



## ON THE STOCHASTIC RESTRICTED LIU-TYPE MAXIMUM LIKELIHOOD ESTIMATOR IN LOGISTIC REGRESSION MODEL

JIBO WU AND YASIN ASAR

**ABSTRACT.** In order to overcome multicollinearity, we propose a stochastic restricted Liu-type maximum likelihood estimator by incorporating Liu-type maximum likelihood estimator to the logistic regression model when the linear restrictions are stochastic. We also discuss the properties of the new estimator. Moreover, we give a method to choose the biasing parameter in the new estimator. Finally, a simulation study is given to show the performance of the new estimator.

### 1. INTRODUCTION

Consider the following logistic regression model

$$y_i = \pi_i + \varepsilon_i, i = 1, \dots, n, \quad (1)$$

where  $\pi_i = \pi(x_i) = E[y_i] = \frac{e^{x_i \beta}}{1 + e^{x_i \beta}}$ ,  $y_i \sim \text{Bernoulli}(\pi_i)$  and  $\beta = (\beta_0, \beta_1, \dots, \beta_p)^\top$  denotes the unknown  $(p+1)$ -vector of regression coefficients,  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  is the  $n \times (p+1)$  data matrix with  $\mathbf{x}_i = (1, x_{1i}, \dots, x_{pi})^\top$  and  $\varepsilon_i$ 's are independent with zero mean and their variance equal to  $w_i = \pi_i(1 - \pi_i)$ .

In the estimation process of the coefficient vector  $\beta$ , one often uses the maximum likelihood (ML) approach. One may use the iteratively re-weighted least squares algorithm (IRLS) as follows ([6]) :

$$\hat{\beta}^{t+1} = \hat{\beta}^t + (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^\top \widehat{\mathbf{W}}^t (\mathbf{y} - \hat{\boldsymbol{\pi}}^t) \quad (2)$$

where  $\hat{\boldsymbol{\pi}}^t$  is the estimated values of  $\hat{\boldsymbol{\pi}}$  using  $\hat{\beta}^t$  and  $\widehat{\mathbf{W}}^t = \text{diag}(\hat{\pi}_i^t(1 - \hat{\pi}_i^t))$  such that  $\hat{\pi}_i^t$  is the  $i$ th element of  $\hat{\boldsymbol{\pi}}^t$ . After some algebra, Equation (2) can be written as follows:

$$\hat{\beta}_{ML} = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} \mathbf{X}^\top \widehat{\mathbf{W}} \hat{\mathbf{z}} \quad (3)$$

Received by the editors: October 18, 2017; Accepted: March 30, 2018.

2010 *Mathematics Subject Classification.* Primary 62J05; Secondary 62J070.

*Key words and phrases.* Liu-type maximum likelihood estimator, stochastic restricted Liu-type maximum likelihood estimator, multicollinearity.

where  $\widehat{\mathbf{z}}^\top = (z_1 \cdots z_n)$  with  $\eta_i = x_i' \beta$  and  $z_i = \eta_i + (y_i - \pi_i)(\partial \eta_i / \partial \pi_i)$ .

As it is well known that when multicollinearity exists in the explanatory variables, the MLE becomes unstable and its variance is inflated. Many methods have been proposed to overcome this problem. We refer to the following papers: [8], [11], [14], [1]. In this article, we also suppose that  $\beta$  is subjected to lie in the sub-space restriction  $\mathbf{H}\beta = \mathbf{h}$ , where  $\mathbf{H}$  is  $q \times (p + 1)$  known matrix and  $\mathbf{h}$  is a  $q \times 1$  vector of pre-specified values. This problem was also studied in [7], [6], [15], [16].

For the stochastic linear restrictions, [16], [17] and [18] discussed the logistic regression model. In this paper, we will discuss the logistic regression model with stochastic linear restrictions. By combining the Liu-type maximum likelihood estimator ([11]) and stochastic linear restrictions, we propose a new estimator called stochastic restricted Liu-type maximum likelihood estimator to overcome multicollinearity. We also discuss the properties of the new estimator and we show that the new estimator can not only overcome multicollinearity, but also has good properties than other exists estimators.

The rest of the paper is organized as follows, in Section 2, the new estimator is proposed and the properties of the new estimator is discussed in Section 3. A method is given to choose the biasing parameters in Section 4 and we also conduct a simulation study to show the performance of the new estimator in Section 5, some conclusive remarks are given in Section 6.

## 2. THE NEW ESTIMATOR

For the model (1), [8] proposed a Liu maximum likelihood estimator (LE) to overcome the multicollinearity, which is defined as

$$\widehat{\beta}_{LE}(d) = (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}}\widehat{\mathbf{z}} + d\widehat{\beta}_{MLE}) = \mathbf{F}_d \widehat{\beta}_{MLE} \quad (4)$$

where  $\mathbf{F}_d = (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{I})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + d\mathbf{I})$ ,  $0 < d < 1$ .

[11] proposed a Liu-type maximum likelihood estimator (LTE), which is defined as

$$\widehat{\beta}_{LTE}(k, d) = (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}}\widehat{\mathbf{z}} - d\widehat{\beta}_{MLE}) = \mathbf{F}_{kd} \widehat{\beta}_{MLE} \quad (5)$$

where  $\mathbf{F}_{kd} = (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} - d\mathbf{I})$ ,  $k > 0$ ,  $-\infty < d < +\infty$ .

Assume that the following general stochastic linear restrictions is given in addition to the general logistic regression model

$$\mathbf{h} = \mathbf{H}\beta + \mathbf{v}, E(\mathbf{v}) = \mathbf{0}, Cov(\mathbf{v}) = \mathbf{\Psi} \quad (6)$$

where  $\mathbf{H}$  is  $q \times (p + 1)$  known matrix and  $\mathbf{h}$  is a  $q \times 1$  vector of pre-specified values.  $\mathbf{v}$  is an random vector. We assumed that  $\mathbf{v}$  is independent of  $\varepsilon$ .

By combining (1) and (6), [16] proposed the following stochastic restricted maximum likelihood estimator (SRE) as follows:

$$\widehat{\beta}_{\text{SRE}} = \left( \mathbf{C} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{h} \right). \quad (7)$$

[18] proposed the stochastic restricted Liu maximum likelihood estimator (SRLE), by combining the LE and SRE, as

$$\widehat{\beta}_{\text{SRLE}}(d) = \mathbf{F}_d \left( \mathbf{C} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{h} \right) = \mathbf{F}_d \widehat{\beta}_{\text{SRE}}. \quad (8)$$

Now incorporating the LTE to the logistic regression model under the stochastic linear restriction, we propose a new biased estimator which is called stochastic restricted Liu-type maximum likelihood estimator (SRLTE)

$$\widehat{\beta}_{\text{SRLTE}}(k, d) = \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{h} \right) = \mathbf{F}_{kd} \widehat{\beta}_{\text{SRE}} \quad (9)$$

where  $\mathbf{F}_{kd} = \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + k\mathbf{I} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} - d\mathbf{I} \right)$ ,  $k > 0$ ,  $-\infty < d < +\infty$  are two biasing parameters. By the definition of SRLTE, we can see that it is a general estimator which includes the MLE, LE, LTE, SRE and SRLE as special cases.

- When  $k = -d$ ,

$$\widehat{\beta}_{\text{SRLTE}}(k, d) = \widehat{\beta}_{\text{SRE}} = \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{h} \right),$$

- When  $k = 1, d = -d$ ,

$$\widehat{\beta}_{\text{SRLTE}}(k, d) = \widehat{\beta}_{\text{SRLE}} = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \widehat{\mathbf{z}} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{h} \right),$$

- When  $\mathbf{H} = \mathbf{0}$ ,  $\widehat{\beta}_{\text{SRLTE}}(k, d) = \widehat{\beta}_{\text{LTE}} = \mathbf{F}_{kd} \widehat{\beta}_{\text{MLE}}$ ,
- When  $\mathbf{H} = \mathbf{0}$ ,  $k = -d$ ,  $\widehat{\beta}_{\text{SRLTE}}(k, d) = \widehat{\beta}_{\text{MLE}}$ ,
- When  $\mathbf{H} = \mathbf{0}$ ,  $k = 1, d = -d$ ,  $\widehat{\beta}_{\text{SRLTE}}(k, d) = \widehat{\beta}_{\text{LE}} = \mathbf{F}_d \widehat{\beta}_{\text{MLE}}$ .

In next section, we will study the performance of the new estimator over the existing estimator given in the literature.

### 3. THE PROPERTIES OF THE NEW ESTIMATOR

In this section, we will present the comparison of the new estimator with other estimators under the mean squared error matrix criteria. Firstly, we present the definition of mean squared error matrix. The mean squared error matrix (MSEM) of estimator  $\widetilde{\beta}$  is defined as

$$MSEM(\widetilde{\beta}) = E \left[ \left( \widetilde{\beta} - \beta \right) \left( \widetilde{\beta} - \beta \right)^\top \right] = Cov \left( \widetilde{\beta} \right) + Bias \left( \widetilde{\beta} \right) Bias \left( \widetilde{\beta} \right)^\top.$$

where  $Cov(\tilde{\beta})$  is the covariance matrix and  $Bias(\tilde{\beta})$  is the bias of the estimator  $\tilde{\beta}$ . The scalar mean squared error (SMSE) is defined as follows:

$$SMSE(\tilde{\beta}) = tr[MSEM(\tilde{\beta})]$$

where  $tr$  is the trace of a matrix. Now, we will present three lemmas and use them to prove the theorems given in the next section.

**Lemma 1.** ([3], [12]) Suppose that  $\mathbf{M}$  be a positive definite matrix, namely  $\mathbf{M} > \mathbf{0}$ ,  $\alpha$  be some vector, then  $\mathbf{M} - \alpha\alpha^\top \geq \mathbf{0}$  if and only if  $\alpha^\top \mathbf{M}^{-1} \alpha \leq 1$ .

**Lemma 2.** ([12]) Let  $\mathbf{M} > \mathbf{0}$ ,  $\mathbf{N} > \mathbf{0}$ , then  $\mathbf{M} > \mathbf{N}$ , if and only if  $\lambda_{\max}(\mathbf{N}\mathbf{M}^{-1}) < 1$ .

**Lemma 3.** ([12]) Let  $\hat{\beta}_j = \mathbf{A}_j \mathbf{y}$ ,  $j = 1, 2$  be two competing estimator of  $\beta$ . Assume that  $\Delta = Cov(\hat{\beta}_1) - Cov(\hat{\beta}_2) > \mathbf{0}$ , then  $MSEM(\hat{\beta}_1) - MSEM(\hat{\beta}_2) > \mathbf{0}$  if and only if  $u_2^\top (\Delta + \mathbf{u}_1 \mathbf{u}_1^\top)^{-1} \mathbf{u}_2 \leq 1$ , where  $\mathbf{u}_j$  denotes the bias of  $\hat{\beta}_j$ .

**3.1. The new estimator versus SRE.** The asymptotic properties of the SRE are given as follows: The expected value and the covariance matrix are respectively given by

$$E[\hat{\beta}_{SRE}] = \beta \text{ and } Cov[\hat{\beta}_{SRE}] = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1}. \quad (10)$$

Then we may have the MSEM of SRE as

$$MSEM[\hat{\beta}_{SRE}] = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1}. \quad (11)$$

The asymptotic properties of the new estimator SRLTE are obtained similarly as follows:

$$E[\hat{\beta}_{SRLTE}(k, d)] = \mathbf{F}_{kd} \beta = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} - d\mathbf{I}) \beta \quad (12)$$

and

$$Cov[\hat{\beta}_{SRLTE}(k, d)] = \mathbf{F}_{kd} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1} \mathbf{F}_{kd}^\top. \quad (13)$$

Then MSEM of SRLTE is given by

$$MSEM[\hat{\beta}_{SRLTE}(k, d)] = \mathbf{F}_{kd} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1} \mathbf{F}_{kd}^\top + \mathbf{b}_1 \mathbf{b}_1^\top \quad (14)$$

where  $\mathbf{b}_1 = -(d+k) (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + k\mathbf{I})^{-1} \beta$ .

Now we consider the following difference

$$\begin{aligned} \Delta_1 &= MSEM[\hat{\beta}_{SRE}] - MSEM[\hat{\beta}_{SRLTE}(k, d)] \\ &= (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1} - \mathbf{F}_{kd} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \Psi^{-1} \mathbf{H})^{-1} \mathbf{F}_{kd}^\top - \mathbf{b}_1 \mathbf{b}_1^\top \\ &= \mathbf{D}_1 - \mathbf{b}_1 \mathbf{b}_1^\top \end{aligned} \quad (15)$$

where  $\mathbf{D}_1 = \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top$ .

Since  $\left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} > \mathbf{0}$ ,  $\mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top > \mathbf{0}$ , then by Lemma 2, when

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right) \right\} < 1,$$

$\mathbf{D}_1 > \mathbf{0}$ . Then by Lemma 1, when  $\mathbf{b}_1^\top \mathbf{D}_1^{-1} \mathbf{b}_1 \leq 1$ ,  $\boldsymbol{\Delta}_1 \geq \mathbf{0}$ .

Based on the arguments above, we may have the following theorem:

**Theorem 4.** *When*

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi}^{-1} \mathbf{H} \right) \right\} < 1,$$

*the new estimator is superior to the SRE if and only if  $\mathbf{b}_1^\top \mathbf{D}_1^{-1} \mathbf{b}_1 \leq 1$ .*

**3.2. The new estimator versus SRLE.** The asymptotic properties of the SRLE are obtained as follows:

$$E \left[ \widehat{\beta}_{\text{SRLE}}(d) \right] = \mathbf{F}_d \boldsymbol{\beta} = \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{I} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + d\mathbf{I} \right) \boldsymbol{\beta} \quad (16)$$

and

$$\text{Cov} \left[ \widehat{\beta}_{\text{SRLE}}(d) \right] = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top. \quad (17)$$

Then we have MSEM of SRLE as

$$\text{MSEM} \left[ \widehat{\beta}_{\text{SRLE}}(d) \right] = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top + \mathbf{b}_2 \mathbf{b}_2^\top \quad (18)$$

where  $\mathbf{b}_2 = (d-1) \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{I} \right)^{-1} \boldsymbol{\beta}$ .

Now we consider the following difference

$$\begin{aligned} \boldsymbol{\Delta}_2 &= \text{MSEM} \left[ \widehat{\beta}_{\text{SRLE}} \right] - \text{MSEM} \left[ \widehat{\beta}_{\text{SRLTE}}(k, d) \right] \\ &= \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \\ &\quad + \mathbf{b}_2 \mathbf{b}_2^\top - \mathbf{b}_1 \mathbf{b}_1^\top \\ &= \mathbf{D}_2 + \mathbf{b}_2 \mathbf{b}_2^\top - \mathbf{b}_1 \mathbf{b}_1^\top \end{aligned} \quad (19)$$

where  $\mathbf{D}_2 = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top$ .

Since  $\mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top > \mathbf{0}$  and  $\mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top > \mathbf{0}$ , then by Lemma 2, when

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left[ \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}}\mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top \right]^{-1} \right\} < 1,$$

$\mathbf{D}_2 > \mathbf{0}$ . Thus by Lemma 3, when  $\mathbf{b}_1^\top (\mathbf{D}_2 + \mathbf{b}_2 \mathbf{b}_2^\top)^{-1} \mathbf{b}_1 \leq 1$ ,  $\mathbf{\Delta}_2 \geq \mathbf{0}$ .  
Based on the computations above, we may write the following theorem:

**Theorem 5.** *When*

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left[ \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_d^\top \right]^{-1} \right\} < 1,$$

the new estimator is superior to the SRLE if and only if  $\mathbf{b}_1^\top (\mathbf{D}_2 + \mathbf{b}_2 \mathbf{b}_2^\top)^{-1} \mathbf{b}_1 \leq 1$ .

**3.3. The new estimator versus LTE.** The asymptotic properties of the LTE are given by

$$E \left[ \widehat{\beta}_{\text{LTE}}(k, d) \right] = \mathbf{F}_{kd} \boldsymbol{\beta} = \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + k \mathbf{I} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} - d \mathbf{I} \right) \boldsymbol{\beta} \quad (20)$$

and

$$\text{Cov} \left[ \widehat{\beta}_{\text{LTE}}(k, d) \right] = \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_{kd}^\top \quad (21)$$

Then we have

$$MSEM \left[ \widehat{\beta}_{\text{LTE}}(k, d) \right] = \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_{kd}^\top + \mathbf{b}_1 \mathbf{b}_1^\top \quad (22)$$

where  $\mathbf{b}_1 = -(d+k) \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + k \mathbf{I} \right)^{-1} \boldsymbol{\beta}$ . Thus, the following difference is obtained

$$\begin{aligned} \Delta_3 &= MSEM \left[ \widehat{\beta}_{\text{LTE}} \right] - MSEM \left[ \widehat{\beta}_{\text{SRLTE}}(k, d) \right] \\ &= \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_{kd}^\top - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \\ &= \mathbf{F}_{kd} \left[ \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} - \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \right] \mathbf{F}_{kd}^\top. \end{aligned} \quad (23)$$

By [12], the following identity holds:

$$\begin{aligned} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} &= \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} - \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \\ &\quad \times \mathbf{H}^\top \left[ \mathbf{H} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{H}^\top \right]^{-1} \mathbf{H} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1}. \end{aligned}$$

Then

$$\begin{aligned} &\left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{H}^\top \left( \mathbf{H} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{H} \right)^{-1} \mathbf{H} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \\ &= \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} - \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \geq \mathbf{0}. \end{aligned}$$

So  $\mathbf{\Delta}_3 = MSEM \left[ \widehat{\beta}_{\text{LTE}} \right] - MSEM \left[ \widehat{\beta}_{\text{SRLTE}}(k, d) \right] \geq \mathbf{0}$ .

Therefore, we may present the following theorem:

**Theorem 6.** *The new estimator is always superior to the LTE.*

3.4. **The new estimator versus LE.** The asymptotic properties of the LE are given as:

$$E \left[ \widehat{\beta}_{LE}(d) \right] = \mathbf{F}_d \boldsymbol{\beta} = \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{I} \right)^{-1} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + d \mathbf{I} \right) \boldsymbol{\beta} \quad (24)$$

and

$$Cov \left[ \widehat{\beta}_{LE}(d) \right] = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top. \quad (25)$$

Then we have

$$MSEM \left[ \widehat{\beta}_{LE}(d) \right] = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top + \mathbf{b}_2 \mathbf{b}_2^\top \quad (26)$$

where  $\mathbf{b}_2 = (d-1) \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{I} \right)^{-1} \boldsymbol{\beta}$ .

Now using (26) and (14) we consider the following difference

$$\begin{aligned} \boldsymbol{\Delta}_4 &= MSEM \left[ \widehat{\beta}_{LE} \right] - MSEM \left[ \widehat{\beta}_{SRLTE}(k, d) \right] \\ &= \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top + \mathbf{b}_2 \mathbf{b}_2^\top - \mathbf{b}_1 \mathbf{b}_1^\top \\ &= \mathbf{D}_3 - \mathbf{b}_1 \mathbf{b}_1^\top. \end{aligned} \quad (27)$$

where  $\mathbf{D}_3 = \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top - \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top$ .

Since  $\mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top > \mathbf{0}$  and  $\mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top > \mathbf{0}$ , then by Lemma 2, when

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left[ \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top \right]^{-1} \right\} < 1,$$

$\mathbf{D}_3 > \mathbf{0}$ . Thus by Lemma 3, when  $\mathbf{b}_1^\top (\mathbf{D}_3 + \mathbf{b}_2 \mathbf{b}_2^\top)^{-1} \mathbf{b}_1 \leq 1$ ,  $\boldsymbol{\Delta}_4 \geq \mathbf{0}$ .

As a result, we can write the following theorem:

**Theorem 7.** *When*

$$\lambda_{\max} \left\{ \mathbf{F}_{kd} \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H} \right)^{-1} \mathbf{F}_{kd}^\top \left[ \mathbf{F}_d \left( \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{F}_d^\top \right]^{-1} \right\} < 1,$$

*the new estimator is superior to the LE if and only if  $\mathbf{b}_1^\top (\mathbf{D}_3 + \mathbf{b}_2 \mathbf{b}_2^\top)^{-1} \mathbf{b}_1 \leq 1$ .*

3.5. **Selection of the parameters  $\mathbf{k}$  and  $\mathbf{d}$ .** We propose an algorithm to choose the biasing parameters  $k$  and  $d$  iteratively to use in SRLTE. Since  $\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} > \mathbf{0}$ , using the spectral decomposition of the matrix  $\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} = \mathbf{Q} \boldsymbol{\Lambda} \mathbf{Q}^\top$ , where  $\mathbf{Q}$  is the matrix whose columns are the eigenvectors of  $\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X}$  and  $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{p+1})$

such that  $\lambda_i$ 's are eigenvalues of  $\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X}$ , then we have

$$SMSE \left[ \widehat{\beta}_{\text{SRLTE}}(k, d) \right] = \sum_{i=1}^{p+1} \frac{(\lambda_i - d)^2 b_{ii} + (d + k)^2 \alpha_i^2}{(\lambda_i + k)^2} \quad (28)$$

where  $(\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} + \mathbf{H}^\top \boldsymbol{\Psi} \mathbf{H})^{-1} = \mathbf{Q} \text{diag}(b_{11}, \dots, b_{(p+1)(p+1)}) \mathbf{Q}^\top$  and  $\hat{\boldsymbol{\alpha}} = \mathbf{Q}^\top \widehat{\boldsymbol{\beta}}_{\text{MLE}}$ .

To estimate  $k$ , we start by taking the derivative of Equation (28) with respect to  $k$  and equating the resulting function to zero then solving for each individual parameter, we obtain the following:

$$k = \frac{(\lambda_i - d)b_{ii} - d}{\alpha_i^2}. \quad (29)$$

Since  $k$  is positive by definition, the numerator of Equation (29) gives us an estimate of  $d$  as

$$\frac{\lambda_i b_{ii}}{b_{ii} + 1} \geq d \quad (30)$$

Thus, we propose the following algorithm:

- (1) Find an initial estimate of  $d$  using (30) by choosing the minimum of those values.
- (2) Compute the individual  $k$  values using (29) and compute their minimum as an estimate of  $k$ .

#### 4. A SIMULATION STUDY

In this section we present the details of a Monte Carlo simulation study which is conducted to evaluate the performances of the estimators MLE, LE, LTE, SRE, SRLE and SRLTE. We consider two criteria which are the simulated mean squared error (MSE) and predictive MSE (PMSE) to compare the performances of listed estimators. Following [10], and [18], we generate the data matrix  $\mathbf{X}$  containing the explanatory variables such that  $\rho^2$  is the degree of collinearity between any two variables as follows:

$$x_{ij} = (1 - \rho^2)^{1/2} w_{ij} + \rho w_{i(p+1)} \quad (31)$$

where  $w_{ij}$  are obtained from the standard normal distribution. The data matrix is centered and standardized before computations.

We consider the following setting:

- The sample size is taken as 50, 100 and 200.
- The number of explanatory variables is  $p = 4$ .
- The parameter values as chosen such that  $\boldsymbol{\beta}^\top \boldsymbol{\beta} = 1$ .
- The degree of correlation  $\rho$  changes as 0.9, 0.99 and 0.999.
- The dependent variable is generated using  $\pi_i = \frac{e^{x_i \boldsymbol{\beta}}}{1 + e^{x_i \boldsymbol{\beta}}}$

- The restrictions are chosen as follows: Following [18], we choose  $\mathbf{H} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$  and  $\mathbf{h} = \mathbf{H}\boldsymbol{\beta} + \mathbf{v}$  where  $\mathbf{v} \sim N(\mathbf{0}, \boldsymbol{\Psi})$  such that  $\boldsymbol{\Psi} = \mathbf{I}_3$ .
- We choose the biasing parameter of LE and SRLE following [8]. For LTE, we follow [2] to choose  $k$  and  $d$ . For SRLTE, we use our proposed algorithm to estimate the parameters.

TABLE 1. The simulated MSE values

n	MLE	LE	LTE	SRMLE	SRLE	SRLTE
$\rho = 0.9$						
50	1.85652	1.36613	1.12453	0.51381	0.66333	0.58042
100	1.78156	1.34392	1.10278	0.58567	0.74037	0.67935
200	1.75873	1.50912	1.16144	0.86342	0.89278	0.91667
$\rho = 0.99$						
50	2.60743	1.03528	1.20035	0.72942	0.93741	0.71122
100	2.18982	1.05415	1.13695	0.67586	0.90969	0.65671
200	1.89717	1.11810	1.08619	0.60522	0.82612	0.58546
$\rho = 0.999$						
50	17.24458	1.00220	3.61286	0.82232	0.99678	0.82139
100	5.79611	1.00791	1.78444	0.81379	0.98176	0.80758
200	3.41868	1.01677	1.36106	0.77453	0.96657	0.76399

TABLE 2. The simulated PMSE values

n	MLE	LE	LTE	SRMLE	SRLE	SRLTE
$\rho = 0.9$						
50	2.66301	2.42390	2.49373	2.35958	2.33175	2.34593
100	4.80013	4.58821	4.65525	4.56813	4.54326	4.55901
200	9.63524	9.22454	9.45762	9.13080	9.11585	9.12089
$\rho = 0.99$						
50	2.32046	2.25402	2.20685	2.22910	2.19648	2.22816
100	4.49735	4.42789	4.38679	4.41898	4.37117	4.41659
200	8.87562	8.78514	8.76574	8.80102	8.73969	8.79209
$\rho = 0.999$						
50	2.30213	2.25291	2.18845	2.19244	2.18757	2.19244
100	4.47697	4.41158	4.36401	4.36879	4.36197	4.36873
200	8.81881	8.76106	8.70836	8.72060	8.70271	8.72036

The simulation is repeated 5000 times. All computations are performed using the R program ([13]). The simulated MSE and PMSE values of estimators are computed respectively as follows:

$$\text{MSE}(\tilde{\beta}) = \frac{\sum_{r=1}^{5000} (\tilde{\beta}_r - \beta)^\top (\tilde{\beta}_r - \beta)}{5000}$$

and

$$\text{PMSE}(\tilde{\beta}) = \frac{\sum_{r=1}^{5000} (\tilde{\pi}_r - \pi)^\top (\tilde{\pi}_r - \pi)}{5000}$$

where  $\tilde{\beta}_r$  is any estimator considered in the study and  $\tilde{\pi}_r$  is the predicted probabilities in the  $r^{\text{th}}$  repetition.

The results of the simulation are presented in Tables 1-4. According to the results, we observe that SRLTE has the best performance in the sense of MSE for most of the situations and MLE has the worst performance in the sense of both MSE and PMSE. When  $\rho = 0.9$ , SRE has the best performance, however, SRLTE becomes the second best estimator. It was expected that SRLE would have a better performance than SRE, but it is not. This may be the result of the selection of the biasing parameter.

Although, MSE of LTE is lower than MSE of LE and MLE when the correlation is low, LE has a better performance than LTE when the correlation is high.

It is also seen from tables that increasing the degree of correlation makes an increase in the MSE values of the estimators except for LTE and LE. However, if we consider PMSE values, this situation is vice versa, namely, if the degree of correlation increases, PMSE values decrease. According to Table 2, it is observed that SRLE has the best predictive performance. SRLTE is the second best estimator in the sense of PMSE.

Moreover, if we consider MSE as a function of the sample size, we conclude that all the estimators have asymptotic properties for most of the cases, in other words, if the sample size increases, MSE values decrease. On the other hand the PMSE values increase for the same situation.

## 5. CONCLUSION

In this paper, we propose a new stochastic restricted estimator SRLTE and study its performance both theoretically and numerically. We obtain some conditions such that SRLTE is superior to SRE, SRLE, LTE and LE using mean square error matrices of estimators. We also propose an algorithm to choose the biasing parameters of SRLTE. Moreover, we conduct a Monte Carlo simulation experiment to compare the performances of estimators numerically. We used both MSE and predictive MSE as comparison criteria. According to the results of the simulation, SRLTE performs better in most of the situations in the sense of MSE.

## REFERENCES

- [1] Asar, Y. and Genç, A. New Shrinkage Parameters for the Liu-Type Logistic Estimators, *Communication in Statistics-Simulation and Computation* 45(12), (2016), 1094–1103.
- [2] Asar, Y. Some New Methods to Solve Multicollinearity in Logistic Regression, *Communications in Statistics-Simulation and Computation* 46(4), (2017), 2576–2586.
- [3] Farebrother, R.W. Further Results on the Mean Square Error of Ridge Regression, *Journal of the Royal Statistical Society B* 38,(1976), 248–250.
- [4] Liu, K. A new class of biased estimate in linear regression, *Communications in Statistics Theory and Method* 22(2), (1993), 393–402.
- [5] Liu, K. Using Liu-type estimator to combat collinearity, *Communications in Statistics Theory and Method* 32(5), (2003), 1009–1020.
- [6] Saleh, A. M. E., and Kibria, B. G. Improved ridge regression estimators for the logistic regression model, *Computational Statistics* 28(6), (2013), 2519–2558.
- [7] Kibria, B. G., and Saleh, A. M. E. Improving the estimators of the parameters of a probit regression model: A ridge regression approach, *Journal of Statistical Planning and Inference* 142(6), (2012), 1421–1435.
- [8] Månsson, K., Kibria, B. G., and Shukur, G. On Liu estimators for the logit regression model, *Economic Modelling* 29(4), (2012), 1483–1488.
- [9] Månsson, K., Kibria, B. G., and Shukur, G. A restricted Liu estimator for binary regression models and its application to an applied demand system, *Journal of Applied Statistics* 43(6), (2016), 1119–1127.
- [10] McDonald, G. C., and Galarneau, D. I. A Monte Carlo evaluation of some ridge-type estimators, *Journal of the American Statistical Association* 70(350), (1975), 407–416.
- [11] İnan, D., and Erdoğan, B. E. Liu-type logistic estimator, *Communications in Statistics-Simulation and Computation* 42(7), (2013), 1578–1586.
- [12] Rao, C. R., Helge Toutenburg, S., Shalabh, and Heumann, C. Linear models and generalizations, least squares and alternatives, 3rd Edition Springer Berlin Heidelberg New York, 2008.
- [13] R Development Core Team R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2016.
- [14] Wu, J., and Asar, Y. On almost unbiased ridge logistic estimator for the logistic regression model, *Hacettepe Journal of Mathematics and Statistics* 45(3), (2016), 989–998.
- [15] Wu, J., and Asar, Y., More on the restricted Liu estimator in the logistic regression model, *Communications in Statistics - Simulation and Computation* 46(5), (2017), 3680–3689
- [16] Varathan, N. and Wijekoon P. Stochastic restricted maximum likelihood estimator in logistic regression model, *Open Journal of Statistics* 5, (2015), 837–851.
- [17] Varathan, N., and Wijekoon, P. Ridge Estimator in Logistic Regression under Stochastic Linear Restrictions, *British Journal of Mathematics and Computer Science* 15(3), (2016), 1.
- [18] Varathan, N., and Wijekoon, P. Logistic Liu Estimator under stochastic linear restrictions, *Statistical Papers* (2016), <https://doi.org/10.1007/s00362-016-0856-6>

*Current address:* Jibo Wu: Key Laboratory of Group & Graph Theories and Applications, Chongqing University of Arts and Sciences, Chongqing, 402160, China.

*E-mail address:* [linfen52@126.com](mailto:linfen52@126.com)

ORCID Address: <http://orcid.org/0000-0001-6233-6704>

*Current address:* Yasin Asar (Corresponding Author): Department of Mathematics-computer Sciences, Necmettin Erbakan University, Konya, Turkey.

*E-mail address:* [yasar@konya.edu.tr](mailto:yasar@konya.edu.tr), [yasinasar@hotmail.com](mailto:yasinasar@hotmail.com)

ORCID Address: <http://orcid.org/0000-0003-1109-8456>