Thailand Statistician
October 2023; 21(4): 767-782.
http://statassoc.or.th
Contributed paper

A New Ridge Parameter Estimator In Poisson Regression With Correlated Predictors: Optimal Design Approach

Salah Ghorbani and Mehrdad Niaparast*
Department of Statistics, Razi University, Kermanshah, Kermanshah,Iran
*Corresponding author; e-mail: niaparast@razi.ac.ir

Received: 12 July 2021
Revised: 15 January 2022
Accepted: 24 January 2022

Abstract

Poisson ridge regression is used as a tool to analyze counting data with linearly dependent predictor variables. Several methods for estimating the ridge parameter have been introduced in this model. In this paper, in addition to obtaining the optimal designs for the Poisson regression model with collinearity in predictor variables, we present a new method based on the theory of optimal designs for estimating the ridge parameter. These estimates are obtained based on two criteria, DM- and AM-optimality. Finally, using simulation, based on the efficiency criteria that we introduce, the performance of new estimates of the ridge parameter is obtained.

Keywords: Optimal design, Poisson regression, Ridge regression, Ridge parameter.

1. Introduction

Count data have been gaining attention in the past researches. The most popular model for studying such data is the Poisson regression model as a member of the generalized linear models family. Despite wide theoretical work on analyzing this data, Optimal design for this type of data has not much history in experimental designs.

Researchers such as Wang et al. (2006) and Russel et al. (2009) studied the optimal design for the Poisson regression model. Most of this research is based on the locally optimal design proposed by Chernov (1953). In most of these studies, the independence of predictor variables is assumed, and hence the model has a full rank design matrix. This condition may be violated in some cases. Dependence on predictor variables leads to the ill-conditioned variance-covariance matrix, which is used for parameter estimations. In this paper, the case of the Poisson regression model is intended, which predictor variables have collinearity. Poisson ridge regression to avoid this problem is presented in the Månsson & Shukur (2011). The first idea of ridge regression was presented by Hoerl & Kannard (1975). Then, other researchers, including Muniz & Kibria (2009) and Lawless & Wang (1976), have developed it.

Optimality criteria for finding optimal designs in Silvey(1980) have been introduced. These criteria are based on the information matrix as the inverse of the variance-covariance matrix of the parameter estimates. Due to the Asymptotic properties of the estimator, and as long as the sample size is large, the relation between the Information matrix and variance-covariance matrix is established. Thus, in small samples, in addition to hesitancy to having this property, considered to be the bias of parameters estimates. In this case, using the optimality criteria introduced in Silvey (1980) does not suffice. Since the estimator of the parameters for the Poisson ridge regression is biased in small

samples, many authors have been used the mean square error (MSE) as a criterion for comparison of estimation of parameters. In this paper, We apply two optimal criteria based on MSE, which are known as DM- and AM-criteria [Mehr Mansour and Niaparast (2018, 2019)], to obtain the Poisson ridge regression parameter estimation. A comprehensive study to compare different estimations of ridge parameter has been done by Ali et al.(2019) and Ismail Shah et al. (2021). In this paper, using simulation, we compare our proposed ridge parameter estimation to the previous estimates. The results confirm that in most cases, the new ridge parameter estimation is better.
This paper unfolds as follows. Section 2 gives the definition of our model. In Section 3, we present the description of the design and optimal criteria which we use. In Section 4, in addition to introducing efficiency measures, we review some well-known existing ridge parameter estimators and our proposed ridge parameter estimators. Finally, we present the main results in this section. Concluding remarks are given in Section 5.

2. Model Specification

Poisson regression is suitable for describing response variable changes in practical problems with count data. This model is defined as follows:

$$y_i \sim P(\mu_i(\beta)); \quad i = 1, 2, \ldots, n, \tag{1}$$

where $\mu_i(\boldsymbol{\beta}) = \exp\{f^T(\mathbf{x}_i)\boldsymbol{\beta}\}$. In other words, the link function is equal to $log(\mu_i) = f^T(\mathbf{x}_i)\boldsymbol{\beta}$ where $\mathbf{x}_i$ is the ith vector of predictor and $\boldsymbol{\beta}_{p\times 1}$ is a vector containing the fixed and unknown parameters. The commonly applied estimation method for the Poisson regression model is the maximum likelihood (ML). The parameters estimate by solving the log-likelihood function and its first derivative, obtained as follows:

$$\begin{aligned} l = Ln(\boldsymbol{\mu}(\boldsymbol{\beta}); y) &= \sum_{i=1}^{n} y_i \log(\mu_i(\boldsymbol{\beta})) - \sum_{i=1}^{n} \mu_i(\boldsymbol{\beta}) - \log(\prod_{i=1}^{n} y_i!) \\ &= \sum_{i=1}^{n} y_i \log(\exp\{f^T(\mathbf{x}_i)\boldsymbol{\beta}\}) - \sum_{i=1}^{n} \exp\{f^T(\mathbf{x}_i)\boldsymbol{\beta}\} - \log(\prod_{i=1}^{n} y_i!). \\ \Rightarrow \frac{\partial l}{\partial \boldsymbol{\beta}} &= \mathbf{F}^T(\mathbf{Y} - \boldsymbol{\mu}(\boldsymbol{\beta})), \end{aligned} \tag{2}$$

where $\boldsymbol{\mu}(\boldsymbol{\beta}) = (\mu_1(\boldsymbol{\beta}), \mu_2(\boldsymbol{\beta}), \ldots, \mu_n(\boldsymbol{\beta}))^T$, $\mathbf{F}_{n\times p} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1(p-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{n(p-1)} \end{pmatrix}$ and $\mathbf{Y}_{n\times 1} = (y_1, \ldots, y_n)^T$.

Since Eqn. (2) is nonlinear in $\boldsymbol{\beta}$, the solution of $\frac{\partial l}{\partial \beta}$ equalling zero is found using the following iterative weighted least square (IWLS) algorithm:

$$\hat{\boldsymbol{\beta}}_{ML} = (\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F})^{-1}\mathbf{F}^T\hat{\mathbf{W}}\hat{\mathbf{G}},$$

where $\hat{\mathbf{W}} = diag(\hat{\boldsymbol{\mu}}(\boldsymbol{\beta}))$ and $\hat{\mathbf{G}}$ is a vector where the ith element equals $\hat{G}_i = \log(\hat{\mu}_i(\boldsymbol{\beta})) + \frac{y - \hat{\mu}_i(\boldsymbol{\beta})}{\hat{\mu}_i(\boldsymbol{\beta})}$.

If there is collinearity in predictor variables, then $\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F}$ is an ill-conditioned matrix, and hence the estimator of parameters are not stable. As a method of estimation for the type of data with linearity correlation, ridge regression was introduced by Hoerl & Kennard (1970).

Månsson & Shukur (2011) applied ridge regression methods to overcome the collinearity problem in the Poisson regression model. Poisson ridge regression estimator is shown with $\hat{\boldsymbol{\beta}}_{PRR}$ and equals to

$$\hat{\boldsymbol{\beta}}_{PRR} = (\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F} + k\mathbf{I})^{-1}\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F}\hat{\boldsymbol{\beta}}_{ML} = \mathbf{Z}\hat{\boldsymbol{\beta}}_{ML},$$

where $\mathbf{Z} = (\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F} + k\mathbf{I})^{-1}\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F}$, $\mathbf{I}$ is identity matrix and k is the ridge parameter. Also, the MSE critertion for $\hat{\boldsymbol{\beta}}_{PRR}$ is obtained as follows:

$$\begin{aligned}
MSE_{PRR} &= E[(\hat{\boldsymbol{\beta}}_{PRR} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_{PRR} - \boldsymbol{\beta})^T] \\
&= E[(\hat{\boldsymbol{\beta}}_{PRR} - \mathbf{Z}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_{PRR} - \mathbf{Z}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})^T] \\
&= E[(\mathbf{Z}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{Z}\boldsymbol{\beta})(\mathbf{Z}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{Z}\boldsymbol{\beta})^T + (\mathbf{Z}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{Z}\boldsymbol{\beta})(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})^T \\
&+ (\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})(\mathbf{Z}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{Z}\boldsymbol{\beta})^T + (\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})^T] \\
&= E[\mathbf{Z}(\hat{\boldsymbol{\beta}}_{ML} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_{ML} - \boldsymbol{\beta})^T\mathbf{Z}^T] + E[\mathbf{Z}(\hat{\boldsymbol{\beta}}_{ML} - \boldsymbol{\beta})(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})^T] \\
&+ E[(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_{ML} - \boldsymbol{\beta})^T\mathbf{Z}^T] + E[(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})(\mathbf{Z}\boldsymbol{\beta} - \boldsymbol{\beta})^T] \\
&= \mathbf{Z}\text{Var}(\hat{\boldsymbol{\beta}}_{ML})\mathbf{Z}^T + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= \mathbf{Z}(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}\mathbf{Z}^T + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= (\mathbf{F}^T\mathbf{W}\mathbf{F} + k\mathbf{I})^{-1}(\mathbf{F}^T\mathbf{W}\mathbf{F})(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}(\mathbf{F}^T\mathbf{W}\mathbf{F})(\mathbf{F}^T\mathbf{W}\mathbf{F} + k\mathbf{I})^{-1} \\
&+ (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= (\mathbf{F}^T\mathbf{W}\mathbf{F} + k\mathbf{I})^{-1}[(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}]^{-1}(\mathbf{F}^T\mathbf{W}\mathbf{F} + k\mathbf{I})^{-1} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= [\mathbf{I} + k(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}]^{-1}(\mathbf{F}^T\mathbf{W}\mathbf{F} + k\mathbf{I})^{-1} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= [\mathbf{F}^T\mathbf{W}\mathbf{F} + 2k\mathbf{I} + k^2(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}]^{-1} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= [(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}(k^2 + 2k(\mathbf{F}^T\mathbf{W}\mathbf{F}) + (\mathbf{F}^T\mathbf{W}\mathbf{F})^2)]^{-1} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= [(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}(k\mathbf{I} + \mathbf{F}^T\mathbf{W}\mathbf{F})^2]^{-1} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= (k\mathbf{I} + \mathbf{F}^T\mathbf{W}\mathbf{F})^{-2}\mathbf{F}^T\mathbf{W}\mathbf{F} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \\
&= (k\mathbf{I} + \mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}\mathbf{Z} + (\mathbf{Z} - \mathbf{I})\boldsymbol{\beta}\boldsymbol{\beta}^T(\mathbf{Z} - \mathbf{I})^T \qquad (3)
\end{aligned}$$

When k equals to zero, MSE_{PRR} will be $(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}$and indicated by MSE_{PR} . Under the regular conditions, $\hat{\boldsymbol{\beta}}_{ML}$ is asymptotically unbiased estimator for $\boldsymbol{\beta}$, and hence $(\mathbf{F}^T\mathbf{W}\mathbf{F})^{-1}$ is the variance-covariance matrix of $\hat{\boldsymbol{\beta}}_{ML}$.

3. DM- and AM-Optimal Designs

A design, ξ, is a collection of points of the predictor variables, $\{x_1, x_2, \ldots, x_n\}$, where n is the size of the design. It can be written taking the m different points(called the support points) and for each one the proportion(weight) than it has in the design, that is

$$\xi = \left\{ \begin{array}{c} x_1, x_2, \ldots, x_m \\ \delta_1, \delta_2, \ldots, \delta_m \end{array} \right\}, \quad \sum_{i=1}^{m} \delta_i = 1\,, \;\; x_i \in \chi\,, \;\; i = 1, \ldots, m,$$

where $m_i = n\delta_i$ is the frequency of point x_i in the design. From this point of view, an approximate design can be defined as any probability measure in χ with finite support. D-optimality is the most commonly used design criterion which is defined as follows.

Definition 1 ξ^* is an D-optimal design if

$$\xi^* = \arg \max_{\xi \in \Xi} \log |\mathbf{M}(\xi)| = \arg \min_{\xi \in \Xi} - \log |\mathbf{M}(\xi)|,$$

where $\mathbf{M}(\xi)$ is the Fisher information matrix from the design ξ and Ξ is the set of all design measures.

Another most commonly design criterion which used for obtain the optimal design is A-optimal.

Definition 2 ξ^* is an A-optimal design if

$$\xi^* = \arg \min_{\xi \in \Xi} \text{tr}(\mathbf{M}^{-1}(\xi)).$$

Remark 3.1 For a Poisson regression model with design matrix $\mathbf{F}$ and model parameter vector $\boldsymbol{\beta}$, the Fisher information matrix is

$$\mathbf{M}(\xi) = \mathbf{F}_\xi^T \mathbf{W} \mathbf{F}_\xi$$

where $\mathbf{W} = ndiag(\delta_i \boldsymbol{\mu}_i(\boldsymbol{\beta}))_{i=1,\cdots,m}$ and $\mathbf{F}_\xi = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1(p-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{m1} & \cdots & x_{m(p-1)} \end{pmatrix}$.

Since that D-and A-optimal criteria are based on the property of unbiasedness of MLE, whereas this property can not be extended to ridge estimators, following Mehr Mansour and Niaparast (2018, 2019), we apply DM- and AM-optimal criteria. Let $MSE(\xi)$ be the MSE matrix from the design ξ.

Definition 3 ξ^* is an DM-optimal design if

$$\xi^* = \arg \min_{\xi \in \Xi} \log |MSE(\xi)|.$$

Also, the way of working to find the AM-optimal design minimizes the total MSE of the model parameters estimator. In other words, the AM-optimal criterion is defined as follows.

Definition 4 ξ^* is an AM-optimal design if

$$\xi^* = \arg \min_{\xi \in \Xi} \text{tr}(MSE(\xi)).$$

To adapt the above two definitions to the concepts of the Poisson ridge regression model, we rewrite these two definitions as follows.

1. $(\xi^*, k_D) = \arg \min_{\xi \in \Xi, k} \log |MSE(\xi)|$
2. $(\xi^*, k_A) = \arg \min_{\xi \in \Xi, k} \text{tr}(MSE(\xi))$

where k_D and k_A are optimal ridge Parameter(k) based on DM- and AM- optimal criteria, respectively.

4. Simulated Example and Main Results

Consider the Poisson regression model with a mean $\mu_i(\boldsymbol{\beta}) = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}}$ which is the special case of the model(1). Then the proposed saturated design is as follows:

$$\xi = \left\{ \begin{array}{ccc} (x_{11}, x_{12}) & (x_{21}, x_{22}) & (x_{31}, x_{32}) \\ \delta_1 & \delta_2 & \delta_3 \end{array} \right\}.$$

So that the design matrix is

$$\mathbf{F}_\xi = \begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \end{pmatrix}.$$

Thus we will have:

$$\mathbf{F}_\xi^T \mathbf{W} \mathbf{F}_\xi = n \begin{pmatrix} \sum_{i=1}^{3} \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i1} \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i2} \delta_i \mu_i(\boldsymbol{\beta}) \\ \sum_{i=1}^{3} x_{i1} \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i1}^2 \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i1} x_{i2} \delta_i \mu_i(\boldsymbol{\beta}) \\ \sum_{i=1}^{3} x_{i2} \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i2} x_{i1} \delta_i \mu_i(\boldsymbol{\beta}) & \sum_{i=1}^{3} x_{i2}^2 \delta_i \mu_i(\boldsymbol{\beta}) \end{pmatrix}.$$

Now, we are ready to generate a set of linear dependent predictor variables, to simulate support points of the above design. To achieve different degrees of collinearity by statistical software R, following McDonald & Galarneau (1975) and Gibbons (1981), the predictor variables were generated using the following system:

$$x_{ij} = \sqrt{1-\rho^2} z_{ij} + \rho z_{ip} \ ; \ i = 1, 2, \ldots, n \ , \ j = 1, 2, \ldots, p,$$

where z_{ij} are pseudo-random numbers generated using the standard normal distribution.

We apply DM- and Am-optimal criteria to find the locally optimal weights (δ_i^*, $i = 1, 2, 3$) and locally optimal ridge parameters for different known $\boldsymbol{\beta}$ and ρ. The results of the locally DM-optimal and AM-optimal are listed in Tables 1-8 for different representative values of parameters and correlations. In these tables, the optimal weights and ridge parameter (k) based on AM- and DM- criteria are obtained for both PR and PRR. For any values of $\boldsymbol{\beta}$, It is clear that the MSE in PR and PRR increased with increasing ρ. But the key point is that the values of $tr(MSE)$ and $logdet(MSE)$ in the PRR have always been lower than the corresponding values in the PR method. Also, when ρ is rising, the increasing rate of $tr(MSE)$ and $logdet(MSE)$ in PRR is always lower than the corresponding values in PR. For each $\boldsymbol{\beta}$ and ρ, unlike the optimal design for the PR, which has equal weights for all three support points, the optimal design for the PRR has different weights.

4.1. Efficiency measures and numerical results

To measure the quality of any arbitrary design such as ξ for the PR model comparing to the corresponding design fo the PRR model , we apply two measures DM-efficiency and AM-efficiency which are defined as follows.

Definition 5 The DM-efficiency of an arbitrary design ξ is defined as

$$DM_{eff} = \left(\frac{det(MSE_{PR}(\xi))}{det(MSE_{PRR}^{(k_D)}(\xi))} \right)^{\frac{1}{p}},$$

where p is the number of parameters and $MSE_{PRR}^{(k_D)}(\xi)$ is the MSE matrix for Poisson ridge regression model based on the design ξ and optimal ridge parameter k_D.

Definition 6 The AM-efficiency of an arbitrary design ξ is defined as

$$AM_{eff} = \left(\frac{\text{tr}\,(MSE_{PR}(\xi))}{\text{tr}\,(MSE_{PRR}^{(k_A)}(\xi))} \right)^{\frac{1}{p}},$$

where $MSE_{PRR}^{(k_A)}(\xi)$ is the MSE matrix for Poisson ridge regression model based on the design ξ and optimal ridge parameter k_A.

Note that if the values of DM_{eff} and AM_{eff} is greater than 1, then our method works better. The results of AM- and DM-efficiency have also been presented in Tables 1-8. The results show that under the condition of the presence of linearity in the predictor variables, in all cases considered in the simulation, the PRR model is better than the PR model. Of course, this result was already predictable. To more clarify, the results of Tables 1-4 and Tables 5-8 are drown in Figures 1(a) and 1(b), respectively.

Table 1 locally DM-optimal weights for PR and PRR models locally DM-optimal k_D and DM-efficiency measure for ($\beta_0 = 2, \beta_1 = 1, \beta_2 = -1$) and 8 different correlations

		DM Optimality							
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	δ_2^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	δ_3^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	$logdet(MSE)$	-18.4658	-18.4700	-18.4451	-18.3786	-18.2468	-17.9964	-17.4577	-15.9713
PRR	δ_1^*	0.3335	0.3335	0.3334	0.3334	0.3333	0.3333	0.3333	0.335
	δ_2^*	0.3326	0.3326	0.3327	0.3327	0.3328	0.3328	0.3328	0.3339
	δ_3^*	0.3339	0.3339	0.334	0.3338	0.3339	0.3339	0.3339	0.3312
	$logdet(MSE)$	-18.4708	-18.4750	-18.4502	-18.3840	-18.2528	-18.0040	-17.4705	-16.0283
	DM_{eff}	1.0017	1.0017	1.0017	1.0018	1.002	1.0025	1.0043	1.0192
	k_D	0.4316	0.4371	0.4404	0.447	0.4581	0.4703	0.485	0.4946

Table 2 locally DM-optimal weights for PR and PRR models, locally DM-optimal k_D and DM-efficiency measure for ($\beta_0 = 1, \beta_1 = 2, \beta_2 = -1$ and 8 different correlations

		DM Optimality							
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	δ_2^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	δ_3^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	$logdet(MSE)$	-17.1358	-17.2684	-17.3720	-17.4339	-17.4306	-17.3087	-16.8984	-15.5148
PRR	δ_1^*	0.3254	0.3242	0.3234	0.3231	0.3242	0.3267	0.3296	0.3298
	δ_2^*	0.3381	0.3388	0.3392	0.3391	0.338	0.3363	0.335	0.3365
	δ_3^*	0.3365	0.3369	0.3374	0.3379	0.3377	0.337	0.3354	0.3337
	$logdet(MSE)$	-17.1596	-17.2932	-17.3975	-17.4592	-17.4537	-17.3277	-16.9144	-15.5521
	DM_{eff}	1.008	1.0083	1.0085	1.0085	1.0077	1.0063	1.0053	1.0125
	k_D	0.8279	0.9164	0.9922	1.017	0.9329	0.7129	0.4312	0.2516

Table 3 locally DM-optimal weights for PR and PRR models, locally DM-optimal k_D and DM-efficiency measure for ($\beta_0 = 0, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

		DM Optimality							
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	δ_2^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	δ_3^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334	0.3334
	$logdet(MSE)$	-18.0685	-17.8157	-17.5339	-17.2105	-16.8218	-16.3145	-15.5189	-13.8270
PRR	δ_1^*	0.3235	0.3247	0.3258	0.3268	0.3277	0.3287	0.3293	0.3263
	δ_2^*	0.337	0.3368	0.3365	0.3363	0.3363	0.3364	0.3371	0.3442
	δ_3^*	0.3395	0.3385	0.3377	0.3369	0.336	0.335	0.3336	0.3295
	$logdet(MSE)$	-18.0995	-17.8446	-17.5610	-17.2363	-16.8471	-16.3410	-15.5526	-13.9296
	DM_{eff}	1.0104	1.0097	1.0091	1.0087	1.0085	1.0089	1.0113	1.0348
	k_D	0.9819	0.8633	0.7523	0.6478	0.5501	0.453	0.3543	0.2682

Table 4 locally DM-optimal weights for PR and PRR models, locally DM-optimal k_D and DM-efficiency measure for ($\beta_0 = -1, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

	DM Optimality								
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333	0.3333
	δ_2^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3333	0.3333	0.3333
	δ_3^*	0.3334	0.3334	0.3334	0.3334	0.3334	0.3333	0.3333	0.3333
	$logdet(MSE)$	-15.0685	-14.8157	-14.5339	-14.2105	-13.8218	-13.3145	-12.5189	-10.8270
PRR	δ_1^*	0.3246	0.3239	0.3231	0.322	0.3204	0.3181	0.3143	0.2993
	δ_2^*	0.3378	0.338	0.3383	0.3388	0.3396	0.3409	0.3445	0.365
	δ_3^*	0.3376	0.3381	0.3386	0.3392	0.34	0.3409	0.3412	0.3357
	$logdet(MSE)$	-15.1079	-14.8568	-14.5775	-14.2580	-13.8756	-13.3802	-12.6156	-11.1187
	DM_{eff}	1.0132	1.0138	1.0146	1.0159	1.0181	1.0221	1.0328	1.1021
	k_D	0.4569	0.4525	0.4482	0.4449	0.4384	0.4265	0.3921	0.3104

Table 5 locally AM-optimal weights for PR and PRR models, locally AM-optimal k_A and DM-efficiency measure for ($\beta_0 = 2, \beta_1 = 1, \beta_2 = -1$) and 8 different correlations

	AM Optimality								
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.3207	0.3181	0.3128	0.3042	0.2911	0.2721	0.2457	0.2217
	δ_2^*	0.4864	0.4745	0.4617	0.4477	0.4315	0.411	0.3805	0.329
	δ_3^*	0.1928	0.2074	0.2255	0.2481	0.2774	0.3169	0.3738	0.4493
	$tr(MSE)$	0.0102	0.0104	0.0107	0.0114	0.0127	0.0158	0.0257	0.1098
PRR	δ_1^*	0.3256	0.3187	0.3124	0.3048	0.2905	0.2738	0.2441	0.2166
	δ_2^*	0.4827	0.4722	0.4607	0.4456	0.4305	0.4094	0.3802	0.3298
	δ_3^*	0.1917	0.2091	0.2269	0.2496	0.2789	0.3169	0.3757	0.4536
	$tr(MSE)$	0.0102	0.0103	0.0107	0.0113	0.0127	0.0157	0.0254	0.1039
	AM_{eff}	1.0005	1.0006	1.0007	1.0008	1.0011	1.0017	1.0038	1.0188
	k_A	0.4475	0.4546	0.439	0.4765	0.5292	0.5584	0.5988	0.5637

Table 6 locally AM-optimal weights for PR and PRR models, locally AM-optimal k_A and DM-efficiency measure for ($\beta_0 = 1, \beta_1 = 2, \beta_2 = -1$ and 8 different correlations

	AM Optimality								
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.5467	0.5572	0.5633	0.5634	0.5553	0.5342	0.492	0.4312
	δ_2^*	0.3611	0.3391	0.3178	0.297	0.2761	0.2535	0.2241	0.1784
	δ_3^*	0.0922	0.1037	0.1189	0.1396	0.1687	0.2123	0.2838	0.3904
	$tr(MSE)$	0.0223	0.0211	0.0203	0.0199	0.0204	0.0228	0.0334	0.1356
PRR	δ_1^*	0.541	0.5502	0.5552	0.5552	0.5446	0.5295	0.4897	0.4298
	δ_2^*	0.3644	0.3437	0.3225	0.3012	0.2872	0.2555	0.2246	0.1785
	δ_3^*	0.0945	0.1061	0.1223	0.1436	0.1682	0.215	0.2857	0.3917
	$tr(MSE)$	0.0222	0.021	0.0202	0.0198	0.0202	0.0227	0.0331	0.1317
	AM_{eff}	1.0022	1.0022	1.0022	1.0021	1.0018	1.0019	1.0024	1.0096
	k_A	0.6279	0.6938	0.7238	0.7193	0.6965	0.4707	0.3221	0.2369

Table 7 locally AM-optimal weights for PR and PRR models, locally AM-optimal k_A and DM-efficiency measure for ($\beta_0 = 0, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

AM Optimality									
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.5342	0.5212	0.5072	0.4919	0.4748	0.4561	0.4382	0.4443
	δ_2^*	0.0477	0.0521	0.0571	0.063	0.0698	0.0779	0.0875	0.0946
	δ_3^*	0.4181	0.4267	0.4357	0.4452	0.4554	0.4661	0.4743	0.4612
	$tr(MSE)$	0.0234	0.025	0.0272	0.0305	0.0357	0.0459	0.0759	0.3186
PRR	δ_1^*	0.527	0.5125	0.501	0.4864	0.471	0.4536	0.4365	0.4422
	δ_2^*	0.0487	0.0532	0.0581	0.0638	0.0704	0.0784	0.0879	0.0952
	δ_3^*	0.4243	0.4343	0.4408	0.4498	0.4586	0.468	0.4757	0.4626
	$tr(MSE)$	0.0232	0.0248	0.027	0.0302	0.0354	0.0455	0.0748	0.2971
	AM_{eff}	1.0032	1.003	1.0028	1.0027	1.0027	1.003	1.005	1.0235
	k_A	0.8689	0.7858	0.6643	0.5515	0.4536	0.3672	0.2966	0.2507

Table 8 locally AM-optimal weights for PR and PRR models, locally AM-optimal k_A and DM-efficiency measure for ($\beta_0 = -1, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

AM Optimality									
		$\rho = 0.65$	$\rho = 0.70$	$\rho = 0.75$	$\rho = 0.80$	$\rho = 0.85$	$\rho = 0.90$	$\rho = 0.95$	$\rho = 0.99$
PR	δ_1^*	0.5342	0.5212	0.5072	0.4919	0.4748	0.456	0.4381	0.4443
	δ_2^*	0.0477	0.0521	0.0572	0.063	0.0698	0.0779	0.0875	0.0946
	δ_3^*	0.4181	0.4267	0.4357	0.4452	0.4554	0.4661	0.4744	0.4612
	$tr(MSE)$	0.0635	0.068	0.074	0.0828	0.0971	0.1248	0.2062	0.8661
PRR	δ_1^*	0.5292	0.516	0.5009	0.4846	0.4656	0.4441	0.4238	0.4245
	δ_2^*	0.0482	0.0528	0.058	0.0639	0.071	0.0795	0.0897	0.0979
	δ_3^*	0.4226	0.4312	0.4411	0.4515	0.4634	0.4764	0.4865	0.4776
	$tr(MSE)$	0.0628	0.0671	0.073	0.0815	0.0952	0.1215	0.1965	0.7093
	AM_{eff}	1.0039	1.0042	1.0046	1.0052	1.0065	1.009	1.0162	1.0689
	k_A	0.3752	0.3776	0.3856	0.396	0.4078	0.409	0.3693	0.2862

The primary purpose of this paper is to compare the use of conventional Ridge parameter estimation with our proposed parameter estimations. For this purpose, we list the estimators proposed in Månson and Shukur (2011). The eight ridge parameter estimator are listed as follows.

$$k_1 = \frac{\hat{\sigma}^2}{\hat{\alpha}^2_{max}} \qquad k_2 = \frac{1}{\hat{\alpha}^2_{max}} \qquad k_3 = \frac{s^2}{(\prod_{i=1}^{p} \hat{\alpha}_i{}^2)^{\frac{1}{p}}} \qquad k_4 = median\{m_i^2\}$$

$$k_5 = max(s_i) \qquad k_6 = max(\frac{1}{m_i}) \qquad k_7 = (\prod_{i=1}^{p} \frac{1}{m_i})^{\frac{1}{p}} \qquad k_8 = median(\frac{1}{m_i})$$

where $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n}(y_i - \hat{\mu}_i)^2}{n-p-1}$, $\hat{\alpha}_i^2 = \boldsymbol{\gamma}_i{}^T \hat{\boldsymbol{\beta}}_{ML}$, $s_i = \dfrac{t_i \hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_i \hat{\alpha}_i^2}$, $m_i = \sqrt{\dfrac{\hat{\sigma}^2}{\hat{\alpha}_i^2}}$.

Here, $\boldsymbol{\gamma}_i$ is the eigenvector corresponding to the eigenvalues λ_i of $\mathbf{F}^T\hat{\mathbf{W}}\mathbf{F}$ matrix and t_i is the eigenvalue of the $\mathbf{F}^T\mathbf{F}$ matrix.

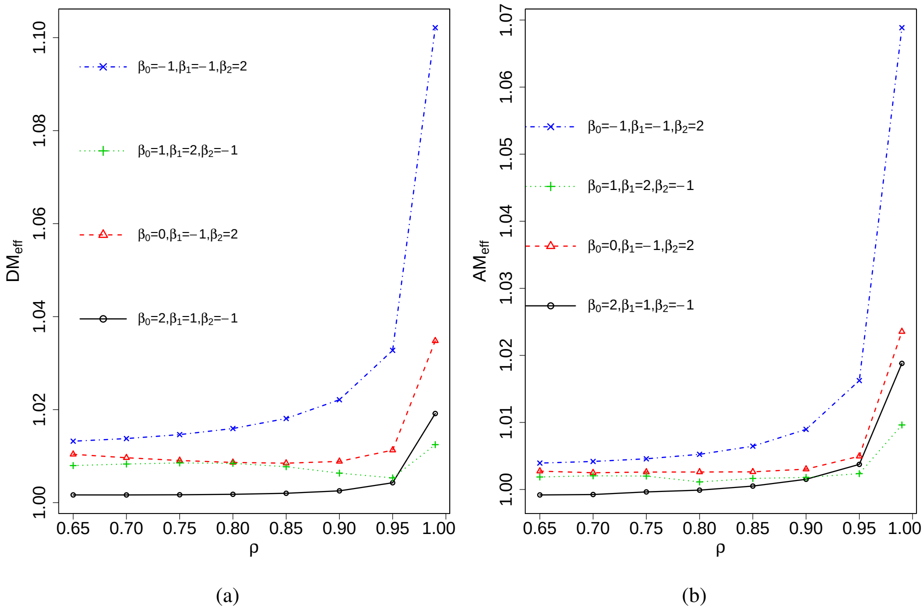

Figure 1 Efficiency of optimal designs; (a) plot of DM-efficiency, (b) plot of AM-efficiency

Table 9 $DM_{k_{eff}}$ to compare k_D and $k_1,\ldots,\ k_8$ for $(\beta_0 = 2, \beta_1 = 1, \beta_2 = -1)$ and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
0.65	$logdet(MSE)$	-18.4000	-18.4700	-16.6530	-18.3860	-18.4680	-18.4690	-18.4680	-18.4710
	k	2.1222	0.2347	17.1820	2.2929	0.1190	0.6865	0.1185	0.4361
	$DM_{k_{eff}}$	1.0238	1.0003	1.8327	1.0285	1.0009	1.0006	1.0009	1.0000
0.70	$logdet(MSE)$	-18.4140	-18.4740	-16.5600	-18.3850	-18.4720	-18.4730	-18.4720	-18.4750
	k	2.0132	0.2268	18.5520	2.3735	0.1226	0.7048	0.1119	0.4213
	$DM_{k_{eff}}$	1.0205	1.0004	1.8936	1.0304	1.0009	1.0006	1.0009	1.0000
0.75	$logdet(MSE)$	-18.3150	-18.4500	-16.2460	-18.4080	-18.4480	-18.4500	-18.4470	-18.4500
	k	2.8595	0.3239	22.8340	1.7431	0.1210	0.5914	0.0957	0.5737
	$DM_{k_{eff}}$	1.0460	1.0001	2.0847	1.0141	1.0009	1.0002	1.0010	1.0002
0.80	$logdet(MSE)$	-18.3330	-18.3820	-16.0720	-18.2600	-18.3810	-18.3820	-18.3810	-18.3840
	k	1.8642	0.2087	24.6170	2.7275	0.1301	0.7324	0.0905	0.3666
	$DM_{k_{eff}}$	1.0170	1.0005	2.1614	1.0421	1.0009	1.0007	1.0011	1.0001
0.85	$logdet(MSE)$	-18.2120	-18.2510	-15.8780	-18.1030	-18.2500	-18.2500	-18.2490	-18.2520
	k	1.6792	0.2001	25.3890	2.9055	0.1341	0.7717	0.0884	0.3442
	$DM_{k_{eff}}$	1.0136	1.0006	2.2068	1.0510	1.0010	1.0009	1.0013	1.0001
0.90	$logdet(MSE)$	-17.9620	-18.0010	-15.6460	-17.7900	-18.0000	-18.0010	-17.9990	-18.0030
	k	1.6157	0.1926	24.0730	3.2409	0.1381	0.7867	0.0920	0.3086
	$DM_{k_{eff}}$	1.0141	1.0009	2.1950	1.0741	1.0013	1.0011	1.0016	1.0003
0.95	$logdet(MSE)$	-17.4260	-17.4660	-15.3410	-17.1240	-17.4640	-17.4640	-17.4630	-17.4680
	k	1.4348	0.1870	17.9500	3.4907	0.1422	0.8348	0.1147	0.2865
	$DM_{k_{eff}}$	1.0151	1.0016	2.0337	1.1225	1.0021	1.0021	1.0025	1.0007
0.99	$logdet(MSE)$	-15.9030	-16.0060	-14.1260	-15.2300	-16.0000	-16.0030	-16.0000	-16.0170
	k	1.3613	0.1844	13.5190	3.7749	0.1455	0.8571	0.1418	0.2649
	$DM_{k_{eff}}$	1.0427	1.0073	1.8852	1.3050	1.0093	1.0086	1.0095	1.0039

To compare our proposed estimation, k_D and k_A, and the above estimators, $k_1, k_2, \ldots, k_8$, we introduce two measures for efficiencies, $DM_{k_{eff}}$ and $AM_{k_{eff}}$, which are defined as follows.

$$DM_{k_{eff}} = \left(\frac{\det \left(MSE_{PRR}^{(k_i)}(\xi)\right)}{\det \left(MSE_{PRR}^{(k_D)}(\xi)\right)} \right)^{\frac{1}{p}} \quad ; \quad i = 1, 2, \ldots, 8 \tag{4}$$

$$AM_{k_{eff}} = \left(\frac{\text{tr} \left(MSE_{PRR}^{(k_i)}(\xi)\right)}{\text{tr} \left(MSE_{PRR}^{(k_A)}(\xi)\right)} \right)^{\frac{1}{p}} \quad ; \quad i = 1, 2, \ldots, 8 \tag{5}$$

where $MSE_{PRR}^{(k_i)}(\xi)$ is the MSE matrix for the Poisson ridge regression model with ridge parameter k_i. The results for $DM_{k_{eff}}$ and $AM_{k_{eff}}$ measures are listed in Tables 9-12 and Tables 13-16 for some representative values of $\boldsymbol{\beta}$ and ρ , respectively.

Table 10 $DM_{k_{eff}}$ to compare k_D and $k_1,\ldots,k_8$ for ($\beta_0 = 1, \beta_1 = 2, \beta_2 = -1$) and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
	$logdet(MSE)$	-16.5480	-17.1550	-16.5140	-17.0680	-17.1420	-17.1530	-17.1470	-17.1530
0.65	k	6.6937	0.4818	6.9553	2.5920	0.1172	0.3865	0.2335	0.3858
	$DM_{k_{eff}}$	1.2263	1.0014	1.2403	1.0310	1.0060	1.0023	1.0042	1.0023
	$logdet(MSE)$	-16.8370	-17.2870	-16.7510	-17.2170	-17.2740	-17.2860	-17.2790	-17.2850
0.70	k	6.0224	0.4498	6.7152	2.6467	0.1202	0.4075	0.2397	0.3778
	$DM_{k_{eff}}$	1.1641	1.0022	1.1979	1.0258	1.0064	1.0026	1.0046	1.0029
	$logdet(MSE)$	-17.3560	-17.3800	-15.4630	-17.1190	-17.3780	-17.3950	-17.3750	-17.3810
0.75	k	2.3248	0.1837	29.2950	4.9030	0.1257	0.6559	0.0794	0.2040
	$DM_{k_{eff}}$	1.0140	1.0058	1.9058	1.0971	1.0067	1.0010	1.0074	1.0055
	$logdet(MSE)$	-17.3190	-17.4470	-17.0230	-17.3830	-17.4390	-17.4530	-17.4440	-17.4480
0.80	k	3.6840	0.3103	6.4580	2.9177	0.1257	0.5210	0.2469	0.3427
	$DM_{k_{eff}}$	1.0478	1.0042	1.1566	1.0257	1.0067	1.0020	1.0049	1.0038
	$logdet(MSE)$	-17.4200	-17.4390	-16.4820	-17.2640	-17.4360	-17.4520	-17.4380	-17.4410
0.85	k	2.1168	0.1979	10.9570	3.9783	0.1336	0.6873	0.1661	0.2514
	$DM_{k_{eff}}$	1.0114	1.0049	1.3824	1.0653	1.0058	1.0005	1.0053	1.0042
	$logdet(MSE)$	-17.2830	-17.3170	-14.6530	-17.0360	-17.3150	-17.3270	-17.3110	-17.3190
0.90	k	1.8568	0.1859	57.3510	4.0185	0.1381	0.7339	0.0480	0.2489
	$DM_{k_{eff}}$	1.0150	1.0036	2.4391	1.1022	1.0042	1.0001	1.0057	1.0028
	$logdet(MSE)$	-16.5400	-16.9140	-16.0780	-16.7440	-16.9070	-16.9130	-16.9130	-16.9140
0.95	k	2.9729	0.3356	5.0690	2.0021	0.1374	0.5800	0.2960	0.4995
	$DM_{k_{eff}}$	1.1327	1.0003	1.3217	1.0584	1.0025	1.0006	1.0005	1.0001
	$logdet(MSE)$	-14.6370	-15.5520	-13.2970	-14.6970	-15.5450	-15.4780	-15.5250	-15.5340
0.99	k	2.4174	0.2445	87.4080	2.2759	0.1417	0.6432	0.0350	0.4394
	$DM_{k_{eff}}$	1.3566	1.0000	2.1205	1.3296	1.0023	1.0249	1.0092	1.0062

Table 11 $DM_{k_{eff}}$ to compare k_D and $k_1,\ldots,k_8$ for ($\beta_0 = 0, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
0.65	$logdet(MSE)$	-17.4560	-18.0830	-16.2680	-17.7110	-18.0750	-18.0870	-18.0730	-18.0790
	k	7.5592	0.2679	27.0740	5.4599	0.1172	0.3637	0.0843	0.1832
	$DM_{k_{eff}}$	1.2394	1.0055	1.8416	1.1382	1.0082	1.0041	1.0088	1.0070
0.70	$logdet(MSE)$	-17.2290	-17.8310	-15.9190	-17.4290	-17.8230	-17.8360	-17.8210	-17.8270
	k	6.5846	0.2621	26.7660	5.1178	0.1204	0.3897	0.0850	0.1954
	$DM_{k_{eff}}$	1.2277	1.0047	1.8999	1.1488	1.0072	1.0029	1.0079	1.0058
0.75	$logdet(MSE)$	-17.0180	-17.5490	-15.4060	-17.0820	-17.5420	-17.5560	-17.5390	-17.5460
	k	5.3873	0.2473	34.0870	4.9691	0.1234	0.4308	0.0709	0.2012
	$DM_{k_{eff}}$	1.1986	1.0041	2.0510	1.1732	1.0064	1.0016	1.0075	1.0049
0.80	$logdet(MSE)$	-16.5120	-17.2290	-15.3230	-16.9220	-17.2200	-17.2330	-17.2180	-17.2290
	k	5.7982	0.3097	19.4380	3.4231	0.1287	0.4153	0.1080	0.2921
	$DM_{k_{eff}}$	1.2733	1.0023	1.8924	1.1105	1.0056	1.0011	1.0060	1.0026
0.85	$logdet(MSE)$	-16.1050	-16.8420	-14.6710	-16.5110	-16.8320	-16.8460	-16.8250	-16.8430
	k	5.0418	0.3100	72.1490	3.0300	0.1306	0.4454	0.0404	0.3300
	$DM_{k_{eff}}$	1.2807	1.0016	2.0657	1.1185	1.0049	1.0003	1.0073	1.0013
0.90	$logdet(MSE)$	-15.5980	-16.3380	-14.2230	-15.9130	-16.3280	-16.3410	-16.3260	-16.3400
	k	4.0997	0.3075	17.9740	2.8234	0.1380	0.4939	0.1146	0.3542
	$DM_{k_{eff}}$	1.2811	1.0009	2.0256	1.1535	1.0043	1.0001	1.0049	1.0004
0.95	$logdet(MSE)$	-14.4860	-15.5530	-13.6350	-15.0040	-15.5410	-15.5480	-15.5440	-15.5510
	k	4.1316	0.3621	10.5430	2.3822	0.1419	0.4920	0.1709	0.4198
	$DM_{k_{eff}}$	1.4271	1.0000	1.8947	1.2005	1.0040	1.0016	1.0030	1.0004
0.99	$logdet(MSE)$	-12.5730	-13.8990	-12.3670	-13.0520	-13.9090	-13.8820	-13.9280	-13.8840
	k	4.2715	0.4383	7.1995	2.0901	0.1446	0.4839	0.2275	0.4785
	$DM_{k_{eff}}$	1.5715	1.0104	1.6833	1.3399	1.0068	1.0161	1.0007	1.0153

Table 12 $DM_{k_{eff}}$ to compare k_D and $k_1,\ldots,k_8$ for ($\beta_0 = -1, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
0.65	$logdet(MSE)$	-14.7990	-15.0950	-13.2550	-14.3390	-15.0850	-15.0980	-15.0800	-15.1010
	k	2.0842	0.1911	34.4170	3.7879	0.1126	0.6927	0.0704	0.2640
	$DM_{k_{eff}}$	1.1083	1.0044	1.8542	1.2921	1.0075	1.0032	1.0095	1.0023
0.70	$logdet(MSE)$	-14.6470	-14.8430	-13.0040	-14.1000	-14.8340	-14.8390	-14.8300	-14.8500
	k	1.6882	0.1880	25.8430	3.6652	0.1143	0.7697	0.0872	0.2728
	$DM_{k_{eff}}$	1.0724	1.0047	1.8547	1.2868	1.0077	1.0060	1.0090	1.0021
0.75	$logdet(MSE)$	-14.4320	-14.5620	-12.8870	-13.6520	-14.5530	-14.5480	-14.5470	-14.5670
	k	1.4061	0.1794	31.3900	4.3245	0.1159	0.8433	0.0754	0.2312
	$DM_{k_{eff}}$	1.0497	1.0052	1.7566	1.3613	1.0080	1.0098	1.0102	1.0034
0.80	$logdet(MSE)$	-14.1680	-14.2400	-13.2730	-12.5510	-14.2320	-14.2100	-14.2190	-14.2220
	k	1.1388	0.1676	60.7960	17.2030	0.1168	0.9371	0.0459	0.0581
	$DM_{k_{eff}}$	1.0305	1.0061	1.3888	1.7663	1.0087	1.0163	1.0130	1.0122
0.85	$logdet(MSE)$	-13.7520	-13.8610	-12.8260	-13.2000	-13.8470	-13.8260	-13.8710	-13.8720
	k	1.2168	0.2096	4.7734	3.0483	0.1208	0.9066	0.3097	0.3281
	$DM_{k_{eff}}$	1.0419	1.0048	1.4188	1.2527	1.0095	1.0168	1.0015	1.0011
0.90	$logdet(MSE)$	-13.2380	-13.3670	-11.9590	-13.0600	-13.3500	-13.3140	-13.3580	-13.3730
	k	1.1906	0.2341	10.1580	1.7239	0.1381	0.9165	0.1758	0.5801
	$DM_{k_{eff}}$	1.0487	1.0044	1.6060	1.1127	1.0100	1.0223	1.0075	1.0025
0.95	$logdet(MSE)$	-12.5690	-12.5960	-11.4060	-12.2230	-12.5770	-12.4020	-12.5960	-12.5920
	k	0.7097	0.2112	7.9707	1.6384	0.1422	1.1870	0.2108	0.6104
	$DM_{k_{eff}}$	1.0156	1.0066	1.4965	1.1398	1.0129	1.0739	1.0066	1.0078
0.99	$logdet(MSE)$	-11.0420	-11.0890	-10.4440	-10.5700	-11.0480	-10.6960	-11.1060	-11.0510
	k	0.5578	0.1981	3.4283	1.8563	0.1446	1.3389	0.3969	0.5387
	$DM_{k_{eff}}$	1.0259	1.0100	1.2520	1.2005	1.0239	1.1515	1.0042	1.0228

Table 13 $AM_{k_{eff}}$ to compare k_A and $k_1,\ldots,k_8$ for $(\beta_0 = 2, \beta_1 = 1, \beta_2 = -1)$ and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
0.65	$tr(MSE)$	0.0103	0.0102	0.3903	0.0131	0.0102	0.0103	0.0102	0.0102
	k	1.1959	0.1747	73.3460	4.9736	0.1191	0.9144	0.0399	0.2011
	$AM_{k_{eff}}$	1.0032	1.0001	3.3664	1.0863	1.0002	1.0014	1.0005	1.0001
0.70	$tr(MSE)$	0.0105	0.0103	0.0203	0.0108	0.0104	0.0104	0.0103	0.0103
	k	1.4524	0.2162	9.2192	2.2649	0.1225	0.8298	0.1890	0.4415
	$AM_{k_{eff}}$	1.0051	1.0001	1.2526	1.0156	1.0003	1.0009	1.0001	1.0000
0.75	$tr(MSE)$	0.0111	0.0107	0.0174	0.0108	0.0107	0.0107	0.0107	0.0107
	k	2.0732	0.3165	7.5278	1.5391	0.1191	0.6945	0.2200	0.6498
	$AM_{k_{eff}}$	1.0123	1.0000	1.1777	1.0057	1.0003	1.0004	1.0001	1.0003
0.80	$tr(MSE)$	0.0115	0.0113	0.0438	0.0123	0.0113	0.0113	0.0113	0.0113
	k	1.5808	0.1987	16.3310	2.9050	0.1301	0.7954	0.1231	0.3442
	$AM_{k_{eff}}$	1.0059	1.0002	1.5701	1.0265	1.0004	1.0006	1.0004	1.0000
0.85	$tr(MSE)$	0.0129	0.0127	0.0809	0.0140	0.0127	0.0127	0.0127	0.0127
	k	1.5563	0.1917	22.7950	3.2431	0.1341	0.8016	0.0959	0.3083
	$AM_{k_{eff}}$	1.0056	1.0004	1.8556	1.0348	1.0006	1.0005	1.0007	1.0001
0.90	$tr(MSE)$	0.0160	0.0157	0.1698	0.0182	0.0157	0.0157	0.0158	0.0157
	k	1.5408	0.1867	31.2120	3.5989	0.1381	0.8056	0.0757	0.2779
	$AM_{k_{eff}}$	1.0058	1.0007	2.2120	1.0501	1.0010	1.0004	1.0013	1.0004
0.95	$tr(MSE)$	0.0259	0.0256	0.3484	0.0324	0.0256	0.0255	0.0257	0.0255
	k	1.3749	0.1831	34.0770	3.7524	0.1422	0.8529	0.0709	0.2665
	$AM_{k_{eff}}$	1.0060	1.0018	2.3927	1.0837	1.0022	1.0007	1.0029	1.0011
0.99	$tr(MSE)$	0.1125	0.1065	0.9457	0.2074	0.1071	0.1053	0.1080	0.1055
	k	1.3388	0.1819	23.9720	3.8667	0.1456	0.8643	0.0923	0.2586
	$AM_{k_{eff}}$	1.0271	1.0084	2.0882	1.2593	1.0102	1.0046	1.0130	1.0053

Table 14 $AM_{k_{eff}}$ to compare k_A and $k_1,\ldots,k_8$ for $(\beta_0 = 1, \beta_1 = 2, \beta_2 = -1)$ and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
0.65	$tr(MSE)$	0.0225	0.0222	0.0230	0.0225	0.0223	0.0222	0.0222	0.0222
	k	1.5277	0.3754	2.0877	1.5428	0.1137	0.8091	0.5758	0.6482
	$AM_{k_{eff}}$	1.0047	1.0003	1.0119	1.0048	1.0015	1.0002	1.0000	1.0000
0.70	$tr(MSE)$	0.0222	0.0210	0.0231	0.0214	0.0211	0.0210	0.0210	0.0210
	k	2.6486	0.4127	3.2758	1.7498	0.1172	0.6145	0.4107	0.5715
	$AM_{k_{eff}}$	1.0193	1.0003	1.0325	1.0059	1.0015	1.0000	1.0003	1.0000
0.75	$tr(MSE)$	0.0202	0.0202	0.0423	0.0208	0.0202	0.0202	0.0202	0.0202
	k	0.9180	0.1978	10.0890	2.2426	0.1232	1.0437	0.1766	0.4459
	$AM_{k_{eff}}$	1.0002	1.0012	1.2802	1.0112	1.0015	1.0006	1.0013	1.0003
0.80	$tr(MSE)$	0.0198	0.0198	0.0336	0.0214	0.0199	0.0198	0.0198	0.0198
	k	1.1667	0.1916	8.0040	3.0579	0.1298	0.9258	0.2101	0.3270
	$AM_{k_{eff}}$	1.0012	1.0011	1.1932	1.0259	1.0014	1.0003	1.0010	1.0006
0.85	$tr(MSE)$	0.0206	0.0203	0.0242	0.0213	0.0203	0.0203	0.0203	0.0203
	k	1.6428	0.2524	4.0864	2.3705	0.1317	0.7802	0.3479	0.4219
	$AM_{k_{eff}}$	1.0060	1.0007	1.0613	1.0170	1.0012	1.0001	1.0003	1.0001
0.90	$tr(MSE)$	0.0230	0.0227	0.0606	0.0313	0.0227	0.0228	0.0227	0.0227
	k	1.2414	0.1836	9.2512	4.4121	0.1372	0.8975	0.1885	0.2267
	$AM_{k_{eff}}$	1.0053	1.0007	1.3876	1.1133	1.0010	1.0017	1.0007	1.0005
0.95	$tr(MSE)$	0.0348	0.0332	0.9545	0.2605	0.0332	0.0339	0.0333	0.0333
	k	1.2105	0.1684	40.7710	13.7680	0.1273	0.9089	0.0620	0.0726
	$AM_{k_{eff}}$	1.0169	1.0005	3.0661	1.9887	1.0009	1.0076	1.0015	1.0014
0.99	$tr(MSE)$	0.2399	0.1317	1.8887	0.2675	0.1323	0.1466	0.1325	0.1358
	k	1.7834	0.2408	15.2690	2.0107	0.1406	0.7488	0.1295	0.4973
	$AM_{k_{eff}}$	1.2212	1.0000	2.4295	1.2663	1.0015	1.0364	1.0019	1.0103

Table 15 $AM_{k_{eff}}$ to compare k_A and $k_1,\ldots,k_8$ for ($\beta_0 = 0, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
	$tr(MSE)$	0.0234	0.0232	0.0249	0.0236	0.0233	0.0232	0.0232	0.0232
0.65	k	1.6497	0.3480	3.2194	1.9851	0.1110	0.7786	0.4161	0.5038
	$AM_{k_{eff}}$	1.0030	1.0011	1.0242	1.0058	1.0025	1.0000	1.0008	1.0005
	$tr(MSE)$	0.0249	0.0248	0.0260	0.0251	0.0249	0.0248	0.0248	0.0248
0.70	k	1.2414	0.3401	2.5488	1.6890	0.1116	0.8975	0.4957	0.5921
	$AM_{k_{eff}}$	1.0013	1.0009	1.0166	1.0047	1.0022	1.0001	1.0003	1.0001
	$tr(MSE)$	0.0271	0.0271	0.0294	0.0276	0.0272	0.0270	0.0270	0.0270
0.75	k	1.1588	0.3170	2.7951	1.6764	0.1139	0.9289	0.4626	0.5965
	$AM_{k_{eff}}$	1.0018	1.0007	1.0290	1.0070	1.0019	1.0005	1.0002	1.0000
	$tr(MSE)$	0.0303	0.0303	0.0444	0.0313	0.0304	0.0305	0.0303	0.0302
0.80	k	0.8247	0.2585	4.9992	1.7279	0.1298	1.1011	0.2991	0.5787
	$AM_{k_{eff}}$	1.0007	1.0007	1.1363	1.0119	1.0016	1.0027	1.0005	1.0000
	$tr(MSE)$	0.0358	0.0355	0.0511	0.0365	0.0356	0.0359	0.0355	0.0355
0.85	k	0.9447	0.3066	4.0569	1.3634	0.1336	1.0289	0.3498	0.7335
	$AM_{k_{eff}}$	1.0031	1.0003	1.1295	1.0103	1.0013	1.0042	1.0001	1.0010
	$tr(MSE)$	0.0488	0.0455	0.0682	0.0484	0.0457	0.0461	0.0455	0.0459
0.90	k	1.4423	0.3742	3.3448	1.3716	0.1287	0.8327	0.4043	0.7291
	$AM_{k_{eff}}$	1.0235	1.0000	1.1447	1.0207	1.0013	1.0047	1.0000	1.0028
	$tr(MSE)$	0.0875	0.0748	0.1549	0.0855	0.0751	0.0783	0.0749	0.0774
0.95	k	1.3799	0.3845	3.2964	1.2858	0.1410	0.8513	0.4088	0.7777
	$AM_{k_{eff}}$	1.0539	1.0004	1.2747	1.0458	1.0013	1.0156	1.0007	1.0119
	$tr(MSE)$	0.5064	0.3046	0.8248	0.4739	0.3008	0.3768	0.3155	0.3653
0.99	k	1.3546	0.4172	2.3958	1.2402	0.1425	0.8592	0.5193	0.8063
	$AM_{k_{eff}}$	1.1945	1.0083	1.4054	1.1683	1.0041	1.0824	1.0202	1.0712

Table 16 $AM_{k_{eff}}$ to compare k_A and $k_1,\ldots,k_8$ for ($\beta_0 = -1, \beta_1 = -1, \beta_2 = 2$) and 8 different correlations

ρ		k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8
	$tr(MSE)$	0.0628	0.0629	0.0808	0.0647	0.0632	0.0695	0.0629	0.0648
0.65	k	0.4152	0.2384	2.3585	0.9978	0.0925	1.5519	0.5254	1.0022
	$AM_{k_{eff}}$	1.0000	1.0005	1.0878	1.0102	1.0022	1.0343	1.0006	1.0104
	$tr(MSE)$	0.0671	0.0672	0.0754	0.0686	0.0676	0.0770	0.0677	0.0701
0.70	k	0.3267	0.2382	1.6225	0.8938	0.0892	1.7495	0.6956	1.1188
	$AM_{k_{eff}}$	1.0001	1.0006	1.0394	1.0073	1.0025	1.0471	1.0028	1.0148
	$tr(MSE)$	0.0731	0.0731	0.0770	0.0737	0.0735	0.0868	0.0746	0.0791
0.75	k	0.2749	0.2576	1.1839	0.7258	0.1019	1.9073	0.8811	1.3778
	$AM_{k_{eff}}$	1.0004	1.0005	1.0179	1.0034	1.0025	1.0595	1.0071	1.0271
	$tr(MSE)$	0.0817	0.0818	0.0852	0.0899	0.0823	0.0996	0.0838	0.0822
0.80	k	0.2536	0.2044	1.0855	1.4601	0.0860	1.9859	0.9403	0.6849
	$AM_{k_{eff}}$	1.0007	1.0012	1.0146	1.0333	1.0032	1.0690	1.0093	1.0027
	$tr(MSE)$	0.0954	0.0954	0.0956	0.0978	0.0964	0.1151	0.1064	0.1003
0.85	k	0.2933	0.2810	0.5993	0.8992	0.0920	1.8464	1.4682	1.1121
	$AM_{k_{eff}}$	1.0005	1.0006	1.0014	1.0088	1.0039	1.0652	1.0376	1.0175
	$tr(MSE)$	0.1217	0.1222	0.8107	0.1280	0.1230	0.1509	0.1226	0.1273
0.90	k	0.3214	0.2183	10.2910	1.0191	0.1381	1.7639	0.1740	0.9813
	$AM_{k_{eff}}$	1.0004	1.0019	1.8826	1.0176	1.0040	1.0748	1.0030	1.0156
	$tr(MSE)$	0.1987	0.1984	0.3521	0.2133	0.2001	0.3457	0.1980	0.2242
0.95	k	0.1891	0.2045	2.3526	0.9177	0.1420	2.2998	0.5264	1.0897
	$AM_{k_{eff}}$	1.0038	1.0031	1.2146	1.0277	1.0061	1.2072	1.0026	1.0450
	$tr(MSE)$	0.7286	0.7214	1.0899	1.0623	0.7411	1.6954	0.9499	0.9543
0.99	k	0.1721	0.1943	1.1497	1.1005	0.1428	2.4107	0.9007	0.9086
	$AM_{k_{eff}}$	1.0090	1.0057	1.1540	1.1441	1.0147	1.3371	1.1023	1.1040

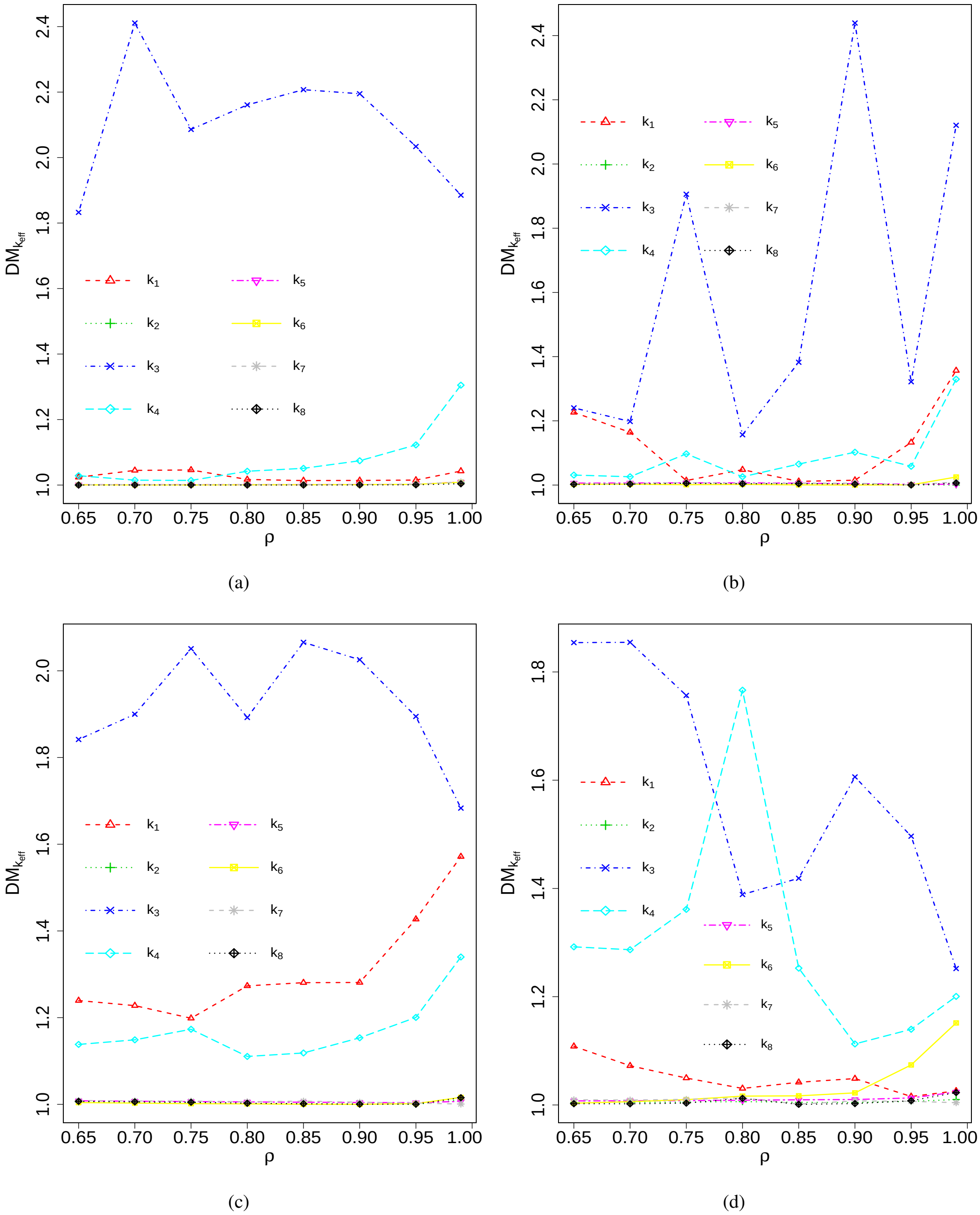

Figure 2 Plot $DM_{k_{eff}}$ versus ρ for k_1,…,k_8 and some representative values of parameters; a:$\boldsymbol{\beta} = (2, 1, -1)$, b:$\boldsymbol{\beta} = (1, 2, -1)$, c:$\boldsymbol{\beta} = (0, -1, 2)$, d:$\boldsymbol{\beta} = (-1, -1, 2)$

Although for some values of the parameters, the efficiency is close to 1, the results show that, in general, our proposed estimates perform better based on both the DM- and AM-optimal criteria. Also, by increasing the value of ρ in the simulated data, for each $k_i, i = 1, \ldots, 8$, the efficiency of k_D and k_A increases. Of course, this increase in efficiency does not have a fixed value and varies depending on k_i.

$DM_{k_{eff}}$ and $AM_{k_{eff}}$ plot versus for different ridge parameters $k_i, i = 1, \ldots, 8$, corresponding to Tables 9-12 and Tables 13-16, in Figure 2 and Figure 3, respectively.

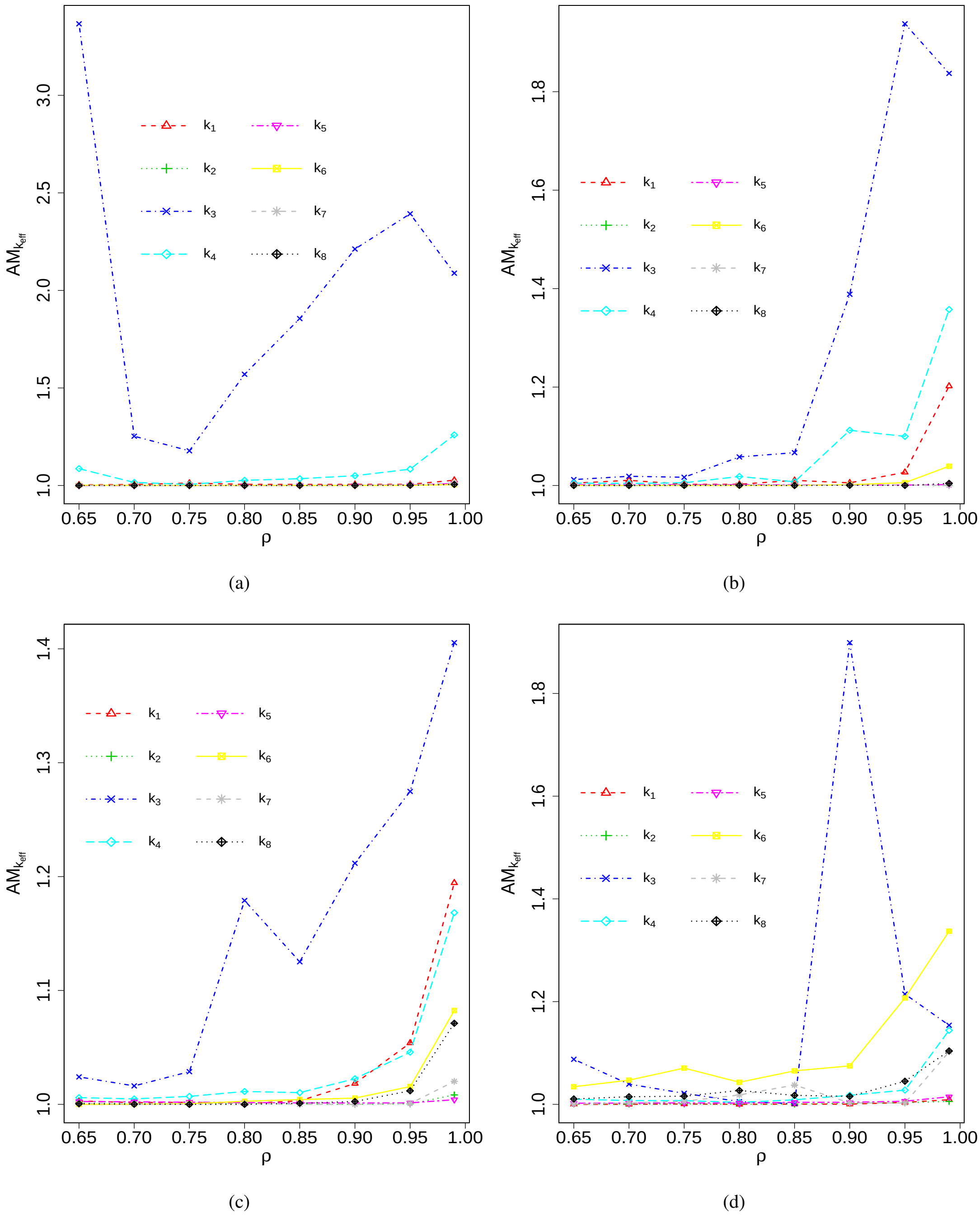

Figure 3 Plot $AM_{k_{eff}}$ versus ρ for $k_1,\ldots,k_8$ and some representative values of parameters; a:$\boldsymbol{\beta} = (2, 1, -1)$, b:$\boldsymbol{\beta} = (1, 2, -1)$, c:$\boldsymbol{\beta} = (0, -1, 2)$, d:$\boldsymbol{\beta} = (-1, -1, 2)$

5. Conclusion

Collinearity in predictor variables causes some problems in analyzing the real data, including count data. Ridge regression is known as a tool to remedy this problem. In this study, we apply an optimal design approach to find the efficient ridge parameter estimations for the Poisson ridge regression model. For different values of correlation, we have also obtained locally DM- and AM-optimal designs for this model and we compare them to corresponding optimal designs in PR model. The results in Tables 1-8 show that the new optimal designs are more efficient. Comparing our proposed results for ridge parameter estimates to the previous estimates, Tables 9-16, indicate that the new estimates are better, and also the new DM- and AM-optimal designs are more efficient. Results show that the efficiency of new DM- and AM-optimal designs and optimal Ridge parameters are increasing when ρ arises.

References

Ali S, Khan H, Shah I, Butt MM, Suhail M. A comparison of some new and old robust ridge regression estimators. Commun Stat Simul Comput. 2021; 50(8): 2213-2231.

Chernoff H. Locally optimal designs for estimating parameters. Ann Math Statist. 1953; 24(4): 586-602.

Gibbons DG. A simulation study of some ridge estimators. J Am Stat Assoc. 1981; 76(373): 131-139.

Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 1970; 12(1): 55-67.

Lawless JF, Wang P. A simulation study of ridge and other regression estimators. Commun Stat Theo Meth. 1976; 5(4): 307-323.

Mehr Mansour S, Niaparast M. The effect of small sample on optimal designs for logistic regression models. Commun Stat Theo Meth. 2018; 48(12): 2893-2903.

Mehr Mansour S, Niaparast M. Optimal designs for small Poisson regression experiments using second-order asymptotic. Commun Stat Appl Methods. 2019; 26(6): 527-538.

Mansson K, Shukur G. A Poisson ridge regression estimator. Econ Model. 2011; 28(4): 1475-1481.

McDonald GC, Galarneau DI. A Monte Carlo evaluation of some ridge-type estimators. J Am Stat Assoc. 1975; 70(350): 407-416.

Muniz G, Kibria BG. On some ridge regression estimators: An empirical comparisons. Commun Stat Simul Comput. 2009; 38(3): 621-630.

Russell KG, Woods DC, Lewis SM, Eccleston JA. D-optimal designs for Poisson regression models. Stat Sinica. 2009; 19: 721-730.

Shah I, Sajid F, Ali S, Rehman A, Bahaj SA, Fati SM (2021). On the Performance of Jackknife Based Estimators for Ridge Regression. IEEE Access. 2021; 9: 68044-68053.

Silvey S D. Optimal design: an introduction to the theory for parameter estimation. UK: Chapman and Hall; 1980.

Wang Y, Myers RH, Smith EP, Ye K. D-optimal designs for Poisson regression models. J Stat Plan Infer. 2006; 136(8): 2831-2845.