Scientific African 15 (2022) e01100

Contents lists available at ScienceDirect

Scientific African

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

# A new biased regression estimator: Theory, simulation and application[☆]

Issam Dawoud[a], Adewale F. Lukman[b,*], Abdul-Rahaman Haadi[c]

[a] Department of Mathematics, Al-Aqsa University, Gaza, Palestine
[b] Department of Epidemiology and Biostatistics, University of Medical Sciences, Nigeria
[c] Department of Statistical Sciences, Tamale Technical University, Ghana

ARTICLE INFO

*Article history:*
Received 11 August 2021
Revised 21 November 2021
Accepted 25 January 2022

Editor DR B Gyampoh

*Keywords:*
Two-parameter estimator
Liu estimator
Multicollinearity
Ridge regression estimator

ABSTRACT

The linear regression model explores the relationship between a response variable and one or more independent variables. The ordinary least squared estimator is usually adopted to estimate the parameters of the model when the independent variables are uncorrelated. However, the estimator performance dropped when the independent variables are correlated- a situation known as multicollinearity. This paper developed a new biased regression estimator based on a one-parameter and two-parameter estimators as an alternative to the ordinary least squares estimator when the independent variables are linearly dependent. Theoretical comparison, simulation and real-life data were carried out. The results revealed that the new estimator dominates other estimators considered in this study.

© 2022 The Author(s). Published by Elsevier B.V. on behalf of African Institute of Mathematical Sciences / Next Einstein Initiative.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

## Introduction

The linear regression model is defined as:

$$y = X\beta + \varepsilon,\ \varepsilon \sim N\left(0,\ \sigma^2 I_n\right) \tag{1}$$

where $y$ is an $n \times 1$ vector of the dependent variable, $X$ is a known $n \times p$ matrix of explanatory variables, $\beta$ is an $p \times 1$ vector of unknown regression parameters. The ordinary least squares estimator (OLS) of $\beta$ in (1) is defined by

$$\hat{\beta} = S^{-1}X'y \tag{2}$$

where $S = X'X$.

In many applications, multiple linear regression models are used for building models of real-life datasets. However, some problems came out, and the most frequent one is called the multicollinearity problem. Variuos regression estimators were proposed by different researchers to tackle this problem [1–14].

The paper aims to develop a new estimator to handle multicollinearity in the linear regression model. Hence, compare its performance with some existing estimators [3,6,8,14].

☆ Editor: DR B Gyampoh

* Corresponding author.
*E-mail address:* adewale.folaranmi@lmu.edu.ng (A.F. Lukman).

https://doi.org/10.1016/j.sciaf.2022.e01100
2468-2276/© 2022 The Author(s). Published by Elsevier B.V. on behalf of African Institute of Mathematical Sciences / Next Einstein Initiative. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

### Some existing biased estimators

The canonical form of Eq. (1) is

$$y = Z\alpha + \varepsilon \tag{3}$$

where $Z = XD$ and $\alpha = D'\beta$. Here, $D$ is an orthogonal matrix such that $Z'Z = D'X'XD = \Gamma = diag(\gamma_1, \gamma_2, \ldots, \gamma_p)$. The OLS estimator of $\alpha$ is then

$$\hat{\alpha} = \Gamma^{-1}Z'y, \tag{4}$$

$$MSEM(\hat{\alpha}) = \sigma^2\Gamma^{-1}. \tag{5}$$

The ORR estimator of $\alpha$ is

$$\hat{\alpha}_k = W\Gamma\hat{\alpha}, \tag{6}$$

where $W = [\Gamma + kI_p]^{-1}$ and $k$ is the biasing parameter.

The mean squared error matrix (MSEM) of the estimator of $\alpha$ would be

$$MSEM(\hat{\alpha}_k) = \sigma^2 W\Gamma W + (W\Gamma - I_p)\alpha\alpha'(W\Gamma - I_p)' \tag{7}$$

The biasing parameter $k$ can be estimated as follows [15]:

$$\hat{k}_{HM} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\hat{\alpha}_i^2} \tag{8}$$

The Liu estimator of $\alpha$ is

$$\hat{\alpha}_d = F\hat{\alpha}, \tag{9}$$

where $F = [\Gamma + I_p]^{-1}[\Gamma + dI_p]$ and the biasing parameter $d$ is estimated as follows:

$$\hat{d}_{opt} = 1 - \hat{\sigma}^2\left[\frac{\sum_{i=1}^{p}(1/(\gamma_i(\gamma_i+1)))}{\sum_{i=1}^{p}\left(\hat{\alpha}_i^2/(\gamma_i+1)^2\right)}\right], \tag{10}$$

$$MSEM(\hat{\alpha}_d) = \sigma^2 F\Gamma^{-1}F' + (1-d)^2(\Gamma + I_p)^{-1}\alpha\alpha'(\Gamma + I_p)^{-1}. \tag{11}$$

Alternatively, $d$ can be estimated as follows especially when its value in Eq. (10) is negative.

$$\hat{d}_{alt} = \min\left[\frac{\hat{\alpha}_i^2}{\left(\hat{\sigma}^2/\gamma_i\right) + \hat{\alpha}_i^2}\right]_{i=1}^{p} \tag{12}$$

The TP estimator of $\alpha$ [8] is

$$\hat{\alpha}_{TP} = R\hat{\alpha}, \tag{13}$$

where $R = (\Gamma + kI_p)^{-1}(\Gamma + kdI_p)$. $k$ and $d$ of the TP estimator are defined by

$$\hat{k}_{\min}(TP) = \min\left[\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2 - d\left(\left(\hat{\sigma}^2/\gamma_i\right) + \hat{\alpha}_i^2\right)}\right] \tag{14}$$

$$\hat{d}_{\min}(TP) = \min\left[\frac{\gamma_i\left(\hat{k}_{\min}\hat{\alpha}_i^2 - \hat{\sigma}^2\right)}{\hat{k}_{\min}\left(\hat{\sigma}^2 + \hat{\alpha}_i^2\gamma_i\right)}\right] \tag{15}$$

$$MSEM(\hat{\alpha}_{TP}) = \sigma^2 R\Gamma^{-1}R' + [R - I_p]\alpha\alpha'[R - I_p]' \tag{16}$$

If $\hat{d}_{opt}$ is negative, we can use $\hat{d}_{alt}$ that was derived by [8].

The KL estimator of $\alpha$ [14] is

$$\hat{\alpha}_{KL} = WM\hat{\alpha}, \tag{17}$$

where $M = (\Gamma - k\,\mathrm{I}_p)$.

$$\hat{k}_{\min}(KL) = \min\left[\frac{\hat{\sigma}^2}{2\hat{\alpha}_i^2 + \left(\hat{\sigma}^2/\lambda_i\right)}\right] \tag{18}$$

$$MSEM(\hat{\alpha}_{KL}) = \sigma^2 WM\Gamma^{-1}M'W' + [WM - I_p]\alpha\alpha'[WM - I_p]' \tag{19}$$

*The proposed nbr estimator*

Following a method similar to that proposed by Liu [6], Yang and Chang [10], and Kaciranlar et al. [16], the proposed new biased regression (NBR) estimator for $\alpha$ is obtained by replacing $\hat{\alpha}_{KL}$ with $\hat{\alpha}$ in the TP estimator and becomes as follows

$$\hat{\alpha}_{NBR} = RWM\hat{\alpha} \tag{20}$$

*Properties of the proposed nbr estimator*

$$E\left(\hat{\alpha}_{NBR}\right) = RWME\left(\hat{\alpha}\right) = RWM\alpha. \tag{21}$$

The proposed NBR estimator is a biased estimator of $\alpha$. The amount of bias and covariance are given respectively,

$$B\left(\hat{\alpha}_{NBR}\right) = [RWM - I_p]\alpha\,. \tag{22}$$

$$D\left(\hat{\alpha}_{NBR}\right) = \sigma^2 RWM\Gamma^{-1}M'W'R' \tag{23}$$

and the MSEM is defined by

$$MSEM\left(\hat{\alpha}_{NBR}\right) = \sigma^2 RWM\Gamma^{-1}M'W'R' + [RWM - I_p]\,\alpha\alpha'\,[RWM - I_p]' \tag{24}$$

The following lemmas are useful to compare the estimarors theoretically.

**Lemma 1.** *Let $G$ be an $n \times n$ positive definite matrix, that is $G > 0$ and $\alpha$ be some vector; then, $G - \alpha\alpha' > 0$ if and only if $\alpha' G^{-1}\alpha < 1$ [17].*

**Lemma 2.** *Let $\alpha_i = H_i y,\ i = 1, 2$ be two linear estimators of $\alpha$. Suppose that $D = Cov\,(\hat{\alpha}_1) - Cov\,(\hat{\alpha}_2) > 0$, where $Cov\,(\hat{\alpha}_i)\ i = 1, 2$ be the covariance matrix of $\hat{\alpha}_i$ and $b_i = Bias(\hat{\alpha}_i) = (H_iX - I)\alpha,\ i = 1, 2$. Consequently,*

$$\Delta\left(\hat{\alpha}_1 - \hat{\alpha}_2\right) = MSEM\left(\hat{\alpha}_1\right) - MSEM\left(\hat{\alpha}_2\right) = \sigma^2 D + b_1b_1' - b_2b_2' > 0 \tag{25}$$

if and only if $b_2'[\sigma^2 D + b'_1 b_1]^{-1} b_2 < 1$ where $MSEM(\hat{\alpha}_i) = Cov\,(\hat{\alpha}_i) + b_ib_i'$ [18].

The article was organized as follows: We compare the proposed NBR estimator with the existing estimators and derive the biasing parameters in Section 2. A simulation study and a real-life example were conducted in Sections 3 and 4, respectively. Finally, concluding remarks are given in section 5.

## Comparison among the estimators

In this section, we compared the performances of the estimator theoretically.

$\hat{\alpha}$ and $\hat{\alpha}_{NBR}$

The difference between $MSEM\ (\hat{\alpha})$ and $MSEM\ (\hat{\alpha}_{NBR})$is given by

$$\begin{aligned} &MSEM\left(\hat{\alpha}\right) - MSEM\left(\hat{\alpha}_{NBR}\right) = \\ &\sigma^2\left(\Gamma^{-1} - RWM\Gamma^{-1}M'W'R'\right) - [RWM - I_p]\alpha\alpha'[RWM - I_p]' \end{aligned} \tag{26}$$

**Theorem 3.1.** *The estimator $\hat{\alpha}_{NBR}$ is superior to estimator $\hat{\alpha}$ using the MSEM criterion, that is, $MSEM\ (\hat{\alpha}) - MSEM\ (\hat{\alpha}_{NBR}) > 0$ if and only if*

$$\alpha'[RWM - I_p]'\left[\sigma^2\left(\Gamma^{-1} - RWM\Gamma^{-1}M'W'R'\right)\right][RWM - I_p]\alpha < 1 \tag{27}$$

*Proof*:

$$\begin{aligned} Difference &= \sigma^2\left(\Gamma^{-1} - RWM\Gamma^{-1}M'W'R'\right) \\ &= \sigma^2\, diag\left\{\frac{1}{\gamma_i} - \frac{(\gamma_i + kd)^2(\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^4}\right\}_{i=1}^{p} \end{aligned} \tag{28}$$

where $\Gamma^{-1} - RWM\Gamma^{-1}M'W'R'$ will be positive definite (pd) if and only if $(\gamma_i + k)^4 - (\gamma_i + kd)^2(\gamma_i - k)^2 > 0$.

$\hat{\alpha}_k$ and $\hat{\alpha}_{NBR}$

The difference between $MSEM\ (\hat{\alpha}_k)$ and $MSEM\ (\hat{\alpha}_{NBR})$is given by

$$\begin{aligned} &MSEM\left(\hat{\alpha}_k\right) - MSEM\left(\hat{\alpha}_{NBR}\right) = \sigma^2\left(W\Gamma W' - RWM\Gamma^{-1}M'W'R'\right) \\ &+[W\Gamma - I_p]\alpha\alpha'[W\Gamma - I_p]' - [RWM - I_p]\alpha\alpha'[RWM - I_p]' \end{aligned} \tag{29}$$

**Theorem 3.2.** *The estimator $\hat{\alpha}_{NBR}$ is superior to estimator $\hat{\alpha}_k$ using the MSEM criterion, that is, $MSEM\ (\hat{\alpha}_k) - MSEM\ (\hat{\alpha}_{NBR}) > 0$ if and only if*

$$\alpha'[RWM - I_p]'\left[V_1 + [W\Gamma - \mathrm{I_p}]\alpha\alpha'[W\Gamma - \mathrm{I_p}]'\right][RWM - I_p]\alpha < 1 \tag{30}$$

where $V_1 = \sigma^2\,(W\Gamma W' - RWM\Gamma^{-1}M'W'R')$

*Proof*:

$$\begin{aligned} V_1 &= \sigma^2 \left( W\,\Gamma W' \; - R W M\,\Gamma^{-1} M' W' R' \right) \\ &= \sigma^2\, diag \left\{ \frac{\gamma_i}{(\gamma_i + k)^2} - \frac{(\gamma_i + k\,d)^2 (\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^4} \right\}_{i=1}^{p} \end{aligned} \tag{31}$$

where $W\,\Gamma W' \; - R W M\,\Gamma^{-1} M' W' R'$ will be pd if and only if $\gamma_i^2 (\gamma_i + k\,)^2 - \; (\gamma_i + k\,d)^2 (\gamma_i - k)^2 > 0$.

$\hat{\alpha}_d$ and $\hat{\alpha}_{NBR}$

The difference between $MSEM\ (\hat{\alpha}_d)$ and $MSEM\ (\hat{\alpha}_{NBR})$is given by

$$\begin{aligned} MSEM\left(\hat{\alpha}_d\right) - MSEM\left(\hat{\alpha}_S\right) = \sigma^2 &\left( F\,\Gamma^{-1} F' - R W M\,\Gamma^{-1} M' W' R' \right) \\ &+ (1-d)^2 (\Gamma + I_p)^{-1} \alpha\,\alpha'\,(\Gamma + I_p)^{-1} \\ &- (R W M \; - I_p)\,\alpha\,\alpha'\,[R W M \; - I_p]' \end{aligned} \tag{32}$$

**Theorem 3.3.** *The estimator $\hat{\alpha}_{NBR}$ is superior to estimator $\hat{\alpha}_d$ using the MSEM criterion, that is, $MSEM\ (\hat{\alpha}_d) - MSEM\ (\hat{\alpha}_{NBR}) > 0$ if and only if*

$$\alpha' [R W M \; - I_p]' \left[ V_2 + (1-d)^2 (\Gamma + I_p)^{-1} \alpha\,\alpha'\,(\Gamma + I_p)^{-1} \right] [R W M - I_p]\,\alpha \; < 1 \tag{33}$$

where $V_2 = \sigma^2\,(F\,\Gamma^{-1} F' - R W M\,\Gamma^{-1} M' W' R')$

*Proof*:

$$\begin{aligned} V_2 &= \sigma^2 \left( F\,\Gamma^{-1} F' - R W M\,\Gamma^{-1} M' W' R' \right) \\ &= \sigma^2\, diag \left\{ \frac{(\gamma_i + d)^2}{\gamma_i (\gamma_i + 1)^2} - \frac{(\gamma_i + k\,d)^2 (\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^4} \right\}_{i=1}^{p} \end{aligned} \tag{34}$$

where $F\,\Gamma^{-1} F' - R W M\,\Gamma^{-1} M' W' R'$ will be pd if and only if $(\gamma_i + k)^4\,(\gamma_i + d)^2 - \; (\gamma_i + k\,d)^2 (\gamma_i - k)^2 (\gamma_i + 1)^2 > 0$.

$\hat{\alpha}_{TP}$ and $\hat{\alpha}_{NBR}$

The difference between $MSEM\ (\hat{\alpha}_{TP})$ and $MSEM\ (\hat{\alpha}_{NBR})$is given by

$$\begin{aligned} MSEM\left(\hat{\alpha}_{TP}\right) - MSEM\left(\hat{\alpha}_{NBR}\right) = \sigma^2 &\left( R\,\Gamma^{-1}\,R' - RWM\,\Gamma^{-1}\,M' W' R' \right) \\ &+ [R - I_p]\,\alpha\,\alpha'\,[R \; - I_p]' - [RWM \; - I_p]\,\alpha\,\alpha'\,[RWM - I_p]' \end{aligned} \tag{35}$$

**Theorem 3.4.** *The estimator $\hat{\alpha}_{NBR}$ is superior to estimator $\hat{\alpha}_{TP}$ using the MSEM criterion, that is, $MSEM\ (\hat{\alpha}_{TP}) - MSEM\ (\hat{\alpha}_{NBR}) > 0$ if and only if*

$$\alpha' [RWM - I_p]' \left[ V_3 + [R \; - I_p]\,\alpha\,\alpha'\,[R \; - I_p]' \right] [RWM - I_p]\,\alpha \; < 1 \tag{36}$$

where $V_3 = \sigma^2\,(R\,\Gamma^{-1}\,R - \; RWM\,\Gamma^{-1} M' W' R')$

*Proof*:

$$\begin{aligned} V_3 &= \sigma^2 \left( R\,\Gamma^{-1}\,R - \; RWM\,\Gamma^{-1}\,M' W' R' \right) \\ &= \sigma^2\, diag \left\{ \frac{(\gamma_i + k\,d)^2}{\gamma_i (\gamma_i + k)^2} - \frac{(\gamma_i + k\,d)^2 (\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^4} \right\}_{i=1}^{p} \end{aligned} \tag{37}$$

where $R\,\Gamma^{-1}\,R - \; RWM\,\Gamma^{-1} M' W' R'$ will be pd if and only if $(\gamma_i + k)^2 \; - (\gamma_i - k)^2 > 0$.

$\hat{\alpha}_{KL}$ and $\hat{\alpha}_{NBR}$

The difference between $MSEM\ (\hat{\alpha}_{KL})$ and $MSEM\ (\hat{\alpha}_{NBR})$is given by

$$\begin{aligned} MSEM\left(\hat{\alpha}_{KL}\right) - MSEM\left(\hat{\alpha}_{NBR}\right) = \sigma^2 &\left( WM\,\Gamma^{-1}\,M' W' - RWM\,\Gamma^{-1}\,M' W' R' \right) \\ &+ [WM - I_p]\,\alpha\,\alpha'\,[WM - I_p]' \; - [RWM \; - I_p]\,\alpha\,\alpha'\,[RWM - I_p]' \end{aligned} \tag{38}$$

**Theorem 3.5.** *The estimator $\hat{\alpha}_{NBR}$ is superior to estimator $\hat{\alpha}_{KL}$ using the MSEM criterion, that is, $MSEM\ (\hat{\alpha}_{KL}) - MSEM\ (\hat{\alpha}_{NBR}) > 0$ if and only if*

$$\alpha' [RWM - I_p]' \left[ V_4 + [WM \; - I_p]\,\alpha\,\alpha'\,[WM \; - I_p]' \right] [RWM - I_p]\,\alpha \; < 1 \tag{39}$$

where $V_4 = \sigma^2\,(WM\,\Gamma^{-1} M' W' - \; RWM\,\Gamma^{-1} M' W' R')$

*Proof*:

$$\begin{aligned} V_4 &= \sigma^2 \left( WM\,\Gamma^{-1} M' W' - \; RWM\,\Gamma^{-1} M' W' R' \right) \\ &= \sigma^2\, diag \left\{ \frac{(\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^2} - \frac{(\gamma_i + k\,d)^2 (\gamma_i - k)^2}{\gamma_i\,(\gamma_i + k)^4} \right\}_{i=1}^{p} \end{aligned} \tag{40}$$

where $WM\,\Gamma^{-1} M' W' - \; RWM\,\Gamma^{-1} M' W' R'$ is pd since $(\gamma_i + k)^2 \; - (\gamma_i + k\,d)^2 > 0$.

*Choice of the biasing parameters*

Researchers have introduced different estimators of $k$ and $d$ for various regression models see [15,19-23]. Firstly, let $d$ be kept constant, then $k$ can be obtained as follows:

$MSEM(\hat{\alpha}_{NBR}) = E\left((\hat{\alpha}_{NBR} - \alpha)(\hat{\alpha}_{NBR} - \alpha)\right),$

$m(k, d) = tr\left(MSEM(\hat{\alpha}_{NBR})\right),$

$$m(k, d) = \sigma^2 \sum_{i=1}^{p} \frac{(\gamma_i + kd)^2(\gamma_i - k)^2}{\gamma_i(\gamma_i + k)^4} + \sum_{i=1}^{p} \frac{\left(k^2(1+d) + \gamma_i k(3-d)\right)^2 \alpha_i^2}{(\gamma_i + k)^4} \tag{41}$$

Differentiating $m(k, d)$ with respect to $k$ and setting $(\partial m(k, d)/\partial k) = 0$, we get

$$k = \frac{-\left(\gamma_i^2\alpha_i^2(3-d) + \sigma^2\gamma_i(1-d)\right)}{2\left(\sigma^2 d + \gamma_i\alpha_i^2(1+d)\right)} \tag{42}$$

$$\pm \frac{\gamma_i\sqrt{\gamma_i^2\left(\alpha_i^2\right)^2(d-3)^2 + 2\gamma_i\sigma^2\alpha_i^2\left(5 - 2d + d^2\right) + \left(\sigma^2\right)^2(1+d)^2}}{2\left(\sigma^2 d + \gamma_i\alpha_i^2(1+d)\right)} \tag{42}$$

Since $k > 0$, so

$$k = \frac{-\left(\gamma_i^2\alpha_i^2(3-d) + \sigma^2\gamma_i(1-d)\right)}{2\left(\sigma^2 d + \gamma_i\alpha_i^2(1+d)\right)} \tag{43}$$

$$+ \frac{\gamma_i\sqrt{\gamma_i^2\left(\alpha_i^2\right)^2(d-3)^2 + 2\gamma_i\sigma^2\alpha_i^2\left(5 - 2d + d^2\right) + \left(\sigma^2\right)^2(1+d)^2}}{2\left(\sigma^2 d + \gamma_i\alpha_i^2(1+d)\right)}. \tag{43}$$

Since, the optimal estimated value of $k$ in (43) is given by

$$\hat{k} = \frac{-\left(\gamma_i^2\hat{\alpha}_i^2(3-d) + \hat{\sigma}^2\gamma_i(1-d)\right)}{2\left(\hat{\sigma}^2 d + \gamma_i\hat{\alpha}_i^2(1+d)\right)} \tag{44}$$

$$+ \frac{\gamma_i\sqrt{\gamma_i^2\left(\hat{\alpha}_i^2\right)^2(d-3)^2 + 2\gamma_i\hat{\sigma}^2\hat{\alpha}_i^2\left(5 - 2d + d^2\right) + \left(\hat{\sigma}^2\right)^2(1+d)^2}}{2\left(\hat{\sigma}^2 d + \gamma_i\hat{\alpha}_i^2(1+d)\right)}. \tag{44}$$

and,

$$\hat{k}_{\min}(NBR) = \min\left\{\begin{array}{c} \hat{k} = \frac{-\left(\gamma_i^2\hat{\alpha}_i^2(3-d) + \hat{\sigma}^2\gamma_i(1-d)\right)}{2\left(\hat{\sigma}^2 d + \gamma_i\hat{\alpha}_i^2(1+d)\right)} \\ + \frac{\gamma_i\sqrt{\gamma_i^2\left(\hat{\alpha}_i^2\right)^2(d-3)^2 + 2\gamma_i\hat{\sigma}^2\hat{\alpha}_i^2\left(5-2d+d^2\right) + \left(\hat{\sigma}^2\right)^2(1+d)^2}}{2\left(\hat{\sigma}^2 d + \gamma_i\hat{\alpha}_i^2(1+d)\right)} \end{array}\right\}_{i=1}^{p}. \tag{45}$$

Then, the optimal value of $d$ can be found by differentiating $m(k, d)$ with respect to $d$ for a fixed $k$ and setting $(\partial m(k, d)/\partial d) = 0$, we obtain

$$d = \frac{\gamma_i^2\left(\sigma^2 - 3\alpha_i^2 k\right) - \gamma_i k\left(\sigma^2 + \alpha_i^2 k\right)}{k(k - \gamma_i)\left(\sigma^2 + \gamma_i\alpha_i^2\right)} \tag{46}$$

and the optimal $d$ with estimated parameters is

$$\hat{d} = \frac{\gamma_i^2\left(\hat{\sigma}^2 - 3\hat{\alpha}_i^2\hat{k}\right) - \gamma_i\hat{k}\left(\hat{\sigma}^2 + \hat{\alpha}_i^2\hat{k}\right)}{\hat{k}\left(\hat{k} - \gamma_i\right)\left(\hat{\sigma}^2 + \gamma_i\hat{\alpha}_i^2\right)} \tag{47}$$

Also,

$$\hat{d}_{\min}(NBR) = \min\left\{\frac{\gamma_i^2\left(\hat{\sigma}^2 - 3\hat{\alpha}_i^2\hat{k}_{\min}(NBR)\right) - \gamma_i\hat{k}_{\min}(())\left(\hat{\sigma}^2 + \hat{\alpha}_i^2\hat{k}_{\min}(NBR)\right)}{\hat{k}_{\min}(NBR)\left(\hat{k}_{\min}(NBR) - \gamma_i\right)\left(\hat{\sigma}^2 + \gamma_i\hat{\alpha}_i^2\right)}\right\}_{i=1}^{p} \tag{48}$$

The parameters $k$ and $d$ in $\hat{\alpha}_{NBR}$ are obtained iteratively as follows:

Step 1: Calculate an initial estimate of $d$ using $\hat{d} = \min\left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2}\right)$.

Step2: Calculate $\hat{k}_{\min}(NBR)$ from (45) using $\hat{d}$ in step 1.
Step 3: Calculate $\hat{d}_{\min}(NBR)$ in (48) by using $\hat{k}_{\min}(NBR)$ in step 2.
Step 4: If $\hat{d}_{\min}(NBR)$ is not between 0 and 1, use $\hat{d}_{\min}(NBR) = \hat{d}$.

**Table 1**
Estimated MSE when $p = 3$ and n=50.

| $k$ $\rho = 0.9$ | $d$ | $\sigma$ | OLS | ORR | Liu | TP | KL | NBR |
|---|---|---|---|---|---|---|---|---|
| 0.3 | 0.2 | 1 | 0.21360 | 0.20050 | 0.18210 | 0.20310 | 0.18790 | 0.17880 |
| | | 5 | 5.33940 | 5.01350 | 4.55070 | 5.07780 | 4.69820 | 4.46930 |
| | | 10 | 21.3576 | 20.0540 | 18.2030 | 20.3113 | 18.7931 | 17.8775 |
| | 0.5 | 1 | 0.21360 | 0.20050 | 0.19360 | 0.20700 | 0.18790 | 0.18220 |
| | | 5 | 5.33940 | 5.01350 | 4.83880 | 5.17510 | 4.69820 | 4.55450 |
| | | 10 | 21.3576 | 20.0540 | 19.3552 | 20.7005 | 18.7931 | 18.2180 |
| | 0.8 | 1 | 0.21360 | 0.20050 | 0.20540 | 0.21090 | 0.18790 | 0.18560 |
| | | 5 | 5.33940 | 5.01350 | 5.13610 | 5.27340 | 4.69820 | 4.64050 |
| | | 10 | 21.3576 | 20.0540 | 20.5443 | 21.0935 | 18.7931 | 18.5620 |
| 0.6 | 0.2 | 1 | 0.21360 | 0.18870 | 0.18210 | 0.19360 | 0.16550 | 0.15030 |
| | | 5 | 5.33940 | 4.71760 | 4.55070 | 4.83880 | 4.13610 | 3.75280 |
| | | 10 | 21.3576 | 18.8707 | 18.2030 | 19.3552 | 16.5443 | 15.0111 |
| | 0.5 | 1 | 0.21360 | 0.18870 | 0.19360 | 0.20090 | 0.16550 | 0.15590 |
| | | 5 | 5.33940 | 4.71760 | 4.83880 | 5.02350 | 4.13610 | 3.89420 |
| | | 10 | 21.3576 | 18.8707 | 19.3552 | 20.0941 | 16.5443 | 15.5769 |
| | 0.8 | 1 | 0.21360 | 0.18870 | 0.20540 | 0.20850 | 0.16550 | 0.16160 |
| | | 5 | 5.33940 | 4.71760 | 5.13610 | 5.21180 | 4.13610 | 4.03840 |
| | | 10 | 21.3576 | 18.8707 | 20.5443 | 20.8474 | 16.5443 | 16.1536 |
| 0.9 | 0.2 | 1 | 0.21360 | 0.17800 | 0.18210 | 0.18480 | 0.14590 | 0.12690 |
| | | 5 | 5.33940 | 4.44830 | 4.55070 | 4.61970 | 3.64220 | 3.16040 |
| | | 10 | 21.3576 | 17.7933 | 18.2030 | 18.4790 | 14.5685 | 12.6409 |
| | 0.5 | 1 | 0.21360 | 0.17800 | 0.19360 | 0.19530 | 0.14590 | 0.13380 |
| | | 5 | 5.33940 | 4.44830 | 4.83880 | 4.88320 | 3.64220 | 3.33680 |
| | | 10 | 21.3576 | 17.7933 | 19.3552 | 19.5330 | 14.5685 | 13.3467 |
| | 0.8 | 1 | 0.21360 | 0.17800 | 0.20540 | 0.20620 | 0.14590 | 0.14100 |
| | | 5 | 5.33940 | 4.44830 | 5.13610 | 5.15440 | 3.64220 | 3.51830 |
| | | 10 | 21.3576 | 17.7933 | 20.5443 | 20.6176 | 14.5685 | 14.0730 |
| $\rho = 0.99$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 1.94520 | 1.12580 | 0.68130 | 1.27170 | 0.53080 | 0.34950 |
| | | 5 | 48.6289 | 28.1455 | 17.0310 | 31.7937 | 13.2686 | 8.73730 |
| | | 10 | 194.515 | 112.582 | 68.1241 | 127.174 | 53.0746 | 34.9493 |
| | 0.5 | 1 | 1.94520 | 1.12580 | 1.07860 | 1.50750 | 0.53080 | 0.41300 |
| | | 5 | 48.6289 | 28.1455 | 26.9655 | 37.6864 | 13.2686 | 10.3236 |
| | | 10 | 194.515 | 112.582 | 107.862 | 150.745 | 53.0746 | 41.2947 |
| | 0.8 | 1 | 1.94520 | 1.12580 | 1.56790 | 1.76330 | 0.53080 | 0.48180 |
| | | 5 | 48.6289 | 28.1455 | 39.1977 | 44.0837 | 13.2686 | 12.0455 |
| | | 10 | 194.515 | 112.582 | 156.790 | 176.334 | 53.0746 | 48.1820 |
| 0.6 | 0.2 | 1 | 1.94520 | 0.73490 | 0.68130 | 0.93040 | 0.10720 | 0.05580 |
| | | 5 | 48.6289 | 18.3723 | 17.0310 | 23.2587 | 2.67820 | 1.39030 |
| | | 10 | 194.515 | 73.4894 | 68.1241 | 93.0347 | 10.7129 | 5.56120 |
| | 0.5 | 1 | 1.94520 | 0.73490 | 1.07860 | 1.26720 | 0.10720 | 0.07290 |
| | | 5 | 48.6289 | 18.3723 | 26.9655 | 31.6803 | 2.67820 | 1.81820 |
| | | 10 | 194.515 | 73.4894 | 107.862 | 126.721 | 10.7129 | 7.27290 |
| | 0.8 | 1 | 1.94520 | 0.73490 | 1.56790 | 1.65650 | 0.10720 | 0.09260 |
| | | 5 | 48.6289 | 18.3723 | 39.1977 | 41.4126 | 2.67820 | 2.31220 |
| | | 10 | 194.515 | 73.4894 | 156.790 | 165.650 | 10.7129 | 9.24890 |
| 0.9 | 0.2 | 1 | 1.94520 | 0.51840 | 0.68130 | 0.73020 | 0.01090 | 0.00960 |
| | | 5 | 48.6289 | 12.9585 | 17.0310 | 18.2542 | 0.26780 | 0.23050 |
| | | 10 | 194.515 | 51.8341 | 68.1241 | 73.0170 | 1.07090 | 0.92150 |
| | 0.5 | 1 | 1.94520 | 0.51840 | 1.07860 | 1.11690 | 0.01090 | 0.01000 |
| | | 5 | 48.6289 | 12.9585 | 26.9655 | 27.9212 | 0.26780 | 0.24260 |
| | | 10 | 194.515 | 51.8341 | 107.862 | 111.685 | 1.07090 | 0.96990 |
| | 0.8 | 1 | 1.94520 | 0.51840 | 1.56790 | 1.58630 | 0.01090 | 0.01050 |
| | | 5 | 48.6289 | 12.9585 | 39.1977 | 39.6565 | 0.26780 | 0.25690 |
| | | 10 | 194.515 | 51.8341 | 156.790 | 158.625 | 1.07090 | 1.02750 |

## Simulation study

A simulation has been carried out in this section to study the performance of the estimators.

### *Simulation technique*

Following [24], the explanatory variables are generated using the equation:

$$x_{ji} = \left(1 - \rho^2\right)^{1/2} z_{ji} + \rho\, z_{j,\, p+1}, \qquad j = 1,\ 2,\ \ldots,\ n, \quad i = 1,\ 2,\ \ldots,\ p \tag{49}$$

**Table 2**
*Estimated MSE when $p = 3$and n=100.*

| $k$ | $d$ | $\sigma$ | OLS | ORR | Liu | TP | KL | NBR |
|---|---|---|---|---|---|---|---|---|
| $\rho = 0.9$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 0.10640 | 0.10320 | 0.09820 | 0.10380 | 0.10000 | 0.09750 |
| | | 5 | 2.66110 | 2.57930 | 2.45380 | 2.59560 | 2.49890 | 2.43760 |
| | | 10 | 10.6443 | 10.3172 | 9.81490 | 10.3822 | 9.99560 | 9.75030 |
| | 0.5 | 1 | 0.10640 | 0.10320 | 0.10120 | 0.10480 | 0.10000 | 0.09840 |
| | | 5 | 2.66110 | 2.57930 | 2.53050 | 2.62000 | 2.49890 | 2.46050 |
| | | 10 | 10.6443 | 10.3172 | 10.1218 | 10.4801 | 9.99560 | 9.84190 |
| | 0.8 | 1 | 0.10640 | 0.10320 | 0.10430 | 0.10580 | 0.10000 | 0.09940 |
| | | 5 | 2.66110 | 2.57930 | 2.60840 | 2.64460 | 2.49890 | 2.48350 |
| | | 10 | 10.6443 | 10.3172 | 10.4336 | 10.5784 | 9.99560 | 9.93400 |
| 0.6 | 0.2 | 1 | 0.10640 | 0.10010 | 0.09820 | 0.10130 | 0.09390 | 0.08940 |
| | | 5 | 2.66110 | 2.50150 | 2.45380 | 2.53300 | 2.34710 | 2.23490 |
| | | 10 | 10.6443 | 10.0058 | 9.81490 | 10.1318 | 9.38820 | 8.93930 |
| | 0.5 | 1 | 0.10640 | 0.10010 | 0.10120 | 0.10320 | 0.09390 | 0.09110 |
| | | 5 | 2.66110 | 2.50150 | 2.53050 | 2.58060 | 2.34710 | 2.27660 |
| | | 10 | 10.6443 | 10.0058 | 10.1218 | 10.3224 | 9.38820 | 9.10620 |
| | 0.8 | 1 | 0.10640 | 0.10010 | 0.10430 | 0.10520 | 0.09390 | 0.09280 |
| | | 5 | 2.66110 | 2.50150 | 2.60840 | 2.62870 | 2.34710 | 2.31880 |
| | | 10 | 10.6443 | 10.0058 | 10.4336 | 10.5149 | 9.38820 | 9.27480 |
| 0.9 | 0.2 | 1 | 0.10640 | 0.09710 | 0.09820 | 0.09890 | 0.08820 | 0.08210 |
| | | 5 | 2.66110 | 2.42730 | 2.45380 | 2.47310 | 2.20480 | 2.05070 |
| | | 10 | 10.6443 | 9.70900 | 9.81490 | 9.89240 | 8.81900 | 8.20240 |
| | 0.5 | 1 | 0.10640 | 0.09710 | 0.10120 | 0.10170 | 0.08820 | 0.08440 |
| | | 5 | 2.66110 | 2.42730 | 2.53050 | 2.54280 | 2.20480 | 2.10780 |
| | | 10 | 10.6443 | 9.70900 | 10.1218 | 10.1710 | 8.81900 | 8.43090 |
| | 0.8 | 1 | 0.10640 | 0.09710 | 0.10430 | 0.10450 | 0.08820 | 0.08670 |
| | | 5 | 2.66110 | 2.42730 | 2.60840 | 2.61340 | 2.20480 | 2.16570 |
| | | 10 | 10.6443 | 9.70900 | 10.4336 | 10.4536 | 8.81900 | 8.66260 |
| $\rho = 0.99$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 0.99130 | 0.74460 | 0.52880 | 0.79110 | 0.53410 | 0.42800 |
| | | 5 | 24.7820 | 18.6158 | 13.2208 | 19.7767 | 13.3531 | 10.6993 |
| | | 10 | 99.1280 | 74.4631 | 52.8829 | 79.1070 | 53.4122 | 42.7969 |
| | 0.5 | 1 | 0.99130 | 0.74460 | 0.68500 | 0.86340 | 0.53410 | 0.46640 |
| | | 5 | 24.7820 | 18.6158 | 17.1256 | 21.5860 | 13.3531 | 11.6590 |
| | | 10 | 99.1280 | 74.4631 | 68.5024 | 86.3438 | 53.4122 | 46.6360 |
| | 0.8 | 1 | 0.99130 | 0.74460 | 0.86190 | 0.93910 | 0.53410 | 0.50650 |
| | | 5 | 24.7820 | 18.6158 | 21.5472 | 23.4765 | 13.3531 | 12.6613 |
| | | 10 | 99.1280 | 74.4631 | 86.1888 | 93.9059 | 53.4122 | 50.6451 |
| 0.6 | 0.2 | 1 | 0.99130 | 0.58110 | 0.52880 | 0.65420 | 0.28240 | 0.18940 |
| | | 5 | 24.7820 | 14.5269 | 13.2208 | 16.3549 | 7.05980 | 4.73520 |
| | | 10 | 99.1280 | 58.1077 | 52.8829 | 65.4197 | 28.2392 | 18.9406 |
| | 0.5 | 1 | 0.99130 | 0.58110 | 0.68500 | 0.77220 | 0.28240 | 0.22210 |
| | | 5 | 24.7820 | 14.5269 | 17.1256 | 19.3060 | 7.05980 | 5.55100 |
| | | 10 | 99.1280 | 58.1077 | 68.5024 | 77.2239 | 28.2392 | 22.2037 |
| | 0.8 | 1 | 0.99130 | 0.58110 | 0.86190 | 0.90030 | 0.28240 | 0.25740 |
| | | 5 | 24.7820 | 14.5269 | 21.5472 | 22.5080 | 7.05980 | 6.43390 |
| | | 10 | 99.1280 | 58.1077 | 86.1888 | 90.0318 | 28.2392 | 25.7354 |
| 0.9 | 0.2 | 1 | 0.99130 | 0.46680 | 0.52880 | 0.55570 | 0.14190 | 0.08290 |
| | | 5 | 24.7820 | 11.6687 | 13.2208 | 13.8921 | 3.54620 | 2.06970 |
| | | 10 | 99.1280 | 46.6749 | 52.8829 | 55.5685 | 14.1846 | 8.27830 |
| | 0.5 | 1 | 0.99130 | 0.46680 | 0.68500 | 0.70410 | 0.14190 | 0.10310 |
| | | 5 | 24.7820 | 11.6687 | 17.1256 | 17.6015 | 3.54620 | 2.57500 |
| | | 10 | 99.1280 | 46.6749 | 68.5024 | 70.4061 | 14.1846 | 10.2998 |
| | 0.8 | 1 | 0.99130 | 0.46680 | 0.86190 | 0.87040 | 0.14190 | 0.12560 |
| | | 5 | 24.7820 | 11.6687 | 21.5472 | 21.7601 | 3.54620 | 3.13840 |
| | | 10 | 99.1280 | 46.6749 | 86.1888 | 87.0404 | 14.1846 | 12.5533 |

where $z_{ji}$ are independent standard normal pseudo-random numbers and here $\rho = 0.90$ and $\rho = 0.99$ are the correlation between any two explanatory variables. The $n$ observations for the response variable $y$ are determined by the following equation:

$$y_j = \beta_1 x_{j1} + \beta_2 x_{j2} + \cdots + \beta_p x_{jp} + e_j, \quad j = 1, 2, \ldots, n \tag{50}$$

where $e_i$ are *i.i.d* $N(0, \sigma^2)$ and $p$ taken to be 3 and 7. The values of $\beta$ are chosen such that $\beta' \beta = 1$ [25]. We choose $k$ between 0 and 1 as Wichern and Churchill [26] suggested for getting better results of the ORR estimator, so $k = 0.3, 0.6, 0.9$ and $d$=0.2, 0.5, 0.8. Sample size $n$ is 50 and 100; $\sigma^2$ =1, 25, and 100. Also, we calculate the mean square error (MSE) of the

**Table 3**
Estimated MSE when $p = 7$ and $n = 50$.

| $k$ | $d$ | $\sigma$ | OLS | ORR | Liu | TP | KL | NBR |
|---|---|---|---|---|---|---|---|---|
| $\rho = 0.9$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 0.74360 | 0.66930 | 0.57860 | 0.68370 | 0.60010 | 0.55480 |
| | | 5 | 18.5888 | 16.7321 | 14.4654 | 17.0934 | 15.0020 | 13.8693 |
| | | 10 | 74.3551 | 66.9286 | 57.8616 | 68.3734 | 60.0079 | 55.4772 |
| | 0.5 | 1 | 0.74360 | 0.66930 | 0.63740 | 0.70580 | 0.60010 | 0.57150 |
| | | 5 | 18.5888 | 16.7321 | 15.9352 | 17.6447 | 15.0020 | 14.2872 |
| | | 10 | 74.3551 | 66.9286 | 63.7407 | 70.5786 | 60.0079 | 57.1488 |
| | 0.8 | 1 | 0.74360 | 0.66930 | 0.69990 | 0.72830 | 0.60010 | 0.58850 |
| | | 5 | 18.5888 | 16.7321 | 17.4967 | 18.2073 | 15.0020 | 14.7133 |
| | | 10 | 74.3551 | 66.9286 | 69.9869 | 72.8293 | 60.0079 | 58.8533 |
| 0.6 | 0.2 | 1 | 0.74360 | 0.60810 | 0.57860 | 0.63380 | 0.49020 | 0.42610 |
| | | 5 | 18.5888 | 15.2030 | 14.4654 | 15.8452 | 12.2549 | 10.6513 |
| | | 10 | 74.3551 | 60.8122 | 57.8616 | 63.3807 | 49.0196 | 42.6050 |
| | 0.5 | 1 | 0.74360 | 0.60810 | 0.63740 | 0.67360 | 0.49020 | 0.44940 |
| | | 5 | 18.5888 | 15.2030 | 15.9352 | 16.8412 | 12.2549 | 11.2353 |
| | | 10 | 74.3551 | 60.8122 | 63.7407 | 67.3648 | 49.0196 | 44.9412 |
| | 0.8 | 1 | 0.74360 | 0.60810 | 0.69990 | 0.71510 | 0.49020 | 0.47360 |
| | | 5 | 18.5888 | 15.2030 | 17.4967 | 17.8766 | 12.2549 | 11.8402 |
| | | 10 | 74.3551 | 60.8122 | 69.9869 | 71.5065 | 49.0196 | 47.3606 |
| 0.9 | 0.2 | 1 | 0.74360 | 0.55680 | 0.57860 | 0.59140 | 0.40450 | 0.33480 |
| | | 5 | 18.5888 | 13.9199 | 14.4654 | 14.7847 | 10.1129 | 8.36910 |
| | | 10 | 74.3551 | 55.6796 | 57.8616 | 59.1389 | 40.4513 | 33.4762 |
| | 0.5 | 1 | 0.74360 | 0.55680 | 0.63740 | 0.64590 | 0.40450 | 0.35990 |
| | | 5 | 18.5888 | 13.9199 | 15.9352 | 16.1466 | 10.1129 | 8.99780 |
| | | 10 | 74.3551 | 55.6796 | 63.7407 | 64.5864 | 40.4513 | 35.9911 |
| | 0.8 | 1 | 0.74360 | 0.55680 | 0.69990 | 0.70340 | 0.40450 | 0.38630 |
| | | 5 | 18.5888 | 13.9199 | 17.4967 | 17.5861 | 10.1129 | 9.65680 |
| | | 10 | 74.3551 | 55.6796 | 69.9869 | 70.3442 | 40.4513 | 38.6269 |
| $\rho = 0.99$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 7.03880 | 3.34170 | 2.04020 | 3.94800 | 1.30900 | 0.90260 |
| | | 5 | 175.971 | 83.5429 | 51.0056 | 98.6998 | 32.7241 | 22.5645 |
| | | 10 | 703.884 | 334.171 | 204.022 | 394.799 | 130.896 | 90.2578 |
| | 0.5 | 1 | 7.03880 | 3.34170 | 3.52670 | 4.98220 | 1.30900 | 1.04530 |
| | | 5 | 175.971 | 83.5429 | 88.1681 | 124.555 | 32.7241 | 26.1322 |
| | | 10 | 703.884 | 334.171 | 352.672 | 498.223 | 130.896 | 104.528 |
| | 0.8 | 1 | 7.03880 | 3.34170 | 5.47880 | 6.16630 | 1.30900 | 1.19960 |
| | | 5 | 175.971 | 83.5429 | 136.970 | 154.156 | 32.7241 | 29.9905 |
| | | 10 | 703.884 | 334.171 | 547.880 | 616.626 | 130.896 | 119.961 |
| 0.6 | 0.2 | 1 | 7.03880 | 2.08350 | 2.04020 | 2.79960 | 0.56550 | 0.25880 |
| | | 5 | 175.971 | 52.0876 | 51.0056 | 69.9897 | 14.1365 | 6.4696 |
| | | 10 | 703.884 | 208.350 | 204.022 | 279.958 | 56.5458 | 25.8782 |
| | 0.5 | 1 | 7.03880 | 2.08350 | 3.52670 | 4.13150 | 0.56550 | 0.35640 |
| | | 5 | 175.971 | 52.0876 | 88.1681 | 103.287 | 14.1365 | 8.90910 |
| | | 10 | 703.884 | 208.350 | 352.672 | 413.151 | 56.5458 | 35.6365 |
| | 0.8 | 1 | 7.03880 | 2.08350 | 5.47880 | 5.77280 | 0.56550 | 0.47490 |
| | | 5 | 175.971 | 52.0876 | 136.970 | 144.319 | 14.1365 | 11.8713 |
| | | 10 | 703.884 | 208.350 | 547.880 | 577.279 | 56.5458 | 47.4852 |
| 0.9 | 0.2 | 1 | 7.03880 | 1.45220 | 2.04020 | 2.18600 | 0.65980 | 0.15470 |
| | | 5 | 175.971 | 36.3057 | 51.0056 | 54.6503 | 16.4952 | 3.86560 |
| | | 10 | 703.884 | 145.222 | 204.022 | 218.601 | 65.9807 | 15.4624 |
| | 0.5 | 1 | 7.03880 | 1.45220 | 3.52670 | 3.64630 | 0.65980 | 0.29900 |
| | | 5 | 175.971 | 36.3057 | 88.1681 | 91.1565 | 16.4952 | 7.47330 |
| | | 10 | 703.884 | 145.222 | 352.672 | 364.625 | 65.9807 | 29.8929 |
| | 0.8 | 1 | 7.03880 | 1.45220 | 5.47880 | 5.53800 | 0.65980 | 0.49740 |
| | | 5 | 175.971 | 36.3057 | 136.970 | 138.449 | 16.4952 | 12.4350 |
| | | 10 | 703.884 | 145.222 | 547.880 | 553.798 | 65.9807 | 49.7401 |

estimators for each replicate by using the equation below:

$$MSE(\alpha^*) = \frac{1}{1000} \sum_{j=1}^{1000} \left(\alpha_{ij}^* - \alpha_i\right)' \left(\alpha_{ij}^* - \alpha_i\right) \tag{51}$$

where $\alpha_{ij}^*$ is the estimator and $\alpha_i$ is the parameter (see Tables 1-4 for the result).

**Table 4**
*Estimated MSE when $p = 7$ and $n = 100$.*

| $k$ | $d$ | $\sigma$ | OLS | ORR | Liu | TP | KL | NBR |
|---|---|---|---|---|---|---|---|---|
| $\rho = 0.9$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 0.32640 | 0.31510 | 0.29800 | 0.31740 | 0.30410 | 0.29570 |
| | | 5 | 8.16010 | 7.87800 | 7.45040 | 7.93390 | 7.60130 | 7.39200 |
| | | 10 | 32.6402 | 31.5118 | 29.8016 | 31.7358 | 30.4052 | 29.5682 |
| | 0.5 | 1 | 0.32640 | 0.31510 | 0.30850 | 0.32070 | 0.30410 | 0.29880 |
| | | 5 | 8.16010 | 7.87800 | 7.71240 | 8.01830 | 7.60130 | 7.47010 |
| | | 10 | 32.6402 | 31.5118 | 30.8496 | 32.0733 | 30.4052 | 29.8806 |
| | 0.8 | 1 | 0.32640 | 0.31510 | 0.31920 | 0.32410 | 0.30410 | 0.30190 |
| | | 5 | 8.16010 | 7.87800 | 7.97930 | 8.10320 | 7.60130 | 7.54870 |
| | | 10 | 32.6402 | 31.5118 | 31.9174 | 32.4128 | 30.4052 | 30.1947 |
| 0.6 | 0.2 | 1 | 0.32640 | 0.30450 | 0.29800 | 0.30880 | 0.28340 | 0.26830 |
| | | 5 | 8.16010 | 7.61150 | 7.45040 | 7.71950 | 7.08380 | 6.70630 |
| | | 10 | 32.6402 | 30.4460 | 29.8016 | 30.8782 | 28.3352 | 26.8253 |
| | 0.5 | 1 | 0.32640 | 0.30450 | 0.30850 | 0.31530 | 0.28340 | 0.27390 |
| | | 5 | 8.16010 | 7.61150 | 7.71240 | 7.88320 | 7.08380 | 6.84650 |
| | | 10 | 32.6402 | 30.4460 | 30.8496 | 31.5327 | 28.3352 | 27.3862 |
| | 0.8 | 1 | 0.32640 | 0.30450 | 0.31920 | 0.32190 | 0.28340 | 0.27950 |
| | | 5 | 8.16010 | 7.61150 | 7.97930 | 8.04870 | 7.08380 | 6.98840 |
| | | 10 | 32.6402 | 30.4460 | 31.9174 | 32.1947 | 28.3352 | 27.9535 |
| 0.9 | 0.2 | 1 | 0.32640 | 0.29440 | 0.29800 | 0.30060 | 0.26420 | 0.24370 |
| | | 5 | 8.16010 | 7.35950 | 7.45040 | 7.51600 | 6.60400 | 6.09260 |
| | | 10 | 32.6402 | 29.4379 | 29.8016 | 30.0639 | 26.4160 | 24.3704 |
| | 0.5 | 1 | 0.32640 | 0.29440 | 0.30850 | 0.31020 | 0.26420 | 0.25130 |
| | | 5 | 8.16010 | 7.35950 | 7.71240 | 7.75410 | 6.60400 | 6.28170 |
| | | 10 | 32.6402 | 29.4379 | 30.8496 | 31.0165 | 26.4160 | 25.1269 |
| | 0.8 | 1 | 0.32640 | 0.29440 | 0.31920 | 0.31990 | 0.26420 | 0.25900 |
| | | 5 | 8.16010 | 7.35950 | 7.97930 | 7.99630 | 6.60400 | 6.47400 |
| | | 10 | 32.6402 | 29.4379 | 31.9174 | 31.9853 | 26.4160 | 25.8961 |
| $\rho = 0.99$ | | | | | | | | |
| 0.3 | 0.2 | 1 | 3.09830 | 2.27350 | 1.59690 | 2.42750 | 1.58610 | 1.25950 |
| | | 5 | 77.4569 | 56.8373 | 39.9236 | 60.6865 | 39.6517 | 31.4875 |
| | | 10 | 309.827 | 227.349 | 159.694 | 242.746 | 158.607 | 125.949 |
| | 0.5 | 1 | 3.09830 | 2.27350 | 2.09910 | 2.66870 | 1.58610 | 1.37720 |
| | | 5 | 77.4569 | 56.8373 | 52.4782 | 66.7179 | 39.6517 | 34.4306 |
| | | 10 | 309.827 | 227.349 | 209.912 | 266.871 | 158.607 | 137.722 |
| | 0.8 | 1 | 3.09830 | 2.27350 | 2.67430 | 2.92230 | 1.58610 | 1.50060 |
| | | 5 | 77.4569 | 56.8373 | 66.8572 | 73.0582 | 39.6517 | 37.5159 |
| | | 10 | 309.827 | 227.349 | 267.428 | 292.233 | 158.607 | 150.063 |
| 0.6 | 0.2 | 1 | 3.09830 | 1.75160 | 1.59690 | 1.98830 | 0.81250 | 0.54570 |
| | | 5 | 77.4569 | 43.7901 | 39.9236 | 49.7084 | 20.3119 | 13.6427 |
| | | 10 | 309.827 | 175.160 | 159.694 | 198.833 | 81.2475 | 54.5709 |
| | 0.5 | 1 | 3.09830 | 1.75160 | 2.09910 | 2.37400 | 0.81250 | 0.63920 |
| | | 5 | 77.4569 | 43.7901 | 52.4782 | 59.3499 | 20.3119 | 15.9796 |
| | | 10 | 309.827 | 175.160 | 209.912 | 237.399 | 81.2475 | 63.9186 |
| | 0.8 | 1 | 3.09830 | 1.75160 | 2.67430 | 2.79630 | 0.81250 | 0.74050 |
| | | 5 | 77.4569 | 43.7901 | 66.8572 | 69.9084 | 20.3119 | 18.5134 |
| | | 10 | 309.827 | 175.160 | 267.428 | 279.633 | 81.2475 | 74.0535 |
| 0.9 | 0.2 | 1 | 3.09830 | 1.39650 | 1.59690 | 1.68000 | 0.40470 | 0.24030 |
| | | 5 | 77.4569 | 34.9119 | 39.9236 | 42.0010 | 10.1164 | 6.00620 |
| | | 10 | 309.827 | 139.647 | 159.694 | 168.003 | 40.4657 | 24.0250 |
| | 0.5 | 1 | 3.09830 | 1.39650 | 2.09910 | 2.15860 | 0.40470 | 0.29670 |
| | | 5 | 77.4569 | 34.9119 | 52.4782 | 53.9657 | 10.1164 | 7.41770 |
| | | 10 | 309.827 | 139.647 | 209.912 | 215.862 | 40.4657 | 29.6710 |
| | 0.8 | 1 | 3.09830 | 1.39650 | 2.67430 | 2.70110 | 0.40470 | 0.35940 |
| | | 5 | 77.4569 | 34.9119 | 66.8572 | 67.5279 | 10.1164 | 8.98500 |
| | | 10 | 309.827 | 139.647 | 267.428 | 270.111 | 40.4657 | 35.9401 |

*Simulation results discussions*

Tables 1-4 show that when $\sigma$, $\rho$ and $p$ are increasing, the estimated MSE values are also increasing. The MSE decreases as the sample size $n$ increase. The OLS estimator performs worst among all the estimators as expected due to multicollinearity. Also, we observe that the proposed NBR estimator dominates other estimators in this study in most cases. Thus, the findings agree with the theoretical results. For each row, the smallest MSE value is bolded.

**Table 5**
The results of regression coefficients and the corresponding MSE values.

| Coef. | $\hat{\alpha}$ | $\hat{\alpha}(\hat{k}_{\min})$ | $\hat{\alpha}(\hat{d}_{alt})$ | $\hat{\alpha}_{TP}$ $(\hat{k}_{\min}, \hat{d}_{\min})$ | $\hat{\alpha}_{KL}(\hat{k}_{\min})$ | $\hat{\alpha}_{NBR}$ $(\hat{k}_{\min}, \hat{d}_{\min})$ |
|---|---|---|---|---|---|---|
| $\alpha_0$ | 62.4053 | 8.58715 | 27.6657 | 27.6200 | 27.6270 | 27.5987 |
| $\alpha_1$ | 1.55110 | 2.10461 | 1.90080 | 1.9089 | 1.90884 | 1.9091 |
| $\alpha_2$ | 0.51016 | 1.06484 | 0.86996 | 0.8687 | 0.86859 | 0.8689 |
| $\alpha_3$ | 0.10190 | 0.66808 | 0.46192 | 0.4679 | 0.46782 | 0.4681 |
| $\alpha_4$ | -0.14406 | 0.39959 | 0.20801 | 0.2073 | 0.20724 | 0.2075 |
| $MSE$ | 4912.0902 | 2989.8202 | 2170.9669 | 2170.9600 | 2170.9604 | 2160.9596 |
| $(k, d)$ | —– | 0.007676 | 0.442224 | (0.0015419, 0.001536) | 0.000471 | (0.000339, 0.001530) |

## Application

The Portland cement data which was originally adopted by Woods et al. [27] is analyzed here to explain the theoretical results. Many authors have adopted this data in their studies [12,14,16]. The regression model for this data is defined as

$$y_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_j, \qquad j = 1, 2, ..., 13 \tag{52}$$

The model was diagnosed using the following diagnostic tools: variance inflation factor, eigenvalues and the condition number. The variance inflation factors (VIFs) of this data are calculated as $VIF_1 = 38.50$, $VIF_2 = 254.42$, $VIF_3 = 46.87$, $VIF_4 = 282.51$. Also, eigenvalues of S are given as $\lambda_1 = 44676.206$, $\lambda_2 = 5965.422$, $\lambda_4 = 105.419$, and the condition number (CN) of S is computed to be 424. The result of all the diagnostic tools indicate the presence multicollinearity in the model. The estimated coefficients and the MSE values of the estimators are given in Table 5. It is obvious from Table 5 that the proposed NBR estimator dominates other considered estimators.

## Conclusion

In this paper, we proposed a new biased regression estimator called NBR as an alternative to OLS to deal with multicollinearity in the linear regression model. We obtained the biasing parameters of the proposed and compared this new method of estimation with some existing estimators (OLS, ORR, Liu, TP and KL) theoretically. The simulation study and the theoretical findings revealed the superiority of the proposed method. Also, real-life data are used and analyzed to support the theoretical finding and simulation result. Furthermore, the performance of the estimators depends on the biasing parameter. Thus, we will examine the generalized cross-validation (GCV) method of selecting the biasing parameter in our future study (see [28–31]).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Funding

No financial assistance was received for this study.

## References

[1] Stein, C., "Inadmissibility of the usual estimator for mean of multivariate normal distribution", In Proceedings of the Third Berkley Symposium on Mathematical and Statistics Probability, J. Neyman, Ed., Springer, Berlin, Germany, 1: 197–206, (1956).
[2] W.F. Massy, "Principal components regression in exploratory statistical research", J. Am. Statist. Assoc. 60 (309) (1965) 234–256.
[3] A.E. Hoerl, R.W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems", Technometrics 12 (1) (1970) 55–67.
[4] L.S. Mayer, T.A. Willke, "On biased estimation in linear models", Technometrics 15 (3) (1973) 497–508.
[5] B.F. Swindel, "Good ridge estimators based on prior information", Commun. Stat.-Theory and Methods 5 (11) (1976) 1065–1075.
[6] K. Liu, "A new class of biased estimate in linear regression", Commun. Stat.-Theory and Methods 22 (1993) 393–402.
[7] F. Akdeniz, S. Kaçiranlar, "On the almost unbiased generalized liu estimator and unbiased estimation of the bias and mse", Commun. Stat.-Theory and Methods 24 (7) (1995) 1789–1797.
[8] M.R. Ozkale, S. Kaçiranlar, "The restricted and unrestricted two-parameter estimators", Commun. Stat.-Theory and Methods 36 (15) (2007) 2707–2725.
[9] S. Sakallıoglu, S. Kaçıranlar, "A new biased estimator based on ridge estimation", Stat. Papers 49 (4) (2008) 669–689.
[10] H. Yang, X. Chang, "A new two-parameter estimator in linear regression", Commun. Stat.-Theory and Methods 39 (6) (2010) 923–934.
[11] M. Roozbeh, Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion, Comput. Stat. Data Anal. 117 (2018) 45–61.
[12] A.F. Lukman, K. Ayinde, S. Binuomote, O.A. Clement, "Modified ridge-type estimator to combat multicollinearity: application to chemical data", J. Chemom. 33 (5) (2019) e3125.
[13] Lukman, A.F., Ayinde, K., Sek, S.K., Adewuyi, E., "A modified new two-parameter estimator in a linear regression model", Model. Simul. Eng., 2019: 1–10, (2019). https://doi.org/10.1155/2019/6342702
[14] Kibria, B.M.G., Lukman, A.F., "A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications", Scientifica, 2020: 1–16, (2020). https://doi.org/10.1155/2020/9758378

[15] A.E. Hoerl, R.W. Kannard, K.F. Baldwin, Ridge regression: some simulations", Commun. Stat. 4 (2) (1975) 105–123.
[16] S. Kaciranlar, S. Sakallioglu, F. Akdeniz, G.P.H. Styan, H.J. Werner, "A new biased estimator in linear regression and a detailed analysis of the widely–analysed dataset on portland cement", Sankhya 61 (1999) 443–459.
[17] R.W. Farebrother, "Further results on the mean square error of ridge regression", J. R. Stat. Soc.: Series B (Methodological) 38 (3) (1976) 248–250.
[18] G. Trenkler, H. Toutenburg, "Mean squared error matrix comparisons between biased estimators-an overview of recent results", Statistical Papers 31 (1) (1990) 165–179.
[19] B.M.G. Kibria, "Performance of some new ridge regression estimators", Commun. Stat.-Simul. Comput. 32 (1) (2003) 419–435.
[20] G. Khalaf, G. Shukur, "Choosing ridge parameter for regression problems", Commun. Stat. Theory Methods 34 (5) (2005) 1177–1182.
[21] B.M.G. Kibria, S. Banik, "Some ridge regression estimators and their performances", J. Mod. Appl. Stat. Methods 15 (1) (2016) 206–238.
[22] A.F. Lukman, K. Ayinde, "Review and classifications of the ridge parameter estimation techniques, Hacettepe J. Math. Stat. 46 (5) (2017) 953–967.
[23] K. Månsson, B.M.G. Kibria, G. Shukur, "Performance of some weighted Liu estimators for logit regression model: an application to Swedish accident data, Commun. Stat. Theory Methods 44 (2) (2015) 363–375.
[24] D.G. Gibbons, "A simulation study of some ridge estimators", J. Am. Statist. Assoc. 76 (373) (1981) 131–139.
[25] Newhouse, J.P. Oman, S.D., "An evaluation of ridge estimators", A report prepared for United States air force project RAND, (1971).
[26] D.W. Wichern, G.A. Churchill, "A comparison of ridge estimators", Technometrics 20 (3) (1978) 301–311.
[27] H. Woods, H.H. Steinour, H.R. Starke, "Effect of composition of Portland cement on heat evolved during hardening", Ind. Eng. Chem. 24 (11) (1932) 1207–1214.
[28] M. Roozbeh, M. Arashi, N.A. Hamzah, "Generalized cross validation for simultaneous optimization of tuning parameters in ridge regression", Iran. J. Sci. Technol., Trans. A: Science 44 (2020) 473–485.
[29] M. Amini, M. Roozbeh, "Optimal partial ridge estimation in restricted semiparametric regression models", J. Multivar. Anal. 136 (2015) 26–40.
[30] Arashi, M., Roozbeh, M., Hamzah, N.A. and Gasparini, M. "Ridge regression and its applications in genetic studies", PLoS One, 16 (4), e0245376, (2021).
[31] Roozbeh, M., Hesamian, G. and Akbari, M.G. "Ridge estimation in semi-parametric regression models under the stochastic restriction and correlated elliptically contoured errors", J. Comput. Appl. Math., 378, 112940, (2020).