scientific reports

OPEN

A new class of Poisson Ridge-type estimator

Esra Ertan ✉ & Kadri Ulaş Akay

The Poisson Regression Model (PRM) is one of the benchmark models when analyzing the count data. The Maximum Likelihood Estimator (MLE) is used to estimate the model parameters in PRMs. However, the MLE may suffer from various drawbacks that arise due to the existence of multicollinearity problems. Many estimators have been proposed as alternatives to each other to alleviate the multicollinearity problem in PRM, such as Poisson Ridge Estimator (PRE), Poisson Liu Estimator (PLE), Poisson Liu-type Estimator (PLTE), and Improvement Liu-Type Estimator (ILTE). In this study, we define a new general class of estimators which is based on the PRE as an alternative to other existing biased estimators in the PRMs. The superiority of the proposed biased estimator over the other existing biased estimators is given under the asymptotic matrix mean square error sense. Furthermore, two separate Monte Carlo simulation studies are implemented to compare the performances of the proposed biased estimators. Finally, the performances of all considered biased estimators are shown in real data.

The Poisson Regression Model (PRM) is one of the benchmark models for count data in much the same way as the normal linear regression model is the benchmark for continuous data[1]. In the PRM, y_i is the response variable and follows a Poisson distribution with mean μ_i, then the probability function is defined as

$$f\left(y_i\right) = \frac{e^{-\mu_i}\mu_i^{y_i}}{y_i!}, \quad i = 1, 2, \ldots, n, y_i = 0, 1, 2, \ldots \tag{1}$$

where μ_i is expressed by using canonical log link function and a linear combination of explanatory variables as follows $\mu_i = \exp\left(x_i'\beta\right)$ where x_i' is the ith row of X, which is an $n \times (p + 1)$ data matrix with p explanatory variables and β is a $(p + 1) \times 1$ vector of coefficients.

The Maximum Likelihood method is the well-known estimation technique to estimate the model parameters in PRMs[2]. The log-likelihood function for PRM is given as follows

$$l(\beta) = \sum_{i=1}^{n} y_i x_i'\beta - \exp\left(x_i'\beta\right) - \log\left(y_i!\right). \tag{2}$$

The Maximum Likelihood Estimator (MLE) of β is obtained by maximizing the log-likelihood function, so the following equations are obtained as

$$S(\beta) = \frac{\partial l(\beta; y)}{\partial \beta} = \sum_{i=1}^{n} \left[y_i - \exp\left(x_i'\beta\right)\right] x_i = 0. \tag{3}$$

Since Eq. (3) is nonlinear function of parameter β, the solution of $S(\beta)$ is obtained using the following iteratively reweighted least squares (IRLS) algorithm

$$\hat{\beta}_{MLE} = \left(X'\hat{W}X\right)^{-1} X'\hat{W}Z, \tag{4}$$

where Z is an n-dimensional vector with the ith element $z_i = \log\left(\hat{\mu}_i\right) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$ and $\hat{W} = \text{diag}\left[\hat{\mu}_i\right]$[3]. The iteration ends when the difference between the old and updated values is less than a given small value, which is usually 10^{-8}[4]. The asymptotic variance–covariance matrix of $\hat{\beta}_{MLE}$ is $cov\left(\hat{\beta}_{MLE}\right) = \left(X'\hat{W}X\right)^{-1}$.

Besides being a widely used estimator of MLE, one of its major disadvantages is that parameter estimates become unstable in the case of multicollinearity[5–13]. The multicollinearity problem, which occurs because of the

Department of Mathematics, Science Faculty, University of Istanbul, Vezneciler, Beyazit, 34134 Istanbul, Turkey. ✉email: eertan@istanbul.edu.tr

approximately linear relationship between the explanatory variables, affects the estimates of model parameters in the PRMs as well as in the linear regression models. One effect of the multicollinearity between explanatory variables is that the variance of the MLE becomes so large that the estimates of the model parameters become unstable[14–20].

In order to reduce the undesirable effects of multicollinearity, the biased estimators that are alternative to the MLE are generalized in a manner similar to that introduced in the linear regression model. For example, Månsson and Shukur[18] proposed the Poisson Ridge Estimator (PRE) as follows:

$$\hat{\beta}_{PRE} = \left(X'\hat{W}X + kI\right)^{-1} X'\hat{W}X\hat{\beta}_{MLE}, \quad k > 0, \tag{5}$$

where k is a biasing parameter. The PRE is the generalization of the Ridge estimator introduced by Hoerl and Kennard[21] for the linear regression model.

Månsson et al.[19], Amin et al.[22] and Qasim et al.[23] defined the Poisson Liu Estimator (PLE) as follows:

$$\hat{\beta}_{PLE} = \left(X'\hat{W}X + I\right)^{-1} \left(X'\hat{W}X + dI\right)\hat{\beta}_{MLE}, \tag{6}$$

where $0 < d < 1$ is a biasing parameter. The PLE is the generalization of the Liu estimator introduced by Liu[24] for the linear regression model.

In recent years, the estimators with two biasing parameters have been proposed as an alternative to PRE and PLE. The purpose of estimators with two biasing parameters obtained by combining several estimators is to obtain more suitable estimators for parameter estimates. In this context, Algamal[25] defined the Poisson Liu-type estimator (PLTE) for the PRMs as follows:

$$\hat{\beta}_{PLTE} = \left(X'\hat{W}X + kI\right)^{-1} \left(X'\hat{W}X - dI\right)\hat{\beta}_{MLE}, \tag{7}$$

where $k > 0$ and $d \in R$ are the biasing parameters. The PLTE is a generalization of the Liu-type estimator, which is firstly introduced by Liu[26]. The PLTE is based on the biasing parameters k and d.

Moreover, Asar and Genç[15] and Çetinkaya and Kaçıranlar[16] proposed another biased estimator with two biasing parameters, defined by Özkale and Kaçıranlar[27] for the linear regression models. The Poisson two-parameter Estimator (PTPE) is defined as:

$$\hat{\beta}_{PTPE} = \left(X'\hat{W}X + kI\right)^{-1} \left(X'\hat{W}X + kdI\right)\hat{\beta}_{MLE}, \tag{8}$$

where $k > 0$ and $0 < d < 1$ are the biasing parameters.

As an alternative to the estimators introduced so far, Akay and Ertan[5] proposed a general Improved Liu-type Estimator (ILTE) which includes MLE, PRE, PLE, PLTE and PTPE as special cases as follows:

$$\hat{\beta}_{ILTE} = \left(X'\hat{W}X + kI\right)^{-1} \left(X'\hat{W}X + f(k)I\right)\hat{\beta}^*, k > 0, \tag{9}$$

where $\hat{\beta}^*$ is any estimator of β and $f(k)$ is a continuous function of the biasing parameter k. The estimator given in (9) is a generalization of the Liu-type estimator proposed by Kurnaz and Akay[28] for linear regression models.

In the literature, many estimators proposed for linear regression models can be generalized to be applied to PRMs. For example, the estimator depending on the Ridge estimator in linear regression models was proposed by Yang and Chang[29]. In this sense, the biased estimator proposed by Yang and Chang[29] is adapted from the PRMs by Asar and Genç[15]. In addition, this estimator is applied to Negative Binomial regression models by Huang and Yang[30]. Depending on the PRE, the estimator given by Huang and Yang[30] in the literature has been as follows:

$$\hat{\beta}_{PHY}(k, d) = \left(X'\hat{W}X + I\right)^{-1} \left(X'\hat{W}X + dI\right)\left(X'\hat{W}X + kI\right)^{-1} X'\hat{W}X\hat{\beta}_{MLE}, \quad k > 0, 0 < d < 1, \tag{10}$$

where k and d are two biasing parameters. Although the estimator given in (10) is depending on the PRE, it is a general estimator which includes the MLE, PRE, and PLE as special cases, too.

From this point of view, another estimator depending on the Ridge estimator in linear regression models was proposed by Sakallıoğlu and Kaçıranlar[31], and is defined by Sakallıoğlu and Kaçıranlar[31] which is defined as:

$$\hat{\beta}_{SK}(k, d) = (X'X + I)^{-1} (X'X + (k + d)I)\hat{\beta}_{RE}, \quad k > 0, -\infty < d < \infty, \tag{11}$$

where k and d are two biasing parameters and $\hat{\beta}_{RE} = (X'X + kI)^{-1} X'Y$. In this context, we can generalize the (11) estimators suggested for PRMs. Based on the PRE, we can generalize the estimator proposed by Sakallıoğlu and Kaçıranlar[31] given in (11) as follows:

$$\hat{\beta}_{PSK}(k, d) = \left(X'\hat{W}X + I\right)^{-1} \left(X'\hat{W}X + (k + d)I\right)\hat{\beta}_{PRE}, \quad k > 0, -\infty < d < \infty, \tag{12}$$

where k and d are two biasing parameters. In this case, the estimator given in (12) is a general estimator which includes the MLE, PRE and PLE as special cases. Best of our knowledge, no study has been conducted about estimator in (12) for the PRMs.

In PRMs, it is known that the performance of biased estimators proposed as an alternative to MLE is generally affected by the value of the biasing parameter. In general, the methods used for the estimation of biasing

parameters have been adapted similarly to those used in linear regression models. On the other hand, the use of estimators with two biasing parameters has become increasingly widespread in recent years. However, one of the most important problems for estimators with two biasing parameters is finding optimal estimates of the biasing parameters is difficult. For this purpose, many iterative techniques have been proposed to estimate these biasing parameters. In these cases, one of the biasing parameters can be estimated depending on the other biasing parameter, or vice versa[15,16,30]. Thus, the idea arises that an unknown functional relationship may exist between these two biasing parameters.

Based on the information mentioned above, our aim in this article is to introduce a new general class of estimators that arises when there is a functional relationship between the biasing parameters. In addition, the proposed general estimator can be defined to specifically include the estimators given by (4), (5), (6), (10) and (12). Thus, this proposed estimator constitutes a general class of estimators like the estimator given in (9). It is a more efficient alternative estimator when compared with the one defined in (9) which can overcome multicollinearity in the PRMs. Another purpose of this article is to compare these two class estimators with a simulation study under some conditions.

The remainder of the article is organized as follows: In "A new general biased estimator", a new biased estimator is defined and some of its properties are given. The superiority of this estimator over the other biased estimators under the matrix mean square error sense are shown in "The superiority of the PRTE in PRMs". In "Determination of function", several rules are proposed to determine the relationship between the biasing parameters. Two separate Monte Carlo simulation studies are executed in "The Monte Carlo simulation studies". In "Numerical example: the aircraft damage data", a real numerical example is provided to evaluate the performances of the proposed biased estimators. Some concluding remarks are given in "Some concluding remarks".

A new general biased estimator

For PRMs, we can define a new general class of estimators including (4), (5), (6), (10) and (12) estimators based on the PRE estimator as follows:

$$\hat{\beta}_{PRTE} = \left(X'\hat{W}X + I\right)^{-1}\left(X'\hat{W}X + g(k)I\right)\hat{\beta}_{PRE}, k > 0, \tag{13}$$

where $g(k)$ is a continuous function of the biasing parameter k. When we select $g(k)$ as a linear function of the biasing parameter k such as $g(k) = ak + b$ where $a, b \in R$, the Poisson Ridge-type estimator (PRTE) is a general estimator which includes the other biased estimators as special cases:

$\hat{\beta}_{PRTE} = \hat{\beta}_{MLE}$ for $g(0) = 1$ where $k = 0$ and $b = 1$.
$\hat{\beta}_{PRTE} = \hat{\beta}_{PRE}$ for $g(k) = 1$ where $a = 0$ and $b = 1$.
$\hat{\beta}_{PRTE} = \hat{\beta}_{PLE}$ for $g(0) = b$ where $a = 0$ and b corresponds to the biasing parameter d.
$\hat{\beta}_{PRTE} = \hat{\beta}_{PHY}(k, d)$ for $g(k) = b$ where b corresponds to the biasing parameter d.
$\hat{\beta}_{PRTE} = \hat{\beta}_{PSK}(k, d)$ for $g(k) = k + b$ where $a = 1$ and b corresponds to the biasing parameter d.

Note that, the proposed estimator given in (13) is different form the biased estimator given in (9). That is, when we use $\hat{\beta}_{PRE}$ instead of $\hat{\beta}^*$ in (9), the resulting estimator $\hat{\beta}_{ILTE(PRE)}$ is given as follows:

$$\hat{\beta}_{ILTE(PRE)} = \left(X'\hat{W}X + kI\right)^{-1}\left(X'\hat{W}X + f(k)I\right)\hat{\beta}_{PRE}, k > 0, \tag{14}$$

where $f(k)$ is a continuous function of the biasing parameter k. Note that the estimator given in (14) does not exactly correspond to the estimators given by (10) and (12), respectively. To show that the estimators given in (13) and (14) are different estimators, let's examine the asymptotic scalar mean square error (SMSE) and asymptotic matrix mean square error (MMSE) of these estimators.

In order to obtain the asymptotic SMSE and the asymptotic MMSE of an estimator, we denote $\alpha = Q'\beta$, $\Lambda = diag(\lambda_1, ..., \lambda_{p+1}) = Q'\left(X'\hat{W}X\right)Q$, where $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{p+1} > 0$ are the ordered eigenvalues of $X'\hat{W}X$, Q is the orthogonal matrix whose columns constitute the eigenvectors of $X'\hat{W}X$ and the ith element of $Q'\beta$ is denoted as $\alpha_j, j = 1, 2, ..., p + 1$.

The asymptotic SMSE and the asymptotic MMSE of an estimator $\hat{\beta} = H\hat{\beta}_{MLE}$, where H is an $(p + 1) \times (p + 1)$ matrix, are defined as:

$$MSEM\left(\hat{\beta}\right) = E\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)' = H\left(\hat{\beta}_{MLE} - \beta\right)\left(\hat{\beta}_{MLE} - \beta\right)'H' + (H\beta - \beta)(H\beta - \beta)'$$
$$SMSE\left(\hat{\beta}\right) = E\left(\hat{\beta} - \beta\right)'\left(\hat{\beta} - \beta\right) = \left(\hat{\beta}_{MLE} - \beta\right)'H'H\left(\hat{\beta}_{MLE} - \beta\right) + (H\beta - \beta)'(H\beta - \beta). \tag{15}$$

Note that there is a relationship $SMSE\left(\hat{\beta}\right) = tr\left(MMSE\left(\hat{\beta}\right)\right)$ between MMSE and SMSE criteria. Because of the relation of $\alpha = Q'\beta$; $\hat{\beta}_{MLE}, \hat{\beta}_{PRE}, \hat{\beta}_{PLE}, \hat{\beta}_{PLTE}, \hat{\beta}_{ILTE}$ and $\hat{\beta}_{PRTE}$ have the same SMSE values as $\hat{\alpha}_{MLE}, \hat{\alpha}_{PRE}, \hat{\alpha}_{PLE}, \hat{\alpha}_{PLTE}, \hat{\alpha}_{ILTE}$ and $\hat{\alpha}_{PRTE}$, respectively.

Using (9), (13) and (14), it is easily computed that

$$MMSE\left(\hat{\beta}_{ILTE}\right) = Q\big((\Lambda + kI)^{-1}(\Lambda + f(k)I)\Lambda^{-1}(\Lambda + f(k)I)(\Lambda + kI)^{-1}$$
$$+ (f(k) - k)^2(\Lambda + kI)^{-1}\alpha\alpha'(\Lambda + kI)^{-1}\big)Q' \tag{16}$$

$$MMSE\left(\hat{\beta}_{ILTE(PRE)}\right) = Q\big((\Lambda + kI)^{-1}(\Lambda + f(k)I)(\Lambda + kI)^{-1}\Lambda(\Lambda + kI)^{-1}(\Lambda + f(k)I)(\Lambda + kI)^{-1}$$
$$+(\Lambda + kI)^{-1}\big(f(k)\Lambda - 2k\Lambda - k^2I\big)(\Lambda + kI)^{-1}\alpha\alpha'(\Lambda + kI)^{-1}$$
$$\big(f(k)\Lambda - 2k\Lambda - k^2I\big)(\Lambda + kI)^{-1}\big)Q'.$$

(17)

$$MMSE\left(\hat{\beta}_{PRTE}\right) = Q\big((\Lambda + I)^{-1}(\Lambda + g(k)I)(\Lambda + kI)^{-1}\Lambda(\Lambda + kI)^{-1}(\Lambda + g(k)I)(\Lambda + I)^{-1}$$
$$+\big((g(k) - k - 1)\Lambda - kI\big)(\Lambda + I)^{-1}(\Lambda + kI)^{-1}\alpha\alpha'(\Lambda + kI)^{-1}(\Lambda + I)^{-1}\big((g(k) - k - 1)\Lambda - kI\big)\big)Q'.$$

(18)

Moreover, we can give the SMSE functions of ILTE, ILTE (PRE) and PRTE as follows:

$$SMSE\left(\hat{\beta}_{ILTE}\right) = \sum_{j=1}^{p+1}\frac{\big(\lambda_j + f(k)\big)^2}{\lambda_j\big(\lambda_j + k\big)^2} + \sum_{j=1}^{p+1}\frac{\big(f(k) - k\big)^2\alpha_j^2}{\big(\lambda_j + k\big)^2}$$

(19)

$$SMSE\left(\hat{\beta}_{ILTE(PRE)}\right) = \sum_{j=1}^{p+1}\frac{\big(\lambda_j + f(k)\big)^2\lambda_j}{\big(\lambda_j + k\big)^4} + \sum_{j=1}^{p+1}\frac{\big(f(k)\lambda_j - 2k\lambda_j - k^2\big)^2\alpha_j^2}{\big(\lambda_j + k\big)^4}$$

(20)

$$SMSE\left(\hat{\beta}_{PRTE}\right) = \sum_{j=1}^{p+1}\frac{\lambda_j\big(\lambda_j + g(k)\big)^2}{\big(\lambda_j + 1\big)^2\big(\lambda_j + k\big)^2} + \sum_{j=1}^{p+1}\frac{\big((g(k) - k - 1)\lambda_j - k\big)^2\alpha_j^2}{\big(\lambda_j + 1\big)^2\big(\lambda_j + k\big)^2}$$

(21)

where the first term is the asymptotic variance and the second term is the squared bias. It should be noted that MMSE and SMSE functions of ILTE (PRE) and PRTE are different. Also, the MMSE and SMSE functions of other existing functions can be obtained according to the appropriate selection of $f(k)$ and $g(k)$.

Let $\hat{\beta}_1$ and $\hat{\beta}_2$ be any two estimators of β parameter. Then, $\hat{\beta}_2$ is superior to $\hat{\beta}_1$ with respect to the MMSE sense if and only if $MMSE\left(\hat{\beta}_1\right) - MMSE\left(\hat{\beta}_2\right)$ is a positive definite (pd) matrix. If $MMSE\left(\hat{\beta}_1\right) - MMSE\left(\hat{\beta}_2\right)$ is a non-negative definite matrix, then $SMSE\left(\hat{\beta}_1\right) - SMSE\left(\hat{\beta}_2\right) \geq 0$. But, the reverse is not always true[32].

In order to compare the MMSEs for the above-mentioned biased estimators, we are using the following theorem.

Theorem 2.1 *Let A be a positive definite matrix, namely $A > 0$, and c nonzero vector. Then, $A - cc'$ is positive definite matrix iff $c'A^{-1}c \leq 1$[33].*

The superiority of the PRTE in PRMs

In this section, we compare the PRTE with the ILTE according to the MMSE criterion. Here, we give a general theorem for comparing estimators with different choices of $g(k)$ and $f(k)$ functions. In this way, a general theorem is obtained for comparing the estimators mentioned above in terms of MMSE sense.

The following theorem is given to show the superiority of PRTE over ILTE.

Theorem 3.1. *Let be $k > 0$ and $-\lambda_j - \frac{(\lambda_j + 1)(\lambda_j + f(k))}{\lambda_j} < g(k) < -\lambda_j + \frac{(\lambda_j + 1)(\lambda_j + f(k))}{\lambda_j}$ where $j=1,2,...,p+1$. Then $MMSE\left(\hat{\beta}_{ILTE}\right) - MMSE\left(\hat{\beta}_{PRTE}\right) > 0$ iff*

$$bias\left(\hat{\beta}_{PRTE}\right)'Q\big((\Lambda + kI)^{-1}(\Lambda + f(k)I)\Lambda^{-1}(\Lambda + kI)^{-1}(\Lambda + f(k)I)$$
$$-(\Lambda + I)^{-1}(\Lambda + g(k)I)(\Lambda + kI)^{-1}\Lambda(\Lambda + kI)^{-1}(\Lambda + g(k)I)(\Lambda + I)^{-1}\big)^{-1}Q'bias\left(\hat{\beta}_{PRTE}\right) < 1$$

(22)

where $bias\left(\hat{\beta}_{PRTE}\right) = \big((g(k) - k - 1)\Lambda - kI\big)Q(\Lambda + I)^{-1}(\Lambda + kI)^{-1}\alpha.$

Proof Using (19) and (21), we obtain

$$MMSE\left(\hat{\beta}_{ILTE}\right) - MMSE\left(\hat{\beta}_{PRTE}\right) = Q\big((\Lambda + kI)^{-1}(\Lambda + f(k)I)\Lambda^{-1}(\Lambda + kI)^{-1}(\Lambda + f(k)I)$$
$$-(\Lambda + I)^{-1}(\Lambda + g(k)I)(\Lambda + kI)^{-1}\Lambda(\Lambda + kI)^{-1}(\Lambda + g(k)I)(\Lambda + I)^{-1}\big)^{-1}Q' - bias\left(\hat{\beta}_{PRTE}\right)bias\left(\hat{\beta}_{PRTE}\right)'$$
$$= Qdiag\left\{\frac{\big(\lambda_j + f(k)\big)^2}{\big(\lambda_j + k\big)^2\lambda_j} - \frac{\lambda_j\big(\lambda_j + g(k)\big)^2}{\big(\lambda_j + 1\big)^2\big(\lambda_j + k\big)^2}\right\}_{j=1}^{p+1}Q' - bias\left(\hat{\beta}_{PRTE}\right)bias\left(\hat{\beta}_{PRTE}\right)'.$$

$D = (\Lambda + kI)^{-1}(\Lambda + f(k)I)\Lambda^{-1}(\Lambda + kI)^{-1}(\Lambda + f(k)I) - (\Lambda + I)^{-1}(\Lambda + g(k)I)$
$(\Lambda + kI)^{-1}\Lambda(\Lambda + kI)^{-1}(\Lambda + g(k)I)(\Lambda + I)^{-1}$ is the pd matrix if $\big(\lambda_j + 1\big)^2\big(\lambda_j + f(k)\big)^2 - \lambda_j^2\big(\lambda_j + g(k)\big)^2 > 0.$

Thus D is the pd matrix if $-\lambda_j - \frac{(\lambda_j+1)(\lambda_j+f(k))}{\lambda_j} < g(k) < -\lambda_j + \frac{(\lambda_j+1)(\lambda_j+f(k))}{\lambda_j}$ and $k > 0$ where $j=1,2,...,p+1$. By Theorem 2.1, the proof is completed.

Determination of $g(k)$ function

Since the performance of the biased estimators is related to the choice of biasing parameters, it is an important problem to find the optimal biasing parameters for the proposed biased estimators. Different techniques for estimating the biasing parameters in the PRE, PLE, PLTE, PSK and PHY are generalized depending on the similarities between linear regression models and PRMs[5,15–19,23,30,34]. The performance of PRTE depends on the function $g(k)$, and therefore only on the biasing parameter k. It should be noted that we have given the appropriate choice of the $g(k)$ function in the introduction that different estimators can be obtained. We may give a method to find the optimal $g(k)$ function that approximately minimizes $SMSE\left(\hat{\beta}_{PRTE}\right)$ according to k. Our aim here is to determine the k and $g(k)$ functions together, which can make the $SMSE\left(\hat{\beta}_{PRTE}\right)$ function approximately minimum. In other words, our goal here is to choose the appropriate k and $g(k)$ functions such that the decrease in the variance term is greater than the increase in squared bias. Note that $SMSE\left(\hat{\beta}_{PRTE}\right)$ is a nonlinear function of the biasing parameter k. So, writing $h(k) = SMSE\left(\hat{\beta}_{PRTE}\right)$, then we find $h'(k)$ as follows differentiating $h(k)$ with respect to k,

$$h'(k) = \sum_{j=1}^{p+1} \frac{2\lambda_j\left(\lambda_j - g'(k)\lambda_j - g'(k)k + g(k)\right)\left[\alpha_j^2\left((k+1-g(k))\lambda_j + k\right) - (\lambda_j + g(k))\right]}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\right)^3}.$$

When $h'(k) = 0$, there are two facts as follows;

Fact 1 $\lambda_j\left(\lambda_j - g'(k)\lambda_j - g'(k)k + g(k)\right) = 0$ differential equation is found. From the solution of this differential equation, we obtain

$$g(k) = ck + (c-1)\lambda_j, \tag{23}$$

where c is the constant of integration.

Fact 2 $\alpha_j^2\left((k+1-g(k))\lambda_j + k\right) - (\lambda_j + g(k)) = 0$ equation is found. We have

$$g(k) = \frac{\alpha_j^2(\lambda_j+1)}{1+\lambda_j\alpha_j^2}k + \frac{\left(\alpha_j^2-1\right)}{1+\lambda_j\alpha_j^2}\lambda_j \;\; or \;\; g(k) = \frac{\alpha_j^2(\lambda_j+1)}{1+\lambda_j\alpha_j^2}k + \left(\frac{\alpha_j^2(\lambda_j+1)}{1+\lambda_j\alpha_j^2} - 1\right)\lambda_j. \tag{24}$$

According to the first and the second facts, it is convenient to choose $g(k)$ as a linear function of the biasing parameter k. Note that, $g(k)$ which is obtained in Fact 2 is a solution of the differential equation which is obtained in Fact 1. According to the results obtained in Fact 1 and Fact 2, we can propose the following generalizations. Firstly, note that the function $g(k)$ given in (23) and (24) makes the $SMSE\left(\hat{\alpha}_{PRTE}\right)$ function approximately minimum for a j value. So, $g(k)$ depends on the eigenvalues of $X'WX$, the unknown parameter α and the estimate of the biasing parameter k. In other words, many functions can be determined depending on the functional relationship given in (23) and (24). For example, the following functional relationships can be proposed for the determination of function $g(k)$:

$$g_1(k) = c_1 k + (c_1 - 1)\lambda_{\min} \text{ where } c_1 \in (0,1), \tag{25}$$

$$g_2(k) = \frac{\alpha_{\min}^2(1 + \lambda_{\min})}{1 + \lambda_{\max}\alpha_{\max}^2}k + \left(\frac{\alpha_{\min}^2(1 + \lambda_{\min})}{1 + \lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}, \tag{26}$$

$$g_3(k) = \frac{\min\left(\alpha_j^2(\lambda_j+1)\right)}{n\max\left(1+\lambda_j\alpha_j^2\right)}k + \left(\frac{\min\left(\alpha_j^2(\lambda_j+1)\right)}{n\max\left(1+\lambda_j\alpha_j^2\right)} - 1\right)\lambda_{\min}, \tag{27}$$

where $\alpha_{\min}^2$ and $\alpha_{\max}^2$ are defined as the minimum and maximum value of $\alpha_j^2, j = 1, 2, ..., p+1$, respectively. Similarly, $\lambda_{\min}$ and $\lambda_{\max}$ indicate the minimum and maximum value of the eigenvalue of $X'WX$, respectively.

In this study, we examined only the first degree polynomial functions given in (25) to (27) for $g(k)$ function. Note that, the function $g(k)$ can be selected as any continuous function of the biasing parameter k. Therefore, the proposed biased estimator depends on a single biasing parameter k. In this case, we should use an appropriate estimate of biasing parameter k, which must be estimated to control the conditioning of the $X'WX$ matrix. Since the proposed estimator depends on a single biasing parameter k, the suitable estimates of k can be used given in Månsson and Shukur[18], Kibria et al.[17], Algamal[25]. In addition to the previously proposed estimators of the biasing parameter, we can also use the following estimators to estimate k:

$$\hat{k}_{PRTE} = \frac{p(\lambda_{\max} - \lambda_{\min})}{n}, \hat{k}_{PRTE} = \frac{\max\left(\lambda_j \hat{\alpha}_j^2\right)}{\sum_{j=1}^{p+1} \hat{\alpha}_j^2}, \hat{k}_{PRTE} = \left(\prod_{j=1}^{p+1} \sqrt{\frac{1}{\hat{\alpha}_j^2}}\right)^{\frac{1}{p+1}}$$

where $m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}, j = 1, 2, ..., p+1$ and $\hat{\sigma}^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2$.

The Monte Carlo simulation studies

In this section, we designed two simulation schemes to compare the performances of different biased estimators in the PRMs. In the first simulation scheme, we discussed the effects of sample size (n), the degree of the collinearity (ρ) and the number of the explanatory variables (p) on the performance of the PRTE, PRE, PLE, PLTE, PSK, PHY estimators and PRTE, based on suggested best biasing estimates. In the second simulation design, we examined the effect of the biasing parameter on the performances of the PRTE and ILTE for each set of the values (n, ρ, p, σ^2). For both simulation designs, we generated the explanatory variables by following Månsson and Shukur[18], Kibria et al.[17], Kibria and Lukman[35] as

$$x_{ij} = \left(1 - \rho^2\right)^{1/2} w_{ij} + \rho w_{ip+1}, i = 1, 2, .., n, j = 1, 2, ..., p, \tag{28}$$

where w_{ij} are independent standard normal pseudo-random numbers and ρ is specified such that the correlation between any two variables is given by ρ^2. Four different sets of correlations are investigated corresponding to $\rho = 0.85, 0.9, 0.99$ and 0.999. Number of explanatory variables is determined as $p = 2, 4, 8$ and 12. For each set of explanatory variables, the parameter β is selected as the normalized eigenvector corresponding to the largest eigenvalue of $X'X$ so that $\beta'\beta = 1$. We used glm function in the R Stats package[4]. We also set the intercept term equal to 0.

In the simulation and application sections, the proposed best biasing parameter estimators for PRE, PLE, PLTE, PSK, and PHY estimators are used based on the works of Månsson and Shukur[18], Månsson et al.[19], Kibria et al.[17], Asar and Genç[15], Alanaz and Algamal[34], Çetinkaya and Kaçıranlar[16], Qasim et al.[23], Huang and Yang[30].

To estimate k in PRE, we used the best estimator of k as $\hat{k}_{PRE} = \max\left(\frac{1}{m_j}\right)$ where $m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}, j = 1, 2, ..., p$ and $\hat{\sigma}^2 = \frac{1}{n-p-1} \sum_{i=1}^{n} \left(y_i - \hat{\mu}_i\right)^2$ which is recommended by Kibria et al.[17].

According the results given by Qasim et al.[23], we used the best estimator of d in PLE as $\hat{d}_{PLE} = \max\left(0, \min\left(\frac{\hat{\alpha}_j^2 - 1}{\max\left(\frac{1}{\lambda_j}\right) + \hat{\alpha}_{\max}^2}\right)\right)$.

For PLTE, the biasing parameters k and d are estimated by grouping them in three different ways as follows:

PLTE I: $\hat{k}_{PLTE} = \max\left(\frac{1}{m_j}\right)$ where $m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}, j = 1, 2, ..., p$ and $\hat{d}_{PLTE} = \frac{\sum_{j=1}^{p} \frac{1 - \hat{k}_{PLTE}\hat{\alpha}_j^2}{\left(\lambda_j + \hat{k}_{PLTE}\right)^2}}{\sum_{j=1}^{p} \frac{1 + \lambda_j \hat{\alpha}_j^2}{\lambda_j\left(\lambda_j + \hat{k}_{PLTE}\right)^2}}$.

PLTE II: $\hat{k}_{PLTE} = \frac{\lambda_1 - 100\lambda_p}{99}$ and $\hat{d}_{PLTE} = \frac{\sum_{j=1}^{p} \frac{1 - \hat{k}_{PLTE}\hat{\alpha}_j^2}{\left(\lambda_j + \hat{k}_{PLTE}\right)^2}}{\sum_{j=1}^{p} \frac{1 + \lambda_j \hat{\alpha}_j^2}{\lambda_j\left(\lambda_j + \hat{k}_{PLTE}\right)^2}}$.

PLTE III: $\hat{d}_{PLTE} = \frac{1}{2} \min\left\{\frac{\lambda_j}{1 + \lambda_j \hat{\alpha}_j^2}\right\}, j = 1, 2, ..., p$ and $\hat{k}_{PLTE} = \frac{1}{p} \sum_{j=1}^{p} \frac{\lambda_j - \hat{d}_{PLTE}^*\left(1 + \lambda_j \hat{\alpha}_j^2\right)}{\lambda_j \hat{\alpha}_j^2}$.

Sakallıoğlu and Kaçıranlar[31] did not provide a specific technique for estimating the biasing parameters k and d for SK estimator. Therefore, we used the following estimator to estimate the biasing parameters k and d in PSK:

PSK: $\hat{k}_{PSK} = \max\left(\frac{1}{m_j}\right)$ where $m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}_j^2}}, j = 1, 2, ..., p$ and $\hat{d}_{PSK} = \frac{\sum_{j=1}^{p} \frac{\lambda_j\left(\hat{\alpha}_j^2 - 1\right)}{\left(\lambda_j + 1\right)^2\left(\lambda_j + \hat{k}_{PSK}\right)^2}}{\sum_{j=1}^{p} \frac{\lambda_j\left(1 + \lambda_j \hat{\alpha}_j^2\right)}{\left(\lambda_j + 1\right)^2\left(\lambda_j + \hat{k}_{PSK}\right)^2}}$.

Moreover, we used the methods proposed by Huang and Yang[30] to estimate the parameters of the PHY estimator. Huang and Yang[30] proposed two methods. We refer to these methods as (K1, D1) and (K2, D2) (see Huang and Yang[30] for details). We used these methods by adapting them for the PHY estimator in PRMs. As a result, the estimator obtained with (K1, D1) indicates PHY I, and the estimator obtained with (K2, D2) with PHY II.

We used the following $g(k)$ functions together with the k estimator to determine the PRTE:

PRTE I: $\hat{k}_{PRTEI} = \frac{1}{n}\left(p\lambda_{\max} - (p+1)\lambda_{\min}\right)$ and $g(k) = \frac{(1+\lambda_{\min})\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2}k + \left(\frac{(1+\lambda_{\min})\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}$.

PRTE II: $\hat{k}_{PRTEII} = \frac{p\lambda_{\max}\alpha_{\text{med}}^2}{n\alpha_{\text{mean}}^2}$ and $g(k) = \frac{(1+\lambda_{\max})\alpha_{\min}^2}{p(1+\lambda_{\max}\alpha_{\max}^2)}k + \left(\frac{(1+\lambda_{\max})\alpha_{\min}^2}{p(1+\lambda_{\max}\alpha_{\max}^2)} - 1\right)\lambda_{\min}$.

PRTE III: $\hat{k}_{PRTEIII} = \frac{p}{n}(\lambda_{\max} - \lambda_{\min})$ and $g(k) = \frac{\min\left((1+\lambda_j)\alpha_j^2\right)}{n\max\left(1+\lambda_j \alpha_j^2\right)}k + \left(\frac{\min\left((1+\lambda_j)\alpha_j^2\right)}{n\max\left(1+\lambda_j \alpha_j^2\right)} - 1\right)\lambda_{\min}$.

PRTE IV: $\hat{k}_{PRTEIV} = \frac{p\max\left(\lambda_j \alpha_j^2\right)}{n\alpha_{\text{mean}}^2}$ and $g(k) = \min\left(\frac{(1+\lambda_j)\alpha_j^2}{n\left(1+\lambda_j \alpha_j^2\right)}\right)k + \left(\min\left(\frac{(1+\lambda_j)\alpha_j^2}{n\left(1+\lambda_j \alpha_j^2\right)}\right) - 1\right)\lambda_{\min}$. where

α_{med}^2 and α_{mean}^2 are defined as the median end mean value of $\alpha_j^2, j = 1, 2, ..., p+1$, respectively.

The performance of the estimated MSEs (EMSEs) is used as basis for comparing the proposed estimators which are calculated for an estimator $\hat{\beta}$ of β as

p	n	ρ	MLE	PRE	PLE	PLTE I	PLTE II	PLTE III	PHY I	PHY II	PSK	PRTE I	PRTE II	PRTE III	PRTE IV
2	50	0.85	3.3505	0.4251	0.6666	1.6658	1.8154	0.8999	1.8294	1.7538	1.4409	0.3999**	0.4042	0.4000***	**0.3997***
2	50	0.9	7.1098	0.3893	0.5615	3.3676	3.5148	0.9444	3.5087	3.4565	2.1748	0.3407***	0.3433	0.3389**	**0.3358***
2	50	0.99	32.3787	0.4784	0.3886	14.7461	14.8312	0.6181	14.8231	14.8011	3.2623	0.3428***	0.3434	0.3425**	**0.3400***
2	50	0.999	396.6289	0.6843	0.3317	189.2896	189.2943	0.4948	189.2885	189.343	3.2163	0.3250***	0.3252	0.3248**	**0.3241***
2	100	0.85	4.1795	0.4290	0.6463	2.0166	2.1865	0.9297	2.1837	2.1022	1.6097	0.3884***	0.3906	0.3863**	**0.3856***
2	100	0.9	5.5753	0.4177	0.6092	2.654	2.8457	1.0138	2.8297	2.7448	1.8903	0.3681***	0.3696	0.3653**	**0.3643***
2	100	0.99	35.2475	0.5059	0.3983	16.5814	16.7079	0.7410	16.684	16.6501	3.3132	0.3415***	0.3418	0.3411**	**0.3396***
2	100	0.999	420.3712	0.7112	0.3395	196.4906	196.4976	0.5258	196.4944	196.5028	2.9325	0.3339***	0.3340	0.3338**	**0.3336***
2	200	0.85	4.3705	0.4327	0.6949	2.1726	2.3697	0.9879	2.3652	2.2690	1.7706	0.3982***	0.4007	0.3941**	**0.3936***
2	200	0.9	5.3801	0.3963	0.6039	2.6009	2.7865	0.9417	2.7773	2.7027	1.9059	0.3485***	0.3507	0.3450**	**0.3444***
2	200	0.99	37.3897	0.5087	0.3891	17.348	17.4741	0.6873	17.4506	17.4194	3.1571	0.3380***	0.3382	0.3377**	**0.3368***
2	200	0.999	392.6338	0.7030	0.3369	182.3938	182.3917	0.5019	182.3881	182.3411	2.7437	0.3320**	0.3320**	0.3320**	**0.3316***
2	500	0.85	4.6132	0.4175	0.6410	2.2544	2.4440	0.9466	2.4340	2.3525	1.7471	0.3704***	0.3719	0.3668**	**0.3667***
2	500	0.9	5.5349	0.4101	0.5948	2.6512	2.8319	0.9460	2.8192	2.7452	1.8788	0.3599***	0.3613	0.3575**	**0.3572***
2	500	0.99	43.9818	0.5336	0.3970	20.7216	20.8506	0.7392	20.8297	20.7898	3.3364	0.3490***	0.3491	0.3487**	**0.3484***
2	500	0.999	506.1048	0.7193	0.3363	235.1907	235.2138	0.5256	235.2081	235.1556	2.7036	0.3319**	0.3319**	0.3319**	**0.3318***

Table 1. The EMSE values of the estimators when $p = 2$.

$$EMSE\left(\hat{\beta}\right) = \frac{1}{N} \sum_{r=1}^{N} \left(\hat{\beta}_r - \beta\right)' \left(\hat{\beta}_r - \beta\right), \tag{29}$$

where $\left(\hat{\beta}_r - \beta\right)$ is the difference between the estimated and true parameter vectors at rth replication and N is the number of replications. For each case of n, p and ρ, the experiment was replicated 2000 times by generating response variables. Our Monte Carlo simulation studies were conducted using the R Programming Language. The results for different n, p and ρ are given in Tables 1, 2, 3 and 4 for $p = 2, 4, 8$ and 12 respectively.

The bold numbers in the tables show the estimators with the smallest EMSE values, and in addition, the signs (*), (**), and (***) represent the first, second, and third smallest EMSE values in each row, respectively. The results from Tables 1, 2, 3 and 4 are listed below:

1. According to the results from Tables 1, 2, 3 and 4, it can be seen that the degree of correlation (ρ), the number of explanatory variables (p) and the sample size (n) have different effects on all estimators in the simulation.
2. It has been observed that the EMSE values of PRTE I, PRTE II, PRTE III and PRTE IV are smaller than the other existing biased estimators. Although our proposed estimators PRTE I, PRTE II, PRTE III, and PRTE IV outperformed other existing estimators in all cases, it is also observed that they outperformed each other in different n, p and ρ values.

p	n	ρ	MLE	PRE	PLE	PLTE I	PLTE II	PLTE III	PHY I	PHY II	PSK	PRTE I	PRTE II	PRTE III	PRTE IV
4	50	0.85	14.8073	0.3343	0.7139	5.2396	8.1800	1.1858	7.8278	5.7486	4.4187	0.2396**	0.2491***	**0.2371***	0.2570
4	50	0.9	14.6284	0.3357	0.7303	5.2304	7.0981	1.1121	7.9068	5.6430	4.5980	0.2350**	0.2411***	**0.2338***	0.2434
4	50	0.99	121.3738	0.4260	0.2938	41.7830	42.6074	1.2557	64.1290	45.5313	15.7597	0.1943**	**0.1934***	0.1943**	0.1974
4	50	0.999	1772.2454	0.7404	0.2133	605.9529	606.2798	0.4316	927.2069	666.0135	17.0897	0.2062***	**0.2053***	0.2061**	0.2061**
4	100	0.85	10.5035	0.3967	0.8442	3.8045	5.2600	1.0593	5.7930	4.0477	3.4920	0.2692***	0.2701	0.2663**	**0.2652***
4	100	0.9	17.2789	0.3340	0.6533	6.1263	7.9786	1.1913	9.2264	6.5948	5.1263	0.2315**	0.2341***	**0.2301***	0.2350
4	100	0.99	170.7186	0.5055	0.2774	59.7278	60.1472	1.0102	90.0369	64.9378	17.5231	0.2072***	**0.2066***	0.2071**	0.2086
4	100	0.999	1151.1888	0.6979	0.2089	387.5967	387.3600	0.4226	607.2847	419.3292	16.4946	0.1994***	**0.1986***	0.1994***	0.199**
4	200	0.85	9.8088	0.4143	0.8649	3.5960	5.1733	1.0801	5.4090	3.8168	3.2808	0.2922***	0.2923	0.2883**	**0.2880***
4	200	0.9	14.3776	0.3539	0.7104	5.0917	6.5090	1.1040	7.7311	5.4392	4.4110	0.2466***	0.2476	**0.2447***	0.2450**
4	200	0.99	159.9949	0.5061	0.2825	55.5575	55.9240	1.0048	84.9881	59.8018	16.3681	**0.2116***	0.2121**	0.2116*	0.2140***
4	200	0.999	1301.5585	0.7337	0.2068	450.2752	450.0915	0.4477	688.5009	489.0979	16.2530	0.1975**	**0.1974***	0.1975**	0.1975**
4	500	0.85	10.1800	0.3855	0.7937	3.6070	4.9723	1.0087	5.5117	3.8471	3.2249	0.2599	0.2594***	0.2571**	**0.2565***
4	500	0.9	15.2034	0.3395	0.6786	5.3519	5.8787	1.1395	8.1412	5.7093	4.5755	0.2242	0.224***	0.2228**	**0.2218***
4	500	0.99	148.5091	0.4908	0.2732	50.9870	50.9130	0.9365	78.1364	54.9269	15.5351	0.2013**	0.2013**	0.2013**	**0.2010***
4	500	0.999	1818.8245	0.7584	0.2015	632.8561	632.6658	0.4061	962.8215	684.2725	15.6989	**0.1942***	0.1943***	0.1942*	0.1945***

Table 2. The EMSE values of the estimators when $p = 4$.

p	n	ρ	MLE	PRE	PLE	PLTE I	PLTE II	PLTE III	PHY I	PHY II	PSK	PRTE I	PRTE II	PRTE III	PRTE IV
8	50	0.85	16.9251	0.5693	1.7384	4.9866	9.3544	0.9601	9.9705	5.1974	4.9074	0.2783**	0.3383	**0.2766***	0.3358***
8	50	0.9	34.2415	0.3185	1.3459	10.2258	14.8173	1.2638	20.3488	10.5736	9.6449	0.1757**	0.2107***	**0.1754***	0.2671
8	50	0.99	376.8382	0.3521	0.2981	112.4872	114.5209	2.4520	222.1473	115.3996	67.5860	0.1270***	**0.1186***	0.1268**	0.1344
8	50	0.999	3550.7387	0.5607	0.1043	1032.0721	1032.1650	0.9530	2101.0805	1057.0848	111.6942	0.0985	0.0882**	0.0982***	**0.0857***
8	100	0.85	28.0655	0.3338	1.3340	8.1175	11.7144	1.0920	16.4358	8.3348	7.8438	0.1865**	0.2236***	**0.1857***	0.2530
8	100	0.9	32.3341	0.3049	1.2851	9.6048	13.0163	1.2299	19.2689	9.7889	9.4039	0.1702**	0.2010***	**0.1697***	0.2345
8	100	0.99	364.0779	0.3548	0.2737	107.2006	108.1535	2.2750	217.2630	109.0771	64.2123	0.1104**	**0.1102***	0.1104**	0.1236***
8	100	0.999	3532.0028	0.6299	0.1251	1049.2453	1049.0598	0.9829	2116.0276	1068.3102	97.2488	0.1080***	**0.1066***	0.1080***	0.1079**
8	200	0.85	27.8812	0.3376	1.3170	8.2821	11.5251	1.0887	16.5757	8.4104	7.9684	0.2047**	0.2381***	**0.2038***	0.2594
8	200	0.9	38.4506	0.2743	1.0961	11.4140	13.2997	1.2850	23.0604	11.5321	11.0466	0.1537**	0.1698***	**0.1533***	0.1851
8	200	0.99	338.5512	0.3522	0.2783	102.2077	102.7632	2.0431	203.3376	103.3862	64.7562	**0.1100***	0.1101**	**0.1100***	0.1224***
8	200	0.999	3801.1621	0.6525	0.1217	1104.5663	1104.3255	0.7988	2250.1080	1117.1855	90.6886	0.1074**	**0.1067***	0.1074**	0.1074**
8	500	0.85	23.9278	0.3604	1.3955	7.2290	8.7036	0.9510	14.4034	7.2898	7.0294	0.1731**	0.1790***	**0.1724***	0.1796
8	500	0.9	36.8726	0.2774	1.1082	11.0530	12.5696	1.1896	22.0890	11.1405	10.6167	0.1557**	0.1638***	**0.1552***	0.1693
8	500	0.99	333.0063	0.3596	0.2557	97.2928	97.3577	1.9152	197.5752	98.3202	59.9565	**0.1048***	0.1053**	**0.1048***	0.1091***
8	500	0.999	3371.3805	0.6615	0.1232	993.3309	993.0760	0.8654	2028.3847	1001.5095	86.1849	0.1064**	**0.1062***	0.1064**	0.1066***

Table 3. The EMSE values of the estimators when $p = 8$.

p	n	ρ	MLE	PRE	PLE	PLTE I	PLTE II	PLTE III	PHY I	PHY II	PSK	PRTE I	PRTE II	PRTE III	PRTE IV
12	50	0.85	53.4955	0.3015	2.0251	14.5174	20.2989	1.0766	32.7296	15.2581	12.7215	0.1246**	0.1642***	**0.1245***	0.2562
12	50	0.9	70.2583	0.2442	1.7893	19.5074	24.4279	1.3157	43.2572	20.2771	16.7916	**0.1204***	0.1368***	**0.1204***	0.2368***
12	50	0.99	548.3644	0.2280	0.4282	152.4782	155.4904	3.3321	336.1475	156.3263	101.6552	0.0906***	**0.0699***	0.0905**	0.0998
12	50	0.999	4667.8801	0.4415	0.1033	1265.9671	1266.5324	2.7146	2861.8004	1295.1068	264.6328	0.0879	0.0658**	0.0875***	**0.0628***
12	100	0.85	36.9427	0.4167	2.2049	10.1231	15.1137	0.9488	23.0251	10.3937	10.0361	0.1688**	0.2419***	**0.1685***	0.2954
12	100	0.9	48.6854	0.2949	1.9479	13.3593	17.9167	1.0852	30.7732	13.6499	13.3409	0.1201**	0.1766***	**0.1200***	0.2466
12	100	0.99	420.8219	0.2308	0.4158	113.6048	115.2467	2.9800	261.4296	115.2886	92.9058	0.0815**	**0.0765***	0.0815**	0.1047***
12	100	0.999	5694.2354	0.5262	0.0988	1561.364	1561.328	2.0811	3598.6116	1583.9902	249.9250	0.0753	**0.0709***	0.0752***	0.0720**
12	200	0.85	35.8940	0.3973	2.1548	9.9192	14.1292	0.9314	22.5629	10.0764	9.8577	0.1831**	0.2617***	**0.1826***	0.2999
12	200	0.9	55.6204	0.2545	1.7140	15.4278	18.5915	1.1692	35.2408	15.5948	15.3354	0.1152**	0.1540***	**0.1150***	0.2043
12	200	0.99	498.4333	0.2532	0.3347	136.5324	137.0078	2.8547	312.8818	137.7243	113.3388	0.0745**	**0.0737***	0.0745**	0.0897***
12	200	0.999	4934.7958	0.5365	0.1022	1376.3954	1376.3458	1.9218	3132.8827	1388.3893	241.3478	0.0723**	**0.0711***	0.0723**	0.0740***
12	500	0.85	36.7089	0.3639	2.0657	10.1674	12.1931	0.8892	23.2058	10.2179	10.1927	0.1459**	0.1713***	**0.1455***	0.1818
12	500	0.9	60.5669	0.2356	1.5392	16.9069	18.6269	1.1705	38.3551	16.9884	16.6802	0.1214**	0.1491***	**0.1212***	0.1782
12	500	0.99	598.8263	0.2989	0.2946	169.3809	169.6310	3.0236	380.3873	170.2031	126.8035	**0.0774***	0.0780**	**0.0774***	0.0934***
12	500	0.999	4728.9493	0.5541	0.1029	1325.9854	1325.7740	1.9088	2999.9883	1332.0518	227.6858	0.0733**	**0.0732***	0.0733**	0.0750

Table 4. The EMSE values of the estimators when $p = 12$.

3. When the number of variables p, and ρ are kept constant, the number of observations in the model did not have a significant effect on the PRTE I, PRTE II, PRTE III, and PRTE IV.
4. Regardless of n and p values, it is observed that PRTE I, PRTE II, PRTE III, and PRTE IV tended to give low EMSE values at high correlation.
5. When the number of observations (n) and correlation (ρ) in the model are kept constant, the EMSE values for PRTE I, PRTE II, PRTE III, and PRTE IV decrease with the increase in the number of explanatory variables (p).

In summary, all our proposed estimators outperformed the other considered estimators in all scenarios. However, it can be seen that there are cases where our estimators outperform each other due to different k and $f(k)$ function choices in different scenarios. As a result, we can observe that the number of observations has a relatively small effect on EMSE values compared to ρ and p. In other words, PRTEs have a robust structure according to the number of observations, therefore it gives very good results in case of high collinearity.

In the second simulation scheme, we examined the effects of the biasing parameter k on ILTEs and PRTE performances when the sample size (n), degree of the collinearity (ρ), and number of explanatory variables (p) are constant. The purpose of this simulation is to examine the performances of ILTE and PRTE at various values of the biasing parameter k according to the EMSE values given in (29). The biasing parameter k was not estimated in the second simulation scheme. Only the EMSE values obtained by increasing the k values in the range [0, 2]

by 0.1 were compared. There are many $f(k)$ and $g(k)$ functions we considered to evaluate the performances of these estimators. In order to compare the performances of these estimators under some n, p and ρ as an example, the ILTEs and PRTE determined by the following $f(k)$ and $g(k)$ functions are considered:

- $\hat{\beta}_{ILTE} = \left(X'\hat{W}X + kI\right)^{-1}\left(X'\hat{W}X + f(k)I\right)\hat{\beta}_{MLE}$ where $f(k) = \frac{\lambda_{\min}\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2}k + \left(\frac{\lambda_{\min}\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}$
- $\hat{\beta}_{ILTE(PRE)} = \left(X'\hat{W}X + kI\right)^{-1}\left(X'\hat{W}X + f(k)I\right)\hat{\beta}_{PRE}$ where $f(k) = \frac{\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2}(k + \lambda_{\min})^2 - (k + \lambda_{\min})$
- $\hat{\beta}_{PRTE} = \left(X'\hat{W}X + I\right)^{-1}\left(X'\hat{W}X + g(k)I\right)\hat{\beta}_{PRE}$ where $g(k) = \frac{(\lambda_{\min}+1)\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2}k + \left(\frac{(\lambda_{\min}+1)\alpha_{\min}^2}{1+\lambda_{\max}\alpha_{\max}^2} - 1\right)\lambda_{\min}$

Note that, when we use $\hat{\beta}_{PRE}$ instead of $\hat{\beta}^*$ in $\hat{\beta}_{ILTE}$, the obtained estimator is shown $\hat{\beta}_{ILTE(PRE)}$. Also, the $f(k)$ functions used in the $\hat{\beta}_{ILTE}$ and $\hat{\beta}_{ILTE(PRE)}$ were determined in accordance with the rules given by Akay and Ertan[5]. Note that when the method given by Akay and Ertan[5] is applied to $\hat{\beta}_{ILTE(PRE)}$, $f(k)$ that minimizes the $SMSE\left(\hat{\beta}_{ILTE(PRE)}\right)$ function is a quadratic function.

We considered the cases $\rho = 0.9, 0.99, 0.999$, $n = 50, 100, 500$, and $p = 4, 8, 12$. Depending on these n, ρ and p values, the explanatory variables are generated according to (28). The simulation is repeated 2000 times for each k value. The results are given graphically in Figs. 1, 2 and 3.

According to Figs. 1, 2 and 3, we can obtain the following results depending on each set of the values (n, ρ, p);

1. At small values of the biasing parameter k, PRTE outperforms other ILTE and ILTE(PRE). Although both the PRTE and ILTE(PRE) include the PRE, the performance of the ILTE(PRE) is quite poor compared to the PRTE at small values of the biasing parameter.
2. When the collinearity between the explanatory variables is relatively low, i.e. $\rho = 0.9$, ILTE(PRE) exhibits quite different behavior from ILTE and PRTE. If the value of correlation of explanatory variables and the number of explanatory variables increases, ILTE, ILTE(PRE) and PRTE show almost the same behavior. However, PRTE exhibits a more consistent behavior at varying values of the biasing parameter k.

As a result of the second simulation design, we recommend the PRTE to the researchers. In general, the performance of these estimators depends on $f(k)$ and $g(k)$ functions, respectively. In practice, we need to replace these functions with suitable functional relationships that can occur between the biasing parameters.

Numerical example: the aircraft damage data

In this section, the aircraft damage data is reanalyzed to demonstrate the benefits of PRTE. This data consists of 30 observations with three explanatory variables. The first variable (x_1) is a dichotomous variable showing the type of the aircraft. The explanatory variables (x_2) and (x_3) are bomb load in tons and total months of aircrew experience, respectively. The count variable y is the number of locations where damage was inflicted on the aircraft[3]. This dataset is also used by Myers et al.[3], Asar and Genç[15], Amin et al.[7], Lukman et al.[36], and Akay and Ertan[5].

Asar and Genc[15], Amin et al.[7] and Akay and Ertan[5] considered the following model $\mu = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)$. Except for the intercept term, the eigenvalues of $X'X$ are 208,522.5106, 374.8961 and 4.3333. Thus, the condition number is 48,120.9495, indicating a high multicollinearity problem among the explanatory variables. Firstly, the variables are standardized and then the intercept term is added to the vector of variables. Also, the eigenvalues of the matrix $X'WX$ are obtained as $\lambda_1 = 47.5850$ $\lambda_2 = 2.2844$, $\lambda_3 = 1.4097$ and $\lambda_4 = 0.3681$. The condition number is 129.2719 which is considerably larger than 30, indicating that MLE is still affected due to multicollinearity. The numerical results are given in Tables 5 to compare the PRTEs with other existing estimators.

In addition, the bootstrap sampling method is used to calculate the SMSE values of the given biased estimators. For this reason, 10,000 bootstrap samples have been created. For each of these samples, the parameter estimates of the given biased estimators are calculated. The mean of the MLE estimates is considered the real parameters. Then the calculated SMSE values are given in Table 5. From Table 5, it can be seen that the estimator with the best SMSE value is PRTE I and PRTE III.

Now, we want to examine the performances of ILTE, ILTE(PRE), and PRTE, which were examined in the previous section. Figure 4 graphically shows the estimated variance values of these estimators based on the value of the biasing parameter k. Also, Fig. 5 shows the SMSE performance of $\hat{\beta}_{ILTE}$, $\hat{\beta}_{ILTE(PRE)}$ and $\hat{\beta}_{PRTE}$ estimators according to the biasing parameter k.

Figures 4 and 5 indicating that the proposed PRTE is a strong alternative to other estimators at small values of the biasing parameter k. This result is also compatible with the second simulation results given in the previous section.

To compare the estimators under the MMSE sense, the parameter estimation obtained with the bootstrap sampling method is used in place of the unknown parameter α. R Programming is used with tolerance 10^{-12} to show the MMSE differences as a positive definite (pd) matrix. That is, if any of the eigenvalues is less than or equal to tolerance, then the matrix is not pd. Otherwise, the considered matrix is pd.

Finally, our aim in this part is to compare the estimators obtained from the choice of various $f(k)$ and $g(k)$ functions as a result of the theorem given in "The superiority of the PRTE in PRMs". To illustrate Theorem 3.1, the function $f(k)$ and $g(k)$ are taken as $f(k) = 0.05k + 0.05$ and $g(k) = 0.5k - 0.05$, respectively. In this case, $\text{cov}\left(\hat{\beta}_{ILTE}\right) - \text{cov}\left(\hat{\beta}_{PRTE}\right)$ is pd matrix for for $0 < k \leq 2.0057$. Also, k values which provide (22) criterion are $0 < k < 2.0054$. Consequently, $MMSE\left(\hat{\beta}_{ILTE}\right) - MMSE\left(\hat{\beta}_{PRTE}\right)$ is the pd matrix where $0 < k < 2.0054$.
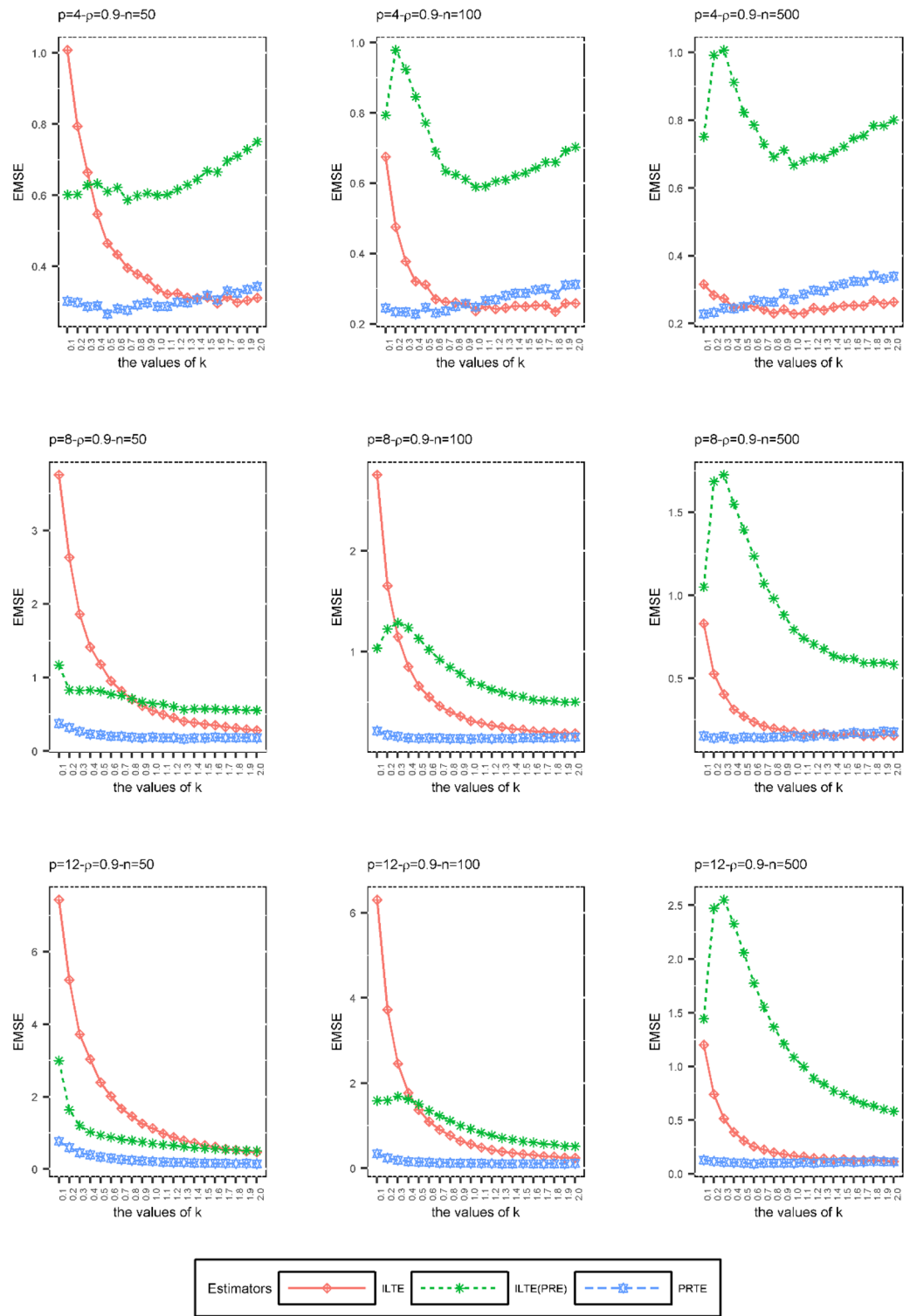
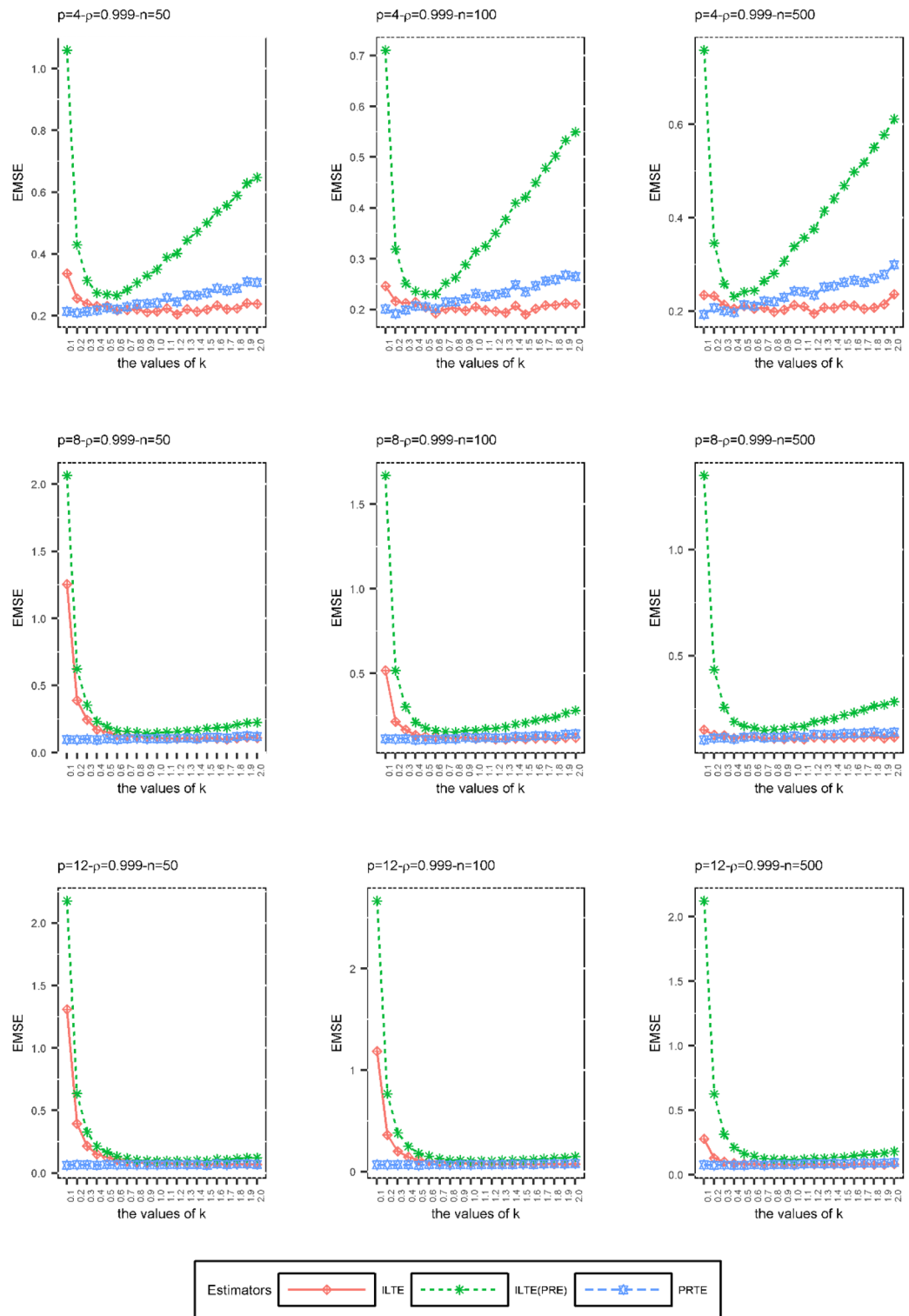**Figure 1.** The EMSE values of ILTE, ILTE(PRE), PRTE as a function of k values where $\rho = 0.9$.

Some concluding remarks

In this article, we defined a new general class of estimator named the PRTE as an alternative to MLE and the other existing biased estimators in the presence of multicollinearity for the PRMs. The PRTE is a general estimator which includes other biased estimators, such as the PRE, PLE, PHY and PSK estimators as special cases. In
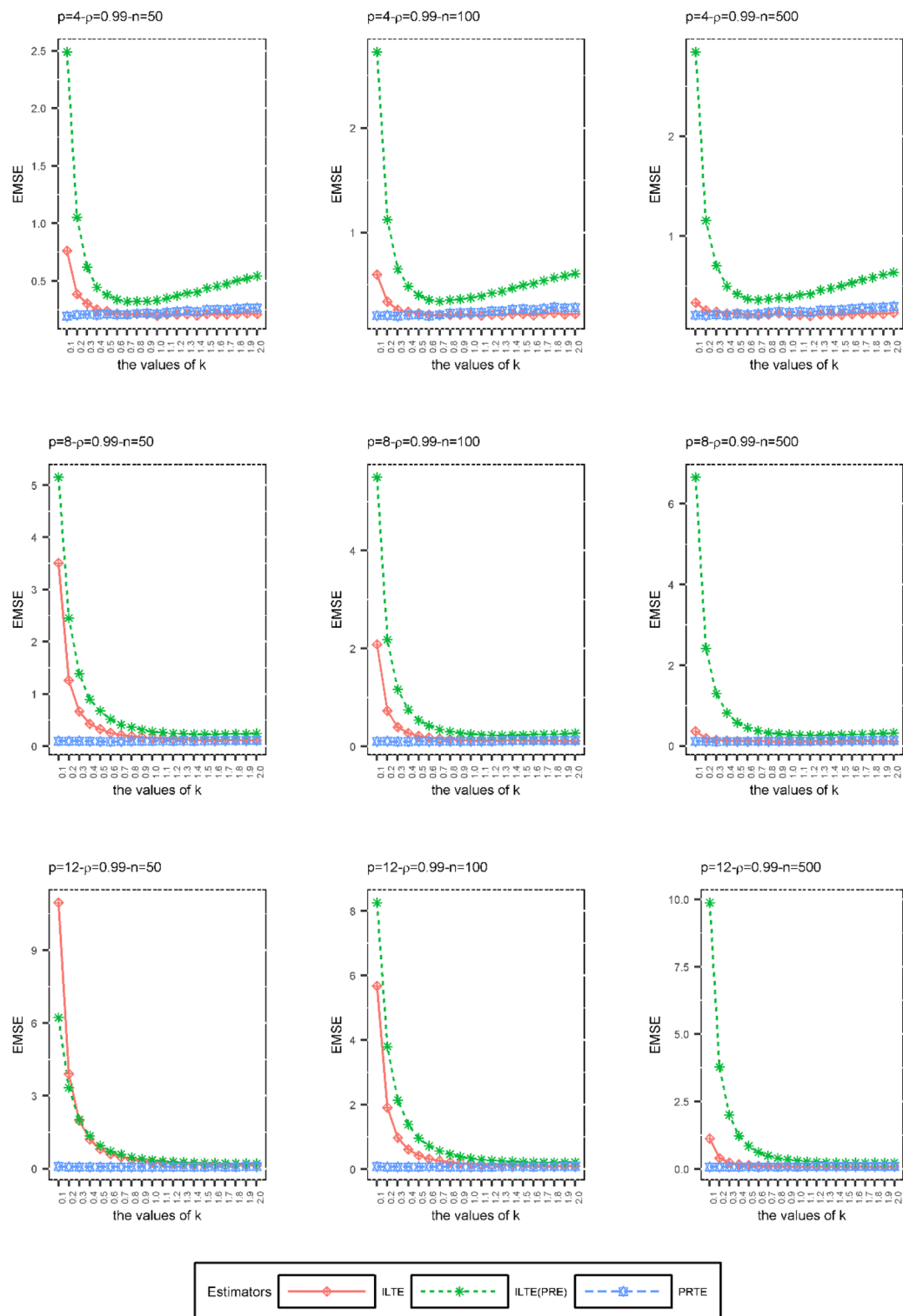
Figure 2. The EMSE values of ILTE, ILTE(PRE), PRTE as a function of k values where $\rho = 0.99$.

this study, we propose several rules for the determination of function $g(k)$. By using Monte Carlo simulations, the performance of the proposed PRTE with the existing estimators is evaluated in the smaller EMSE sense. The results show that the proposed PRTE outperforms the existing estimators in case of high multicollinearity. In addition, the comparison of ILTEs and PRTE is given with a general simulation study. In this simulation study, these two general estimators are compared according to the values of the biasing parameter k. It is observed that the PRTE is superior at small values of the biasing parameter k. Although the PRTE and ILTE(PRE) are both depending on the PRE, the main advantage of PRTE over ILTE(PRE) is that it can minimize the SMSE function

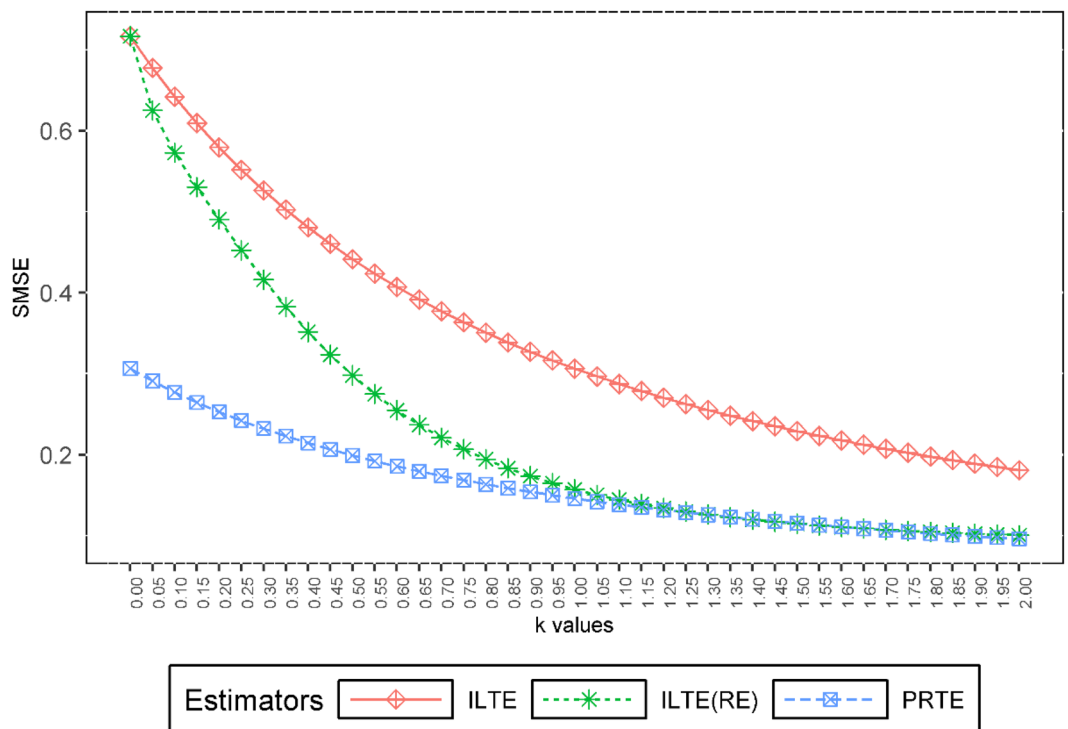**Figure 3.** The EMSE values of ILTE, ILTE(PRE), PRTE as a function of *k* values where $\rho = 0.999$.

with the help of a liner function of the biasing parameter *k*. Also, the estimators are applied to real dataset and it is observed that the results are consistent with simulation study. Depending on the experimental conditions examined, the proposed biased estimator outperforms the other existing biased estimators. Therefore, based on the results of the simulations and example, the PRTEs are recommended to the practitioners when there is multicollinearity problem in the PRMs.

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$var\left[\hat{\beta}\right]$	$SMSE\left[\hat{\beta}\right]$
$\hat{\beta}_{MLE}$	0.1262	1.5576	2.6710	− 1.4157	3.8847	
$\hat{\beta}_{PRE}\left(\hat{k}_{PRE} = 2.7044\right)$	0.4200	0.7970	1.2620	− 0.5504	0.2329	0.3707
$\hat{\beta}_{PLE}\left(\hat{d}_{PLE} = 0\right)$	0.2994	1.1554	1.8704	− 0.8852	0.6714	0.6739
$\hat{\beta}_{PLTEI}\left(\hat{k} = 2.7044, \hat{d} = -1.2946\right)$	0.2793	1.1611	1.9365	− 0.9646	1.3472	1.3494
$\hat{\beta}_{PLTEII}\left(\hat{k} = 0.1088, \hat{d} = 0.3268\right)$	0.2329	1.3552	2.1716	− 1.0640	0.6947	0.6956
$\hat{\beta}_{PLTEIII}\left(\hat{k} = 36.3139, \hat{d} = 0.0484\right)$	0.3820	0.1383	0.2081	− 0.0787	0.0093	0.0556
$\hat{\beta}_{PHYI}\left(\hat{K}_1 = 0.0637, \hat{D}_1 = 0.9485\right)$	0.1507	1.5078	2.5561	− 1.3366	2.8723	2.8723
$\hat{\beta}_{PHYII}\left(\hat{K}_2 = 0.0506, \hat{D}_2 = 0.2048\right)$	0.2734	1.2170	1.9907	− 0.9645	0.9259	0.9277
$\hat{\beta}_{PSK}\left(\hat{k}_{SK} = 2.7044, \hat{d}_{SK} = 1.2396\right)$	0.2117	1.4161	2.3129	− 1.1287	1.1497	1.1502
$\hat{\beta}_{PRTEI}\left(g(k) = 0.1766 \times 10^{-4}k - 0.3681\right)$ $\hat{k}_{PRTEI} = 6.2833$	0.5387	0.3238	0.4858	− 0.1687	0.0305	0.0513
$\hat{\beta}_{PRTEII}\left(g(k) = 0.1568 \times 10^{-3}k - 0.3680\right)$ $\hat{k}_{PRTEII} = 1.8500$	0.4952	0.5991	0.9146	− 0.3420	0.0887	0.1011
$\hat{\beta}_{PRTEIII}\left(g(k) = 0.1003 \times 10^{-4}k - 0.3681\right)$ $\hat{k}_{PRTEIII} = 6.2956$	0.5387	0.3234	0.4852	− 0.1684	0.0305	0.0512
$\hat{\beta}_{PRTEIV}\left(g(k) = 0.2769 \times 10^{-3}k - 0.3680\right)$ $\hat{k}_{PRTEIV} = 1.2282$	0.4695	0.6907	1.0597	− 0.4083	0.1198	0.1301

Table 5. The estimated parameter values and the SMSE values of the estimators.
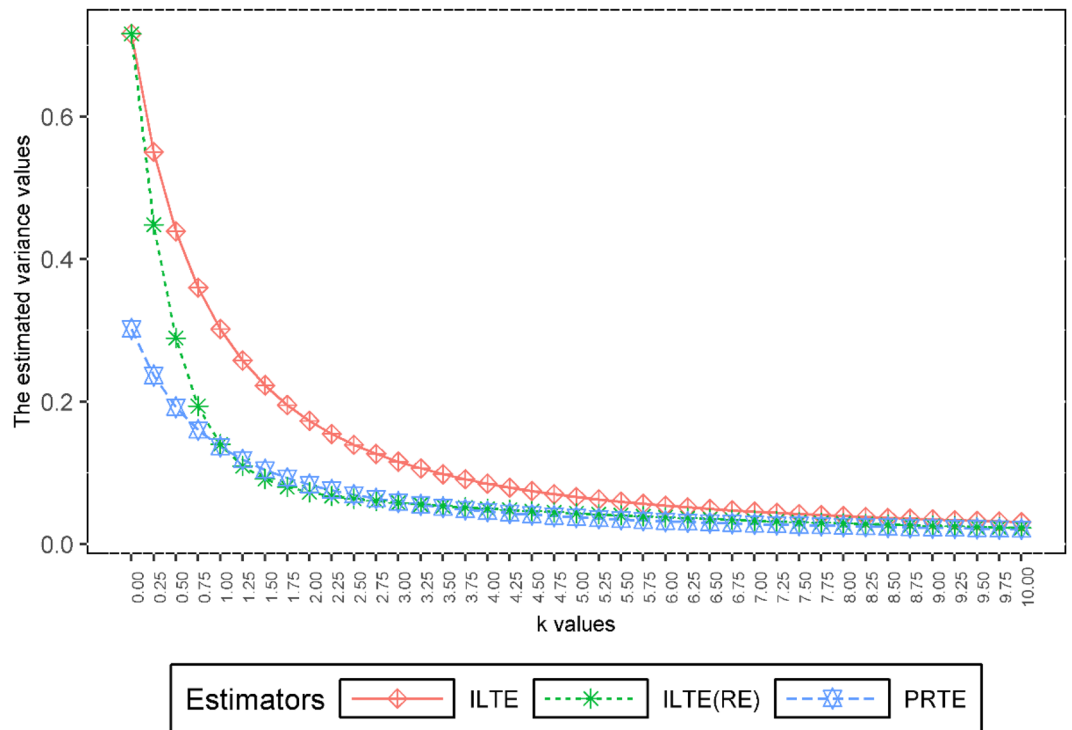
Figure 4. The estimated variance values of ILTE, ILTE(PRE) and PRTE as a function of k.

Figure 5. The SMSE values of ILTE, ILTE(PRE) and PRTE as a function of *k*.

Data availability

All data generated or analyzed during this study are used by the given reference in this article. The data analyzed/generated are available upon the reasonable request by the E. E.

References

1. Winkelmann, R. *Econometric Analysis of Count Data*. (Springer Science & Business Media, 2008).
2. Hilbe, J. M. *Modeling Count Data* (Cambridge University Press, 2014).
3. Myers, R. H., Montgomery, D. C., Vining, G. G. & Robinson, T. J. *Generalized Linear Models: With Applications in Engineering and the Sciences* (Wiley, 2012).
4. Dunn, P. K. & Smyth, G. K. *Generalized Linear Models with Examples in R* (Springer, 2018).
5. Akay, K. U. & Ertan, E. A new Liu-type estimator in Poisson regression models. *Hacet. J. Math. Stat.* **51**(5), 1484–1503 (2022).
6. Alkhateeb, A. & Algamal, Z. Jackknifed Liu-type estimator in Poisson regression model. *J. Iran. Stat. Soc. (JIRSS)* **19**(1), 21–37 (2020).
7. Amin, M., Akram, M. N. & Amanullah, M. On the James–Stein estimator for the poisson regression model. *Commun. Stat. Simul. Comput.* **51**(10), 5596–5608 (2022).
8. Jadhav, N. H. A new linearized ridge Poisson estimator in the presence of multicollinearity. *J. Appl. Stat.* **49**(8), 2016–2034 (2022).
9. Kibria, B. M. G., Månsson, K. & Shukur, G. Some ridge regression estimators for the zero-inflated Poisson model. *J. Appl. Stat.* **40**(4), 721–735 (2013).
10. Lukman, A. F., Aladeitan, B., Ayinde, K. & Abonazel, M. R. Modified ridge-type for the Poisson regression model: Simulation and application. *J. Appl. Stat.* **49**(8), 2124–2136 (2022).
11. Månsson, K. & Kibria, B. M. G. Estimating the unrestricted and restricted liu estimators for the poisson regression model: Method and application. *Comput. Econ.* **58**, 311–326 (2021).
12. Rashad, N. K. & Algamal, Z. Y. A new ridge estimator for the Poisson regression model. *Iran. J. Sci. Technol. Trans. A Sci.* **43**(6), 2921–2928 (2019).
13. Türkan, S. & Özel, G. A new modified Jackknifed estimator for the Poisson regression model. *J. Appl. Stat.* **43**(10), 1892–1905 (2016).
14. Alheety, M. I., Qasim, M., Månsson, K. & Kibria, B. M. G. Modified almost unbiased two-parameter estimator for the Poisson regression model with an application to accident data. *SORT* **45**(2), 121–142 (2021).
15. Asar, Y. & Genç, A. A new two-parameter estimator for the poisson regression model. *Iran. J. Sci. Technol. Trans. A Sci.* **42**(2), 793–803 (2018).
16. Çetinkaya, M. K. & Kaçıranlar, S. Improved two-parameter estimators for the negative binomial and Poisson regression models. *J. Stat. Comput. Simul.* **89**(14), 2645–2660 (2019).
17. Kibria, B. M. G., Månsson, K. & Shukur, G. A simulation study of some biasing parameters for the ridge type estimation of Poisson regression. *Commun. Stat. Simul. Comput.* **44**(4), 943–957 (2015).
18. Månsson, K. & Shukur, G. A Poisson ridge regression estimator. *Econ. Model.* **28**(4), 1475–1481 (2011).
19. Månsson, K., Kibria, B. G., Sjolander, P. & Shukur, G. Improved Liu estimators for the Poisson regression model. *Int. J. Probab. Stat.* **1**(1), 2–6 (2012).

20. Qasim, M., Månsson, K., Amin, M., Kibria, B. M. G. & Sjölander, P. Biased adjusted Poisson ridge estimators-method and application. *Iran. J. Sci. Technol. Trans. A Sci.* **44**(6), 1775–1789 (2020).
21. Hoerl, A. E. & Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**(1), 55–67 (1970).
22. Amin, M., Akram, M. N. & Kibria, B. M. G. A new adjusted Liu estimator for the Poisson regression model. *Concurr. Comput.* **33**(20), e6340 (2021).
23. Qasim, M., Kibria, B. M. G., Månsson, K. & Sjölander, P. A new Poisson Liu regression estimator: Method and application. *J. Appl. Stat.* **47**(12), 2258–2271 (2020).
24. Liu, K. A new class of biased estimate in linear regression. *Commun. Stat. Theory Methods* **22**(2), 393–402 (1993).
25. Algamal, Z. Y. Biased estimators in Poisson regression model in the presence of multicollinearity: A subject review. *Al-Qadisiyah J. Adm. Econ. Sci.* **20**(1), 37–43 (2018).
26. Liu, K. Using Liu-type estimator to combat collinearity. *Commun. Stat. Theory Methods* **32**(5), 1009–1020 (2003).
27. Özkale, M. R. & Kaçıranlar, S. The restricted and unrestricted two-parameter estimators. *Commun. Stat. Theory Methods* **36**, 2707–2725 (2007).
28. Kurnaz, F. S. & Akay, K. U. A new Liu-type estimator. *Stat. Papers* **56**, 495–517 (2015).
29. Yang, H. & Chang, X. A new two-parameter estimator in linear regression. *Commun. Stat. Theory Methods* **39**(6), 923–934 (2010).
30. Huang, J. & Yang, H. A two-parameter estimator in the negative binomial regression model. *J. Stat. Comput. Simul.* **84**(1), 124–134 (2014).
31. Sakallıoğlu, S. & Kaçıranlar, S. A new biased estimator based on ridge estimation. *Stat. Papers* **49**, 669–689 (2008).
32. Theobald, C. M. Generalizations of mean square error applied to ridge regression. *J. R. Stat. Soc. B* **36**, 103–106 (1974).
33. Farebrother, R. W. Further results on the mean square error of ridge regression. *J. R. Stat. Soc. B* **28**, 248–250 (1976).
34. Alanaz, M. M. & Algamal, Z. Y. Proposed methods in estimating the ridge regression parameter in Poisson regression model. *Electron. J. Appl. Stat. Anal.* **11**(2), 506–515 (2018).
35. Kibria, B. M. G. & Lukman, A. F. A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica* https://doi.org/10.1155/2020/9758378 (2020).
36. Lukman, A. F., Adewuyi, E., Månsson, K. & Kibria, B. M. G. A new estimator for the multicollinear Poisson regression model: Simulation and application. *Sci. Rep.* **11**(1), 1–11 (2021).

Acknowledgements

Author contributions

All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.