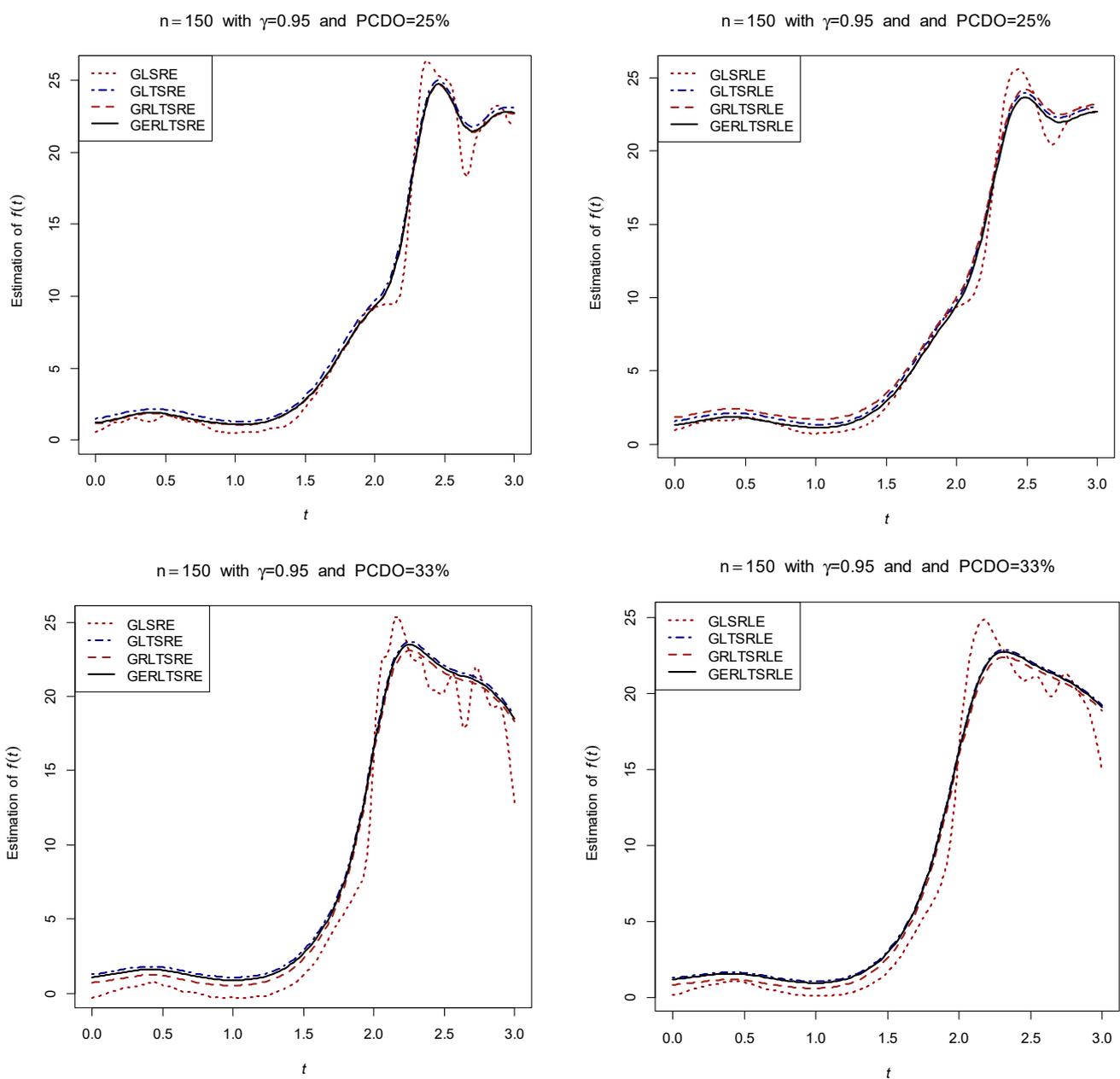
Figure 2. *Cont.*
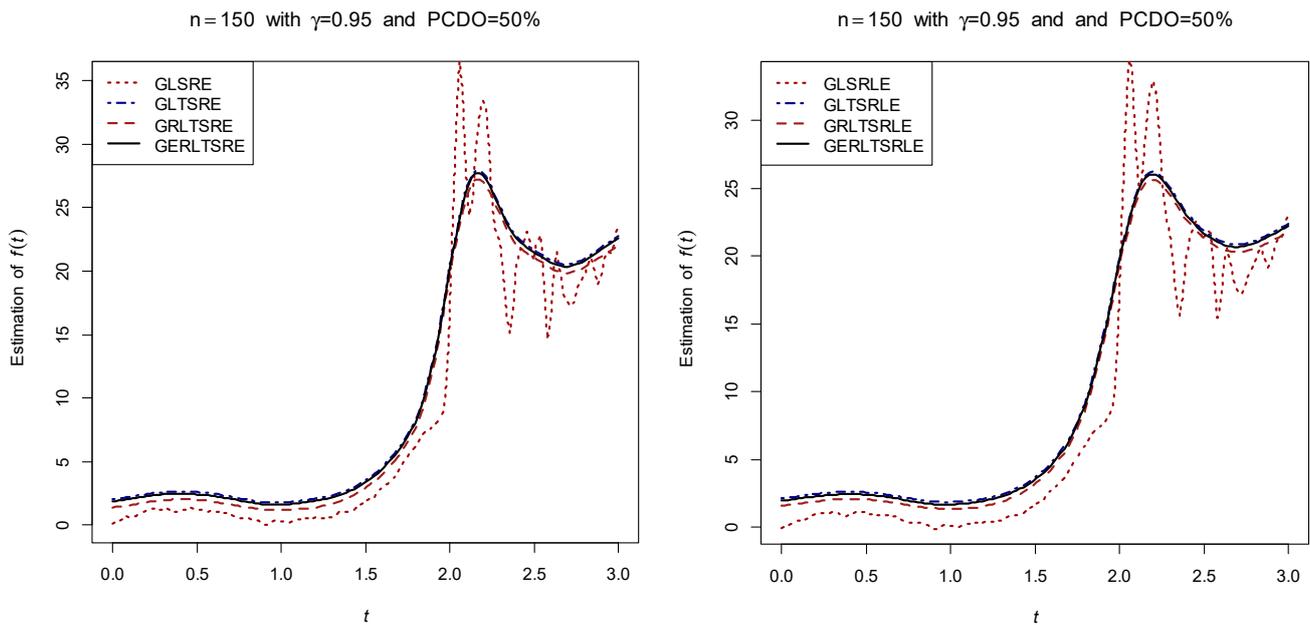
Figure 2. Kernel prediction of the function under study for $n = 150$ and $\gamma = 0.95$. Non-Liu and Liu estimators are plotted for a PCDO = 25% (**low**), PCDO = 33% (**middle**) and PCDO = 50% (**high**).

6.2. Real-World Data Example

To evaluate the proposed estimation approaches using the partially linear model, we examine the hedonic pricing of house features [45]. The data consist of 92 detached homes in the Ottawa area that were sold in 1987. Here is how the variables are defined: the predictor variable is the sale price (SP); the explanatory variables are lot size (lot area = LT), the square footage of housing (SFH), average neighborhood income (ANI), distance to the highway (DHW), the presence of a garage (GAR), and a fireplace (FP). We first consider the pure parametric model

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

We use added-variable charts to intuitively determine the parametric and nonparametric components of the model (see [46] 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 its added-variable plot (Figure 3), we identify ANI as a nonparametric component. Therefore, the partially linear model is specified as follows:

$$(SP)_i = \beta_0 + \beta_1(LT)_i + \beta_2(SFH)_i + \beta_3(FP)_i + \beta_4(DHW)_i + \beta_5(GAR)_i + f(ANI)_i + \varepsilon_i \tag{31}$$

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.
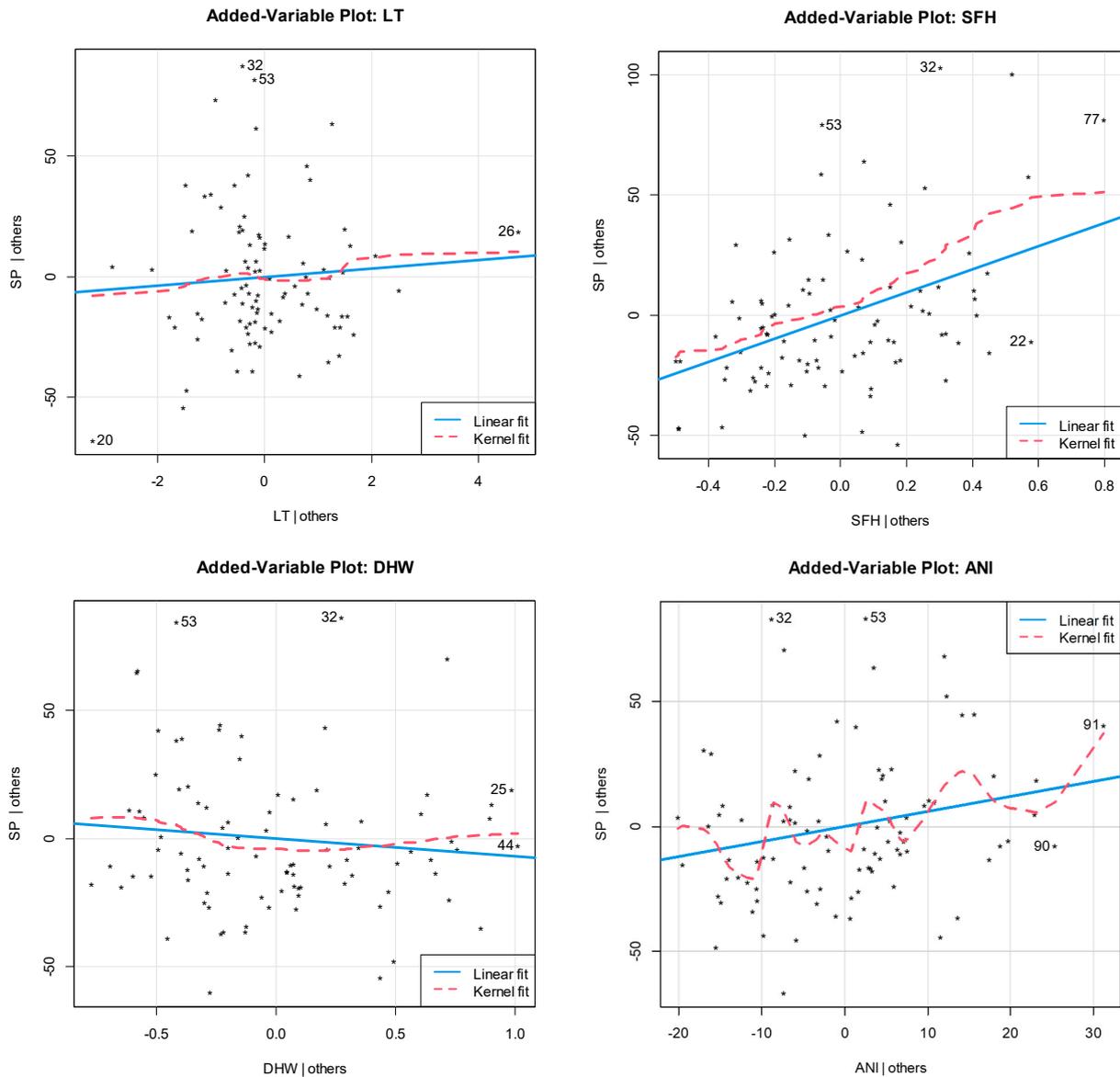
Figure 3. Added-variable plots of individual explanatory variables vs. dependent variable, using linear fit (blue solid line) and kernel fit (red dashed line). "*" symbol shows the outlier points for the linear scheme.

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

The restriction $\boldsymbol{R\beta} = \boldsymbol{r}$ may be identified as follows, based on a basic investigation of the partially linear model (31) using a robust Liu estimator:

$$\boldsymbol{R} = \begin{pmatrix} -1 & 0 & -1 & -1 & 1 \\ 1 & 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & -2 & 8 \end{pmatrix}, \boldsymbol{r} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
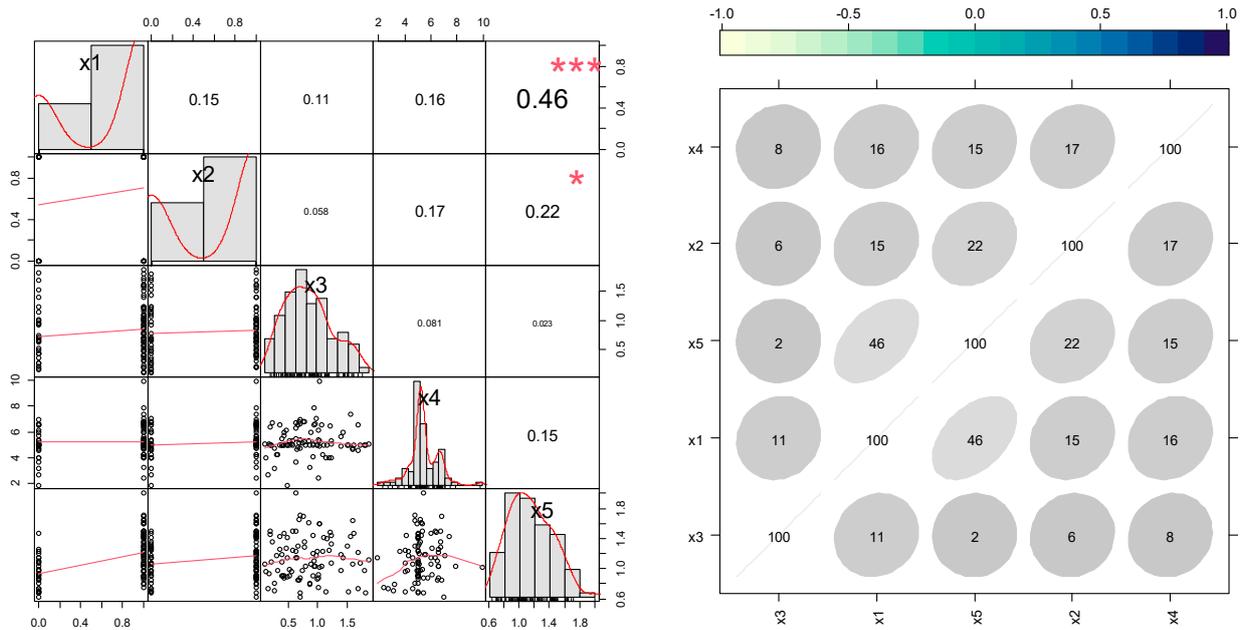
Figure 4. Visualization of the correlation plots of the explanatory variables in the real data set. Each significance level is associated to a symbol: symbols ("***", "*", ".", " ") <=> *p*-values (0, 0.001, 0.01, 0.05, 0.1, 1).

Now, the linear hypothesis $R\beta = r$ is examined within the framework of the restricted partially linear model (31). The test statistic is computed as follows under $R\beta = r$:

$$\chi^2_{rank(R)} = \left(R\widehat{\beta}_{GLS} - r\right)^{\top} \left(R\widehat{\Sigma}_{\widehat{\beta}} R^{\top}\right)^{-1} \left(R\widehat{\beta}_{GLS} - r\right) = 0.4781,$$

where $\widehat{\Sigma}_{\widehat{\beta}} = \widehat{s}^2 \left(\widetilde{X}^{\top} \widehat{V}^{-1} \widetilde{X}\right)^{-1}$, in which $\widehat{s}^2 = \frac{1}{n-p} \left(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GLS}\right)^{\top} \widehat{V}^{-1} \left(\widetilde{y} - \widetilde{X}\widehat{\beta}_{GLS}\right)$.

Consequently, the null hypothesis H_0 is not rejected. Table 15 shows a brief evaluation of the proposed estimators. Compared to the other estimators, the GERLTSRLE seems to be effective based on the results that were obtained. Following the estimate of the linear component of the model using the suggested estimators for model (31), the function fitted by kernel smoothing is shown in Figure 5.
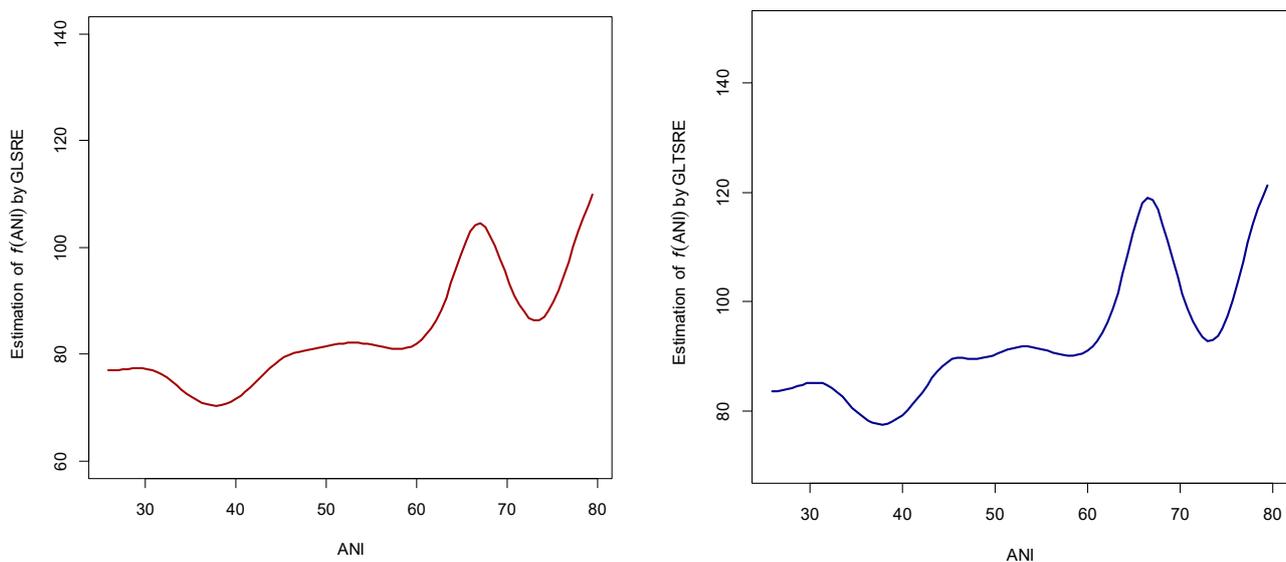
Figure 5. *Cont.*
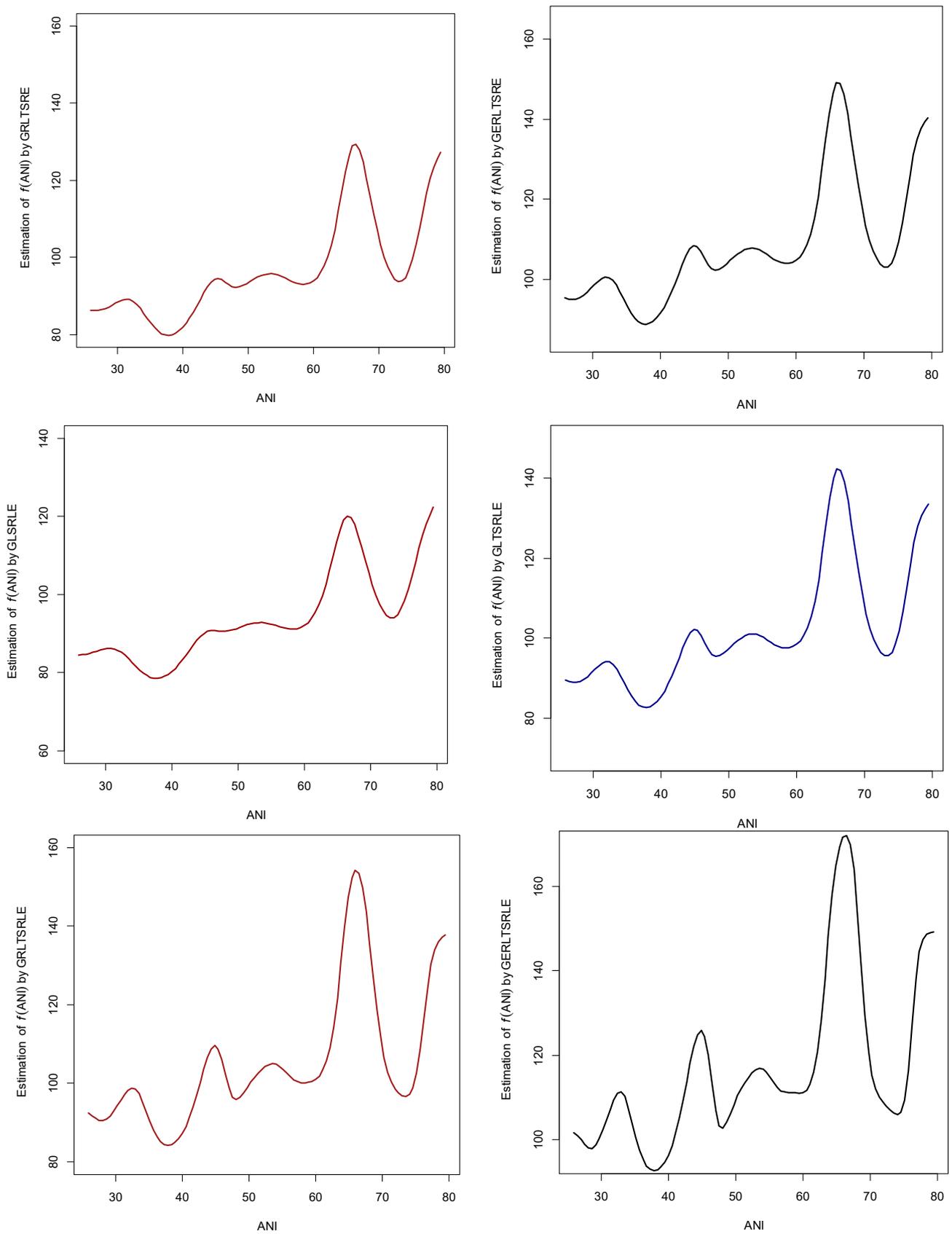
Figure 5. Estimations of the nonparametric part of model (31).

Table 15. The evaluation of the proposed estimators using the hedonic prices of house attributes.

Coefficients	Method							
	GLSRE	GLTSRE	GRLTSRE	GERLTSRE	GLSRLE	GLTSRLE	GRLTSRLE	GERLTSRLE
LT	0.7018	1.0509	0.1522	0.1507	0.8514	1.1235	0.2569	0.2550
SFH	46.7515	33.5686	30.9487	30.2085	38.9154	26.3747	25.0010	24.1247
FP	3.9311	2.5740	2.7443	2.6777	3.5210	1.9568	2.1235	2.0474
DHW	−1.6147	−0.7616	−1.2961	−1.2635	−0.9952	−0.4125	−0.9389	−0.9541
GAR	6.2476	4.3865	4.1926	4.0919	5.2015	2.9958	3.3021	3.2354
$e^\top$ ($z \vee z^* \vee z^{**}$)	92.0000	86.0000	71.7541	74.1542	92.0000	86.0000	71.7541	74.1542
$\hat{d}$	1.0000	1.0000	1.0000	1.0000	0.1542	0.6741	0.6857	0.7154
$\widehat{\text{MSE}}$	926.80	456.41	386.63	301.63	809.59	335.17	279.25	**198.09**
R^2	0.2346	0.6156	0.6956	0.7509	0.3354	0.7325	0.7801	**0.8409**

Table 15 shows a brief evaluation of the proposed estimators. In this table, the values of $\widehat{\text{MSE}}$ and R^2 are calculated, in which $R^2 = 1 - \frac{\text{RSS}}{S_{YY}}$ is the coefficient of determination of the model, where RSS $= \sum_{i=1}^{n}(y_i - \widehat{y_i})^2$ is the residual sum of squares and $\widehat{y_i} = x_i\hat{\beta} + \hat{f}(t_i)$, both of which were calculated for the eight suggested approaches. For the estimation of the nonparametric effect, we first estimated the parametric effects using one of the proposed methods and then a kernel smoother was applied to fit $SP_i - x_i^\top \hat{\beta}$ onto ANI_i, $i = 1, \ldots, n$ for all proposed linear estimators, where $x_i^\top = (LT_i, SFH_i, FP_i, DHW_i, GAR_i)$.

As can be seen from Table 15 and Figure 5, because of the existence of multicollinearity between the explanatory variables, the LTS–Liu and Liu estimators perform better than the non-Liu ones in both the parametric and nonparametric fittings. Furthermore, since there are some outliers in the real data set that were detected in Figure 3, robust estimators outperform non-robust ones in terms of the goodness-of-fit of their models. Hence, developing efficient robust Liu estimation strategies is required for data modeling. From Table 15 and Figure 5, it can be concluded that the GERLTSRLE performs better than the other methods since it offers the smallest $\widehat{\text{MSE}}$ and biggest R^2 values (bold values) in the presence of both multicollinearity and outlier difficulties, while the performances of the non-Liu or non-robust types of estimators are quite poor.

7. Conclusions

In this paper, we suggested Liu and non-Liu types of generalized restricted robust estimators of a partially linear model with dependent errors and some additional linear constraints on the whole parameter space. In the presence of multicollinearity and outliers, an extended robust Liu estimator was introduced based on semidefinite programming to improve the classical least trimmed squares approach used in the partially linear model. Also, we proposed some robust estimations of the biasing parameter of the Liu estimator, which plays an important role in model prediction. In both our simulation studies and a real-life example, it can be found that Liu estimators (both robust and non-robust) perform better than non-Liu ones in both parametric and nonparametric fittings under intense multicollinearity. Moreover, since there are some outliers in the real data set, robust estimators are more efficient than non-robust estimators during model fitting. The GERLTSRLE outperforms the others in terms of the mean squared error and R^2 criteria, making it the most reliable method. Also, it can be deduced that the GERLTSRLE performs effectively in predicting the dependent variable of restricted PLMs without being affected by the corruptive impacts of multicollinearity and outlier observations. As was mentioned earlier, the proposed estimators for the Liu parameter d are not the best and thus a good topic for future research could be obtaining alternative estimators for this parameter based on other suitable criteria, such as cross-validation [47], instead of the mean squared error criterion.

References

1. Engle, R.F.; Granger, C.W.J.; Rice, J.; Weiss, A. Semiparametric estimates of the relation between weather and electricity sales. *J. Am. Stat. Assoc.* **1986**, *81*, 310–320. [CrossRef]
2. Eubank, R.L.; Kambour, E.L.; Kim, J.T.; Klipple, K.; Reese, C.S.; Schimek, M. Estimation in partially linear models. *Comput. Statist. Data. Anal.* **1988**, *29*, 27–34. [CrossRef]
3. Green, P.; Jennison, C.; Seheult, A. Analysis of field experiments by least squares smoothing. *J. Roy. Statist. Soc. Ser. B* **1985**, *47*, 299–315. [CrossRef]
4. Speckman, P. Kernel somoothing in partial linear models. *J. R. Stat. Soc. Ser. B* **1988**, *50*, 413–436. [CrossRef]
5. Eubank, R.L. *Nonparametric Regression and Spline Smoothing*; Marcel Dekker: New York, NY, USA, 1999.
6. Ruppert, D.; Wand, M.P.; Carroll, R.C. *Semiparametric Regression*; Cambridge University Press: Cambridge, UK, 2003.
7. Härdle, W.; Müller, M.; Sperlich, S.; Werwatz, A. *Nonparametric and Semiparmetric Models*; Springer: Berlin/Heidelberg, Germany, 2004.
8. Yatchew, A. *Semiparametric Regression for the Applied Econometrican*; Cambridge University Press: Cambridge, UK, 2003.
9. Akdeniz, F.; Tabakan, G. Restricted ridge estimators of the parameters in semiparametric regression model. *Commun. Stat. Theory Methods* **2009**, *38*, 1852–1869.
10. Akdeniz, D.E.; Hardle, W.K.; Osipenko, M. Difference based ridge and Liu type estimators in semiparametric regression models. *J. Multivar. Anal.* **2012**, *105*, 164–175. [CrossRef]
11. Arashi, M.; Valizadeh, T. Performance of Kibria's methods in partial linear ridge regression model. *Stat. Pap.* **2015**, *56*, 231–246.
12. Liu, K.J. A new class of biased estimate in linear regression. *Commun. Stat. Theo. Meth.* **1993**, *22*, 393–402.
13. Akram, M.N.; Kibria, B.M.G.; Arashi, M.; Lukman, A.F. A new improved Liu estimator for the QSAR model with inverse Gaussian response. *Commun. Stat. Simul. Comput.* **2024**, *53*, 1873–1888. [CrossRef]
14. Kaçıranlar, S.; Ozbay, N.; Ozkan, E.; Guler, H. Comparison of Liu and two parameter principal component estimator to combat multicollinearity. *Commun. Stat. Theo. Meth.* **2022**, *34*, e6737. [CrossRef]
15. Arashi, M.; Lukman, A.F.; Algamal, Z.Y. Liu regression after random forest for prediction and modeling in high dimension. *J. Chemom.* **2022**, *36*, e3393. [CrossRef]
16. Arashi, M.; Roozbeh, M.; Niroumand, H.A. A note on Stein-type shrinkage estimator in partial linear models. *Statistics* **2012**, *46*, 673–685.
17. Roozbeh, M.; Hamzah, N.A. Uncertain stochastic ridge estimation in partially linear regression models with elliptically distributed errors. *Statistics* **2020**, *54*, 494–523. [CrossRef]
18. Roozbeh, M.; Rouhi, A.; Mohamed, N.A.; Jahadi, F. Generalized support vector regression and symmetry functional regression approaches to model the high-dimensional data. *Symmetry* **2023**, *15*, 1262. [CrossRef]
19. Roozbeh, M.; Babaie-Kafaki, S.; Aminifard, Z. A nonlinear mixed–integer programming approach for variable selection in linear regression model. *Commun. Stat. Simul. Comput.* **2023**, *11*, 5434–5445. [CrossRef]
20. Rousseeuw, P.J. Least median of squares regression. *J. Am. Stat. Assoc.* **1984**, *79*, 871–880. [CrossRef]
21. Rousseeuw, P.J.; Leroy, A.M. *Robust Regression and Outlier Detection*; John Wiley: New York, NY, USA, 1987.
22. Rousseeuw, P.J.; van Driessen, K. Computing LTS regression for large data sets. *Data Min. Knowl. Discov.* **2006**, *12*, 29–45. [CrossRef]
23. Arashi, M.; Kibria, B.M.G.; Valizadeh, T. On ridge parameter estimators under stochastic subspace hypothesis. *J. Stat. Comput. Simul.* **2017**, *87*, 966–983. [CrossRef]

24. Karbalaee, M.H.; Arashi, M.; Tabatabaey, S.M.M. Performance analysis of the preliminary test estimator with series of stochastic restrictions. *Commun. Stat. Theory Methods* **2018**, *47*, 1–17. [CrossRef]

25. Roozbeh, M.; Hesamian, G.; Akbari, M.G. Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors. *J. Comput. Appl. Math.* **2020**, *378*, 112940. [CrossRef]

26. Fallah, R.; Arashi, M.; Tabatabaey, S.M.M. On the ridge regression estimator with sub-space restriction. *Commun. Stat. Theory Methods* **2017**, *46*, 11854–11865. [CrossRef]

27. Fallah, R.; Arashi, M.; Tabatabaey, S.M.M. Shrinkage estimation in restricted elliptical regression model. *J. Iran. Stat. Soc.* **2018**, *17*, 49–61. [CrossRef]

28. Toutenburg, H. *Prior Information in Linear Models*; John Wiley: New York, NY, USA, 1982.

29. Durbin, J. A note on regression when there is extraneous information about one of the coefficients. *J. Am. Stat. Assoc.* **1990**, *48*, 799–808. [CrossRef]

30. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics* **1970**, *12*, 69–82. [CrossRef]

31. Akdeniz, F.; Roozbeh, M.; Akdeniz, E.; Khan, M.N. Generalized difference-based weighted mixed almost unbiased liu estimator in semiparametric regression models. *Commun. Stat. Theory Methods* **2022**, *51*, 4395–4416. [CrossRef]

32. Kibria, B.M.G. Some Liu and ridge type estimators and their properties under the ill- conditioned Gaussian linear regression model. *J. Stat. Comput. Simul.* **2012**, *82*, 1–17. [CrossRef]

33. Månsson, K.; Kibria, B.M.G.; Shukur, G. A restricted Liu estimator for binary regression models and its application to an applied demand system. *J. Appl. Stat.* **2016**, *43*, 1119–1127. [CrossRef]

34. Månsson, K.; Kibria, B.M.G. Estimating the unrestricted and restricted Liu estimators for the Poisson regression model: Method and application. *Comput. Econ.* **2021**, *58*, 311–326. [CrossRef]

35. Rousseeuw, P.J. Multivariate estimation with high breakdown point. *Math. Stat. Appl.* **1985**, *8*, 283–297.

36. Alfons, A.; Croux, C.; Gelper, S. Sparse least trimmed squares regression for analyzing high-dimensional large data sets. *Ann. Appl. Stat.* **2013**, *7*, 226–248. [CrossRef]

37. Nguyen, T.D.; Welsch, R. Outlier detection and least trimmed squares approximation using semi-definite programming. *Comput. Stat. Data Anal.* **2010**, *54*, 3212–3226. [CrossRef]

38. Roozbeh, M.; Babaie-Kafaki, S. Extended least trimmed squares estimator in semiparametric regression models with correlated errors. *J. Stat. Comput. Simul.* **2016**, *86*, 357–372. [CrossRef]

39. Kan, B.; Alpu, O.; Yazici, B. Robust ridge and robust Liu estimator for regression based on the LTS estimator. *J. Appl. Stat.* **2013**, *40*, 644–655. [CrossRef]

40. Zellner, A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *J. Am. Stat. Assoc.* **1962**, *57*, 348–368. [CrossRef]

41. Taavoni, M.; Arashi, M. Kernel estimation in semiparametric mixed effect longitudinal modeling. *Stat. Pap.* **2021**, *62*, 1095–1116.

42. McDonald, G.C.; Galarneau, D.I. A Monte Carlo evaluation of some ridge-type estimators. *J. Am. Stat. Assoc.* **1975**, *70*, 407–416. [CrossRef]

43. Gibbons, D.G. A simulation study of some ridge estimators. *J. Am. Stat. Assoc.* **1981**, *76*, 131–139. [CrossRef]

44. Priestley, M.B.; Chao, M.T. Non-Parametric Function Fitting. *J. R. Stat. Soc. Ser. B* **1972**, *34*, 385–392. [CrossRef]

45. Ho, M. Essays on the Housing Market. Ph.D. Dissertation, University of Toronto, Toronto, ON, Canada, 1995.

46. Sheather, S.J. *A Modern Approach to Regression with R*; Springer: New York, NY, USA, 2009.

47. Roozbeh, M.; Najarian, M. Efficiency of the QR class estimator in semiparametric regression models to combat multicollinearity. *J. Stat. Comput. Simul.* **2018**, *88*, 1804–1825. [CrossRef]