REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000-000
https://doi.org/00.00000/revstat.v00i0.000

Novel Robust Estimators for the Linear Regression Model with Multicollinearity and Outlier Problems

Authors: Murat Erisoglu
–
merisoglu@erbakan.edu.tr

Aydin Karakoca ✉
–
akarakoca@erbakan.edu.tr

Abdullah Yurtaslan
–
abdullahyurtaslan@gmail.com

Necmettin Erbakan University, Faculty of Science,
Statistics Department, Konya,Türkiye

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- In this study, we introduce new robust M estimators based on ridge estimation (M-Ridge) for data sets with both multicollinearity and outlier problems in multiple linear regression analysis. In the proposed approach, the iterative re-weighted least squares (IRLS) algorithm for parameter estimation is implemented based on ridge estimation.The proposed approach also provides a solution to the problem of the optimal ridge estimator selection with M-type estimators. The performance of the proposed estimators is evaluated against other estimators using a Monte Carlo simulation study and a real data application. The estimated mean square error (MSE) and k-fold cross validation are used as performance measures in the Monte Carlo simulation study and the real data application, respectively. The proposed M-Ridge estimators outperformed the other estimators considered in many evaluated instances in both the simulation study and the real data application.

Keywords:

- *K-fold cross-validation; M-estimator; MSE; Multicollinearity; Outlier; Ridge regression.*

AMS Subject Classification:

- 62F35, 62J07.

✉ Corresponding author

1. INTRODUCTION

Regression analysis is a versatile statistical tool that can be used to model the relationship between a dependent variable and one or more independent variables. It is one of the most widely used statistical analyses due to its simplicity and effectiveness. Multiple linear regression model in matrix form is given by

$$\mathbf{y} = X\beta + \varepsilon. \tag{1.1}$$

In Eq. (1.1), $\mathbf{y}$ is an $n \times 1$ response vector, X is an $n \times p$ design matrix of known constants, β is an $p \times 1$ unknown parameters vector and ε is an $n \times 1$ stochastic error vector such that $E(\varepsilon) = 0$ and $\mathrm{Cov}(\varepsilon) = \sigma^2\mathbf{I}$. The ordinary least squares (OLS) method is widely used for estimation of β. OLS aims to minimize the sum of squares of the error terms. The objective function and OLS estimates are given by respectively.

$$\widehat{\beta} = \underset{\beta}{\operatorname{argmin}} \left(\sum_{i=1}^{n} \varepsilon_i^2 \right) = \underset{\beta}{\operatorname{argmin}} \left(\sum_{i=1}^{n} (y_i - \mathbf{x}_i\beta)^2 \right) \tag{1.2}$$

$$\widehat{\beta} = (X'X)^{-1}X'\mathbf{y} \tag{1.3}$$

The covariance matrix of OLS estimator is $\sigma^2(X'X)^{-1}$. The OLS estimator has the minimum variance property among all linear unbiased estimators under the model assumptions. However, the high correlations between columns of X or the nonexistence of the full rank property can result multicollinearity problem. In the presence of multicollinearity, the OLS estimator loses its minimum variance property. The ridge estimator is a widely used alternative to the OLS estimator in the presence of the multicollinearity problem. The ridge estimator that proposed by [8] is biased but has lower variance than OLS estimator. The normality of the error terms is another important assumption that must be met in the multiple linear regression model for statistical significance testing of the model. Under the normality assumption, the ML and OLS estimators of β are equal. Outliers are one of the factors that can lead to a violation of the normality assumption. M-estimators are commonly used as an alternative to OLS estimators when the normality assumption is violated. [11] proposed M-estimators, which are a generalization of ML estimators and are robust to departures from the normality assumption. Many M-estimators have been proposed in the literature (see for example [4]).

In multiple linear regression analysis, outliers and multicollinearity problems can be encountered simultaneously. In the presence of both outliers and multicollinearity problems in the model, The ridge estimator based on M estimator

Ridge-M (RM) is proposed by [24] and [5]. The RM is given by

$$\widehat{\beta}_{\mathrm{RM}} = \left(X^{'}X + kI\right)^{-1} X^{'}X\widehat{\beta}_{\mathbf{M}}. \tag{1.4}$$

where $\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}}$ denotes OLS based M estimator. The use of OLS estimators in RM may adversely affect the performance of the estimator. From this point of view, we propose M estimators based on ridge estimation instead of OLS estimation when outlier and multicollinearity problems coexist. The proposed M estimator based on ridge estimation is obtained by rearranging the Iteratively Reweighted Least Squares (IRLS) algorithm. The proposed approach also provides a solution to the problem of the optimal ridge estimator selection with M-type estimators.

This paper is organized as follows. The proposed M estimators based on ridge estimation are defined after examining the ridge and M estimators. The performance of the proposed estimator is examined through by a simulation study and a real data application. In the simulation study, the mean square error (MSE) is used as a comparison criterion. The proposed estimators are compared with the ridge and classical M estimators, considering different sample sizes, outlier rates, and correlation structures. The performance of the proposed estimators is also compared in the simulation study for lognormal distributed errors. In the real data application, the Tobacco dataset, which has both multicollinearity and outlier problems, is used. Iterative k-fold cross validation is used the comparison criterion in the real data application.

2. Material and Method

2.1. Ridge Estimators

The ridge estimator of unknown parameter vector is given by

$$\widehat{\beta}_{\mathbf{R}} = \left(X^{'}X + kI\right)^{-1} X^{'}\mathbf{y}. \tag{2.1}$$

where k is known as ridge (shrinkage, bias) parameter which tunes the variance -bias trade off.

The canonical form of Eq. (1.1) with $Z = XD$ and $\boldsymbol{\alpha} = D\boldsymbol{\beta}$ as

$$\boldsymbol{y} = Z\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \tag{2.2}$$

D is defined as an orthogonal eigenvector matrix such that $D'\left(X'X\right)D = \Lambda$ where $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \ \lambda_p)$ is eigenvalues of $X'X$. The ridge estimator of $\boldsymbol{\alpha}$ is given by

$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{R}} = \left(Z'Z + kI\right)^{-1} Z'\boldsymbol{y}. \tag{2.3}$$

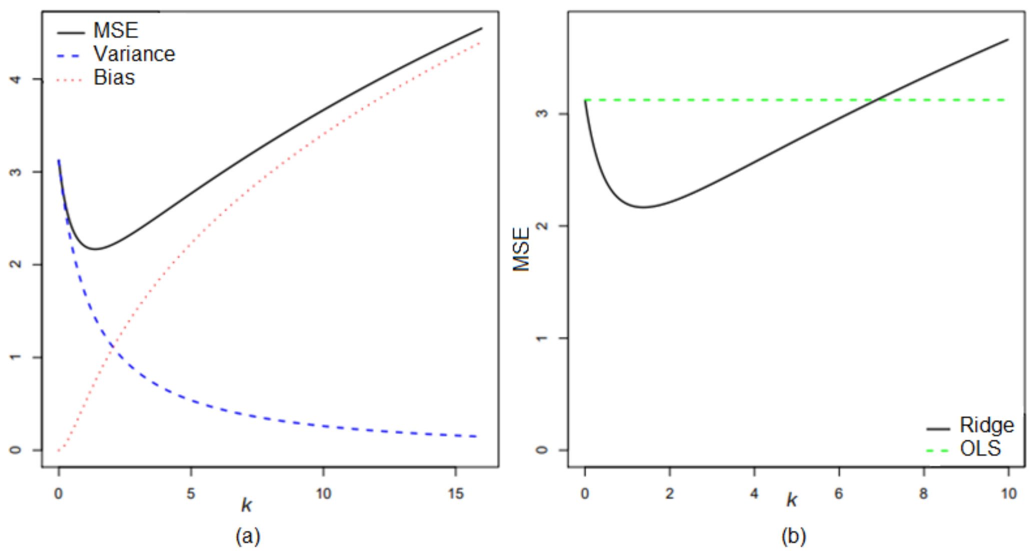

Figure 1: (a) The components of MSE for ridge estimator (b) The MSE comparisons of ridge and OLS estimators

The scalar MSE of the Ridge estimator is equal to

$$\text{SMSE}\left(\widehat{\beta}_{\mathbf{R}}\right)=\sum_{i=1}^{p}\frac{\sigma^2\lambda_i}{(\lambda_i+k)^2}+k^2\beta'\left(X'X+kI\right)^{-2}\beta \tag{2.4}$$

or

$$\text{SMSE}\left(\widehat{\alpha}_{R}\right)=\sum_{i=1}^{p}\frac{\sigma^2\lambda_i}{(\lambda_i+k)^2}+\sum_{i=1}^{p}\frac{k^2\alpha_i^2}{(\lambda_i+k)^2}. \tag{2.5}$$

The first term on the right-hand side of the Eq. (2.4)-(2.5) is the contribution of the total variance to the MSE. The second term on the right-hand side of the Eq. (2.4)-(2.5) is the contribution of the sum of the square of the bias to the MSE.

Figure 1 shows the MSE values as a function of k for ridge estimation, and the comparison of MSE values between ridge and OLS. The choice of the ridge parameter has a vital role in ridge regression, so there are different ridge estimators proposed by many authors in the literature (see [16] and [17] for instance). Table 1 provides a summary of the ridge estimators investigated in this study.

2.2. M Estimators

The M estimators are a class of robust estimators used as an alternative to OLS estimators for outlier or non-normality problems in regression analysis. M estimators were first described by [11] as a generalization of MLE. [12] extended the idea of using M estimators for solution of regression problems. In regression

Table 1: Selected ridge parameter estimators from the literature.

Estimators	Reference
$\widehat{k}_{\mathrm{HK}} = \frac{\widehat{\sigma}^2}{\widehat{\alpha}^2_{\max}}$	[8]
$\widehat{k}_{\mathrm{HKB}} = \frac{p\widehat{\sigma}^2}{\sum_{i=1}^{p}\widehat{\alpha}_i^2}$	[9]
$\widehat{k}_{\mathrm{LW}} = \frac{p\sigma^2}{\sum_{i=1}^{p}\lambda_i\widehat{\alpha}_i^2}$	[14]
$\widehat{k}_{\mathrm{HSL}} = \widehat{\sigma}^2\frac{\sum_{i=1}^{p}(\lambda_i\widehat{\alpha}_i)^2}{(\sum_{i=1}^{p}\lambda_i\widehat{\alpha}_i^2)^2}$	[7]
$\widehat{k}_{\mathrm{AM}} = \frac{1}{p}\sum_{i=1}^{p}\frac{\widehat{\sigma}^2}{\widehat{\alpha}_i^2}$	[15]
$\widehat{k}_{\mathrm{GM}} = \frac{\widehat{\sigma}^2}{(\prod_{i=1}^{p}\widehat{\alpha}_i^2)^{\frac{1}{p}}}$	[15]
$\widehat{k}_{\mathrm{MED}} = Median\left\{\frac{\widehat{\sigma}^2}{\widehat{\alpha}_i^2}\right\},\ i=1,\ldots,p$	[15]
$\widehat{k}_{\mathrm{KS}} = \frac{\lambda_{\max}\widehat{\sigma}^2}{(n-p)\widehat{\sigma}^2+\lambda_{\max}\widehat{\alpha}^2_{\max}}$	[13]

analysis, M estimates are obtained by minimizing a certain objective function in Eq. (2.6), such as the sum of the squared error terms.

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{M}} = \underset{\boldsymbol{\beta}}{argmin}\left(\sum_{i=1}^{n}\rho(\varepsilon_i)\right) \tag{2.6}$$

In M estimators, the objective function is the negative form of the natural logarithm of the likelihood function of the distribution of errors. The objective functions $\rho(.)$ have properties following below ([19]):

1. $\rho(0) = 0$
2. $\rho(\varepsilon) \geq 0$
3. $\rho(\varepsilon) = \rho(-\varepsilon)$
4. $\rho(\varepsilon_1) < \rho(\varepsilon_2)$ for $0 < \varepsilon_1 < \varepsilon_2$
5. $\rho(\varepsilon)$ is continuous and differentiable.

The influence function first described by [6] is a measure of the qualitative robustness. The influence function measures the marginal effect of the data on the parameter estimator. The influence function is defined by

$$\psi(\varepsilon) = \frac{d\rho(\varepsilon)}{d\varepsilon} \tag{2.7}$$

The influence function $\psi(.)$ has properties following ([2]):

1. $\psi(\varepsilon) \geq 0$ for $\varepsilon \geq 0$
2. $\psi(-\varepsilon) = -\psi(\varepsilon)$

Table 2: Objective and influence functions for the selected M-Estimators from literature.

Estimators	Objective Function	Influence Function
Huber	$\begin{cases} \frac{\varepsilon^2}{2}, & -k \le \varepsilon \le k \\ k\lvert\varepsilon\rvert - \frac{k^2}{2}, & \varepsilon < -k \ or \ \varepsilon > k \end{cases}$	$\begin{cases} \varepsilon, & -k \le \varepsilon \le k \\ ksign(\varepsilon), & \varepsilon < -k \ or \ \varepsilon > k \end{cases}$
Fair	$k_F^2 \ln\left(\frac{\lvert\varepsilon\rvert}{k_F} - ln\left(1 + \frac{\lvert\varepsilon\rvert}{k_F}\right)\right), \ \lvert\varepsilon\rvert < \infty$	$\frac{\varepsilon}{1+\frac{\lvert\varepsilon\rvert}{k_F}}$
Hampel	$\begin{cases} \frac{\varepsilon^2}{2} & \lvert\varepsilon\rvert \le a \\ a\lvert\varepsilon\rvert - \frac{a^2}{2} & a < \lvert\varepsilon\rvert \le b \\ ab - \frac{a^2}{2} + \frac{a(c-b)}{2}\left[1 - \left(\frac{c-\lvert\varepsilon\rvert}{c-b}\right)^2\right] & b < \lvert\varepsilon\rvert \le c \\ ab - -\frac{a^2}{2} + \frac{a(c-b)}{2} & \lvert\varepsilon\rvert > c \end{cases}$	$\begin{cases} \varepsilon & \lvert\varepsilon\rvert \le a \\ asign(\varepsilon) & a < \lvert\varepsilon\rvert \le b \\ \frac{asign(\varepsilon)(c-\lvert\varepsilon\rvert)}{c-b} & b < \lvert\varepsilon\rvert \le c \\ 0 & \lvert\varepsilon\rvert > c \end{cases}$
Tukey	$\begin{cases} \frac{k_T^2}{6}\left(1 - \left\{1 - (\varepsilon/k_T)^2\right\}^3\right) & \lvert\varepsilon\rvert \le k_T \\ \frac{k_T^2}{6} & \lvert\varepsilon\rvert > k_T \end{cases}$	$\begin{cases} \varepsilon\left(1 - (\varepsilon/k_T)^2\right)^2 & \lvert\varepsilon\rvert \le k_T \\ 0 & \lvert\varepsilon\rvert > k_T \end{cases}$
Andrew	$\begin{cases} k_A^2\left\{1 - cos\left(\frac{\varepsilon}{k_A}\right)\right\} & \lvert\varepsilon\rvert \le k_A\pi \\ 2k_A^2 & \lvert\varepsilon\rvert > k_A\pi \end{cases}$	$\begin{cases} k_A \sin\left(\frac{\varepsilon}{k_A}\right) & \lvert\varepsilon\rvert \le k_A\pi \\ 0 & \lvert\varepsilon\rvert > k_A\pi \end{cases}$
Welsch	$\frac{k_w^2}{2}\left(1 - e^{-\left(\frac{\varepsilon}{k_W}\right)^2}\right), \ \lvert\varepsilon\rvert < \infty$	$\varepsilon e^{-\left(\frac{\varepsilon}{k_W}\right)^2}$
Cauchy	$\frac{k_C^2}{2} \ln\left(1 + \left(\frac{\varepsilon}{k_C}\right)^2\right), \ \lvert\varepsilon\rvert < \infty$	$\frac{\varepsilon}{1+\left(\frac{\varepsilon}{k_C}\right)^2}$
Talwar	$\begin{cases} \frac{\varepsilon^2}{2} & \lvert\varepsilon\rvert \le k_{\mathrm{Tw}} \\ \frac{k_{\mathrm{Tw}}^2}{2} & \lvert\varepsilon\rvert > k_{\mathrm{Tw}} \end{cases}$	$\begin{cases} \varepsilon & \lvert\varepsilon\rvert \le k_{\mathrm{Tw}} \\ 0 & \lvert\varepsilon\rvert > k_{\mathrm{Tw}} \end{cases}$
Ramsay	$\frac{1-(1+k_R\lvert\varepsilon\rvert)e^{-k_R\lvert\varepsilon\rvert}}{k_R^2}, \ \lvert\varepsilon\rvert < \infty$	$\varepsilon e^{-k_R\lvert\varepsilon\rvert}$
Geman-McClure	$\frac{\varepsilon^2}{k_{\mathrm{GM}}^2+\varepsilon^2}, \ \lvert\varepsilon\rvert < \infty$	$\frac{2\varepsilon k_{\mathrm{GM}}^2}{\left(k_{\mathrm{GM}}^2+\varepsilon^2\right)^2}$

3. $\psi^{'}(0) = 1$

4. $\psi^{'}{'}(0) = 0$

5. $\psi^{'''}(0) < 0$

6. $\psi(\varepsilon)$ is continuous and partially differentiable

Another important function in M estimator is the weight function obtained by dividing the influence function by ε.

$$\omega(\varepsilon) = \frac{\psi(\varepsilon)}{\varepsilon} \tag{2.8}$$

commonly used objective functions, influence functions, and weight functions for selected M estimators are given in Tables 2–3.

Iteratively reweighted least squares (IRLS) algorithm is widely used for parameter estimations in M regression. IRLS algorithm is given below.

- **Step 1:** Initial estimates of the parameters vector $\widehat{\beta}^0$ is obtained by OLS

$$\widehat{\beta}^0 = (X'WX)^{-1}X'W\mathbf{y}$$

where $W = diag(\mathbf{1})$.

Table 3: Weight functions and tuning parameters for the selected M-Estimators from literature.

Estimators	Weight Function	Tuning Parameter	Reference
Huber	$\begin{cases} 1, & -k \leq \varepsilon \leq k \\ \frac{k}{\lvert\varepsilon\rvert}, & \varepsilon < -k \text{ or } \varepsilon > k \end{cases}$	$k = 1.5\widehat{\sigma}$ $\widehat{\sigma} = 1.4826MAD$	[3]
Fair	$\frac{1}{1+\frac{\lvert\varepsilon\rvert}{k_F}}$	$k_F = 1.3998$	[22]
Hampel	$\begin{cases} 1 & \lvert\varepsilon\rvert \leq a \\ \frac{a}{\lvert\varepsilon\rvert} & a < \lvert\varepsilon\rvert \leq b \\ \frac{a(c-\lvert\varepsilon\rvert)}{\lvert\varepsilon\rvert(c-b)} & b < \lvert\varepsilon\rvert \leq c \\ 0 & \lvert\varepsilon\rvert > c \end{cases}$	$a = 1.35$ $b = 2.7$ $c = 5.4$	[22]
Tukey	$\begin{cases} \left(1-\left(\frac{\varepsilon}{k_T}\right)^2\right)^2 & \lvert\varepsilon\rvert \leq k_T \\ 0 & \lvert\varepsilon\rvert > k_T \end{cases}$	$k_T = 2$	[1]
Andrew	$\begin{cases} \frac{\text{ksin}\left(\frac{\varepsilon}{k_A}\right)}{\varepsilon} & \lvert\varepsilon\rvert \leq k_A\pi \\ 0 & \lvert\varepsilon\rvert > k_A\pi \end{cases}$	$k_A = 1.339$	[23]
Welsch	$e^{-\left(\frac{\varepsilon}{k_W}\right)^2}$	$k_W = 2.9846$	[23]
Cauchy	$\frac{1}{1+\left(\frac{\varepsilon}{k_C}\right)^2}$	$k_C = 2.3849$	[22]
Talwar	$\begin{cases} 1 & \lvert\varepsilon\rvert \leq k_{\text{Tw}} \\ 0 & \lvert\varepsilon\rvert > k_{\text{Tw}} \end{cases}$	$k_{\text{Tw}} = 2.7955$	[10]
Ramsay	$e^{-k_R\lvert\varepsilon\rvert}$	$k_R = 0.3569$	[4]
Geman-McClure	$\frac{2k_{\text{GM}}^2}{\left(k_{\text{GM}}^2+\varepsilon^2\right)^2}$	$k_{\text{GM}} = 3.787376$	[4]

- **Step 2:** The estimates of stochastic error vector ε is calculated by

$$\widehat{\varepsilon} = \mathbf{y} - \mathbf{X}\widehat{\beta}^0.$$

- **Step 3:** The weights denoted by W^{new} are updated based on the weight function of selected M-estimator by using $\widehat{\varepsilon}$.

- **Step 4:** The estimates of the parameters vector are updated by

$$\widehat{\beta}^{\text{new}} = \left(X'W^{\text{new}}X\right)^{-1} X'W^{\text{new}}\mathbf{y}.$$

- **Step 5:** The convergence condition is examined for the termination of the algorithm. The convergence condition can be used as

$$\begin{cases} \widehat{\beta}_M = \widehat{\beta}^{\text{new}} \quad \text{and stop algorithm} & \text{if } \sum_{i=0}^{p} \left|\widehat{\beta}_i^{\text{new}} - \widehat{\beta}_i^0\right| \leq tol \\ \widehat{\beta}^0 = \widehat{\beta}^{\text{new}} \quad \text{and repeat Step 2 - 5} & \text{if } \sum_{i=0}^{p} \left|\widehat{\beta}_i^{\text{new}} - \widehat{\beta}_i^0\right| > tol \end{cases}$$

where tol is tolerance value and set to 10^{-4} in this study.

2.3. Proposed M-Ridge Estimators

In this study, we propose a novel M-estimators based on ridge estimators when multicollinearity and outlier problems coexist in the multiple linear regression model. Eq. (1.4) is a common approach for dealing with multicollinearity and non-normality. In many studies, the IRLS algorithm is used to obtain estimates by using Eq. (1.4). M-estimators are typically obtained by IRLS algorithm using OLS errors. However, it is clear that OLS estimators are unreliable in case of multicollinearity problem. Therefore, we propose to modify the IRLS algorithm by obtaining M-estimators using the errors from ridge estimators instead of OLS. OLS method gives equal weights ($w_i = 1$) to each observation vector. On the other hand, M estimators provide more robust estimates in case of outlier or non-normality by weighting the observation vectors according to the size of the errors. M-estimates using weights based on the errors of OLS can adversely affect the performance of RM estimators since OLS estimates are unstable in existence of multicollinearity. Based on this idea obtaining M-estimates from ridge estimates instead of OLS can increase the performance of the estimator in cases where both multicollinearity and outlier problems coexist.

In this study, M estimators based on ridge estimators (M-Ridge) are proposed to address the coexistence of outliers and multicollinearity problems in multiple linear regression models. The estimates of the stochastic error vector are computed based on ridge estimates in the proposed approach. The IRLS algorithm is arranged as follows to obtain the proposed M-Ridge estimates:

- **Step 1:** Initial estimates of the parameter vector $\widehat{\beta}^0_{M-Ridge}$ is obtained based on ridge estimator with the selected ridge parameter estimate

$$\widehat{\beta}^0_{M-Ridge} = \left(X^{'}WX + \widehat{k}I\right)^{-1} X^{'}W\mathbf{y}$$

where $W = diag(\mathbf{1})$.

- **Step 2:** The estimates of stochastic error vector (ε) is calculated by

$$\widehat{\varepsilon} = \mathbf{y} - \mathbf{X}\widehat{\beta}^0_{M-Ridge}.$$

- **Step 3:** The weights are updated based on the weight function of selected M-estimator by using $\widehat{\varepsilon}$ and denoted by W^{new}.

- **Step 4:** The estimates of the parameter vector is updated by

$$\widehat{\beta}^{\text{new}}_{M-Ridge} = \left(X^{'}W^{\text{new}}X + \widehat{k}I\right)^{-1} X^{'}W^{\text{new}}\mathbf{y}.$$

- **Step 5:** The convergence condition for termination of the algorithm used as

$$\begin{cases} \widehat{\beta}_{M-Ridge} = \widehat{\beta}^0_{M-Ridge} & \text{and stop algorithm} \quad \text{if } \sum_{i=0}^{p} \left|\widehat{\beta}^{\text{new}}_{M-Ridge} - \widehat{\beta}^0_{M-Ridge}\right| \leq tol \\ \widehat{\beta}^0_{M-Ridge} = \widehat{\beta}^{\text{new}}_{M-Ridge} & \text{and repeat Step 2 - 5} \quad \text{if } \sum_{i=0}^{p} \left|\widehat{\beta}^{\text{new}}_{M-Ridge} - \widehat{\beta}^0_{M-Ridge}\right| > tol \end{cases}$$

where tol represents tolerance value and set to 10^{-4} in this study.

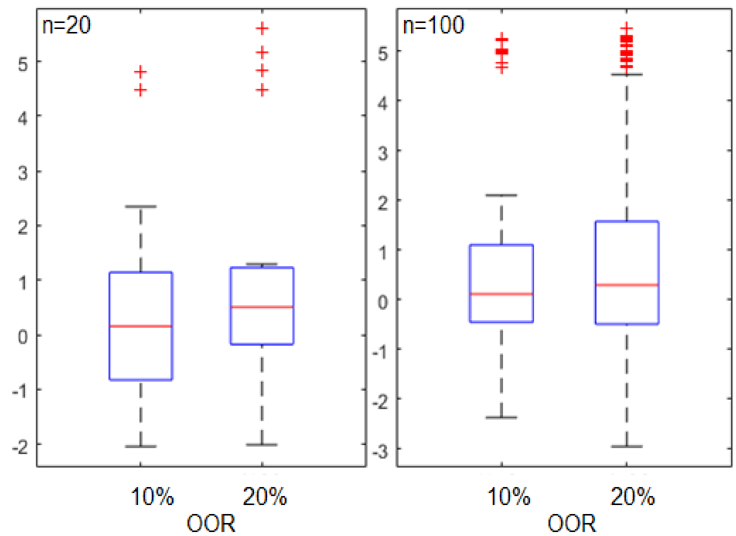

Figure 2: Box-plots of the dependent variable for a selection of artificial data sets with different sample sizes and outlier rates.

3. Simulation Study

In the simulation study, the performance of the proposed method for artificial data sets that contain outliers and multicollinearity problems is examined. The explanatory variables in the artificial dataset are generated as

$$x_{\mathrm{ij}} = \left(1-\rho^2\right)^{1/2} z_{\mathrm{ij}} + \rho z_{\mathrm{i(p+1)}} \tag{3.1}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, p$ where ρ^2 denotes the correlation between explanatory variables and z_{ij}'s are random numbers from standard normal distribution [20]. Coefficient vector β in Eq. (3.2) is selected as the normalized eigenvector corresponding to the largest eigenvalue of $X'X$ so that $\beta'\beta = 1$. The dependent variable vector is determined by

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{\mathrm{ip}} + \varepsilon_i \tag{3.2}$$

where ε_i generated from two component mixture normal distribution in order to create an outlier problem. According to the selected outlier observation rate (OOR), the errors corresponding outlier observations are produced from the normal distribution with $N(5, 0.1)$ while rest of are from a standard normal distribution. Some selected box plots of dependent variables for different OOR's are shown in Figure 2.

Condition index (CI) is a measure of multicollinearity and values of CI greater than 30 are indicative of strong multicollinearity [18]. In the simulation study, CI's are computed as 41.728 and 39.844 for $n = 20$, $\rho^2 = 0.99$ and 0.999 respectively. For $n = 100$, $\rho^2 = 0.99$ and 0.999, CI's are 130.174 and 127.231 respectively.

The performance of the proposed estimator is evaluated according to the selected values of parameters for $\rho^2 = 0.99$ and 0.99, $n = 20$ and 100, OOR = 10% and 20% and $p = 4$. A Monte Carlo simulation with 5000 replications is performed to show the performance of the proposed estimator. M-Ridge is compared to OLS, Ridge, M, and RM estimators in terms of the estimated MSE. Monte Carlo simulation results are given in Tables 4–11.

$$\widehat{\text{MSE}}\left(\widehat{\beta}\right) = \frac{1}{5000}\sum_{t=1}^{5000}\sum_{i=1}^{p}\left(\widehat{\beta}_{i,t} - \beta_i\right)^2 \tag{3.3}$$

According to simulation results for OOR=10%, $n = 20$ and $\rho^2 = 0.99$, the proposed M-Ridge estimators are generally more successful than others. Cauchy M-Ridge estimator with k_{AM} ridge parameter is the best estimator with the smallest estimated MSE value of 0.3469. The M-Ridge estimators are successful 79 out of 80 comparisons against the RM estimators. Among the selected ridge parameters in robust estimators, the proposed k_{AM} ridge parameter by [15] is the most successful ridge parameter with smaller estimated MSE. The M-Ridge estimators are successful 9 out of 10 comparisons against the RM estimators for k_{AM} ridge parameter.

When the simulation results for OOR=20%, $n = 20$ and $\rho^2 = 0.99$ is analyzed, it is observed that the estimated MSE values are increased when the OOR's are increased as expected. Talwar M-Ridge estimator with k_{AM} ridge parameter is the best estimator with the smallest estimated MSE value of 0.400. The M-Ridge estimators are successfull 78 out of the 80 comparisons against the RM estimators.

According to results of the simulation for OOR=10%, $n = 100$ and $\rho^2 = 0.99$, the proposed M-Ridge estimators are found more successful than RM estimators in all comparisons with RM estimators. Geman-McClure M-Ridge estimator with k_{AM} ridge parameter is the best estimator with the smallest estimated MSE value of 0.293.

According to results of the simulation for OOR=20%, $n = 100$ and $\rho^2 = 0.99$, the proposed M-Ridge estimators are found more successful than RM estimators in all comparisons with RM estimators. Geman-McClure M-Ridge estimator with k_{HSL} ridge parameter is the best estimator with the smallest estimated MSE value of 0.198.

When the simulation results for OOR=10%, $n = 20$ and $\rho^2 = 0.999$ are analyzed, it is observed that the estimated MSE values increase as expected when the ρ^2 increases. The proposed M-Ridge estimators are generally more successful than others. The M-Ridge estimators are successful 72 out of the 80 comparisons against the RM estimators. According to results of the simulation for OOR=20%, $n = 20$ and $\rho^2 = 0.999$, the proposed M-Ridge estimators are successful 76 out of the 80 comparisons against the RM estimators.

According to results of the simulation for OOR=10%, $n = 100$ and $\rho^2 =$

0.999, the proposed M-Ridge estimators are found more successful than RM estimators in all comparisons with RM estimators. Geman-McClure M-Ridge estimator with k_{HSL} ridge parameter iss the best estimator with the smallest estimated MSE value of 0.224.

According to results of the simulation performed for OOR=20%, $n = 100$ and $\rho^2 = 0.999$, the proposed M-Ridge estimators are found more successful than RM estimators in all comparisons with RM estimators. Geman-McClure M-Ridge estimator with k_{HSL} ridge parameter is the best estimator with the smallest estimated MSE value of 0.186. Among the selected ridge parameters in M-Ridge estimators, the proposed k_{HSL} ridge parameter by Hocking, [7] is the most successful ridge parameter with smaller estimated MSE values.

The results from Tables 4–11 can be summarized as follows. In most cases, MSE values increased with increasing OOR and ρ^2, while controlling for other factors. The M-Ridge and RM estimators are found to be more successful than the OLS, ridge, and M estimators according to the MSE criteria. The M-Ridge estimator is found to be superior to the RM estimator according to the MSE criteria.

Box plots of MSE values for M-Ridge and RM estimators with different ridge parameters given in Figure 3 shows that the M-Ridge estimators are more robust to the changes in ridge parameter than the RM estimators, and have smaller MSE values. When the average MSE values obtained with different ridge parameters are examined, the M-Ridge estimates are found to be more successful than the RM estimates in all cases.

The proposed estimator M-Ridge has smaller MSE than the RM estimator in all cases examined except for the bias parameter k_{AM}. In comparison M-Ridge and RM estimators with ridge parameter k_{AM}, the proposed M-Ridge estimator has smaller MSE than the RM estimator in 65 of the 80 cases. Box plot of MSE values for M-Ridge and RM estimators with ridge parameter k_{AM} is given Figure 4. As evident from Figure 4, the proposed M-Ridge estimator outperforms the RM estimator in terms of mean squared error (MSE) for the bias parameter k_{AM}.

In the second stage of the simulation study, the performance of the estimators is examined for the errors from log-normal distribution with parameters $\mu = 0$ and $\sigma^2 = 1$. The simulation study involves specific parameter settings, including $\rho^2 = 0.99$ and 0.999, $n = 20$ and 100, and $p = 4$. Monte Carlo simulation results based on 5000 replications are presented in Tables 12–15.

Robust estimators are found more successful in the simulation study for the log-normal distributed errors. Among the robust estimators, the proposed M-Ridge estimators generally outperformed the RM estimators. Box plots of MSE values for M-Ridge and RM estimators with different ridge parameters are given Figure 5.

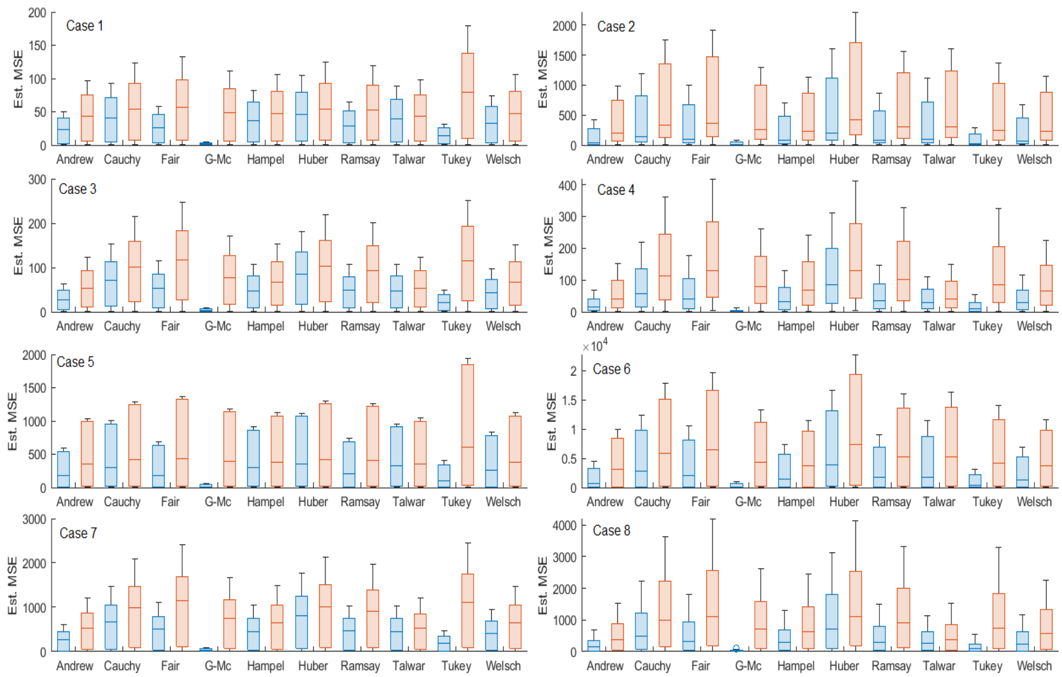

Figure 3: Box plots of MSE values for M-Ridge and RM estimators with different ridge estimators(Cases 1-8)

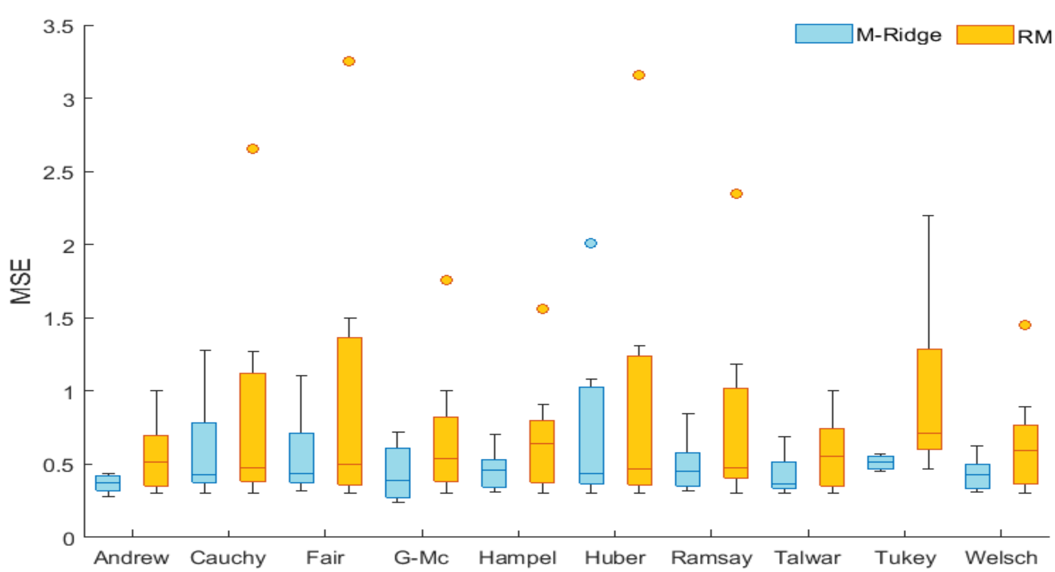

Figure 4: Box plots of MSE values for M-Ridge and RM estimators with ridge estimator k_{AM}

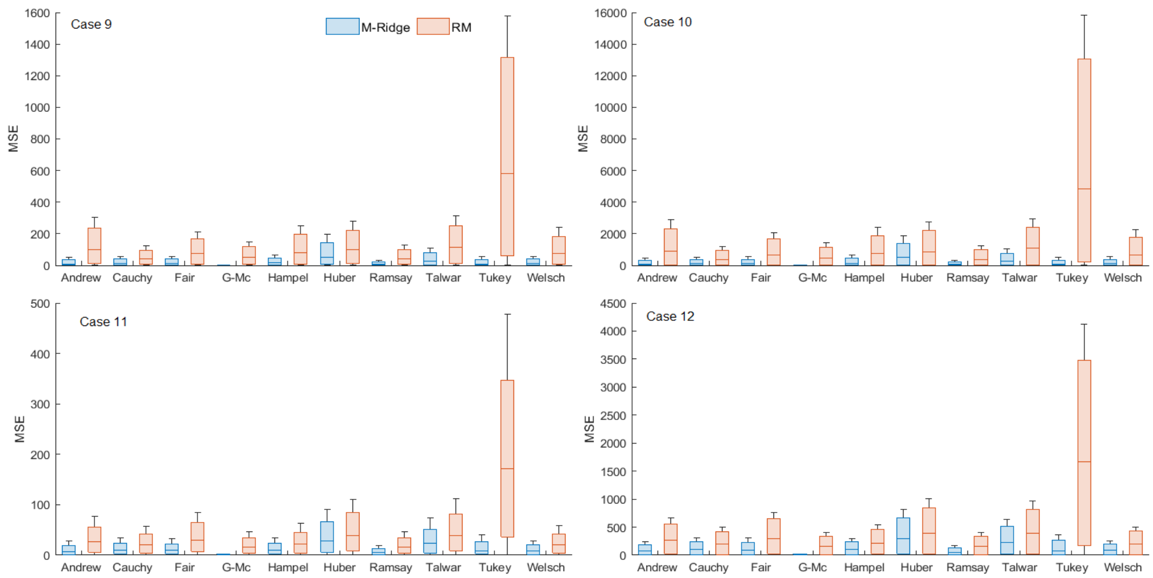

Figure 5: Box plots of MSE values for M-Ridge and RM estimators with different ridge parameters (Cases 9-12)

4. Real Data Application

In this section, we consider Tobacco data [21], which contains both multicollinearity and outlier problems. In application, the dataset centralized and standardized by proportioning the square root of the overall sum of squares. The scatter matrix for the independent variables and the box-plot of the error terms for OLS estimates are given in Figure 6.

According to Figure 6, it is clearly seen that the data set contains outlier and multicollinearity problems. The calculated CI value of 43.0758 for Tobacco dataset indicates the presence of a strong multicollinearity issue.

The performance of the proposed M-Ridge estimators evaluated with the k-fold cross validation technique in the Tobacco dataset. The original data set is randomly partitioned into k equal sized subgroups in k-fold cross validation. Of the k subgroups, a single subgroup is retained as the validation data for testing the model, and the remaining k-1 subgroups are used as training data. This process is then repeated k times, with each of the k subgroups used exactly once as the validation data. In k-fold cross validation, the cross validation(CV) as the mean of sum of errors for test data sets is given below.

$$E_i = \sum_{j=1}^{m} (y_{j,test} - \widehat{y}_{j,test})^2 \tag{4.1}$$

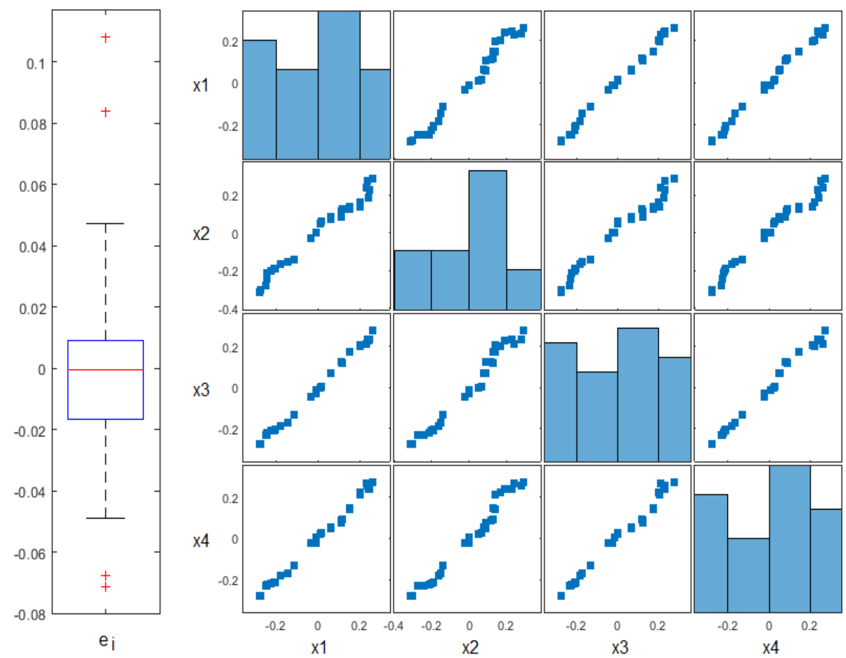

Figure 6: The scatter matrix for the independent variables and the box-plot of the error terms for the OLS estimates

$$CV = \frac{1}{k}\sum_{i=1}^{k} E_i \tag{4.2}$$

The performances of the estimators are tested with k-fold cross validation by choosing $k = 10$. The k-fold cross validation results for the proposed estimator and other investigated estimators in the Tobacco dataset are given in Table 16.

According to results of k-fold cross validation, M-ridge and RM estimators as robust estimator are found more successful than OLS, ridge, and M estimators. In most cases M-Ridge found superior to RM estimators in comparisons between the proposed M-Ridge with RM estimators. The proposed Huber's M-ridge estimator by k_{AM} bias parameter is found the most successful estimator which has the smallest CV.

5. Conclusions

In this study, new robust estimators which named by M-Ridge are improved in case of coexistence of multicollinearity and outlier problems. The M-Ridge estimators are based on the notion that rather than converting M-estimates derived from OLS estimates into ridge estimates as proposed by [24] and [5], M-estimates are obtained directly from errors based on ridge estimates. The performance of M-Ridge estimators is examined according to estimated MSE and k-fold cross validation in simulations and real data application, respectively. in the simulation study, M-ridge estimators with generally smaller estimated MSE values outperformed OLS, ridge, M, and RM estimators. In the comparison of robust M-Ridge and RM estimators according to the MSE criterion, the success rates of M-Ridge estimators for normal distributed errors are 95.3125% in the small sample sizes and 100% in the large sample sizes. The success rates of M-Ridge estimators are found 96.875% in small sample sizes and 89.375% in large sample sizes for log-normal distributed errors. In the real data application, the performances of the estimators are tested with k-fold cross validation by choosing k=10. According to k-fold cross validation, Huber's M-ridge estimator based on k_{AM} is found the best estimator. In future studies, new estimators can be obtained by using different biased estimators instead of ridge estimators with the proposed new approach.

ACKNOWLEDGMENTS

The authors received no financial support for this article.

REFERENCES

[1] Beaton, A.E., and Tukey, J.W. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data, *Technometrics*, **16**, 2, 147–185.

[2] Bell, R.M. (1980). *An adaptive choice of the scale parameter for M-estimators*, Stanfod univ ca dept of statistics.

[3] Birkes, D. and Dodge, Y. (2011). *Alternative methods of regression*, John Wiley & Sons.

[4] De Menezes, D., Prata, D.M., Secchi, A.R. and Pinto, J.C. (2021). A review on robust M-estimators for regression analysis, *Computers & Chemical Engineering*, **147**, 107254.

[5] Ertaş, H. (2018). A modified ridge m-estimator for linear regression model with multicollinearity and outliers, *Communications in Statistics-Simulation and Computation*, **47**, 4, 1240–1250.

[6] Hampel, F.R. (1968). *Contributions to the theory of robust estimation*, University of California, Berkeley.

[7] Hocking, R.R., Speed, F. and Lynn, M. (1976). A class of biased estimators in linear regression, *Technometrics*, **18**, 4, 425–437.

[8] Hoerl, A.E., Kennard, R.W. (1970). Ridge regression: applications to nonorthogonal problems, *Technometrics*, **12**, 1, 69–82.

[9] HOERL, A.E., KENNARD, R.W. and BALDWIN, K.F. (1975). Ridge regression: some simulations, *Communications in Statistics-Theory and Methods*, **4**, 2, 105–123.

[10] HOLLAND, P.W. and WELSCH, R.E. (1977). Robust regression using iteratively reweighted least-squares, *Communications in Statistics-theory and Methods*, **6**, 9, 813–827.

[11] HUBER, P.J. (1964). Robust Estimation of a Location Parameter, *The Annals of Mathematical Statistics*, **35**, 1, 73 – 101.

[12] HUBER, P.J. (1973). Robust regression: asymptotics, conjectures and Monte Carlo, *The annals of statistics*, **1**, 5, 799–821.

[13] KHALAF, G. and SHUKUR, G (2005). Choosing ridge parameter for regression problems, *Communications in Statistics - Theory and Methods*, **35**, 5, 1177–1182.

[14] LAWLESS, J., and WANG, P. (1976). A simulation study of ridge and other regression estimators, *Communications in Statistics-theory and Methods*, **5**, 4, 307–323.

[15] KIBRIA B, M. G. (2003). Performance of some new ridge regression estimators, *Communications in Statistics-Simulation and Computation*, **(**32), 2, 419–435.

[16] KIBRIA B, M. G. (2022). More than hundred(100) estimators for estimating the shrinkage parameter in a linear and generalized linear ridge regression models, *Journal of Econometrics and Statistics*, **(**2), 2, 233–252.

[17] KIBRIA B, M. G.and BANIK, S. (2016). Some Ridge Regression Estimators and Their Performances, *Journal of Modern Applied Statistical Methods*, **(**15), 1, 206–238.

[18] KIM, J.H. (2019). Multicollinearity and misleading statistical results. *Korean journal of anesthesiology*, **72**, 6, 558–569.

[19] MARTIN, R.D. and YOHAI, V.J. (2006). *Robust statistics: Theory and Methods*, John Wiley New York.

[20] MCDONALD, G.C. and GALARNEAU, D.I. (1975). A Monte Carlo evaluation of some ridge-type estimators, *Journal of the American Statistical Association*, **70**, 350, 407–416.

[21] MYERS, R.H. (1990). *Classical and modern regression with applications*, Duxbury press, Belmont, CA.

[22] ÖZYURT, D.B. and PIKE, R.W. (2004). Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes, *Computers & chemical engineering*, **28**, 3, 381–402.

[23] REY, W.J. (1983). *Introduction to Robust and Quasi-Robust Statistical Methods*, Springer. 131–133.

[24] SILVAPULLE, M.J. (1991). Robust ridge regression based on an M-estimator, *Australian Journal of Statistics*, **33**; 3, 319–333.

Table 4: Scaler MSE values of the estimators for OOR 10%, $n = 20$ and $\rho^2 = 0.99$ (Case 1).

OLS	1507.06	Ridge	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
			117.891	203.820	175.044	8.077	0.389	14.856	55.152	121.791
M		Robust	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
Huber	893.860	M-Ridge	65.026	105.234	92.032	4.379	0.353	7.331	27.022	67.014
		RM	76.065	124.456	107.344	5.159	0.362	9.094	32.482	78.446
Fair	943.136	M-Ridge	38.283	57.969	53.438	2.584	0.355	3.023	14.198	39.404
		RM	79.520	132.394	114.473	5.438	0.366	9.708	34.836	82.030
Hampel	836.801	M-Ridge	54.173	82.300	73.597	3.397	0.349	5.069	19.628	55.864
		RM	67.543	105.384	91.648	4.259	0.361	7.597	25.959	69.763
Tukey	1725.782	M-Ridge	22.249	30.281	30.816	1.812	0.478	1.482	6.477	22.991
		RM	113.010	179.605	158.260	7.309	0.547	12.915	44.879	117.495
Andrew	796.919	M-Ridge	34.516	49.614	46.638	2.183	0.350	2.647	11.226	35.558
		RM	62.959	96.741	84.938	3.931	0.356	6.884	23.387	65.036
Welsch	823.663	M-Ridge	48.872	73.823	66.670	3.091	0.3471	4.369	17.552	50.371
		RM	67.208	105.509	91.881	4.288	0.359	7.621	26.180	69.392
Cauchy	883.687	M-Ridge	58.091	92.384	81.816	3.817	0.3469	5.739	23.240	59.837
		RM	75.384	123.741	106.886	5.054	0.361	9.069	32.235	77.746
Talwar	798.563	M-Ridge	58.062	88.502	78.570	3.723	0.349	5.993	21.085	59.900
		RM	63.332	97.581	85.610	4.462	0.356	6.962	23.633	65.399
Ramsay	873.291	M-Ridge	42.236	64.156	58.646	2.781	0.350	3.506	15.571	43.483
		RM	73.695	119.985	103.895	4.921	0.363	8.779	31.011	76.031
G-Mc	838.096	M-Ridge	3.511	4.043	4.623	0.620	0.533	0.365	1.089	3.496
		RM	69.923	111.871	97.097	4.574	0.360	8.140	28.382	72.158

Table 5: Scaler MSE values of the estimators for OOR 20%, $n = 20$ and $\rho^2 = 0.99$ (Case 2).

OLS	4694.981	Ridge	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
			3023.756	1657.074	676.982	327.444	0.458	106.551	443.482	3031.023
M		Robust	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
Huber	3499.616	M-Ridge	1599.733	637.071	251.808	128.696	0.435	28.668	147.896	1605.432
		RM	2215.030	1208.167	513.362	251.770	0.446	75.058	320.177	2220.244
Fair	3013.929	M-Ridge	1002.236	337.370	121.444	69.689	0.459	10.292	68.019	1006.376
		RM	1907.099	1045.375	452.743	211.948	0.432	66.114	275.936	1911.399
Hampel	1905.393	M-Ridge	706.030	247.065	104.786	41.117	0.487	8.286	53.669	708.909
		RM	1133.244	611.446	292.873	111.803	0.685	38.936	159.321	1136.115
Tukey	2541.135	M-Ridge	286.908	68.632	27.276	15.501	0.574	2.080	13.644	288.771
		RM	1369.543	689.342	306.027	123.549	0.691	42.663	180.762	1373.767
Andrew	1689.642	M-Ridge	418.640	116.646	50.478	22.527	0.432	3.639	25.613	420.770
		RM	983.336	525.215	258.749	96.927	0.622	33.425	138.235	985.961
Welsch	1903.509	M-Ridge	665.134	223.833	93.211	39.497	0.468	7.313	47.753	667.858
		RM	1141.321	619.021	296.636	116.229	0.645	39.414	161.378	1144.114
Cauchy	2760.425	M-Ridge	1191.597	464.814	179.835	88.974	0.449	17.053	100.373	1195.715
		RM	1744.399	958.343	424.448	195.682	0.469	60.379	251.544	1748.220
Talwar	2636.496	M-Ridge	1110.995	316.013	125.896	53.185	0.400	8.258	55.988	1114.771
		RM	1602.653	856.312	392.798	192.809	0.700	51.032	224.708	1606.681
Ramsay	2493.603	M-Ridge	857.939	289.901	109.017	56.928	0.469	8.972	58.796	861.333
		RM	1562.552	857.858	385.871	173.938	0.491	54.120	224.474	1565.977
G-Mc	2105.865	M-Ridge	86.130	13.987	4.905	4.436	0.722	0.614	2.669	86.714
		RM	1297.936	711.442	332.535	139.783	0.552	45.168	185.207	1300.832

Table 6: Scaler MSE values of the estimators for OOR 10%, $n = 100$ and $\rho^2 = 0.99$ (Case 3).

OLS	1900.100	Ridge	$\widehat{k}_{\text{HK}}$ 258.580	$\widehat{k}_{\text{HKB}}$ 486.807	$\widehat{k}_{\text{LW}}$ 407.134	$\widehat{k}_{\text{HSL}}$ 6.339	$\widehat{k}_{\text{AM}}$ 3.333	$\widehat{k}_{\text{GM}}$ 107.302	$\widehat{k}_{\text{MED}}$ 222.727	$\widehat{k}_{\text{KS}}$ 321.897
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	891.006	M-Ridge	93.567	181.536	154.375	2.118	0.968	33.451	75.540	114.005
		RM	113.235	219.419	185.531	2.959	1.308	44.842	93.393	138.377
Fair	988.946	M-Ridge	59.672	114.154	99.039	1.018	0.568	16.256	45.757	72.816
		RM	128.024	247.708	209.377	3.323	1.502	51.467	107.126	156.737
Hampel	772.782	M-Ridge	54.855	107.133	93.165	1.088	0.581	16.220	37.948	67.716
		RM	77.652	152.433	130.182	1.912	0.911	27.604	56.031	96.922
Tukey	1492.272	M-Ridge	25.028	48.464	46.163	0.591	0.513	5.186	16.207	31.567
		RM	129.960	251.541	216.397	2.900	1.676	46.762	98.990	167.889
Andrew	662.853	M-Ridge	33.579	63.891	57.262	0.624	0.403	8.336	21.429	41.319
		RM	63.863	123.986	106.262	1.512	0.767	21.807	44.028	80.121
Welsch	744.195	M-Ridge	50.271	97.880	85.458	0.979	0.526	14.488	35.122	61.654
		RM	77.018	151.056	128.898	1.927	0.892	27.679	56.528	95.555
Cauchy	885.819	M-Ridge	78.338	152.276	130.334	1.580	0.758	25.550	61.696	95.331
		RM	110.798	215.614	182.538	2.901	1.274	43.720	90.783	135.441
Talwar	660.464	M-Ridge	54.922	106.275	91.784	1.178	0.633	17.414	37.127	68.444
		RM	63.666	123.667	105.832	1.516	0.771	21.831	44.165	79.918
Ramsay	856.196	M-Ridge	55.031	106.159	92.334	0.992	0.531	15.402	41.355	66.862
		RM	103.387	201.764	171.137	2.703	1.183	40.237	83.454	126.539
G-Mc	777.916	M-Ridge	4.829	8.354	8.553	0.247	0.293	0.832	2.949	5.822
		RM	87.590	171.532	145.969	2.259	0.999	32.884	67.883	107.598

Table 7: Scaler MSE values of the estimators for OOR 20%, $n = 100$ and $\rho^2 = 0.99$ (Case 4).

OLS	3140.552	Ridge	$\widehat{k}_{\text{HK}}$ 226.692	$\widehat{k}_{\text{HKB}}$ 811.123	$\widehat{k}_{\text{LW}}$ 874.977	$\widehat{k}_{\text{HSL}}$ 9.332	$\widehat{k}_{\text{AM}}$ 7.970	$\widehat{k}_{\text{GM}}$ 181.913	$\widehat{k}_{\text{MED}}$ 294.674	$\widehat{k}_{\text{KS}}$ 359.045
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	1474.408	M-Ridge	75.450	283.226	310.723	2.560	2.013	51.429	96.867	114.645
		RM	114.687	385.839	410.733	4.178	3.159	85.005	143.605	171.957
Fair	1507.492	M-Ridge	35.669	152.533	178.194	1.234	1.104	20.904	46.666	54.648
		RM	114.665	391.136	418.459	4.239	3.249	85.926	144.340	172.800
Hampel	1189.339	M-Ridge	29.586	111.818	128.957	0.825	0.704	15.732	34.737	43.583
		RM	63.598	221.697	240.768	2.063	1.565	42.226	74.015	95.084
Tukey	1790.016	M-Ridge	9.576	41.978	54.622	0.552	0.565	4.569	11.537	14.676
		RM	75.537	289.926	323.985	2.523	2.201	53.629	92.767	121.165
Andrew	783.499	M-Ridge	15.158	56.914	67.810	0.465	0.440	7.418	17.193	22.266
		RM	38.859	137.467	150.653	1.236	1.005	25.053	44.108	58.605
Welsch	1070.609	M-Ridge	26.181	99.490	114.792	0.736	0.628	13.944	31.093	38.397
		RM	60.273	206.505	223.232	1.974	1.452	39.910	70.273	88.976
Cauchy	1352.270	M-Ridge	50.324	196.219	219.783	1.574	1.280	31.148	64.321	75.687
		RM	101.800	338.651	360.149	3.611	2.653	73.538	125.615	150.779
Talwar	778.506	M-Ridge	27.514	99.066	111.324	0.802	0.689	15.950	31.236	40.935
		RM	38.689	136.494	149.389	1.230	1.002	24.974	44.057	58.376
Ramsay	1287.711	M-Ridge	31.230	127.685	147.766	0.970	0.842	17.486	39.782	46.736
		RM	92.675	308.265	328.386	3.238	2.349	65.859	113.328	136.630
G-Mc	1120.644	M-Ridge	1.808	7.754	11.174	0.198	0.269	0.910	2.177	2.587
		RM	72.654	243.053	260.381	2.461	1.758	49.745	86.907	106.340

Table 8: Scaler MSE values of the estimators for OOR 10%, $n = 20$ and $\rho^2 = 0.999$ (Case 5).

OLS	14807.131	Ridge	$\widehat{k}_{\text{HK}}$ 2068.689	$\widehat{k}_{\text{HKB}}$ 1999.721	$\widehat{k}_{\text{LW}}$ 706.270	$\widehat{k}_{\text{HSL}}$ 21.945	$\widehat{k}_{\text{AM}}$ 0.300	$\widehat{k}_{\text{GM}}$ 56.045	$\widehat{k}_{\text{MED}}$ 543.728	$\widehat{k}_{\text{KS}}$ 2073.417
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	8764.191	M-Ridge	1111.019	1031.415	439.101	11.680	0.303	27.145	264.628	1113.431
		RM	1292.304	1221.216	520.768	13.549	0.302	34.240	319.383	1295.113
Fair	9249.708	M-Ridge	688.926	566.519	221.816	7.872	0.320	9.988	138.747	690.468
		RM	1366.085	1299.036	532.820	13.916	0.303	36.592	342.872	1369.059
Hampel	8199.155	M-Ridge	911.395	806.542	397.144	9.601	0.309	17.856	191.799	913.506
		RM	1124.288	1033.059	507.277	11.099	0.306	28.453	254.572	1126.910
Tukey	16930.813	M-Ridge	400.689	292.637	133.467	4.608	0.453	3.838	60.458	401.800
		RM	1939.736	1760.914	763.987	16.661	0.470	47.636	437.201	1945.317
Andrew	7813.175	M-Ridge	588.172	484.797	243.485	6.487	0.318	8.721	108.747	589.560
		RM	1035.535	948.363	490.190	10.272	0.305	25.642	228.576	1038.009
Welsch	8071.554	M-Ridge	830.188	723.092	348.570	8.912	0.310	15.163	171.399	832.089
		RM	1121.485	1034.472	501.415	11.180	0.305	28.555	256.846	1124.063
Cauchy	8662.866	M-Ridge	1006.607	905.638	380.991	10.934	0.305	20.375	228.035	1008.800
		RM	1285.447	1213.984	516.863	13.268	0.301	34.186	317.189	1288.227
Talwar	7831.714	M-Ridge	957.074	867.037	449.919	9.688	0.302	21.840	205.884	959.290
		RM	1039.385	956.961	483.800	11.357	0.304	25.956	230.884	1041.855
Ramsay	8559.872	M-Ridge	744.164	627.712	266.277	8.316	0.316	11.811	152.213	745.821
		RM	1252.891	1177.020	514.515	12.637	0.304	33.048	304.977	1255.644
G-Mc	8213.268	M-Ridge	62.305	37.282	16.202	1.275	0.485	0.608	8.142	62.445
		RM	1177.158	1097.163	504.450	11.864	0.304	30.579	278.798	1179.791

Table 9: Scaler MSE values of the estimators for OOR 20%, $n = 20$ and $\rho^2 = 0.999$ (Case 6).

OLS	47401.887	Ridge	$\widehat{k}_{\text{HK}}$ 31219.846	$\widehat{k}_{\text{HKB}}$ 16777.823	$\widehat{k}_{\text{LW}}$ 21574.414	$\widehat{k}_{\text{HSL}}$ 510.123	$\widehat{k}_{\text{AM}}$ 0.297	$\widehat{k}_{\text{GM}}$ 387.819	$\widehat{k}_{\text{MED}}$ 3709.695	$\widehat{k}_{\text{KS}}$ 31226.605
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	35046.937	M-Ridge	16614.515	6433.902	9701.346	203.245	0.383	100.656	1201.710	16620.019
		RM	22737.765	12175.911	15828.326	398.921	0.357	270.606	2655.248	22742.587
Fair	30145.911	M-Ridge	10515.510	3417.909	5621.556	120.061	0.411	32.969	542.742	10519.477
		RM	19559.568	10531.888	13638.694	332.774	0.350	238.470	2283.153	19563.543
Hampel	18720.207	M-Ridge	7303.376	2496.267	3983.442	64.521	0.470	26.347	428.332	7306.089
		RM	11468.182	6084.823	7868.771	166.890	0.685	136.012	1282.201	11470.814
Tukey	25115.230	M-Ridge	3072.891	694.490	1377.559	29.169	0.546	5.185	102.614	3074.634
		RM	14038.131	6925.769	9098.242	191.214	0.664	150.676	1473.750	14042.051
Andrew	16616.234	M-Ridge	4398.887	1184.785	2131.911	37.422	0.403	11.202	203.191	4400.891
		RM	9958.532	5222.772	6792.274	159.023	0.616	116.154	1105.293	9960.951
Welsch	18732.179	M-Ridge	6903.999	2264.845	3692.730	62.878	0.446	23.223	380.758	6906.578
		RM	11563.878	6168.397	7977.234	176.626	0.637	138.219	1301.755	11566.440
Cauchy	27512.585	M-Ridge	12372.348	4701.652	7178.307	146.249	0.411	56.566	809.476	12376.232
		RM	17835.668	9632.454	12482.338	305.635	0.408	216.826	2071.882	17839.191
Talwar	26024.879	M-Ridge	11460.867	3189.382	5901.231	90.671	0.374	28.123	442.991	11463.500
		RM	16231.667	8519.122	11150.957	344.609	0.711	177.636	1815.415	16235.367
Ramsay	24797.853	M-Ridge	8964.764	2935.818	4811.464	95.923	0.434	28.542	468.760	8967.998
		RM	15955.039	8613.469	11160.167	270.261	0.441	193.793	1843.388	15958.193
G-Mc	20833.458	M-Ridge	943.263	138.798	339.125	7.597	0.684	1.123	17.713	943.858
		RM	13200.980	7118.483	9213.394	214.553	0.524	160.277	1507.855	13203.641

Table 10: Scaler MSE values of the estimators for OOR 10%, $n = 100$ and $\rho^2 = 0.999$ (Case 7).

OLS	18517.999	Ridge	$\widehat{k}_{\mathrm{HK}}$ 2504.533	$\widehat{k}_{\mathrm{HKB}}$ 4706.645	$\widehat{k}_{\mathrm{LW}}$ 4097.869	$\widehat{k}_{\mathrm{HSL}}$ 1.899	$\widehat{k}_{\mathrm{AM}}$ 1.235	$\widehat{k}_{\mathrm{GM}}$ 467.179	$\widehat{k}_{\mathrm{MED}}$ 2140.604	$\widehat{k}_{\mathrm{KS}}$ 2560.969
M		Robust	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
Huber	8890.015	M-Ridge	910.818	1762.354	1564.003	0.669	0.443	139.555	729.165	928.705
		RM	1104.600	2134.951	1877.057	0.880	0.497	196.037	903.696	1126.708
Fair	9836.618	M-Ridge	577.118	1101.013	1006.757	0.419	0.386	60.238	438.211	588.751
		RM	1247.491	2407.877	2116.585	0.964	0.562	225.078	1035.408	1272.883
Hampel	7803.645	M-Ridge	536.871	1042.389	949.718	0.405	0.333	64.471	368.317	547.574
		RM	763.262	1490.965	1325.432	0.604	0.386	118.841	546.865	779.340
Tukey	15075.198	M-Ridge	239.013	462.022	470.656	0.438	0.458	16.656	152.421	244.621
		RM	1272.080	2454.062	2210.714	0.957	0.732	196.359	960.346	1304.124
Andrew	6695.389	M-Ridge	327.636	619.131	584.560	0.303	0.283	31.200	207.327	334.089
		RM	628.710	1213.243	1082.776	0.499	0.343	93.179	430.567	642.018
Welsch	7511.262	M-Ridge	491.194	951.042	870.874	0.379	0.316	56.855	340.229	500.762
		RM	756.604	1476.726	1311.232	0.600	0.377	119.490	551.231	772.158
Cauchy	8855.387	M-Ridge	761.912	1476.605	1322.699	0.532	0.398	102.518	594.711	776.775
		RM	1081.718	2099.460	1847.911	0.845	0.485	191.164	879.030	1103.333
Talwar	6670.480	M-Ridge	539.118	1037.643	935.352	0.429	0.323	72.074	361.953	550.108
		RM	626.080	1209.639	1078.217	0.522	0.345	93.196	431.408	639.413
Ramsay	8581.543	M-Ridge	534.111	1026.813	939.240	0.392	0.343	58.365	397.546	544.460
		RM	1010.555	1966.145	1734.052	0.793	0.458	175.670	809.017	1030.735
G-Mc	7829.529	M-Ridge	44.168	77.362	85.725	0.224	0.266	2.147	26.096	45.037
		RM	858.215	1674.177	1481.715	0.679	0.404	142.978	659.846	875.358

Table 11: Scaler MSE values of the estimators for OOR 20%, $n = 100$ and $\rho^2 = 0.999$ (Case 8).

OLS	30238.893	Ridge	$\widehat{k}_{\mathrm{HK}}$ 2118.929	$\widehat{k}_{\mathrm{HKB}}$ 7749.269	$\widehat{k}_{\mathrm{LW}}$ 8724.782	$\widehat{k}_{\mathrm{HSL}}$ 3.342	$\widehat{k}_{\mathrm{AM}}$ 3.450	$\widehat{k}_{\mathrm{GM}}$ 749.439	$\widehat{k}_{\mathrm{MED}}$ 2715.282	$\widehat{k}_{\mathrm{KS}}$ 2227.345
M		Robust	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
Huber	14573.073	M-Ridge	704.728	2705.611	3119.209	1.042	1.078	196.682	891.648	736.316
		RM	1082.727	3714.399	4123.085	1.172	1.176	354.481	1334.789	1130.734
Fair	14871.272	M-Ridge	327.168	1439.940	1792.199	0.771	0.860	70.318	421.709	342.186
		RM	1081.582	3761.587	4197.642	1.224	1.234	357.826	1340.583	1130.142
Hampel	12033.327	M-Ridge	277.873	1066.741	1307.909	0.420	0.445	56.344	321.465	288.670
		RM	604.230	2141.156	2440.996	0.590	0.591	173.613	693.384	629.139
Tukey	18102.683	M-Ridge	85.479	389.947	552.495	0.474	0.511	13.505	101.956	89.391
		RM	714.260	2790.553	3290.897	0.868	0.901	214.551	865.647	749.799
Andrew	7924.098	M-Ridge	141.920	540.689	688.279	0.291	0.313	25.735	158.922	147.372
		RM	369.850	1326.828	1528.770	0.399	0.411	102.101	414.734	384.879
Welsch	10821.232	M-Ridge	245.473	948.332	1162.733	0.378	0.406	49.678	287.261	254.978
		RM	572.652	1994.584	2260.556	0.554	0.550	164.785	658.226	595.591
Cauchy	13458.812	M-Ridge	468.079	1868.520	2210.937	0.752	0.801	112.682	589.157	488.358
		RM	962.844	3264.931	3622.010	0.971	0.964	307.095	1169.258	1003.969
Talwar	7871.997	M-Ridge	260.574	950.086	1129.894	0.343	0.360	61.100	291.902	270.608
		RM	368.039	1316.858	1515.506	0.398	0.410	101.647	413.844	383.041
Ramsay	12868.287	M-Ridge	288.666	1209.728	1489.077	0.558	0.621	59.709	362.044	300.966
		RM	877.358	2973.707	3306.614	0.862	0.851	274.816	1055.887	914.148
G-Mc	11272.675	M-Ridge	13.782	67.987	109.560	0.186	0.243	2.080	16.860	14.357
		RM	689.126	2346.527	2628.418	0.657	0.646	206.975	811.716	716.889

Table 12: Scaler MSE of the estimators for lognormal distributed errors, $n = 20$ and $\rho^2 = 0.99$ (Case 9).

OLS	3394.857	Ridge	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
			916.252	603.731	599.784	37.128	1.277	41.289	114.869	923.979
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	1162.188	M-Ridge	192.488	89.897	98.023	8.009	0.442	4.535	16.129	194.639
		RM	278.480	167.505	164.175	12.759	0.562	11.868	33.977	281.253
Fair	870.719	M-Ridge	56.820	19.709	23.938	2.026	0.435	0.941	3.327	57.369
		RM	211.251	128.710	128.574	9.966	0.494	9.090	25.274	213.322
Hampel	1569.666	M-Ridge	65.604	25.721	30.567	2.646	0.397	1.206	4.050	66.543
		RM	248.562	139.487	144.229	9.720	0.552	8.410	24.123	252.698
Tukey	7023.009	M-Ridge	53.023	14.939	19.346	2.260	0.693	0.975	2.382	53.898
		RM	1563.448	1014.971	1070.368	63.501	1.962	61.597	150.988	1579.596
Andrew	1857.532	M-Ridge	49.237	16.610	21.006	1.975	0.448	0.765	2.504	49.894
		RM	296.768	171.919	179.478	10.872	0.631	10.951	31.587	301.625
Welsch	1449.204	M-Ridge	55.090	20.680	25.154	2.233	0.404	0.961	3.236	55.842
		RM	235.004	129.488	134.193	9.115	0.529	7.785	21.807	238.809
Cauchy	657.616	M-Ridge	53.355	23.447	25.361	2.451	0.381	1.237	4.320	54.083
		RM	121.514	70.437	69.866	5.810	0.397	4.793	13.782	123.219
Talwar	1647.434	M-Ridge	109.115	48.123	53.427	4.316	0.438	2.546	8.368	110.804
		RM	307.538	196.182	187.911	14.753	0.696	12.769	39.202	311.609
Ramsay	720.429	M-Ridge	33.217	12.415	14.805	1.386	0.407	0.622	2.124	33.619
		RM	126.110	73.255	75.165	5.622	0.405	4.944	13.988	127.997
G-Mc	924.968	M-Ridge	3.387	0.809	1.191	0.549	0.674	0.260	0.372	3.320
		RM	148.018	84.571	88.603	6.265	0.438	5.548	15.472	150.460

Table 13: Scaler MSE of the estimators for lognormal distributed errors, $n = 20$ and $\rho^2 = 0.999$ (Case 10).

OLS	33306.964	Ridge	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
			8997.876	5948.933	4765.248	104.566	1.083	162.260	1206.539	9005.269
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	11407.251	M-Ridge	1887.833	881.303	843.324	25.901	0.402	15.669	160.412	1889.963
		RM	2740.443	1647.381	1379.774	40.759	0.472	45.442	346.777	2743.180
Fair	8550.317	M-Ridge	550.887	191.837	194.284	5.752	0.405	2.707	31.811	551.423
		RM	2081.149	1265.936	1050.649	31.524	0.420	34.752	257.867	2083.184
Hampel	15477.152	M-Ridge	637.706	250.228	252.280	8.641	0.362	3.724	39.164	638.621
		RM	2394.337	1374.564	1225.932	30.642	0.476	31.099	251.450	2398.295
Tukey	70204.772	M-Ridge	506.971	137.156	156.075	5.348	0.667	2.064	18.025	507.562
		RM	15823.483	10261.280	7822.940	190.337	1.673	261.502	1878.946	15839.540
Andrew	18307.094	M-Ridge	475.212	160.095	167.463	5.641	0.414	2.086	23.054	475.849
		RM	2908.757	1689.845	1508.914	32.021	0.533	40.246	322.296	2913.510
Welsch	14280.645	M-Ridge	534.852	200.797	204.341	7.167	0.369	2.873	30.971	535.587
		RM	2259.062	1275.468	1120.455	29.871	0.455	28.744	222.713	2262.735
Cauchy	6459.583	M-Ridge	514.559	228.322	238.338	7.214	0.347	3.905	41.604	515.262
		RM	1188.789	691.444	614.623	19.014	0.340	17.444	136.780	1190.450
Talwar	16168.063	M-Ridge	1042.757	469.180	483.516	14.245	0.397	7.998	81.662	1044.302
		RM	2932.190	1915.777	1739.678	45.062	0.599	48.285	404.205	2936.126
Ramsay	7079.695	M-Ridge	320.713	120.290	127.999	3.859	0.374	1.738	19.700	321.104
		RM	1241.382	719.259	623.470	18.542	0.346	17.830	139.794	1243.242
G-Mc	9098.966	M-Ridge	30.482	6.632	7.986	0.689	0.636	0.205	0.983	30.427
		RM	1444.830	831.317	722.369	20.078	0.377	19.933	155.790	1447.224

Table 14: Scaler MSE of the estimators for lognormal distributed errors, $n = 100$ and $\rho^2 = 0.99$ (Case 11).

OLS	2591.635	Ridge	$\widehat{k}_{\text{HK}}$ 573.003	$\widehat{k}_{\text{HKB}}$ 448.774	$\widehat{k}_{\text{LW}}$ 413.554	$\widehat{k}_{\text{HSL}}$ 102.524	$\widehat{k}_{\text{AM}}$ 1.180	$\widehat{k}_{\text{GM}}$ 26.348	$\widehat{k}_{\text{MED}}$ 88.479	$\widehat{k}_{\text{KS}}$ 629.044
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	551.834	M-Ridge	80.668	51.510	48.476	9.118	0.440	2.737	9.414	90.462
		RM	98.618	69.915	63.961	12.238	0.455	4.278	13.492	111.049
Fair	397.395	M-Ridge	29.977	15.077	15.355	3.252	0.463	0.668	2.580	33.079
		RM	75.175	53.887	49.261	9.928	0.412	3.331	10.552	84.039
Hampel	423.251	M-Ridge	29.755	17.699	18.068	3.317	0.422	0.846	2.862	34.648
		RM	52.986	37.996	36.471	6.806	0.386	2.249	6.932	63.021
Tukey	3245.259	M-Ridge	36.679	13.949	16.491	3.707	0.790	0.907	2.166	40.897
		RM	405.126	288.511	278.345	64.671	1.090	17.262	53.500	477.177
Andrew	513.681	M-Ridge	23.992	12.479	13.303	2.506	0.481	0.561	1.868	27.491
		RM	65.498	46.550	44.483	8.062	0.406	2.733	8.438	77.534
Welsch	397.044	M-Ridge	24.905	14.398	14.920	2.771	0.430	0.673	2.276	28.891
		RM	49.503	35.653	34.395	6.445	0.380	2.110	6.500	58.918
Cauchy	336.273	M-Ridge	30.534	17.975	17.966	3.347	0.406	0.897	3.131	34.839
		RM	49.484	34.726	32.953	6.161	0.377	2.163	6.694	57.297
Talwar	742.333	M-Ridge	63.221	38.948	39.548	7.171	0.476	1.927	6.432	73.963
		RM	95.081	67.508	64.573	11.798	0.476	4.040	12.171	112.141
Ramsay	298.564	M-Ridge	17.229	9.359	9.892	1.915	0.439	0.450	1.534	19.699
		RM	39.871	28.284	27.394	5.229	0.361	1.740	5.398	46.880
G-Mc	314.456	M-Ridge	2.236	0.775	1.022	0.576	0.725	0.264	0.364	2.344
		RM	39.419	28.255	27.378	5.192	0.361	1.709	5.291	46.883

Table 15: Scaler MSE of the estimators for lognormal distributed errors, $n = 100$ and $\rho^2 = 0.999$ (Case 12).

OLS	26187.711	Ridge	$\widehat{k}_{\text{HK}}$ 5765.011	$\widehat{k}_{\text{HKB}}$ 4521.224	$\widehat{k}_{\text{LW}}$ 4171.784	$\widehat{k}_{\text{HSL}}$ 363.889	$\widehat{k}_{\text{AM}}$ 1.006	$\widehat{k}_{\text{GM}}$ 100.677	$\widehat{k}_{\text{MED}}$ 886.798	$\widehat{k}_{\text{KS}}$ 5813.761
M		Robust	$\widehat{k}_{\text{HK}}$	$\widehat{k}_{\text{HKB}}$	$\widehat{k}_{\text{LW}}$	$\widehat{k}_{\text{HSL}}$	$\widehat{k}_{\text{AM}}$	$\widehat{k}_{\text{GM}}$	$\widehat{k}_{\text{MED}}$	$\widehat{k}_{\text{KS}}$
Huber	5573.883	M-Ridge	812.268	517.892	487.361	25.690	0.397	9.624	93.055	820.752
		RM	992.387	702.943	643.330	33.857	0.395	15.928	134.248	1002.994
Fair	4014.625	M-Ridge	300.611	150.789	153.540	10.356	0.430	1.866	24.279	303.397
		RM	756.714	541.853	495.544	30.149	0.359	12.483	104.993	764.332
Hampel	4273.551	M-Ridge	297.335	177.075	180.812	9.781	0.382	2.597	27.063	301.513
		RM	530.388	381.596	366.397	19.924	0.334	8.043	68.137	538.783
Tukey	32900.989	M-Ridge	362.298	133.322	160.017	11.241	0.772	1.675	15.586	366.051
		RM	4065.108	2903.287	2800.391	286.532	0.901	62.973	530.283	4125.723
Andrew	5184.126	M-Ridge	238.865	124.184	132.466	7.460	0.441	1.471	16.755	241.894
		RM	655.177	467.279	446.635	24.000	0.349	9.799	83.137	665.252
Welsch	4009.178	M-Ridge	248.654	143.897	149.126	7.927	0.391	1.976	21.212	252.068
		RM	495.659	358.134	345.584	18.386	0.329	7.543	63.862	503.526
Cauchy	3398.437	M-Ridge	306.533	180.189	180.040	9.248	0.370	2.842	30.068	310.248
		RM	497.619	349.030	331.268	16.575	0.329	7.888	66.151	504.205
Talwar	7500.811	M-Ridge	633.598	390.162	394.856	22.666	0.429	6.276	61.806	642.454
		RM	953.262	678.261	648.862	34.612	0.405	14.541	119.698	967.474
Ramsay	3017.670	M-Ridge	172.115	93.414	98.636	5.096	0.401	1.217	13.926	174.256
		RM	400.455	284.226	275.344	13.882	0.315	6.269	53.136	406.326
G-Mc	3177.000	M-Ridge	20.322	6.647	8.365	0.931	0.685	0.173	0.812	20.436
		RM	395.365	283.874	275.088	14.118	0.315	6.112	51.940	401.605

Table 16: Comparison of estimation methods with k-fold cross validation method in Tobacco data set ($\times 10^{-4}$)

OLS	68.1432	Ridge	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
			68.0701	65.9109	68.7738	59.4768	60.2571	62.8117	60.8897	68.0576
M		Robust	$\widehat{k}_{\mathrm{HK}}$	$\widehat{k}_{\mathrm{HKB}}$	$\widehat{k}_{\mathrm{LW}}$	$\widehat{k}_{\mathrm{HSL}}$	$\widehat{k}_{\mathrm{AM}}$	$\widehat{k}_{\mathrm{GM}}$	$\widehat{k}_{\mathrm{MED}}$	$\widehat{k}_{\mathrm{KS}}$
Huber	63.1867	M-Ridge	60.6909	58.3774	64.6720	57.5810	**56.3181**	56.9301	56.9256	60.7143
		RM	60.9831	60.5311	65.6347	57.4911	57.1628	58.6564	57.9275	60.9912
Fair	67.9456	M-Ridge	67.7549	65.6060	68.6012	59.3509	59.9869	62.4962	60.6256	67.7434
		RM	67.7889	65.6930	68.6143	59.3478	60.1035	62.6280	60.7455	67.7773
Hampel	68.1432	M-Ridge	68.0701	65.9109	68.7738	59.4768	60.2571	62.8117	60.8897	68.0576
		RM	68.0701	65.9109	68.7738	59.4768	60.2571	62.8117	60.8897	68.0576
Tukey	68.1401	M-Ridge	68.0502	65.8924	68.7567	59.4634	60.2362	62.7901	60.8719	68.0379
		RM	68.0540	65.9005	68.7632	59.4698	60.2493	62.8027	60.8843	68.0416
Andrew	68.1426	M-Ridge	67.9074	65.5226	68.6910	59.8668	59.9523	62.2556	60.4219	67.8935
		RM	68.0671	65.9089	68.7719	59.4755	60.2556	62.8100	60.8887	68.0546
Welsch	68.1425	M-Ridge	68.0656	65.9067	68.7700	59.4738	60.2524	62.8069	60.8857	68.0532
		RM	68.0665	65.9085	68.7714	59.4753	60.2553	62.8097	60.8885	68.0540
Cauchy	68.1425	M-Ridge	68.0656	65.9067	68.7700	59.4738	60.2524	62.8069	60.8857	68.0532
		RM	68.0665	65.9085	68.7714	59.4753	60.2553	62.8097	60.8885	68.0540
Talwar	68.1432	M-Ridge	68.0701	65.9109	68.7738	59.4768	60.2571	62.8117	60.8897	68.0576
		RM	68.0701	65.9109	68.7738	59.4768	60.2571	62.8117	60.8897	68.0576
Ramsay	68.0439	M-Ridge	67.9084	65.7540	68.6823	59.4092	60.1156	62.6480	60.7522	67.8964
		RM	67.9261	65.7997	68.6917	59.4105	60.1784	62.7178	60.8162	67.9141
G-Mc	68.0172	M-Ridge	67.0976	64.8476	73.2344	65.1371	62.7903	63.0754	62.4144	67.0648
		RM	68.0656	65.9080	68.7709	59.4749	60.2549	62.8092	60.8882	68.0532