ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Atif ABBASI

RESTRICTED ESTIMATION METHODS IN GENERALIZED LINEAR MODELS

DEPARTMENT OF STATISTICS

ADANA-2022

ABSTRACT

PhD THESIS

RESTRICTED ESTIMATION METHODS IN GENERALIZED LINEAR MODELS

Atif ABBASI

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF STATISTICS

Supervisor	:	Prof. Dr. Mahmude Revan ÖZKALE
		Year: 2022, Pages: 192
Jury	:	Prof. Dr. Mahmude Revan ÖZKALE
	:	Prof. Dr. Sadullah SAKALLIOĞLU
	:	Assoc. Prof. Dr. Hüseyin GÜLER
	:	Prof. Dr. Yeliz MERT KANTAR
	:	Assoc. Prof. Dr. Bahadır YÜZBAŞI

Multicollinearity has been shown to have a detrimental effect on parameter estimation in generalized linear models (GLMs). The most frequent approach for estimating parameters in GLMs is the maximum likelihood (ML) method, but it is severely affected by the problem of multicollinearity consequently, the variance-covariance matrix of the ML estimator becomes large and prediction is also inefficient. It is, therefore, preferable to employ methods other than ML when there is a problem of multicollinearity among the explanatory variables.

In this dissertation, four iterative restricted biased estimators are proposed to address the problem of multicollinearity by imposing exact and stochastic restrictions on the parameters. These are respectively, the $r-k$ class estimator, stochastic restricted Liu estimator, restricted OK estimator, and stochastic restricted OK estimator. The performances of these estimators are evaluated by the numerical illustrations and simulation studies when a response variable belongs to binomial, Poisson, gamma, and negative binomial distributions. The performance evaluation criteria are the scalar mean square error (SMSE), expected mean square error (EMSE), and prediction mean square error (PMSE). The results demonstrated that under certain conditions, the proposed estimators outperformed the ML and many other estimators considered in this study.

Key Words: Generalized linear models, Multicollinearity, Exact restrictions, Stochastic restrictions, Binomial distribution, Poisson distribution, Gamma distribution, Negative binomial distribution

ÖZ

DOKTORA TEZİ

GENELLEŞTİRİLMİŞ LİNEER MODELLERDE KISITLI TAHMIN METOTLARI

Atif ABBASI

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK

Danışman : Prof. Dr. Mahmude Revan ÖZKALE
Yıl: 2022, Sayfa: 192
Jüri : Prof. Dr. Mahmude Revan ÖZKALE
: Prof. Dr. Sadullah SAKALLIOĞLU
: Assoc. Prof. Dr. Hüseyin GÜLER
: Prof. Dr. Yeliz MERT KANTAR
: Assoc. Prof. Dr. Bahadır YÜZBAŞI

Çoklu doğrusal bağlantının genelleştirilmiş doğrusal modellerde (GLMs) parametre tahmini üzerinde zararlı bir etkiye sahip oldugu gösterilmiştir. GLM'lerde parametreleri tahmin etmek için en sık kullanılan yaklaşım maksimum olabilirlik (ML) yöntemidir, ancak çoklu bağlantı probleminden ciddi şekildeetkilenir, sonuç olarak ML tahmin edicisinin varyans-kovaryans matrisi büyür ve ön tahminde verimsizdir. Bu nedenle, açıklayıcı değişkenler arasında çoklu bağlantı sorunu olduğunda ML dışındaki yöntemlerin kullanılması tercih edilir.

Bu tezde parametrelere tam ve stokastik kısıtlar getirerek çoklu bağlantı sorununu çözmek için dört yinelemeli kısıtlı tahmin edici önerilmiştir. Bunlar sırasıyla r-k sınıf tahmin edicisi, stokastik kısıtlı Liu tahmin edicisi, kısıtlı OK tahmin edicisi ve stokastik kısıtlı OK tahmin edicisidir. Bu tahmin edicilerin performansları, yanıt değişkeni binom, Poisson, gamma ve negatif binom dağılımlarına ait olduğunda sayısal örneklemeler simülasyon çalışmaları ile değerler değerlendirilir. Performans değerlendirme kriterleri sırasıyla skaler hata kareler ortalaması (SMSE), beklenen hata karaler ortalaması (EMSE) ve ön tahmin hata karaler ortalaması (PMSE)'dır. Sonuçları belirli koşullar altında, önerilen tahmin edicilerin ML'den ve bu çalışmada ele alınan diğer bir çok tahmin ediciden daha iyi performans gösterdiğini ortaya koymuştur.

Anahtar Kelimeler: Genelleştirilmiş doğrusal modeller, çoklu doğrusal bağlantı, Tam kısıtlamalar, stokastik kısıtlamalar, Binom dağılımı, Poisson dağılımı, Gamma dağılımı, Negatif binom dağılımı

EXTENDED ABSTRACT

Generalized linear models (GLMs) were first introduced by (Nelder and Wedderburn, 1972). One of the main reasons for GLMs' popularity is that they have brought significant harmony to statistical data analysis. Discrete and continuous response variables can be handled without taking into account any of the assumptions that are required for the application of conventional normal linear models, such as the response variable's normality or homoscedasticity. As a result, GLMs include linear regression, Poisson regression, binary logistic regression, gamma regression, negative binomial, log-linear models, and so on.

Finding a better estimation of the parameters is one of the main goals of any statistical analysis. For estimating the parameter in GLMs, maximum likelihood (ML) approaches which are implemented by the iteratively reweighted least squares (IRLS) algorithm are often utilized (Mackinnon and Puterman, 1989). In the computation, estimation, and interpretation of parameters in linear models' collinearity has always been considered a potential source of problems (Belsley et al., 1980). When estimating GLMs, (Segerstedt, 1992) discovered that an ill-conditioned data matrix has similar effects on the parameter estimator as when estimating a linear regression model. Collinearity among the explanatory variables has a major impact on the conditional ML estimator since its variance is inflated in the same way that collinearity in multiple regression increases the variance of the least squares estimator (Schaefer, 1986). As a result, it is clear that collinearity affects parameter estimation not just in normal linear models but also in GLMs.

Several studies have been undertaken to enhance parameter estimations and ease the problem of multicollinearity in linear regression models. The most popular are the ridge regression estimator proposed by (Hoerl and Kennard, 1970) and the

Liu estimator proposed by (Liu, 1993). Some studies have used prior information in addition to sample information to improve parameter estimates for example:

Theil and Goldberger (1961), Theil (1961), Groß (2003) and Zhong and Yang (2007), Fallaha et al. (2017), Li and Yang (2010), Özkale and Kaçıranlar (2007).

To avoid the problem of multicollinearity in GLMs some of these ideas have been applied to GLMs, such as; Nyquist (1991), Segerstedt (1992), Özkale and Nyquist (2021), Kurtoğlu and Özkale (2019a) and Kurtoğlu and Özkale (2019b), etc.

To address the problem of multicollinearity this dissertation aims; to introduce various biased estimators in GLMs by putting exact and stochastic restrictions on the parameters and to comparing the proposed estimators to some of those that have already been established in the literature. Furthermore, Fisher's scoring algorithm is used to obtain all of the estimators in an iterative procedure. The four innovative estimators we obtained to ease the problem of multicollinearity are; the $r-k$ class estimator, stochastic restricted Liu estimator, restricted OK estimator, and stochastic restricted OK estimator. The first-order approximated (FOA) form of the estimators is described, in addition to the iterative version. The properties of these estimators are provided as well as the techniques for selecting the tuning parameters (k,d).

To evaluate the performance of the estimators' various simulation experiments are performed in addition to the numerical illustrations when a response variable belongs to binomial, Poisson, gamma, and negative binomial distributions, respectively. The performance evaluation criteria are SMSE, EMSE, and PMSE. The results revealed that the performance of the proposed estimators is better than their counterparts for the appropriate choice of the values of the tuning parameters in addition to the type of restrictions imposed in the numerical illustrations. The simulation results showed that the novel estimators outperformed their counterparts in a variety of situations including; the restrictions imposed, the degree of multicollinearity, and the appropriate tuning parameters values.

It is observed that the most important factors which affect the performance of the estimators are the values of tuning parameters (k,d), the type of restriction imposed, and the degree of multicollinearity under consideration. To address the problem of multicollinearity it is suggested that the newly developed estimators can be used instead of ML and many other estimators discussed in this dissertation.

ACKNOWLEDGEMENTS

This thesis is the outcome of research conducted at the Department of Statistics, Çukurova University, Institute of Natural and Applied Sciences, Adana Turkey. It was a real pleasure for me to work in the field of generalized linear models, which is both scientifically and practically fascinating. Now I would like to thank some of the people who supported and encouraged me throughout the research work.

First and foremost I would like to express my profound gratitude to my thesis advisor Prof. Dr. M. Revan Özkale for her encouragement, support, and numerous helpful recommendations and comments for this study. She helped and guided me through every step of this research work, without her guidance, I would not have been able to complete it. I am really lucky to be her student and working under her supervision was a great honor for me.

I would like to thank the rest of my thesis committee members, Prof. Dr. Sadullah Sakallıoğlu and Assoc. Prof. Dr. Hüseyin Güler for their encouragement and insightful comments.

I would also like to thank YÖK (Higher Education Council of Turkey), which has provided us with such a wonderful opportunity to study in Turkey through the YÖK scholarship program. Learning the Turkish language before starting my doctoral studies was a fantastic experience. I am highly obliged and thankful for every facility which YÖK provided.

I am also grateful to Çukurova University's Department of Statistics to provide me with the space and conducive environment I needed to do my thesis work in peace and with full concentration.

Thanks also go to my teachers, who taught and guided me throughout my academic career. Last but not the least, I want to thank my family for their unwavering support during my studies. My mother (who died) deserves special thanks for her never-ending prayers for me. Without my family's assistance, finishing this thesis could be challenging.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

LIST OF ABBREVIATONS

GLMs	: Generalized linear models
LRMs	: Linear regression models
ML	: Maximum likelihood
PCs	: Principal components
PCR	:Principal component regression
PTV	: Percentage of total variation
VIF	: Variance inflation factor
CN	: Condition number
GA	: Genetic algorithm
MAE	: Mean absolute error
SMSE/smse	: Scalar mean square error
MMSE	: Matrix mean square error
MSE	: Mean square error
PMSE	: Prediction mean square error
EMSE	: Expected mean square error
RR	: Reduction rate
tr	: Trace of matrix
$\mathbf{I}_p$	: $p \times p$ identity matrix
FOA	: First-order approximated
var	: Variance
SRL	: Stochastic restricted Liu
OK	: Özkale, Kaçıranlar

Eq	: Equation
diag	: Diagonal
nnd	: Non negative definite
OLS	: Ordinary least square
Fig	: Figure
SRest/SR	: Stochastic restricted
IRLS	: Iterative reweighted least square
SVD	: Singular value decomposition

1. INTRODUCTION

This chapter gives an overview of generalized linear models, with an emphasis on model construction. The exponential distribution family, commonly used link functions, and their inverses as well as the canonical link function, and variance functions are given and discussed.

1.1. Statistical Modelling

Models are simplified, abstract representations of facts that are frequently utilized in technology and research. There are two types of models: deterministic and probabilistic. In the first type, the outcomes are precisely defined, whereas in the second, they are subject to variation due to unknown random causes. Models with a probabilistic component are known as statistical models. The generalized linear models are the most significant class, with which we are concerned. They get their name from the fact that they generalize traditional linear models based on the normal distribution (Lindsey, 2000).

The notion of Generalized linear models (GLMs) was first introduced by (Nelder and Wedderburn, 1972). They discovered the underlying unity for entire class of regression models. This class included models in which the variance of the single response variable, the variable that the model is supposed to explain, is assumed to be represented by a member of the one-parameter exponential family of probability distributions. The Gaussian or normal, binomial, Poisson, gamma, inverse Gaussian, geometric, and negative binomial are all members of this family of distributions. The normal linear regression model is based on a number of assumptions, including the following:

1. The normal or Gaussian distribution describes each observation of the response

variable i.e. $Y_i \sim N(\mu_i, \sigma_i^2)$.

2. All of the observations' distributions have the same variance; $\sigma_i^2 = \sigma^2$ for all i.

3. The linear predictor (linear combination of covariate values and related parameters) and the model's expected values have a direct or "identical" relationship.

The purpose of GLMs is to define the relationship between the response variable and set of covariates. The assumptions of the ordinary linear regression model are relaxed to develop GLMs. Through the linear predictor and the mean, a nonlinear relationship can be reconstructed into a linear relationship (Hardin and Hilbe, 2012).

As an extension of the class of linear models, GLMs were established. GLMs can accept both discrete and continuous response variables and the standard assumptions of normality and constant variance on the response under consideration are not always applied. The use of GLMs has grown extremely common in biological assays, reliability and quality engineering, survival analysis, and a range of other applied biomedical disciplines (Khuri, 2010). According to (Breslow, 1996), the critical assumptions in the GLMs framework are as follows:

- The observations are statistically independent.
- The variance function $var(\mu)$, link function, and scale factor $a(\phi)$ are correctly described, while $a(\phi)$ is (1, for Poisson, binomial, and negative binomial).
- The form of the explanatory variable is correctly specified.
- Individual observations have no disproportionate influence on the fit.

1.1.1. Model Building

The GLMs contain three components (Fox, 2016):

1. A random component that specifies the conditional distribution of Y_i (for the i-th of

n independently selected observations) given the values of the explanatory variables in the model. The distribution of Y_i in the original formulation of GLMs is a member of an exponential family such as; the Gaussian (normal), binomial, Poisson, gamma, or inverse-Gaussian families.

2. A linear predictor which is a linear function of the explanatory variables such as:

$$\eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik}$$

3. A smooth and linearizing link function $g(.)$ that converts the mean or expectation of the response variable, $\mu_i = E(Y_i)$ to the linear predictor:

$$g(\mu_i) = \eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik}$$

since the link function is invertible, it can also be written as:

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik}).$$

As a result, the GLMs can be viewed as either a linear model for transforming the predicted response or a nonlinear regression model for the response. The inverse link function $g^{-1}(.)$ is also known as mean function. The conditional variance of Y_i is a function of its mean μ_i and probably a dispersion parameter ϕ that is $var(Y_i) = a(\phi)var(\mu_i)$. GLMs are generally formulated using the exponential family of distributions, which is outlined in the next section.

1.1.2. Exponential Family

The exponential family of distributions is the foundation for the development of GLMs theory. This formulation basically defines a well-known function into a formula that is more valuable theoretically and shows similarities across seemingly different mathematical forms (Gill, 2001). The exponential family is commonly written as (although there are various algebraically similar forms in the literature)

$$f_{Y_i}(y_i, \theta_i, \phi) = \exp\left[\frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right], \quad i = 1, 2, ..., n \tag{1.1.}$$

where θ_i is a canonical (natural) parameter of location and $\phi > 0$ is a scale or dispersion parameter, while a(.), b(.) and c(.) are the known functions correspond to the type of probability distribution. For members of exponential family of distribution the location parameter (also called canonical link function) refers to the mean, while the scale parameter relates to the variances. Models for continuous, discrete, proportional, count, and binary outcomes can all be specified using the exponential family's notation. Each of the observations y_i is specified in terms of the parameter θ in the exponential family presentation. As the observations y_i are independent, the product of the density over the individual observations defines the joint density of the sample of observations, given the parameters ϕ and θ.

Conveniently the joint probability density function can be expressed as a function of θ and ϕ given the observations y_i. This function is known as the likelihood

L and it's written like:

$$L(\theta, \phi; y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} \exp\left\{ \frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\}.$$

It is easier to work with the log-likelihood to obtain estimates of θ and ϕ that maximize the likelihood function,

$$logL(\theta, \phi; y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n} \left\{ \frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\} \quad (1.2.)$$

because the log-likelihood is maximized by the same values that maximize the likelihood (Hardin and Hilbe, 2012). To express the mean and variance $E(y_i)$ and $var(y_i)$ of random component y_i let

$$l_i = [\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)] \quad (1.3.)$$

which denotes the contribution of the i-th observation y_i to the log-likelihood function. The first and second derivatives can be obtained by differentiating Eq. (1.3.) with respect to θ_i:

$$\frac{\partial l_i}{\partial \theta_i} = \frac{[y_i - b^{'}(\theta_i)]}{a(\phi)}, \quad \frac{\partial^2 l_i}{\partial \theta_i^2} = \frac{-b^{''}(\theta_i)}{a(\phi)}$$

where $b^{'}(\theta_i)$ and $b^{''}(\theta_i)$ indicate the first and second derivatives of $b(\theta_i)$. Now we apply the general likelihood results

$$E(\frac{\partial L}{\partial \theta}) = 0, \text{ and} - E(\frac{\partial^2 L}{\partial \theta^2}) = E(\frac{\partial L}{\partial \theta})^2$$

which hold under regularity conditions satisfied by the exponential family. Thus, the first formula for the single observation gives,

$$E[\frac{y_i - b^{'}(\theta_i)}{a(\phi)}] = 0, \text{ so that } E(y_i) = \mu_i = b^{'}(\theta_i),$$

similarly from the second formula,

$$\begin{aligned}
\frac{b^{''}(\theta_i)}{a(\phi)} &= E\left[\frac{(y_i - b^{'}(\theta_i))}{a(\phi)}\right]^2 \\
&= \frac{var(y_i)}{[a(\phi)]^2} \\
var(y_i) &= a(\phi)\, b^{''}(\theta_i).
\end{aligned}$$

Thus, the function $b(.)$ evaluates the moments of y_i and called the cumulant function as its derivatives provide the cumulant of the distribution, when $a(\phi) = 1$ (Agresti, 2003).

1.1.3. Link Function

Each exponential family distribution has a set of compatible link functions that can be employed in a variety of situations. The identity link is used in the Gaussian family model, which assumes a normally distributed response variable. The identity link considers a continuous response that can have either positive or negative values. For modeling outcomes where the response can take values greater than or equal to

0 is accomplished through gamma family. This model is most commonly used with continuous response data, but it can also be utilized with count data that has a gamma distribution. The gamma model works with the reciprocal link for rate modeling and the log link for log-rate modeling. When modeling a non-negative response with a high initial peak, quick decline, and lengthy right tail, or modelling discrete data, the inverse Gaussian distribution is the best choice. The log and identity links which are similar to the gamma model are frequently utilized with such outcomes. Modeling discrete or proportional responses is done with the binomial-logit family, which includes the Bernoulli/binomial distributions. This family can be used to model the number of trials that result in a certain number of successes. Logit, probit, log-log, complementary log-log, identity, log, inverse, and log-complement are the commonly used links with this family. Logit and probit are two links that are frequently used with this family. To model the response that are counts or rate a Poisson family is used. Generally, a log link function is used with Poisson family of distribution. Count outcomes can also be modelled by using the negative-binomial family distribution, with overdispersed (compared to the Poisson) count data. Similar to the Poisson family a log link function is utilized with negative-binomial family (Deroche, 2015).

1.1.4. Canonical Link Function

The random component and the linear predictor are connected by a link function of GLMs. For a link function g, GLMs assert that a linear predictor $\eta_i = \sum_{j=1}^{p} \beta_j x_{ij}$ is related to the mean μ_i by $\eta_i = g(\mu_i)$, where $(\beta_1, ..., \beta_p)$ are the unknown parameters. The response function g^{-1} maps linear predictor values to the mean in the same way. The canonical link is the link function that converts the mean μ_i to the natural parameter θ_i. The natural parameter is related to the linear predictor by a direct relationship i.e. $\theta_i = \sum_{j=1}^{p} \beta_j x_{ij}$ (Agresti, 2015). Table 1.1., lists the most commonly used link functions and their inverses, while Table 1.2., shows the canonical link functions, range of the response variable, and variance functions for the most commonly used exponential families.

Table 1.1. Some Common Link Functions and Their Inverses

Link	$\eta_i = g(\mu_i)$	$\mu_i = g^{-1}(\eta_i)$
Identity	μ_i	η_i
Log	$\log_e \mu_i$	e^{η_i}
Inverse	μ_i^{-1}	η_i^{-1}
Inverse-square	μ_i^{-2}	$\eta_i^{-1/2}$
Square-root	$\sqrt{\mu_i}$	η_i^2
Logit	$\log_e \frac{\mu_i}{1-\mu_i}$	$\frac{1}{1+e^{-\eta_i}}$
Probit	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$

In Table 1.1., μ_i denotes the expected value of the response variable, η_i is the linear predictor, and Φ denotes the cumulative distribution function of the standard normal distribution.

Table 1.2. Canonical Link, Range of Response and Conditional Variance Function of Exponential Families

Family	Canonical Link	Range of Y_i	$var(Y_i\|\eta_i)$
Gaussian	Identity ($\eta_i = \mu_i$)	$(-\infty, +\infty)$	ϕ
Binomial	Logit ($\eta_i = log[\frac{\mu_i}{1-\mu_i}]$)	$0,1,\ldots,n_i$	$\mu_i(1-\mu_i)/n_i$
Poisson	Log ($\eta_i = log(\mu_i)$)	$0,1,2,\ldots$	μ_i
Gamma	Inverse ($\eta_i = \frac{1}{\mu_i}$)	$(0,\infty)$	$\phi\mu_i^2$
Inverse-Gaussian	Inverse-Square ($\eta_i = \frac{1}{\mu_i^2}$)	$(0,\infty)$	$\phi\mu_i^3$
Negative binomial	Log ($\eta_i = log(\mu_i)$)	$(0,\infty)$	$\mu_i + \alpha\mu_i^2$

In Table 1.2., n_i is the number of trials in the binomial family.

This introduction summarizes that GLMs are a class of standard linear models in which a linear predictor is mapped to the mean of response described by any member of the exponential family of distributions via a link function. However, when it is unreasonable to assume that data are normally distributed or when the response variable has a small set of possible outcomes, the typical linear model is not suitable. Moreover, the linear model is inappropriate in many cases where homoscedasticity is an essential condition. These extensions to the linear model are possible with the GLMs. The first step in creating a GLMs is to choose explanatory variables for the response variable of interest. A probability distribution from the exponential family is chosen and an appropriate link function is defined with the mapped range of values supporting the distribution's implied variance function (Hardin and Hilbe, 2012).

One of the primary goals of statistical modeling is to improve the estimation of parameters. Maximum likelihood (ML) methods, which are implemented by the iteratively reweighted least squares (IRLS) algorithm are commonly used for estimating the parameter in GLMs (Mackinnon and Puterman, 1989). The ML estimator has an approximate asymptotically normal distribution having mean β and variance-covariance matrix $(X^T WX)^{-1}$, where W is a diagonal matrix with elements $w_i = \frac{1}{var(y_i)}(\frac{\partial\mu_i}{\partial\eta_i})^2$ (Agresti, 2015).

Collinearity has always been considered as a potential source of problem in the computation, estimation, and interpretation of parameters in the linear models (Belsley et al., 1980). Segerstedt (1992) found that when estimating GLMs an ill-conditioned data matrix has similar effects on the parameter estimator as when estimating the linear regression model. Schaefer (1986) explored that collinearity among the explanatory variables has significant effects on the conditional ML estimator, since its variance is inflated in the same fashion that collinearity increases the variance of the least squares estimator in multiple regression. Thus, it is obvious that, the collinearity not only affects the estimation of parameters in normal linear models, but also in GLMs as well. Numerous studies have been conducted to handle the problem of collinearity in the linear regression and GLMs.

The objectives of this dissertation are: To introduce some biased estimators in GLMs by applying iterative approaches and imposing exact or stochastic restrictions on the parameters to address the problem of multicollinearity. To make comparisons of the proposed estimators with already existing estimators in the literature.

The remainder of the dissertation is organized as follows:

In Chapter 2, the iterative methods which are; Newton-Raphson and Fisher's scoring is given as well as the method of ML based on iterative methods for estimating the parameter β in GLMs is presented. In addition, an overview of the topic of multicollinearity, its effects, sources, and detection are briefly explained in the context of linear regression models and GLMs. The prior information in the form of exact linear restrictions and stochastic linear restrictions is also described. Some of the studies that have been done to address the issue of multicollinearity in linear regression and GLMs are also mentioned.

In Chapter 3, four newly developed restricted iterative biased estimators are given by using exact linear restrictions and stochastic linear restrictions on the parameters such as; the $r-k$ class estimator, stochastic restricted Liu estimator, restricted OK estimator, and stochastic restricted OK estimators. In addition, the properties of these estimators are also mentioned. Apart from the procedure for selecting tuning parameters is given.

In Chapter 4, different numerical illustrations are provided to figure out the performance of the newly developed estimators by a scalar mean square error (SMSE) criterion.

In Chapter 5, various simulation studies are conducted to make comparisons of the proposed estimators with already existing estimators by using the expected mean square error (EMSE) and prediction mean square error (PMSE) criteria.

In Chapter 6, concluding remarks are given based on the findings obtained in Chapters 4 and 5. The recommendations are also given at the end of this chapter.

2. ESTIMATION AND MULTICOLLINEARITY IN GLMS

This chapter presents the methods which are used to estimate the parameters of GLMs, such as the method of ML, which is commonly used to estimate the parameters in GLMs and it is implemented through iterative techniques like; Newton-Raphson method and the method of Fisher's scoring. Different convergence criteria that might be used in iterative methods, estimation of the dispersion parameter, the problem of multicollinearity its effects, sources, and its detection are briefly discussed. Prior information in the form of exact and stochastic linear restrictions is described as a potential solution to the multicollinearity problem. There are also some earlier studies that have been done to address the problem of multicollinearity.

2.1. Newton-Raphson Method

The Newton-Raphson method is used to iteratively solve nonlinear equations in order to find the maximum point of a function. It starts with a rough estimate of the solution. It gets a second approximation by using a second-degree polynomial to approximate the function in a neighborhood of the previous approximation and then finds the highest value of that polynomial. It then uses another second-degree polynomial to approximate the function in the vicinity of the second guess, with the third guess being the maximum location. The approach generates a series of estimates in this way. When the function is appropriate or the initial approximation is correct these converge to the maximum location. In further detail, here's how Newton-Raphson figures out the value $\hat{\beta}$ at which a function is maximized.
Let $u = \left(\frac{\partial L(\beta)}{\partial \beta_1}, \frac{\partial L(\beta)}{\partial \beta_2}, \ldots, \frac{\partial L(\beta)}{\partial \beta_p}\right)^T$. Let H stands for the Hessian matrix which has entries $h_{jk} = \frac{\partial^2 L(\beta)}{\partial \beta_j \beta_k}$. Let $u^{(t)}$ and $H^{(t)}$ denote u and H respectively computed at $\beta^{(t)}$, the guess t for $\hat{\beta}$. The terms up to the second order in the Taylor series expansion

approximates $L(\beta)$ near $\beta^{(t)}$ at step t in the iterative process $(t = 0, 1, 2, ...)$

$$L(\beta) \approx L(\beta^{(t)}) + u^{(t)^T}(\beta - \beta^{(T)}) + \frac{1}{2}(\beta - \beta^{(t)})^T H^{(t)}(\beta - \beta^{(t)}).$$

Solution of $\frac{\partial L(\beta)}{\partial \beta} \approx u^{(t)} + H^{(t)}(\beta - \beta^{(t)}) = 0$ for β gives the next approximation. This approximation might be expressed as,

$$\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{-1}u^{(t)},$$

$H^{(t)}$ is assumed to be nonsingular. Iterations continue until the differences in $L(\beta^{(t)})$ between subsequent cycles are small enough. The limit of $\beta^{(t)}$ is the ML estimator as $t \to \infty$, however this does not have to happen if $L(\beta)$ has other local maxima when $u(\beta) = 0$. A strong first approximation is necessary in this scenario (Agresti, 2003).

2.1.1. Fisher's Scoring Method

For solving likelihood equations Fisher's scoring is an alternative iterative method. The way it employs the Hessian matrix differs from Newton-Raphson. The expected information in Fisher's scoring is the expected value of this matrix, whereas the observed information used in Newton-Raphson is the Hessian matrix itself. Let $J^{(t)}$ be the ML estimate of the expected information matrix with approximation t that is $J^{(t)}$ has the elements $-E[\dfrac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_k}]$ computed at $\beta^{(t)}$. The Fisher's scoring formula is as follows:

$$\beta^{(t+1)} = \beta^{(t)} + (J^{(t)})^{-1}u^{(t)}$$

or

$$J^{(t)}\beta^{(t+1)} = J^{(t)}\beta^{(t)} + u^{(t)}. \tag{2.1.}$$

Agresti (2003) showed that $J = X^T WX$, where W is a diagonal matrix having elements $w_i = \frac{1}{var(y_i)}(\frac{\partial\mu_i}{\partial\eta_i})^2$. Likewise, $J^{(t)} = X^T W^{(t)} X$ and $W^{(t)}$ is a weight matrix W computed at $\beta^{(t)}$. The observed and expected information are the same for GLMs having the canonical link function.

2.1.2. Maximum Likelihood Estimation

Consider the independent random variables $Y_1,\ldots,Y_n$ having observations $y_1,...,y_n$ which satisfy the properties of GLMs. Let, $\eta_i = x_i^T\beta = \sum_{j=1}^{p} x_{ij}\beta_j$ is the predictor for the i-th observation of the response variable. Our objective is to estimate the parameters β that are associated to the y_i's by $E(y_i) = \mu_i$ and $g(\mu_i) = x_i^T\beta$. As we are interested in estimating the parameter β we consider the likelihood of the density function defined earlier in Chapter 1 given by Eq. (1.1.) and it is convenient to work with the log-likelihood to obtain estimates of β that maximize the likelihood function given by Eq. (1.2.). We write log-likelihood in Eq. (1.2.) as a function of β as we are interested in estimating β so

$$l(\beta) = \sum_{i=1}^{n} l_i(\beta).$$

The derivatives $\frac{\partial l_i(\beta)}{\partial\beta_j}$ can be find according to the chain rule:

$$\frac{\partial l_i(\beta)}{\partial\beta_j} = \sum_{i=1}^{n}\left[\frac{\partial l_i}{\partial\theta_i}\frac{\partial\theta_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\beta_j}\right].$$

The results of the partial derivatives are as follows:

$$\begin{aligned}\frac{\partial l_i}{\partial \theta_i} &= \frac{[y_i-\mu_i]}{a(\phi)}, \quad \mu_i = b^{'}(\theta_i)\\ \frac{\partial \mu_i}{\partial \theta_i} &= b^{''}(\theta_i) = \frac{var(y_i)}{a(\phi)}\\ \frac{\partial \eta_i}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j}\sum_{j=1}^{p} x_{ij}\beta_j = x_{ij}.\end{aligned}$$

After substituting the results of chain rule and simplifying we get,

$$\frac{\partial l_i(\beta)}{\partial \beta_j} = U_j = \sum_{i=1}^{n}\left[\frac{(y_i-\mu_i)x_{ij}}{var(y_i)}\frac{\partial \mu_i}{\partial \eta_i}\right]. \qquad (2.2.)$$

The matrix form of Eq. (2.2.) can be written like $U = X^T WD(y-\mu)$, where $W = diag(w_{ii})$, $w_{ii} = \frac{1}{var(y_i)}(\frac{\partial \mu_i}{\partial \eta_i})^2$ and $D = diag(\frac{\partial \eta_i}{\partial \mu_i})$. Similar to Eq. (2.2.), the second-order derivative with respect to component β_k we have:

$$\begin{aligned}E(\frac{\partial^2 l_i}{\partial \beta_j \partial \beta k}) &= -E(\frac{\partial l_i}{\partial \beta_j})(\frac{\partial l_i}{\partial \beta_k})\\ &= -E\sum_{i=1}^{n}\left[\frac{(y_i-\mu_i)(y_i-\mu_i)x_{ij}x_{ik}}{(var(y_i))^2}(\frac{\partial \mu_i}{\partial \eta_i})^2\right]\\ &= -\sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{var(y_i)}(\frac{\partial \mu_i}{\partial \eta_i})^2\\ E(-\frac{\partial^2 l_i}{\partial \beta_j \partial \beta k}) &= J_{jk} = \sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{var(y_i)}(\frac{\partial \mu_i}{\partial \eta_i})^2. \qquad (2.3.)\end{aligned}$$

The matrix notation of Eq. (2.3.) can be written as: $J = X^T WX$, where W is defined as above. As, the likelihood equations are nonlinear functions of β, an iterative method of Fisher's scoring is used to find the estimate of β. The equation of Fisher's scoring method is given as:

$$\beta^{(t+1)} = \beta^{(t)} + \left\{ [J^{(t)}]|_{\beta=\beta^{(t)}} \right\}^{-1} U^{(t)}|_{\beta=\beta^{(t)}}. \tag{2.4.}$$

Computation can be carried out by replacing β by $\beta^{(t)}$ (Khuri, 2009) thus, after substitute the values in Eq. (2.4.) we obtained:

$$X^T W^{(t)} X \beta^{(t+1)} = X^T W^{(t)} z^{(t)} \tag{2.5.}$$

which is the likelihood equation of a GLMs, where z is the $n \times 1$ working response consists of the elements: $z_i^{(t)} = \sum_{j=1}^{p} x_{ij} \beta_j^{(t)} + D^{(t)}(y_i - \mu_i^{(t)})$. When $t \to \infty$, we get the ML estimate that is

$$\hat{\beta}^{(t+1)} = (X^T W^{(t)} X)^{-1} X^T W^{(t)} z^{(t)}$$

having the asymptotic variance-covariance $(X^T \hat{W} X)^{-1}$ where $\hat{W}$ is evaluated at $\hat{\beta}^{(t)}$. When a solution is obtained then, $\hat{\beta}$ is a asymptotically normal as well as efficient and consistent estimator of β. Thus $\hat{\beta} \sim N(\beta, (X^T \hat{W} X)^{-1})$ (Rao and Toutenburg, 1995).

2.1.3. Convergence Criteria

In the literature there are several convergence criteria. For the Newton-Raphson method the convergence of $\beta^{(t)}$ toward $\hat{\beta}$ is generally fast. For any j, the convergence satisfies when t is large,

$$\left|\beta_j^{(t+1)} - \hat{\beta}_j\right| \leq c\left|\beta_j^{(t)} - \hat{\beta}_j\right|^2,$$

where $c > 0$ and considered as a second-order. This means that after a certain number of iterations the number of correct decimals in the approximation generally doubles. In practice, good convergence frequently requires only a few iterations (Agresti, 2003).

For the least squares estimation of parameters in nonlinear regression by a Gauss-Newton iterative method (Myers et al., 2010) recommended the following convergence criterion:

$$\left|\frac{\beta_{j,t+1} - \beta_{jt}}{\beta_{jt}}\right| < \delta, \quad j = 1, ..., p$$

where δ is any small number like 10^{-6} and t denotes the iteration.

Another convergence criterion given by (Fahrmeir and Tutz, 1994) is defined as

$$\frac{||\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}||}{||\hat{\beta}^{(t)}||} < \varepsilon$$

where ε is assumed to be pre-selected small number.

For the iterative methods, (Özkale, 2016) used the following convergence criterion

$$||\hat{\beta}^{(t)} - \hat{\beta}^{(t-1)}|| \leq 1 \times 10^{-6}$$

where $||.||$ denotes the norm of a vector.

2.2. Multicollinearity in GLMs

This section provides a brief review of multicollinearity, its effects, origins, and detection in the context of linear regression models (LRMs) and GLMs.

2.2.1. Background

The full rank condition in a multiple linear regression model suggests that the explanatory variables are independent of one another. When this assumption is broken, a phenomenon known as multicollinearity emerges (Das, 2019). This phenomenon is described by using a variety of terminologies such as collinearity, multicollinearity, and ill conditioning (Blesly et al., 1980). The term multicollinearity was first introduced by (Frisch, 1934).

2.2.2. Definition

There are various definitions of multicollinearity in the literature, a few of them are mentioned below.

Farrar and Glauber (1964) stated that multicollinearity is viewed as interdependence condition or it is defined by a lack of independence or the presence of interdependence, which is indicated by high intercorrelations $(X^T X)$ within a set of variables.

Gunst (1983) defined multicollinearity as a situation in which two or more explanatory variables are highly correlated.

The term "multicollinearity" refers to a high degree of correlation between multiple explanatory variables. In a multiple linear regression model multicollinearity occurs when two or more explanatory variables are significantly correlated (Das, 2019).

2.2.3. Effects of Multicollinearity

Collinearity has long been recognized as a potential source of issues in linear model parameter estimate, computation, and interpretation (Mackinnon and Puterman, 1989). When multicollinearity presents, the variances of coefficients are inflated. Multicollinearity diminishes the precision of the estimate, which weakens the statistical power of the regression model, and the p-values fail to identify the explanatory variables that are statistically significant. As a result, rejecting the null hypothesis and interpreting the regression coefficients becomes more difficult. It decreases the model's ability to find statistically significant explanatory variables. Because of multicollinearity determining the effect of a single explanatory variable on the response variable is difficult. This is because the predicted regression coefficient of any one variable is dependent on the model's other factors. In the presence of multicollinearity the matrix $X^T X$ becomes ill-conditioned. The ordinary least squares (OLS) estimator remains unbiased in the presence of multicollinearity, but its sampling variance increases dramatically. As a result, the OLS estimator loses precision and the BLUE property is no longer valid (Das, 2019).

Multicollinearity is a problem in GLMs as well and it affects parameter estimation in the same way as it does in traditional LRMs. Schaffer (1986) shown that collinearity among explanatory variables has significant effects on the conditional ML estimator as the variance of this estimator is inflated in the same way that collinearity inflates the variance of the least squares estimator in multiple regression. Segerstedt (1992) revealed that when estimating GLMs an ill-conditioned data matrix has similar consequences on the parameter estimator as when estimating LRMs. In addition, as the conditioning of the covariance matrix deteriorates, the average length of the ML estimator of a parameter vector increases asymptotically. Weissfeld and Sereika (1991) described that, the inverse of observed information matrix is often difficult or impossible due to collinearity, causing the parameter estimates to become unstable and increasing the estimated variances. Lesaffre and Marx (1993) investigated the ML-collinearity in GLMs and reported the same detrimental effects on the covariance matrix such as; inflating the standard error of the certain coefficients.

2.2.4. Sources of Multicollinearity

Multicollinearities can occur due to a number of factors. Montgomery et al. (2012) described four main factors of multicollinearity, that are:

1. The method of data collecting used
2. Constraints in the population or on the model
3. Specification of the model
4. An over defined model

2.2.5. Detection of Multicollinearity in GLMs

A few studies have been undertaken in the context of GLMs to identify the problem of multicollinearity and they are as follows: Schaefer et al. (1984) considered three criteria for collinearity diagnosis i) the value of R^2 coefficient of determination, while regressing one variable on others ii) the residual sum of squared in (i) and iii) the smallest eigenvalue of X^TWX matrix. It is assumed that multicollinearity is present, when R^2 approaches to 1, the residual sum of square approaches to 0, and smallest eigenvalue approaches to 0.

Lesaffre and Marx (1993) deemed the condition number (CN) to be the most essential instrument for detecting collinearity and recommended a threshold value of $\kappa_X > 30$ to indicate the presence of multicollinearity, where κ_X is defined as a condition number of $X^T X$.

The singular value decomposition (SVD) technique has been used by some of the authors to establish collinearity diagnostics such as; (Mackinon and Puterman, 1989), and (Weissfeld and Sereika, 1991) advised a condition number of $X^*(X^*)^T$ as a diagnostic measure for the existence of multicollinearity, where $X^* = \hat{W}^{1/2}X$ and the CN is defined as $\varsigma_{\max} = (\frac{\lambda_{\max}}{\lambda_{\min}})^{1/2}$. Mackinon and Puterman (1989) suggested that if the CN excceds 10, then multicollinearity is present. In addition, (Weissfeld and Sereika, 1991) proposed variance decomposition proportion to identify the problem of multicollinearity, which is defined as $\pi_{ij} = \frac{\phi_{ij}}{\phi_i}$, $i,\ j = 1,\ldots,p$ and recommended that if $\pi_{ij} > 0.5$, then linear dependencies might exist among explanatory variables. Huang et al. (2016) developed a weighted VIF (WVIF) to detect the problem of multicollinearity and it is defined as the diagonal elements of the matrix $(X^T L^T LX)^{-1}$, where $L^T L$ is a matrix such as $\hat{W} = L^T L$, the value of WVIF lies in the interval (0, ∞). Smith and Marx (1990) suggested that center and scale the matrix $X^* = \hat{W}^{1/2}X$ and VIFs of the diagonal elements of $(X^{*^T} X^*)^{-1}$ matrix can be used as a diagnostic criterion for multicollinearity in GLMs. Özkale (2021) also provided various measures to identify the problem of multicollinearity in GLMs, including VIF, Corrected VIF (CVIF), modfified CVIF (MCVIF), and red indicator.

2.3. Utilization of Prior Information

This section gives an overview of prior information in the form of exact linear restrictions and stochastic linear restrictions.

Many attempts have been made in statistical research to develop improved estimators. The use of extraneous information or prior information is one way to improve the estimators. Such prior information regarding the regression coefficients may be available in applied work. Constant returns to scale, for instance, imply that the exponents of a Cobb-Douglas production function should sum to one. In another example, if customers have no money illusion, the total of their money income and price elasticities in a demand function should be zero (Shalabh, 2013). These restrictions or previous information may be derived from:

i) theoretical consideration.

ii) the experimenter's previous experience.

iii) empirical studies.

iv) a few extraneous sources, and so on.

To use such information to improve the estimation of parameters, it can be expressed in the form of:

i) exact linear restrictions

ii) stochastic linear restrictions

iii) inequality restrictions

2.3.1. Exact Linear Restrictions

Assume that the prior knowledge or information that binds the regression coefficients comes from some external sources and can be presented in the form of exact linear constraints as follows (Shalabh, 2013):

$$r = R\beta$$

where R is a $(q \times k)$ matrix having rank(R)=$q(q < k)$, and r is a $(q \times 1)$ vector. It is assumed that r and R are known.

2.3.2. Stochastic Linear Restrictions

The exact linear constraints presume that the auxiliary or prior information contains no randomness. In many practical scenarios this assumption may not hold true and some randomness may be present. In such circumstances the prior information can be expressed as

$$r = R\beta + G, \quad G \sim (0, \quad \sigma^2\Sigma) \tag{2.6.}$$

where G is a $(q \times 1)$ vector of random errors, while vector r and matrix R are defined above in subsection 2.3.1., and are assumed to be known. The randomness involved in the prior information $r = R\beta$ is reflected by the term G (Shalabh, 2013). Rao and Toutenburg (1995) described that, Σ is likely to be known and supposed $\Sigma > 0$ and, so is regular. The vector r can be considered as a random variable with $E(r) = R\beta$. As a result, except in the mean, the constraint does not hold exactly in Eq. (2.6.). As, r is supposed to be known (that it is a realized value of the random vector) and that all expectations are conditional on r such as $E(\hat{\beta}|r)$. The following are some of the possible reasons that might be for such a stochastic linear constraint:
i) Estimation instability is demonstrated with stochastic linear constraints. The standard error of an unbiased estimate may show stability.
ii) When the accuracy of an exact linear restriction such as $r = R\beta$ is ambiguous, an element of uncertainty is introduced. For example, one could say that 95% of the restrictions are hold. As a result, there is a degree of uncertainty (Shalabh, 2013).

Several studies have been carried out to alleviate the problem of multicollinearity and improve parameter estimation in LRMs by using prior information in addition to sample data. Some of them are as follows:

Theil and Goldberger (1961), Theil (1961), Groß (2003) and Zhong and Yang (2007), Fallaha et al. (2017), Li and Yang (2010), Özkale and Kaçıranlar (2007) have acquired a two parameter estimator under exact linear restrictions in linear regression.

To avoid the problem of multicollinearity in GLMs some of these ideas have been applied to GLMs such as; Nyquist (1991) provided a restricted ML approximation of the GLMs through exact linear restrictions on the parameters to encounter the problem of multicollinearity. Segerstedt (1992) provided a ridge estimator in GLMs to deal with multicollinearity. Özkale and Nyquist (2021) have developed a stochastic restricted ridge estimator by applying stochastic restrictions on the parameters against the multicollinearity. Kurtoğlu and Özkale (2019a) and Kurtoğlu and Özkale (2019b) have developed an iterative restricted ridge and iterative restricted Liu estimators, respectively under exact linear restrictions to cope with multicollinearity in GLMs.

Using the ML estimator obtained at the final iteration, certain researches have been conducted using exact or stochastic restrictions for logistic regression. These are: Varathan and Wijekoon (2019), Varathan and Wijekoon (2015), Asar et al. (2017), Zuo and Li (2018).

Furthermore, on the basis of ML estimator in the final iteration, two parameter estimators related to specific cases of GLMs have been constructed to address the problem of multicollinearity. These are the following: Huang and Yang (2014), Amin et al. (2019), Asar and Genç (2017), Asar and Genç (2018). In addition, Asar et al. (2017) and Varathan and Wijekoon (2018) have established a restricted two parameter Liu type estimator in the logistic regression under exact linear restrictions and two parameter Liu-type logistic estimator under stochastic linear restrictions.

2.3.3. Estimation of Dispersion Parameter

The ML estimator of β follows the normal distribution, according to asymptotic theory: $\hat{\beta} \sim N(\beta, a(\phi)(X^TWX)^{-1})$, where $a(\phi)(X^TWX)^{-1}$ is the Fisher's information matrix. The diagonal elements of the inverse Fisher's information matrix which is a function of the dispersion parameter, represent the associated variance. It's worth noting that estimating the regression coefficients in GLMs doesn't necessitate a dispersion parameter estimate. However, the effect of the dispersion parameter is apparent, when we make any kind of statistical inference about the model parameter. Since the dispersion parameter is included in the variance-covariance of the model parameter, $var(\hat{\beta}) = a(\phi)(X^TWX)^{-1}$ apart from $var(y_i) = a(\phi)var(\mu_i)$ as a result, the dispersion parameter influences the variance of the random variable (Du, 2007). Although estimation by ML is theoretically conceivable, it is rarely used (Fox, 2015). Nevertheless, when the sample size n or the total Fisher's information is small, ML estimates are known to be biased (Cordeiro and McCullagh, 1991). The method of moment estimator is most commonly used to estimate the dispersion parameter and it is defined as:

$$\hat{\phi} = \frac{1}{n-p}\sum_{i=1}^{n}\left[\frac{(y_i - \hat{\mu}_i)}{\hat{\mu}_i}\right]^2 = \frac{\chi^2}{n-p}$$

where p denotes the number of parameters, n is the sample size, and χ^2 represents the Pearson's Chi-Square statistic (McCullagh and Nelder, 1989). Another estimate of the dispersion parameter $a(\phi)$ for the gamma distribution is based on the ML estimator and it is defined as

$$\hat{\phi}_{ML} = \frac{\bar{D}(6+\bar{D})}{6+2\bar{D}}$$

where $\bar{D} = D(y;\hat{\mu})/n$ and $D(y;\hat{\mu})$ is the deviance of the model (Du, 2007, McCullagh and Nelder, 1989).

2.4. Method for Selecting Principal Components

In this section the method for determining the number of principal components is described.

As it is well known that for selecting the number of principal components (PCs), numerous strategies have been proposed in the literature. Among those one of the most commonly used strategy is the percentage of total variation (PTV), which we used in numerical example given in the section 4.1.1., and is defined as:

$$PTV = \frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{p} \lambda_i} \times 100$$

where r denotes the number of PCs.

2.5. Some Biased Estimators in GLMs

In this section we described some of the biased estimators in GLMs that have been developed in the literature to address the problem of multicollinearity.

2.5.1. Ridge Estimation in GLMs

Similar to linear regression (Segerstedt, 1992) has established a ridge estimator in GLMs to tackle the multicollinearity problem by considering the objective function

$$p(\beta, k) = \beta^T \beta - \frac{2}{k}\left[\sum_{i=1}^{n} \ln c(y_i) + \sum_{i=1}^{n}\{\theta_i y_i - b(\theta_i)\} + (\varphi_0 - l_{\max})\right]$$

where $\frac{2}{k}$ is a lagrange multiplier and $\varphi_0 = l_{\max} - l(\beta)$ is the difference between values of the log-likelihood function evaluated at the ML estimate. He acquired the ridge estimator as:

$$(X^T W^{(m)} X + kI_p)\hat{\beta}_k^{(m+1)} = X^T W^{(m)} z^{(m)}$$

2.5.2. Principal Component Regression in GLMs

Smith and Marx (1990) provided a principal component regression (PCR) estimator as an alternative to ML estimator in the presence of multicollinearity. By considering the GLMs

$$g(\mu_i) = x_i^T \beta = \eta_i$$

the PCs for each observation are defined as

$$Z = XM$$

where the $(i,\ j)$th elements of Z is the score of the j-th principal component for the i-th observation. Since $X\beta = XMM^T\beta$ and define $\alpha = M^T\beta$, which leads to the transformed full PCs model as

$$\eta_i = z_i^T\alpha$$

where z_i^T is row vector of the Z matrix. The reduced principal components can also be useful and is defined as

$$\eta_{i,s} = z_{i,s}^T\alpha_s$$

where $z_{i,s}^T$ is the subset row vector of Z_s, a $(p+1\times s)$ subset matrix of Z. α_s is a subset vector of α associated with high information. Thus, an iterative PCs estimator can be acquired in the form

$$\hat{\beta}_{t,s}^{pc} = \hat{\beta}_{t-1,s}^{pc} + \sum_{j=0}^{s-1}\tilde{\lambda}_j^{-1}\tilde{m}_j\tilde{m}_j\left[\sum_{i=1}^{N}x_i\hat{k}_{ii}^{-1}\hat{e}_{i,s}\frac{\partial\eta_i}{\partial\mu_i}\right]_{t-1}$$

where $\tilde{\lambda}_j$ and $\tilde{m}_j$ are corresponding eigenvalues and eigenvectors of ML estimate of information $X^T\tilde{K}^{-1}X$ and $\hat{e}_{i,s} = y_i - h(\tilde{\eta}_{i,s})$. Note that $\lambda_s, \lambda_{s+1,} \ldots, \lambda_p$ are usually the $r = p+1-s$ sufficiently small eigenvalues. Alternatively $\hat{\beta}_{t,s}^{pc}$ can also be expressed in terms of $\hat{\alpha}^{pc}$

$$\begin{aligned}
\hat{\alpha}^{pc}_{t,s} &= \hat{\alpha}^{pc}_{t-1,s} + \tilde{\Lambda}_s^{-1}\left[\sum_{i=1}^{N} z_{i,s}\hat{k}_{ii}^{-1}\hat{e}_{i,s}\frac{\partial\eta_i}{\partial\mu_i}\right]_{t-1} \\
&= \tilde{\Lambda}_s^{-1} Z_s^T \tilde{K}^{-1} y^{*pc}_{t-c},
\end{aligned}$$

where $z_{i,s}$ are the columns of Z_s^T and $y^{*pc}_{t-c} = \eta_{i,s} + \hat{e}_{i,s}\frac{\partial\eta_i}{\partial\mu_i}$ evaluated at $(t-1)$ iteration. Naturally PCs estimator can be constructed based on the convergence of $\hat{\alpha}^{pc}_{t,s}$ as

$$\hat{\beta}^{pc}_{t,s} = \tilde{M}_s\hat{\alpha}^{pc}_s.$$

2.5.3. Mixed Estimator in GLMs

The mixed estimator in GLMs was introduced by (Özkal and Nyquist, 2021) by imposing stochastic restrictions on the parameters in addition to the sample information to deal with multicollinearity. To acquired the mixed estimator they used the following log-likelihood function

$$\begin{aligned}
l(\theta; y, r) &= \sum_{i=1}^{n}\left\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i,\phi)\right\} - \frac{1}{2}\ln((2\pi)^q \det(\Sigma)) \\
&\quad -\frac{1}{2}(r - R\beta)^T\Sigma^{-1}(r - R\beta),
\end{aligned}$$

assuming that the random vector y and r are independent. Thus, they acquired the iterative mixed estimator in GLMs as

$$\begin{aligned}
\hat{\beta}_c^{(m+1)} &= (X^T W^{(m)} X)^{-1} X^T W^{(m)} z^{(m)} + (X^T W^{(m)} X)^{-1} \\
&\quad \times R^T\left\{\Sigma + R(X^T W^{(m)} X)^{-1} R^T\right\}^{-1} \\
&\quad \times\left\{r - R(X^T W^{(m)} X)^{-1} X^T W^{(m)} z^{(m)}\right\}.
\end{aligned}$$

This is called the mixed estimator in GLMs which carries the idea of mixed estimator in LRMs.

2.5.4. Restricted ML Estimator in GLMs

Nyquist (1991) proposed restricted ML estimator in GLMs when there is a set of q linear and independent restrictions on the parameters along with the sample information like $R_j^T \beta = r_j$, $j = 1,\dots,q$, where R_j, $j = 1,\dots,q$ are $p \times 1$ vector and r_j $j = 1,\dots,q$ are the scalars. To get the restricted ML estimator in GLMs he considered the following penalty function

$$P(\beta,\lambda) = \sum_{i=1}^{n} \ln c(y_i) + \sum_{i=1}^{n} \{\theta_i y_i - b(\theta_i)\} - \frac{1}{2} \sum_{j=1}^{q} \lambda_j (r_j - R_j^T \beta)^2$$

and find a solution to the unconditional problem $\max_\beta P(\beta,\lambda)$ for fixed and positive values of λ_j, $j = 1,\dots,q$ which yields the restricted iterative ML estimator as

$$\hat{\beta}_r^{(m+1)} = \hat{\beta}^{(m+1)} + \left\{ (X^T W^{(m)} X)^{-1} R^T [R(X^T W^{(m)} X)^{-1}] \right\} (r - R\hat{\beta}^{(m+1)}),$$

where $\hat{\beta}^{(m+1)}$ is the ML estimator.

When multicollinearity is established on the basis of aforementioned criteria, there is no simple solution. However, while dealing with extreme multicollinearity, one of the best options is to utilize biased estimators that are less affected by multicollinearity (Thiart, 1990). A lot of biased estimators have been proposed to address the problem of multicollinearity in the context of LRMs. The most popular are the ridge estimator proposed by (Hoerl and Kennard, 1970) and the Liu estimator proposed by (Liu, 1993) in the linear regression. According to (Toutenburg, 1982) and (Rao and Toutenburg, 1995), there are several strategies to avoid multicollinearity such as: i) use an experimental design strategy to reduce the correlation between explanatory variables values ii) modify the variables (use, for example, first difference, etc.) iii) by changing the list of explanatory variables iv) by including previous knowledge in the form of exact or stochastic linear restrictions. We'll focus on technique (iv) because it's related to prior knowledge.

3. RESTRICTED BIASED ESTIMATION METHODS IN GLMS

This chapter introduces four novel iterative restricted biased estimators under exact and stochastic linear restrictions. These are; the $r-k$ class estimator, stochastic restricted Liu estimator, restricted OK estimator, and stochastic restricted OK estimator. The properties of these estimators are discussed in addition to their derivation. Aside from that the methods for selecting the tuning parameters are discussed.

3.1. The $r-k$ Class Estimator in GLMs

In this section, we develop an iterative restricted $r-k$ class estimator, as it is a combination of ridge and principal component regression (PCR) estimators. Apart from it is obvious that when applying principal components (PCs) the smaller set of components is used thus, the $r-k$ class estimator belongs to the class of restricted estimators. Furthermore, in linear regression (Johnson et al., 1973) have shown that the PCR estimators are equivalent to the restricted least squares estimator. Similar arguments can be found in (Fomby and Hill, 1978) and (Thiart, 1990).

Here, we utilize the SVD which was used by (Smith and Marx, 1990) for proposing the PCR in GLMs. The linear predictor $\eta = X\beta$ can be expressed as $\eta = XMM^{T}\beta = V\alpha$, where $V = XM$, $\alpha = M^{T}\beta$ and $M = [m_1, \ldots, m_q]$ is a $q \times q$ orthogonal matrix through $V^T \hat{W}_{ML} V = \Lambda = diag(\lambda_j)$ is a $q \times q$ diagonal matrix of the eigenvalues of $X^T \hat{W}_{ML} X$ $\left(\lambda_1 = \lambda_{\max} \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_q = \lambda_{\min}\right)$ and m_j are the corresponding eigenvectors accompanied by their eigenvalues λ_j. The linear predictor for the i-th component of the V matrix can be written as $\eta_i = x_i^T MM^{T}\beta = v_i^T\alpha$ where, $v_i^T = x_i^T M$ is the i-th row vector of the V matrix. To overcome the problem of multicollinearity sometimes it is worthwhile to use the reduced set of PCs. Therefore, V matrix and the α vector can be divided into separate parts such as $V = \begin{bmatrix} V_r & V_{q-r} \end{bmatrix}$ and

$\alpha = \begin{bmatrix} \alpha_r^T & \alpha_{q-r}^T \end{bmatrix}$ where $V_r = XM_r$ $(r \leqslant q)$ consisting of the PCs having large eigenvalues so that it will be retained in the model. The M and Λ can also similarly be partitioned into separate parts that are $M = \begin{bmatrix} M_r & M_{q-r} \end{bmatrix}$ and $\Lambda = \begin{bmatrix} \Lambda_r & 0 \\ 0 & \Lambda_{q-r} \end{bmatrix}$, where $\Lambda_r = V_r^T \hat{W}_{ML} V_r = M_r^T X^T \hat{W}_{ML} X M_r$ and $\Lambda_{q-r} = V_{q-r}^T \hat{W}_{ML} V_{q-r} = M_{q-r}^T X^T \hat{W}_{ML} X M_{q-r}$. Thus, the linear predictor for the reduced PCs can be defined as $\eta_{r,i} = v_{r,i}^T \alpha_r$, where $v_{r,i}^T$ is the i-th row vector of the V_r matrix.

In order to obtain the $r-k$ class estimator in GLMs, we maximize the following objective function

$$F(\alpha_r, k) = \underset{\alpha_r}{Max}\{\frac{2}{k} l(\alpha_r) - \alpha_r^T \alpha_r\}$$

where $l(\alpha_r) = \sum_{i=1}^{n} l_i$ is the log-likelihood function of the reduced model.

To find the parameter α_r, we differentiate the objective function $F(\alpha_r, k)$ with respect to parameter $\alpha_{r,j}$, which are the elements of α_r. Then we have:

$$\begin{aligned} \frac{\partial F(\alpha_r, k)}{\partial \alpha_{r,j}} &= \frac{2}{k} \sum_{i=1}^{n} \left[\frac{\partial l_i}{\partial \alpha_{r,j}} - \frac{\partial(\alpha_r^T \alpha_r)}{\partial \alpha_{r,j}} \right] \\ &= \frac{2}{k} \sum_{i=1}^{n} \left[\frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \alpha_{r,j}} - 2\alpha_{r,j} \right] \\ &= \frac{2}{k} \frac{1}{a(\phi)} \sum_{i=1}^{n} \left[\frac{y_i - \mu_i}{var(y_i)} v_{r,ij} \frac{\partial \mu_i}{\partial \eta_{r,i}} - 2\alpha_{r,j} \right] \end{aligned} \tag{3.1.}$$

In matrix notation Eq. (3.1.) can be written as

$$\frac{\partial F(\alpha_r,k)}{\partial \alpha_{r,j}} = U(\alpha_r,k) = \frac{2}{k}\frac{1}{a(\phi)}V_r^T \hat{W}_{ML} D^{-1}(y-\mu) - 2\alpha_r$$

where $D = diag\left(\frac{\partial \mu_i}{\partial \eta_i}\right) = diag(\frac{1}{g'(\mu_i)})$. Differentiate Eq. (3.1.) w.r.t. $\alpha_{r,u}$ we get

$$\frac{\partial^2 F(\alpha_r,k)}{\partial \alpha_{r,j}\partial \alpha_{r,u}} = \frac{-2}{k}\frac{1}{a(\phi)}\sum_{i=1}^{n}\left[(\frac{y_i-\mu_i}{var(y_i)}v_{r,ij}\frac{\partial \mu_i}{\partial \eta_{r,i}})(\frac{y_i-\mu_i}{var(y_i)}v_{r,iu}\frac{\partial \mu_i}{\partial \eta_{r,i}}) - 2\delta_{ju}\right] \tag{3.2.}$$

where $\delta_{ju} = 1$ if $j = u$ and 0 otherwise. Taking the expectation of Eq. (3.2.) we have the following form

$$E[\frac{\partial^2 F(\alpha_r,k)}{\partial \alpha_{r,j}\partial \alpha_{r,u}}] = \frac{-2}{k}\frac{1}{a(\phi)}\sum_{i=1}^{n}\frac{v_{r,ij}v_{r,iu}}{var(y_i)}(\frac{\partial \mu_i}{\partial \eta_i})^2 - 2. \tag{3.3.}$$

Then, by virtue of Eq. (3.3.) we get

$$Q(\alpha_r,k) = -E[\frac{\partial^2 F(\alpha_r,k)}{\partial \alpha_{r,j}\partial \alpha_{r,u}}] = 2\left(\frac{1}{k}\frac{1}{a(\phi)}V_r^T \hat{W}_{ML} V_r + I_r\right).$$

By applying the Fisher's scoring method we get the estimator as

$$\hat{\alpha}_{rk}^{(t+1)} = \hat{\alpha}_{rk}^{(t)} + \{(Q(\alpha_r,k))^{-1}(U(\alpha_r,k))^{(t)}\}$$

On making substitution we have

$$\begin{aligned}
\hat{\alpha}_{rk}^{(t+1)} &= \hat{\alpha}_{rk}^{(t)} + \frac{1}{2}(\frac{1}{k}\frac{1}{a(\phi)}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}(\frac{2}{k}\frac{1}{a(\phi)}V_r^T\hat{W}_{ML}(D^{(t)})^{-1} \\
&\quad \times(y-\mu^{(t)}) - 2\hat{\alpha}_{rk}^{(t)}) \\
&= \hat{\alpha}_{rk}^{(t)} + (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}[\frac{1}{k}V_r^T\hat{W}_{ML}(D^{(t)})^{-1}(y-\mu^{(t)}) - \hat{\alpha}_{rk}^{(t)}] \\
&= \left[I - (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}\right]\hat{\alpha}_{rk}^{(t)} + (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1} \\
&\quad \times(\frac{1}{k}V_r^T\hat{W}_{ML}(D^{(t)})^{-1}(y-\mu^{(t)}).
\end{aligned}$$

Using the fact $(\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}(\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r) = I$ which gives

$$I - (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1} = (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}\frac{1}{k}V_r^T\hat{W}_{ML}V_r,$$

we obtain

$$\begin{aligned}
\hat{\alpha}_{rk}^{(t+1)} &= (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1}\frac{1}{k}V_r^T\hat{W}_{ML}V_r\hat{\alpha}_{rk}^{(t)} + (\frac{1}{k}V_r^T\hat{W}_{ML}V_r + I_r)^{-1} \\
&\quad \times\frac{1}{k}V_r^T\hat{W}_{ML}(D^{(t)})^{-1}(y-\mu^{(t)}) \\
&= (V_r^T\hat{W}_{ML}V_r + kI_r)^{-1}\left[V_r^T\hat{W}_{ML}V_r\hat{\alpha}_{rk}^{(t)} + V_r^T\hat{W}_{ML}(D^{(t)})^{-1}(y-\mu^{(t)})\right] \\
&= (V_r^T\hat{W}_{ML}V_r + kI_r)^{-1}V_r^T\hat{W}_{ML}\left[V_r\hat{\alpha}_{rk}^{(t)} + (D^{(t)})^{-1}(y-\mu^{(t)})\right]
\end{aligned}$$

where D and μ are evaluated at $\hat{\alpha}_{rk}^{(t)}$. Converting back to the original parameters given as $\hat{\beta}_{rk}^{(t+1)} = M_r \hat{\alpha}_{rk}^{(t+1)}$, we obtain the iterative $r-k$ class estimator in GLMs as

$$\begin{aligned}
\hat{\beta}_{rk}^{(t+1)} &= Mr(M_r^T X^T \hat{W}_{ML} X M_r + kI_r)^{-1} M_r^T X^T \hat{W}_{ML} \\
&\quad \times \left(X M_r M_r^T \hat{\beta}_{rk}^{(t)} + \left(D^{(t)}\right)^{-1} \left(y - \mu^{(t)}\right)\right) \\
&= M_r \left(M_r^T X^T \hat{W}_{ML} X M_r + kI_r\right)^{-1} M_r^T X^T \hat{W}_{ML} z_{rk}^{(t)} \qquad (3.4.)
\end{aligned}$$

where $z_{rk}^{(t)} = X M_r M_r^T \hat{\beta}_{rk}^{(t)} + (D^{(t)})^{-1}(y - \mu^{(t)})$. The iteration proceeds until the difference between successive values of the estimates is sufficiently small. The inequality $||\hat{\beta}_{rk}^{(t+1)} - \hat{\beta}_{rk}^{(t)}|| \leq \delta$ can be selected as a convergence criterion where δ is some small number, say 1×10^{-6}. Let us denote the $r-k$ class estimator in GLMs as $\hat{\beta}_{rk} = M_r \left(M_r^T X^T \hat{W}_{ML} X M_r + kI_r\right)^{-1} M_r^T X^T \hat{W}_{ML} \hat{z}_{rk}$. When, $k = 0$, $\hat{\beta}_{rk}^{(t+1)}$ in Eq. (3.4.) gives the iterative PCR estimator given by (Smith and Marx, 1990) as

$$\hat{\beta}_{r}^{(t+1)} = M_r \left(M_r^T X^T \hat{W}_{ML} X M_r\right)^{-1} M_r^T X^T \hat{W}_{ML} z_r^{(t)},$$

where $z_r^{(t)} = X M_r M_r^T \hat{\beta}_r^{(t)} + (D^{(t)})^{-1}(y - \mu^{(t)})$. The first-order approximated (FOA) form of Eq. (3.4.) is

$$\hat{\beta}_{rk}^{(1)} = M_r (M_r^T X^T \hat{W}^{(0)} X M_r + kI_r)^{-1} M_r^T X^T \hat{W}^{(0)} z_r^{(0)} \qquad (3.5.)$$

where $z_r^{(0)} = X M_r M_r^T \beta^{(0)} + (D^{(0)})^{-1}(y - \mu^{(0)})$, with $\beta^{(0)}$ as the initial value of the regression coefficients and $D^{(0)}$ and $\mu^{(0)}$ are evaluated at $\beta^{(0)}$ [1]. We call the estimator in Eq. (3.5.) as the FOA $r-k$ class estimator in GLMs.

[1] One can also use $\hat{W}_{ML}$ instead of $\hat{W}^{(0)}$ which was used by (Özkale and Arıcan, 2016)

Eq. (3.5.) implies that:

1) the FOA ML estimator:

$$\hat{\beta}^{(1)}_{r=p,k=0} = \hat{\beta}^{(1)} = \left(X^T W^{(0)} X\right)^{-1} X^T W^{(0)} z^{(0)}$$

where $z^{(0)} = X\beta^{(0)} + (D^{(0)})^{-1}\left(y - \mu^{(0)}\right)$

2) the FOA PCR estimator:

$$\hat{\beta}^{(1)}_{r,k=0} = \hat{\beta}^{(1)}_r = M_r\left(M_r^T X^T W^{(0)} X M_r\right)^{-1} M_r^T X^T W^{(0)} z_r^{(0)} = M_r M_r^T \hat{\beta}^{(1)} \quad (3.6.)$$

3) the FOA ridge estimator:

$$\hat{\beta}^{(1)}_{r=p,k} = \hat{\beta}^{(1)}(k) = \left(X^T W^{(0)} X + kI_p\right)^{-1} X^T W^{(0)} z^{(0)}$$

$\hat{\beta}^{(1)}_{rk}$ can also be written in terms of $\hat{\beta}^{(1)}$ defined by Eq. (3.6.):

$$\hat{\beta}^{(1)}_{rk} = M_r\left(M_r^T X^T W^{(0)} X M_r + kI_r\right)^{-1} M_r^T X^T W^{(0)} X M_r M_r^T \hat{\beta}^{(1)}_r. \quad (3.7.)$$

3.1.1. Mean Square Error of FOA $r-k$ Class Estimator

There are many criteria to assess the performance of a good estimator among those the most widely used criterion is the mean square error (MSE). In this section, we develop the MSE of the proposed $r-k$ class estimator. As we know from the literature that asymptotically (by Agresti, 2003)

$$E\left(\hat{\beta}^{(1)}\right) = \beta^{(0)},\ var(\hat{\beta}^{(1)}) = a(\phi)\left(X^T W^{(0)} X\right)^{-1}.$$

From (Smith and Marx, 1990) we approximately write the expected value and variance of the FOA PCR estimator as

$$E(\hat{\beta}_r^{(1)}) = M_r \alpha^{(0)}, \quad var(\hat{\beta}_r^{(1)}) = a(\phi)\left(M_r \Lambda_r^{-1} M_r^T\right),$$

where $\alpha^{(0)} = M_r^T \beta^{(0)}$. Then, by Eq. (3.7.) the asymptotic bias and variance of $\hat{\beta}_{rk}^{(1)}$ can be found. Since

$$E(\hat{\beta}_{rk}^{(1)}) = M_r(\Lambda_r^{(0)} + kI_r)^{-1} \Lambda_r^{(0)} M_r^T \beta^{(0)}$$

where $\Lambda_r^{(0)} = M_r^T X^T W^{(0)} X M_r$ we obtain

$$Bias(\hat{\beta}_{rk}^{(1)}) = E(\hat{\beta}_{rk}^{(1)}) - \beta^{(0)} = [M_r\left(\Lambda_r^{(0)} + kI_r\right)^{-1} \Lambda_r^{(0)} M_r^T - I_p]\beta^{(0)}.$$

Since $I_p = M_r M_r^T + M_{p-r} M_{p-r}^T$ on making substitution and simplifying we get

$$Bias(\hat{\beta}_{rk}^{(1)}) = \left[-kM_r\left(\Lambda_r^{(0)} + kI_r\right)^{-1} M_r^T - M_{p-r} M_{p-r}^T\right]\beta^{(0)}.$$

The asymptotic variance of $\hat{\beta}_{rk}^{(1)}$ is given as

$$var(\hat{\beta}_{rk}^{(1)}) = a(\phi) M_r \left(\Lambda_r^{(0)} + kI_r\right)^{-1} \Lambda_r^{(0)} \left(\Lambda_r^{(0)} + kI_r\right)^{-1} M_r^T.$$

Then, by using

$$MSE(\hat{\beta}_{rk}^{(1)}) = var\left(\hat{\beta}_{rk}^{(1)}\right) + [Bias\left(\hat{\beta}_{rk}^{(1)}\right)][Bias\left(\hat{\beta}_{rk}^{(1)}\right)]^T,$$

the asymptotic matrix MSE of $\hat{\beta}_{rk}^{(1)}$ is defined as

$$\begin{aligned} MSE(\hat{\beta}_{rk}^{(1)}) &= a(\phi) M_r \left(\Lambda_r^{(0)} + kI_r\right)^{-1} \Lambda_r^{(0)} \left(\Lambda_r^{(0)} + kI_r\right)^{-1} M_r^T \\ &\quad + [-kM_r \left(\Lambda_r^{(0)} + kI_r\right)^{-1} M_r^T - M_{p-r} M_{p-r}^T] \beta^{(0)} \\ &\quad \times (\beta^{(0)})^T [-kM_r \left(\Lambda_r^{(0)} + kI_r\right)^{-1} M_r^T - M_{p-r} M_{p-r}^T]. \end{aligned} \quad (3.8.)$$

As we know that the scalar MSE (SMSE) is obtained from $SMSE(\hat{\beta}_{rk}^{(1)}) = tr(MSE(\hat{\beta}_{rk}^{(1)}))$ and also $M_r^T M_{p-r} = 0$, by using these results and properties of trace we get the SMSE as

$$SMSE(\hat{\beta}_{rk}^{(1)}) = \sum_{i=1}^{r} \frac{a(\phi)\lambda_i + k^2 \left(\alpha_i^{(0)}\right)^2}{(\lambda_i + k)^2} + \sum_{i=r+1}^{p} \left(\alpha_i^{(0)}\right)^2. \quad (3.9.)$$

Here, we discuss some of the special cases of the $r-k$ class estimator. It can be seen that if we replace $r = p$ in Eqs. (3.8.) and (3.9.) the matrix MSE and SMSE of the

FOA ridge estimator $\hat{\beta}^{(1)}(k)$ are obtained, respectively as

$$\begin{aligned} MSE(\hat{\beta}^{(1)}(k)) &= a(\phi)M\left(\Lambda^{(0)}+kI_p\right)^{-1}\Lambda^{(0)}\left(\Lambda^{(0)}+kI_p\right)^{-1}M^T \\ &\quad +k^2M\left(\Lambda^{(0)}+kI_p\right)^{-1}M^T\beta^{(0)}\left(\beta^{(0)}\right)^T M\left(\Lambda^{(0)}+kI_p\right)^{-1}M^T \end{aligned} \quad (3.10.)$$

and

$$SMSE(\hat{\beta}^{(1)}(k)) = \sum_{i=1}^{p}\frac{a(\phi)\lambda_i + k^2\left(\alpha_i^{(0)}\right)^2}{(\lambda_i+k)^2}. \quad (3.11.)$$

Similarly, the matrix MSE and SMSE for the FOA PCR estimator $\hat{\beta}_r^{(1)}$ can be obtained by substituting $k=0$ in Eqs. (3.8.) and (3.9.) that is:

$$MSE(\hat{\beta}_r^{(1)}) = a(\phi)M_r(\Lambda_r^{(0)})^{-1}M_r^T + M_{p-r}M_{p-r}^T\beta^{(0)}(\beta^{(0)})^T M_{p-r}M_{p-r}^T \quad (3.12.)$$

and

$$SMSE(\hat{\beta}_r^{(1)}) = \sum_{i=1}^{r}\frac{a(\phi)}{\lambda_i} + \sum_{i=r+1}^{p}\left(\alpha_i^{(0)}\right)^2. \quad (3.13.)$$

Whereas, the SMSE for the FOA ML estimator is given below:

$$SMSE(\hat{\beta}_{ML}^{(1)}) = \sum_{i=1}^{p}\frac{a(\phi)}{\lambda_i}. \quad (3.14.)$$

3.1.2. Estimation of the Parameter k

In this section, we estimate the value of parameter k in the $r-k$ class estimator. A plausible value of k can be obtained from the SMSE of $\hat{\beta}_{rk}^{(1)}$. Thus, to find the value of k by differentiating Eq. (3.9.) with respect to k for a specified r we get:

$$\frac{dSMSE(\hat{\beta}_{rk}^{(1)})}{dk} = \sum_{i=1}^{r} \frac{-2a(\phi)\lambda_i + 2\lambda_i k \left(\alpha_i^{(0)}\right)^2}{(\lambda_i + k)^3} = 0. \qquad (3.15.)$$

Considering the right hand side of Eq. (3.15.) and solving it for k we get the required value of k for $i = 1,2,...,r$, as $k = \frac{a(\phi)}{\left(\alpha_i^{(0)}\right)^2}$, which is minimum at that value of k. Although an overwhelming number of choices are available in the literature for the selection of ridge parameter k, we might use the same procedure suggested by (Hoerl and Kennard, 1970) and (Hoerl et al., 1975); the value of k in $r-k$ class estimator can be found by using $k = \frac{1}{\max(m_i^T \beta^{(0)})^2}$ or by taking the harmonic mean of $\frac{a(\phi)}{(\alpha_i^{(0)})^2}$ for $i = 1,2,...,r$, that is $k = \frac{ra(\phi)}{\beta^{(0)T} M_r M_r^T \beta^{(0)}}$.

3.2. Restricted OK Estimator in GLMs

This section explains and establishes the restricted OK estimator in GLMs under exact linear restrictions.

In various applications for the estimation of parameters along with the sample information prior or auxiliary information about the parameter vector β are also available. This information can be utilized in the form of exact linear restrictions or stochastic linear restrictions. The prior information in the form of exact linear restrictions might be expressed as: $r = R\beta$, where r is a $q \times 1$ vector and R is $q \times p$ matrix having rank $(R) = q$ with a set of q linear and independent restrictions on the parameters such as:

$$R = \begin{bmatrix} R_1 \\ \vdots \\ R_q \end{bmatrix}$$

where $R_i = [R_{i1}\ R_{i2} \cdots R_{ip}]$, $i = 1, 2, \cdots, q$ as well as r and R are supposed to be known.

We want to estimate parameter vector β by applying exact linear restrictions on the parameter vector β. Since each element of the random variable Y is a member of an exponential family of distribution, by imposing a set of q linear restrictions on the parameter vector β as $r_i = R_i\beta$, the log-likelihood function by following (Nyquist, 1991) is given as follows:

$$l(\theta, \beta; y, r) = \sum_{i=1}^{n} \{ \frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \} - \frac{1}{2} \sum_{i=1}^{q} \lambda_i (r_i - R_{ij}\beta_j)^2.$$

where $\lambda_1, \ldots, \lambda_q$ are the lagrange multipliers. To find the restricted OK estimator in GLMs we also consider the penalty function by (Kurtoğlu and Özkale, 2016) as:

$$p(\beta, d) = \frac{1}{2}(\beta - d\hat{\beta}_{ML})^T (\beta - d\hat{\beta}_{ML}).$$

Thus, we combine these two log-likelihood functions for estimating the restricted OK estimator and replacing $\frac{1}{2}$ with $\frac{k}{2}$ in the penalty function as follows:

$$\begin{aligned} l(\theta,\beta,k,d;y,r) &= l(\theta,\beta;y,r)+p(\beta,d) \\ &= \sum_{i=1}^{n}\{\frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i,\phi)\} - \frac{k}{2}\sum_{j=1}^{p}(\beta_j - d\hat{\beta}_{MLj})^2 \\ &\quad -\frac{1}{2}\sum_{i=1}^{q}\lambda_i(r_i - R_{ij}\beta_j)^2. \end{aligned}$$

where $k > 0$, $0 < d < 1$ are the tuning parameters.

Differentiating $l(\theta,\beta,k,d;y,r)$ with respect to β_j and using the chain rule we get

$$\begin{aligned} \frac{\partial}{\partial\beta_j} l(\theta,\beta,k,d;y,r) &= \sum_{i=1}^{n}\frac{y_i-\mu_i}{a(\phi)}\frac{1}{b''(\theta_i)}\frac{1}{g'(\mu_i)}x_{ij} - k\sum_{j=1}^{p}(\beta_j - d\hat{\beta}_{MLj}) \\ &\quad +\sum_{i=1}^{q}\lambda_i R_{ij}(r_i - R_{ij}\beta_j) \\ &= \sum_{i=1}^{n}\frac{y_i-\mu_i}{a(\phi)}\frac{1}{b''(\theta_i)}\frac{g'(\mu_i)}{(g'(\mu_i))^2}x_{ij} - k\sum_{j=1}^{p}(\beta_j - d\hat{\beta}_{MLj}) \\ &\quad +\sum_{i=1}^{q}\lambda_i R_{ij}(r_i - R_{ij}\beta_j) \\ &= \sum_{i=1}^{n}\frac{y_i-\mu_i}{w_{ii}}g'(\mu_i)x_{ij} - k\sum_{j=1}^{p}\beta_j + kd\sum_{j=1}^{p}\hat{\beta}_{MLj} \\ &\quad +\sum_{i=1}^{q}\lambda_i R_{ij}(r_i - R_{ij}\beta_j) \qquad (3.16.) \end{aligned}$$

where $w_{ii} = var(y_i)(g'(\mu_i))^2$. The matrix form of Eq. (3.16.) can be written as

$$S(\beta,d,k) = X^T WD(y-\mu) - k\beta + kd\hat{\beta}_{ML} + R^T\Lambda(r - R\beta)$$

where $W = \text{diag}(w_{ii}^{-1})$, $D = \text{diag}(g'(\mu_i))$ and $\Lambda = \text{diag}(\lambda_1,\cdots,\lambda_q)$.

Now taking the derivative of Eq. (3.16.) with respect to β_k we have

$$\begin{aligned}\frac{\partial^2}{\partial\beta_j\partial\beta_k}l(\theta,\beta,k,d;y,r) &= \sum_{i=1}^{n}(y_i-\mu_i)\frac{\partial}{\partial\beta_k}\frac{1}{w_{ii}}g^{'}(\mu_i)x_{ij}+\sum_{i=1}^{n}\frac{g^{'}(\mu_i)x_{ij}}{w_{ii}}\\ &\quad\times(\frac{-1}{g^{'}(\mu_i)}x_{ik})-k\delta_{jk}-\sum_{i=1}^{q}\lambda_iR_{ij}R_{ik} \qquad (3.17.)\end{aligned}$$

where $\delta_{jk}=1$, if $j=k$, and 0 otherwise.

Minus times the expected value of Eq. (3.17.) gives us:

$$\begin{aligned}Q_{jk}(\beta,d,k) &= -E[\frac{\partial^2}{\partial\beta_j\partial\beta_k}l(\theta,\beta,k,d;y,r)]\\ &= \sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{w_{ii}}+k\delta_{jk}+\sum_{i=1}^{q}\lambda_iR_{ij}R_{ik}.\end{aligned}$$

In matrix notation we can write it as

$$Q(\beta,d,k)=(X^TWX+kI_p)+R^T\Lambda R.$$

By using the method of Fisher's scoring we have

$$\hat{\beta}_{kd-R}^{(t+1)}=\hat{\beta}_{kd-R}^{(t)}+\{[Q(\beta,d,k)]^{-1}S(\beta,d,k)\}_{\beta=\hat{\beta}_{kd-R}^{(t)}}.$$

Premultiplying both sides by $Q(\beta,\,d,\,k)$, we get

$$[Q(\beta,d,k)]_{\beta=\hat{\beta}_{kd-R}^{(t)}}\hat{\beta}_{kd-R}^{(t+1)}=[Q(\beta,d,k)]_{\beta=\hat{\beta}_{kd-R}^{(t)}}\hat{\beta}_{kd-R}^{(t)}+[S(\beta,d,k)]_{\beta=\hat{\beta}_{kd-R}^{(t)}}.$$

By substituting the values we have

$$\begin{aligned}\left[(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)+R^T\Lambda R\right]\hat{\beta}^{(t+1)}_{kd-R} &= \{[(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)+R^T\Lambda R]\hat{\beta}^{(t)}_{kd-R}\\ &\quad +X^T\hat{W}^{(t)}_{kd-R}D(y-\hat{\mu}^{(t)}_{kd-R})-k\hat{\beta}^{(t)}_{kd-R}\\ &\quad +kd\hat{\beta}_{ML}+R^T\Lambda(r-R\hat{\beta}^{(t)}_{kd-R})\}\end{aligned}$$

where t denotes the iteration step and $\hat{W}^{(t)}_{kd-R}$ is a weight matrix evaluated at $\hat{\beta}^{(t)}_{kd-R}$. Then we obtain

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-R} &= \left[(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)+R^T\Lambda R\right]^{-1}\{(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)\hat{\beta}^{(t)}_{kd-R}\\ &\quad +X^T\hat{W}^{(t)}_{kd-R}D(y-\hat{\mu}^{(t)}_{kd-R})-k\hat{\beta}^{(t)}_{kd-R}+kd\hat{\beta}_{ML}+R^T\Lambda r\}. \qquad (3.18.)\end{aligned}$$

By using inverse formula, we have

$$\begin{aligned}\left[(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)+R^T\Lambda R\right]^{-1} &= (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\\ &\quad \times R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1}\\ &\quad \times R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}.\end{aligned}$$

Then, Eq. (3.18.) be converted to

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-R} &= (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-R}X\hat{\beta}^{(t)}_{kd-R}+X^T\hat{W}^{(t)}_{kd-R}D(y-\hat{\mu}^{(t)}_{kd-R})\\ &\quad +kd\hat{\beta}_{ML}\}-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1}\\ &\quad \times R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-R}X\hat{\beta}^{(t)}_{kd-R}+X^T\hat{W}^{(t)}_{kd-R}D(y-\hat{\mu}^{(t)}_{kd-R})\\ &\quad +kd\hat{\beta}_{ML}\}+\{(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\\ &\quad \times[\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T]^{-1}R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\}R^T\Lambda r.\end{aligned}$$

(3.19.)

By considering:

$$
\begin{aligned}
& \{(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T \\
& \times[\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T]^{-1}R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\}R^T\Lambda \\
= & \{(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T \\
& \times[\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T]^{-1}R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}\Lambda \\
= & (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{I_p-[\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T]^{-1} \\
& \times R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}\Lambda \\
= & (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1},
\end{aligned}
$$

Eq. (3.19.) becomes

$$
\begin{aligned}
\hat{\beta}^{(t+1)}_{kd-R} = & (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-R}X\hat{\beta}^{(t)}_{kd-R}+X^T\hat{W}^{(t)}_{kd-r}D(y-\hat{\mu}^{(t)}_{kd-r}) \\
& +kd\hat{\beta}_{ML}\}-(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-r}X+kI_p)^{-1} \\
& \times R^T\}^{-1}R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-r}X\hat{\beta}^{(t)}_{kd-R} \\
& +X^T\hat{W}^{(t)}_{kd-R}D(y-\hat{\mu}^{(t)}_{kd-R})+kd\hat{\beta}_{ML}\}+(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1} \\
& \times R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1}r.
\end{aligned}
$$

On further simplification we get

$$
\begin{aligned}
\hat{\beta}^{(t+1)}_{kd-R} = & (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-R}\{X\hat{\beta}^{(t)}_{kd-R}+D(y-\hat{\mu}^{(t)}_{kd-R})\}+kd\hat{\beta}_{ML}] \\
& -(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1} \\
& \times\{R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-r}\{X\hat{\beta}^{(t)}_{kd-R}+D(y-\hat{\mu}^{(t)}_{kd-R})\} \\
& +kd\hat{\beta}_{ML}]-r\}. \qquad (3.20.)
\end{aligned}
$$

Hence, we obtain an iterative restricted OK estimator in GLMs as follows:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-R} &= (X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-R}\hat{z}^{(t)}_{kd-R}+kd\hat{\beta}_{ML}] \\ &\quad -(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\{\Lambda^{-1}+R(X^T\hat{W}^{(t)}_{kd-R}X+kI_p)^{-1}R^T\}^{-1} \\ &\quad \times\{R(X^T\hat{W}^{(t)}_{kd-r}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-R}z^{(t)}_{kd-R}+kd\hat{\beta}_{ML}]-r\} \qquad (3.21.)\end{aligned}$$

where $\hat{z}^{(t)}_{kd-R}=X\hat{\beta}^{(t)}_{kd-R}+D(y-\hat{\mu}^{(t)}_{kd-R})$ is the working response evaluated at $\hat{\beta}^{(t)}_{kd-R}$. As t goes to ∞ and $\lambda_1,\ldots,\lambda_q\to\infty$, then we obtain the asymptotic form of $\hat{\beta}^{(t+1)}_{kd-R}$ in such a way:

$$\begin{aligned}\hat{\beta}_{kd-R} &= (X^T\hat{W}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}_{kd-R}\hat{z}_{kd-R}+kd\hat{\beta}_{ML}] \\ &\quad -(X^T\hat{W}_{kd-R}X+kI_p)^{-1}R^T[R(X^T\hat{W}_{kd-R}X+kI_p)^{-1}R^T]^{-1} \\ &\quad \times\{R(X^T\hat{W}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}_{kd-R}\hat{z}_{kd-R}+kd\hat{\beta}_{ML}]-r\}. \end{aligned} \qquad (3.22.)$$

In Eq. (3.22.), $(X^T\hat{W}_{kd-R}X+kI_p)^{-1}[X^T\hat{W}_{kd-R}\hat{z}_{kd-R}+kd\hat{\beta}_{ML}]$ is in the form of OK estimator that (Özkale and Kaçıranlar, 2007) proposed in linear regression, when the response variable follows a normal distribution[2]. $\hat{\beta}_{kd-R}$ in Eq. (3.22.) presents an estimator in the form of exact restrictions. Therefore, we obtained the iterative restricted OK estimator in GLMs under exact restrictions. With the notions of (Özkale and Kaçıranlar, 2007) and (Özkale, 2014), the FOA form of $\hat{\beta}^{(t+1)}_{kd-R}$ can be written as

[2]We call this new two parameter estimator as the restricted OK estimator because there are many alternative two parameter estimators in the literature. It is based on the restricted two parameter proposed by (Ozkale and Kacranlar, 2007) in linear regression and the abbreviation OK indicates the first letter of author's names.

follows:

$$\begin{aligned}
\hat{\beta}^{(1)}_{kd-R} &= (X^TW^{(0)}X+kI_p)^{-1}[X^TW^{(0)}z^{(0)}+kd\hat{\beta}^{(1)}] \\
&\quad -(XW^{(0)}X+kI_p)^{-1}R^T\{R(X^TW^{(0)}X+kI_p)^{-1}R^T\}^{-1} \\
&\quad \times\{R(X^TW^{(0)}X+kI_p)^{-1}[X^TW^{(0)}z^{(0)}+kd\hat{\beta}^{(1)}]-r\} \\
&= (S_k^{(0)})^{-1}S_{kd}^{(0)}\hat{\beta}^{(1)}-(S_k^{(0)})^{-1}R^T\{R(S_k^{(0)})^{-1}R^T\}^{-1}\{R(S_k^{(0)})^{-1}S_{kd}^{(0)}\hat{\beta}^{(1)}-r\}
\end{aligned}$$

where $\hat{\beta}^{(1)}=(X^TW^{(0)}X)^{-1}X^TW^{(0)}z^{(0)}$, $S_k^{(0)}=X^TW^{(0)}X+kI_p$, $S_{kd}^{(0)}=X^TW^{(0)}X+kdI_p$ and $z^{(0)}=X\beta^{(0)}+D(y-\mu^{(0)})$. Here, $\beta^{(0)}$ is the initial value and both $W^{(0)}$ and $z^{(0)}$ are evaluated at $\beta^{(0)}$.

After rearrangements, $\hat{\beta}^{(1)}_{kd-R}$ can also be written as

$$\hat{\beta}^{(1)}_{kd-R}=M_k^{(0)}S_{kd}^{(0)}\hat{\beta}^{(1)}+(S_k^{(0)})^{-1}R^T\{R(S_k^{(0)})^{-1}R^T\}^{-1}r$$

where $M_k^{(0)}=(S_k^{(0)})^{-1}-(S_k^{(0)})^{-1}R^T\{R(S_k^{(0)})^{-1}R^T\}^{-1}R(S_k^{(0)})^{-1}$.

It is to be noted that the proposed estimator $\hat{\beta}_{kd-R}$ is a general estimator, which includes the following estimators under exceptional conditions of k, d, and R in GLMs if the weight matrix and working response of the estimators are same.

i) If $k=0$ and no restrictions then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}^{(1)}$

ii) If $k=1$ and no restrictions then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}^{(1)}_d$ (see Kurtoğlu and Özkale, 2016)

iii) If $k=1$, then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}_R(d)^{(1)}$ (see Kurtoğlu and Özkale, 2019b)

iv) If $d=0$ and no restrictions then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}(k)^{(1)}$ (see Segerstedt, 1992)

v) If $d=0$, then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}_R(k)^{(1)}$ (see Kurtoğlu and Özkale, 2019a)

vi) If $k=0$ and $d=0$, then $\hat{\beta}^{(1)}_{kd-R}\to\hat{\beta}^{(1)}_R$ (see Nyquist, 1991)

3.2.1. MSE Comparisons of the FOA Restricted OK Estimator to Other Estimators

In this section, the comparisons of the FOA restricted OK estimator with the FOA ML, FOA restricted ML, and FOA restricted ridge estimators in GLMs are given. One of the difficulties encountered in comparing the estimators in GLMs is that the weight matrices and working responses of the estimators are changed in every iteration, whereas weight matrices and working responses should be the same for comparisons. Therefore, we consider the FOA form of the aforementioned estimators, where the weight matrix and working response are the same. The comparison is given through an approximated matrix MSE criterion, when the exact linear restrictions $r = R\beta$ are assumed to hold. The reason for using this criterion is that it contains all the relevant information that an estimator should have. Suppose that we have two estimators $\hat{\beta}_1$ and $\hat{\beta}_2$, and $\hat{\beta}_1$ is superior to $\hat{\beta}_2$ if and only if the difference $MSE(\hat{\beta}_1) - MSE(\hat{\beta}_2)$ is non negative definite (nnd) (Toutenburg, 1982).

We provide the approximated matrix MSE of the restricted FOA OK estimator as:

$$MSE\left(\hat{\beta}^{(1)}_{kd-R}\right) = var(\hat{\beta}^{(1)}_{kd-R}) + [Bias(\hat{\beta}^{(1)}_{kd-R})][Bias(\hat{\beta}^{(1)}_{kd-R})]^T$$

where

$$var(\hat{\beta}^{(1)}_{kd-R}) = a(\phi)M^{(0)}_k S^{(0)}_{kd}(S^{(0)})^{-1}S^{(0)}_{kd}M^{(0)}_k,$$

and

$$\begin{aligned}Bias(\hat{\beta}^{(1)}_{kd-R}) &= (M_k^{(0)}S_{kd}^{(0)} - I_p)\beta + (I_p - M_k^{(0)}S_k^{(0)})\beta \\ &= M_k^{(0)}[S_{kd}^{(0)} - S_k^{(0)}]\beta = k(d-1)M_k^{(0)}\beta\end{aligned}$$

where $S^{(0)} = X^T W^{(0)} X$, when $r = R\beta$ is assumed to be true. If the restrictions are not true i.e. $r \neq R\beta$ then,

$$Bias(\hat{\beta}^{(1)}_{kd-R}) = (M_k^{(0)}S_{kd}^{(0)} - I_p)\beta + (S_k^{(0)})^{-1}R^T[R(S_k^{(0)})^{-1}R^T]^{-1}r.$$

The approximated matrix MSE of $\hat{\beta}^{(1)}_{kd-R}$, approximated matrix MSE of $\hat{\beta}^{(1)}$, approximated matrix MSE of $\hat{\beta}^{(1)}_R$ (by Nyquist, 1991), and approximated matrix MSE of $\hat{\beta}_R(k)^{(1)}$ (by Kurtoğlu and Özkale, 2019a) are given as follows assuming that the restrictions are true i.e. $r = R\beta$:

$$MSE(\hat{\beta}^{(1)}_{kd-R}) = a(\phi)M_k^{(0)}S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}M_k^{(0)} + k^2(d-1)^2M_k^{(0)}\beta\beta^T M_k^{(0)} \tag{3.23.}$$

$$MSE[(\hat{\beta}^{(1)})] = a(\phi)(S^{(0)})^{-1} \tag{3.24.}$$

$$MSE[(\hat{\beta}^{(1)}_R)] = a(\phi)M^{(0)} \tag{3.25.}$$

$$MSE[\hat{\beta}_R(k)^{(1)}] = a(\phi)M_k^{(0)}S^{(0)}M_k^{(0)} + k^2M_k^{(0)}\beta\beta^T M_k^{(0)} \tag{3.26.}$$

where $M^{(0)} = (S^{(0)})^{-1} - (S^{(0)})^{-1}R^T[R(S^{(0)})^{-1}R^T]^{-1}R(S^{(0)})^{-1}$.

Here, we make comparisons of $\hat{\beta}^{(1)}_{kd-R}$ against $\hat{\beta}^{(1)}$, $\hat{\beta}^{(1)}_R$, and $\hat{\beta}_R(k)^{(1)}$ in the sense of approximated matrix MSE criterion. Results of the comparisons are stated by Theorems 3.1., 3.2., and 3.3., respectively. For the proof of Theorems 3.1., 3.2., and

3.3., the paper of (Özkale, 2014) can be followed, because theoretically the closed-form expressions of Eqs. (3.23.), (3.24.), (3.25.), and (3.26.) are the same as that of the formula given by (Özkale, 2014) in the linear regression model. Furthermore, the comparisons are not easy because matrix algebra is needed. Therefore, we modify Theorems 3.1., 3.2., and 3.3., from (Özkale, 2014).

3.2.1.1. Comparison Between FOA Restricted OK and FOA ML Estimators

To show that $\hat{\beta}^{(1)}_{kd-R}$ is dominant over $\hat{\beta}^{(1)}$we take the difference $\Delta_1 = MSE[(\hat{\beta}^{(1)})] - MSE[(\hat{\beta}^{(1)}_{kd-R})]$. In view of Eqs. (3.23.) and (3.24.) Δ_1 becomes

$$\Delta_1 = a(\phi)(S^{(0)})^{-1} - a(\phi)M_k^{(0)}S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}M_k^{(0)} - k^2(d-1)^2M_k^{(0)}\beta\beta^T M_k^{(0)}$$

We can illustrate the result by Theorem 3.1.

Theorem 3.1. Under the condition that the difference $a(\phi)[(S^{(0)})^{-1} - M_k^{(0)}S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}M_k^{(0)}]$ is nnd then, $\hat{\beta}^{(1)}_{kd-R}$ is better than $\hat{\beta}^{(1)}$ according to approximated matrix MSE criterion if and only if

$$k^2(d-1)^2\beta^T M_k^{(0)}[(S^{(0)})^{-1} - M_k^{(0)}S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}M_k^{(0)}]^- M_k^{(0)}\beta \leq a(\phi)$$

where $[(S^{(0)})^{-1} - M_k^{(0)}S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}M_k^{(0)}]^-$ denotes the generalized inverse of the matrix.

3.2.1.2. Comparison Between FOA Restricted OK and FOA Restricted ML Estimators

To see that $\hat{\beta}^{(1)}_{kd-R}$ is better than $\hat{\beta}^{(1)}_{R}$ in the sense of approximated matrix MSE criterion we consider the difference $\Delta_2 = MSE[(\hat{\beta}^{(1)}_{R})] - MSE[(\hat{\beta}^{(1)}_{kd-R})]$. From Eqs. (3.23.) and (3.25.) Δ_2 implies that

$$\Delta_2 = a(\phi)M_{(0)} - a(\phi)M^{(0)}_{k}S^{(0)}_{kd}(S^{(0)})^{-1}S^{(0)}_{kd}M^{(0)}_{k} - k^2(d-1)^2M^{(0)}_{k}\beta\beta^T M^{(0)}_{k}.$$

The result is demonstrated by Theorem 3.2.

Theorem 3.2. Under the scenario of $a(\phi)[M_0 - M^{(0)}_{k}S^{(0)}_{kd}(S^{(0)})^{-1}S^{(0)}_{kd}M^{(0)}_{k}]$ is nnd then, the estimator $\hat{\beta}^{(1)}_{kd-R}$ is considered to be better as compared to $\hat{\beta}^{(1)}_{R}$ by the approximated matrix MSE criterion if and only if

$$k^2(d-1)^2\beta^T M^{(0)}_{k}[M_{(0)} - M^{(0)}_{k}S^{(0)}_{kd}(S^{(0)})^{-1}S^{(0)}_{kd}M^{(0)}_{k}]^{-}M^{(0)}_{k}\beta \leq a(\phi)$$

where $[M_{(0)} - M^{(0)}_{k}S^{(0)}_{kd}(S^{(0)})^{-1}S^{(0)}_{kd}M^{(0)}_{k}]^{-}$ represents the generalized inverse of the matrix.

3.2.1.3. Comparison Between FOA Restricted OK and FOA Restricted Ridge Estimators

For the superiority of $\hat{\beta}^{(1)}_{kd-R}$ over $\hat{\beta}_{R}(k)^{(1)}$ the difference $\Delta_3 = MSE[\hat{\beta}_{R}(k)^{(1)}] - MSE[(\hat{\beta}^{(1)}_{kd-R})]$ is considered, which can be written from Eqs. (3.23.)

and (3.26.) as

$$\Delta_3 = M_k^{(0)}[a(\phi)(S^{(0)} - S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}) + k^2\beta\beta^T - k^2(d-1)^2\beta\beta^T]M_k^{(0)}.$$

The result is presented in Theorem 3.3.

Theorem 3.3. Under the condition of $a(\phi)(S^{(0)} - S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)})$ is nnd then, $\hat{\beta}_{kd-R}^{(1)}$ is superior as compared to $\hat{\beta}_R(k)^{(1)}$ by approximated matrix MSE criterion if and only if

$$k^2(d-1)^2\beta^T M_k^{(0)}[(S^{(0)} - S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}]^- M_k^{(0)}\beta \leq a(\phi)$$

where $[(S^{(0)} - S_{kd}^{(0)}(S^{(0)})^{-1}S_{kd}^{(0)}]^-$ is the generalized inverse of the matrix.

3.2.2. Selection of the Tuning Parameters

This section explains how to choose tuning parameters of the restricted OK estimator by minimizing MSE and employing the genetic algorithm technique.

3.2.2.1. Selection of the Tuning parameters by Minimizing MSE

In this section, we give the method for selecting the tuning parameters of the restricted OK estimator by minimizing MSE. Here, we apply the technique of spectral decomposition to find the values of tuning parameters k and d. As the LRMs can be converted into canonical form through an orthogonal transformation, similarly the information matrix $X^T\hat{W}^{(0)}X$ in the first iteration for GLMs might be expressed as $X^T\hat{W}^{(0)}X = V\Lambda V^T$, where $V = [V_1, \cdots, V_p]$ is a $p \times p$ diagonal matrix whose columns compose the eigenvectors of $X^T\hat{W}^{(0)}X$. Therefore, $V^T X^T \hat{W}^{(0)} XV = Z^T\hat{W}^{(0)}Z = \Lambda =$

$diag(\lambda_j)$ is a $p \times p$ diagonal matrix, which contains the eigenvalues of $X^T\hat{W}^{(0)}X$ on its main diagonals and $\left(\lambda_1 = \lambda_{\max} \geqslant \lambda_2, \cdots, \geq \lambda_p = \lambda_{\min} \geq 0\right)$. Thus, in order to write the canonical form of the MSE of FOA restricted OK estimator we consider Eq. (3.23.) when the restrictions are true and we write the matrix $M_k^{(0)}$ by following (Groß, 2003) as:

$$M_k^{(0)} = (QS_kQ)^+ = V \begin{bmatrix} (\Lambda_{p-q} + kI_{p-q})^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T$$

where $Q = I_p - R^T(RR^T)^{-1}R$ and Λ_{p-q} is a diagonal matrix consists of $p-q$ nonzero eigenvalues of $QS^{(0)}Q$ on its main diagonal and V is an orthogonal matrix, whose columns form the eigenvectors of $QS^{(0)}Q$. Moreover, $QS^{(0)}Q$ and Q are commutative matrices, which can simultaneously diagonalized through the same orthogonal matrix such that $QS^{(0)}Q = V \begin{bmatrix} \Lambda_{p-q} & 0 \\ 0 & 0 \end{bmatrix} V^T$ and $Q = V \begin{bmatrix} I_{p-q} & 0 \\ 0 & 0 \end{bmatrix} V^T$, (see Özkale, 2014).

By considering Eq. (3.23.) and using the spectral decomposition technique as well as the result of (Groß, 2003) we have:

$$\begin{aligned} SMSE(\hat{\beta}_{kd-R}^{(1)}) &= tr[MSE(\hat{\beta}_{kd-R}^{(1)})] \\ &= tr\left\{ V \begin{bmatrix} \begin{matrix} a(\phi)(\Lambda_{p-q} + kI_{p-q})^{-1}(\Lambda_{p-q} + kdI_{p-q})^{-1}\Lambda^{-1} \\ \times(\Lambda_{p-q} + kdI_{p-q})^{-1}(\Lambda_{p-q} + kI_{p-q})^{-1} \end{matrix} & 0 \\ 0 & 0 \end{bmatrix} V^T \right\} \\ &\quad + tr\left\{ V \begin{bmatrix} k^2(d-1)^2(\Lambda_{p-q} + kI_{p-q})^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T \beta\beta^T V \right. \\ &\quad \left. \times \begin{bmatrix} (\Lambda_{p-q} + kdI_{p-q})^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T \right\}. \end{aligned}$$

After simplification we get

$$\begin{aligned} SMSE(\hat{\beta}^{(1)}_{kd-R}) &= a(\phi)\left[\sum_{j=1}^{p-q}\frac{(\lambda_j+kd)^2}{\lambda_j(\lambda_j+k)^2}\right]+\left[\sum_{j=1}^{p-q}\frac{k^2(d-1)^2\alpha_j^2}{(\lambda_j+k)^2}\right] \\ &= \sum_{j=1}^{p-q}\left[\frac{a(\phi)(\lambda_j+kd)^2+k^2(d-1)^2\lambda_j\alpha_j^2}{\lambda_j(\lambda_j+k)^2}\right] \end{aligned} \qquad (3.27.)$$

where $\alpha = V^T\beta$. To find the optimum value of d for fixed k we differentiate Eq. (3.27.) with respect to d and equate it to 0:

$$\frac{\partial}{\partial d}[SMSE(\hat{\beta}^{(1)}_{kd-R})] = \sum_{j=1}^{p-q}\frac{2a(\phi)k(\lambda_j+kd)+2k^2(d-1)\lambda_j\alpha_j^2}{\lambda_j(\lambda_j+k)^2} = 0$$

which leads to

$$d = \frac{\sum_{j=1}^{p-q}\frac{k\alpha_j^2-a(\phi)}{(\lambda_j+k)^2}}{\sum_{j=1}^{p-q}\frac{k(a(\phi)+\lambda_j\alpha_j^2)}{\lambda_j(\lambda_j+k)^2}}.$$

The unknown parameters $a(\phi)$ and α_j^2 can be replaced by their estimates $\hat{\phi}$ and $\hat{\alpha}_j^2$ to get the optimum value of d:

$$d_{opt} = \frac{\sum_{j=1}^{p-q}\frac{k\hat{\alpha}_j^2-\hat{\phi}}{(\lambda_j+k)^2}}{\sum_{j=1}^{p-q}\frac{k(\hat{\phi}+\lambda_j\hat{\alpha}_j^2)}{\lambda_j(\lambda_j+k)^2}}.$$

which is positive for $k > \max\left\{\frac{\hat{\alpha}_j^2}{\hat{\phi}}\right\}$. Similarly, the value of k can be determined by taking the derivative of Eq. (3.27.) with respect to k for fixed d:

$$\frac{\partial}{\partial k}[SMSE(\hat{\beta}_{kd-R}^{(1)})] = \sum_{j=1}^{p-q} \frac{2a(\phi)(d-1)(\lambda_j + kd) + 2k(d-1)^2\lambda_j\alpha_j^2}{(\lambda_j + k)^3}. \quad (3.28.)$$

The value of k can be obtained by equating the numerator of Eq. (3.28.) to 0 that is

$$k = \frac{a(\phi)}{\alpha_j^2 - d(\alpha_j^2 + \frac{a(\phi)}{\lambda_j})} \text{ for all } j = 1, \cdots, p-q.$$

Therefore, the optimum k value is obtained by substituting the unknown parameters $a(\phi)$ and α_j^2 by their estimates $\hat{\phi}$ and $\hat{\alpha}_j^2$ and by following (Kibria, 2003) we give an estimate of k that is:

$$\hat{k}_{AM} = \frac{1}{p-q}\sum_{j=1}^{p-q} \frac{\hat{\phi}}{\hat{\alpha}_j^2 - d(\hat{\alpha}_j^2 + \frac{\hat{\phi}}{\lambda_j})}$$

which is the arithmetic mean of k values and

$$\hat{k}_{GM} = \frac{\hat{\phi}}{\left[\prod_{j=1}^{p-q}(\hat{\alpha}_j^2 - d(\hat{\alpha}_j^2 + \frac{\hat{\phi}}{\lambda_j})\right]^{1/p-q}}$$

which is the geometric mean of k values and by following (Hoerl et al., 1975)

$$\hat{k}_{HM} = \frac{(p-q)\hat{\phi}}{\sum_{j=1}^{p-q}\left(\hat{\alpha}_j^2 - d(\hat{\alpha}_j^2 + \frac{\hat{\phi}}{\lambda_j})\right)}$$

which is the harmonic mean of k values.

It is obvious that the value of k should be positive therefore, to make sure that the value of k is always positive we impose a restraint on the k values by following (Özkale and Kaçıranlar, 2007) which is:

$$\hat{d} < \min\left\{\frac{\hat{\alpha}_j^2}{\hat{\alpha}_j^2 + \frac{\hat{\phi}}{\lambda_j}}\right\}. \tag{3.29.}$$

Therefore, the restraint given in Eq. (3.29.) yields to always positive values of $\hat{k}_{AM}$, $\hat{k}_{GM}$ or $\hat{k}_{HM}$.

The detailed procedure for the selection of k and d is shown in the flowchart given by Figure 3.1.:

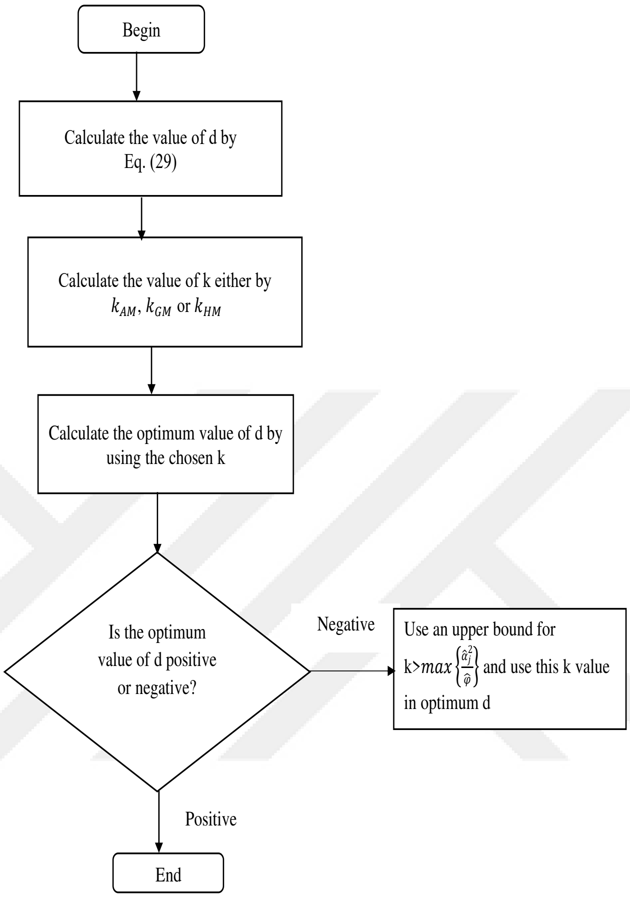


Figure 3.1. Flowchart for the computation of tuning parameters k and d for restricted OK estimator

3.2.2.2. Selection of the Tuning Parameters by Using Genetic Algorithm

In this section, we give a brief overview of genetic algorithms (GA) and its application for the estimation of tuning parameters (k, d).

Genetic algorithms (GAs) were first introduced by (Holland, 1975), which are the stochastic search tools that depend on Darwin's evolutionary theory of natural selection. This algorithm describes the natural selection mechanism where the suitable individuals are chosen for the purpose of reproduction by which the next generation offspring is produced. This natural process of selection begins through the choice of fittest from a population of individuals. Through, this process offspring (children) inheriting characteristics from their parents are added to the next generation. If parents attain better fitness, then their offspring will be better compared to them and have a better chance of survival. This process continues until suitable individuals are found. There are five phases in the process of GA these are; Initial Population, Fitness function, Selection, Crossover, and Mutation (Mallawaarachchi, 2017) and (Siriwardene and Perera, 2006) .

There are several advantages of GA for estimating parameters over conventional methods for example; It does not require taking the derivative of a function and it is more efficient and faster than the traditional methods. Continuous and discrete functions can both be optimized in GA. Thus, GA is a more useful optimization technique than conventional methods for the estimation of tuning parameters because of the derivative involved in these conventional methods and some functions are more complicated and their solution is not easy. On the other hand, GA has the ability to simultaneously optimize more than one tuning parameter in a single algorithm.

The GA has been used by many authors in the context of linear regression for the estimation of parameters due to its advantages over conventional methods. Praga-Alejo et al. (2008) evaluated the optimal ridge tuning parameter k value by minimizing the distance on the basis of variance inflation factor (VIF). Iquebal et al. (2012) have applied a stochastic GA technique in estimating regression parameters. Hamed et al. (2013) suggested a method for the choice of the ridge tuning parameter by using mathematical programming. Tekeli et al. (2019) have applied GA for estimating the tuning parameters of Liu, ridge as well as by simultaneously the tuning parameters of the two parameter estimator in linear regression. Ndabashinze and Şiray (2020) presented a method for the selection of ridge tuning parameter based on GA.

The mean square error prediction (MSEP) and mean absolute error (MAE) are minimized for the simultaneous estimation of parameters of the restricted OK estimator under the condition that VIF and condition number (CN) should be less than or equal to 10 and 30, respectively otherwise a penalty, which is a maximum of VIFs or CN is added to these fitness functions. The MSEP and MAE are defined as follows:

$$MSEP = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{\mu}_i)^2$$
$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \tilde{\mu}_i)|$$

where y_i is the i-th component of y and $\hat{y}_i = \widehat{E(y_i)} = \tilde{\mu}_i = g^{-1}(x_i^T\tilde{\beta})$ (Lamari et al., 2020) and $\tilde{\beta}$ is any estimator of β.

Application of derivatives based techniques is not easy for the optimization of aforementioned measures, therefore, we use GA to find the values of tuning parameters by minimizing these measures. By following (Abdeslam et al., 2014) the pseudo-code of GA is presented in the following Table:

Table 3.1. The pseudo-code of genetic algorithm

1 Generate randomly the initial population of n chromosomes
2 Determine the Fitness function (FF) value for each individual(chromosomes) available in the population.
3 Generate a new population by using the operators:
a Selection : On the basis of FF value
b Recombination : Cross-over chromosomes
c Mutation : Mutate chromosomes
4 Substitute old population with new population on acceptance by (elitism operator)
5 Check for the desired convergence criterion
6 Until the required convergence criterion is satisfied repeat the steps: 1-5.

In order to find the values of tuning parameters k and d first we define the initial population in the GA processing. Then, a fitness function (FF) is used to determine each chromosome. The value of FF determines the reproduction performance of individuals. Therefore, individuals who are more effective in resolving the problem are chosen to transfer their genes to the new generation. Chosen individuals produce a new generation by operators like crossover, mutation and elitism. This process carries on until the prescribed convergence criterion is met. We seek to find optimum values of tuning parameters k and d of the iterative restricted OK estimator simultaneously by means of GA. Individual chromosomes are constructed naturally through the parameters k and d, respectively. Using a certain number of chromosomes containing randomly chosen tuning parameters k and d we generate the initial population. In order to find the values of the tuning parameters we define the following FF:

$$FF(\beta^{(1)}_{kd-R}) = MSEP + \psi_i(k,d) + CN$$

$$FF(\beta^{(1)}_{kd-R}) = MAE + \psi_i(k,d) + CN$$

where

$$\psi_i(k,d) = \begin{cases} 0 \quad if \quad \forall\, VIF_i(k,d) \leq 10,\, i = 1,\cdots,p \\ \max(VIF_i(k,d)) \quad \text{otherwise} \end{cases}$$

and

$$CN(k,d) = \begin{cases} 0 & \text{if } CN \leq 30 \\ CN & \text{otherwise} \end{cases}$$

The measures CN and $VIF_i(k,d)$ of the restricted OK estimator are defined as follows:

$$CN = \sqrt{\frac{\lambda_{\max}[(X^T\hat{W}_{ML}X)(X^T\hat{W}_{ML}X + kdI_p)^{-1}(X^T\hat{W}_{ML}X + kI_p)]}{\lambda_{\min}[(X^T\hat{W}_{ML}X)(X^T\hat{W}_{ML}X + kdI_p)^{-1}(X^T\hat{W}_{ML}X + kI_p)]}}$$

and $VIF_i(k,d)$ is the i-th diagonal element of $(X^T\hat{W}_{ML}X + kI_p)^{-1}(X^T\hat{W}_{ML}X + kdI_p)(X^T\hat{W}_{ML}X)^{-1}(X^T\hat{W}_{ML}X + kdI_p)(X^T\hat{W}_{ML}X + kI_p)^{-1}$ with $\hat{W}_{ML}$ being a weight matrix obtained at the final iteration of the ML estimator. By implementing these GA rules those chromosomes are considered, which make the FF value minimum as the objective is to minimize the FF.

In addition, we have also defined the CN and VIF of the FOA stochastic restricted OK estimator given in Section 3.4., for the particular case when $r = 0$ in

GLMs as follows:

$$CN = \sqrt{\frac{\lambda_{\max}[(X^{*T}X^{*})(X^{*T}X^{*}+kdI_p)^{-1}(X^{*T}X^{*}+kI_p+R^T\Sigma^{-1}R)]}{\lambda_{\min}[(X^{*T}X^{*})(X^{*T}X^{*}+kdI_p)^{-1}(X^{*T}X^{*}+kI_p+R^T\Sigma^{-1}R)]}}$$

where X^* is the matrix obtained by unit length scaling to $\sqrt{\hat{W}^{(0)}}X$ with $\hat{W}^{(0)}$ is a weight matrix acquired in the initial value of β parameter and $VIF_j(k,d)$ is the j-th diagonal elements of $var(\beta^{(1)}_{kd-sr})$, where $var(\beta^{(1)}_{kd-sr})$ is given in subsection 3.4.1.

3.3. Stochastic Restricted Liu Estimator in GLMs

In this section, we demonstrate an iterative stochastic restricted Liu estimator by considering stochastic restrictions on the parameters. The stochastic linear restrictions can be written as follows

$$r = R\beta + G, \quad G \sim (0, a(\phi)\Sigma) \tag{3.30.}$$

where r is an $q \times 1$ vector and R is an $q \times p$ matrix with $rank(R) = q$ including a set of q linear and independent restrictions on the parameter vector β such as

$$R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_q \end{bmatrix},$$

where $R_i = \left[R_{i1}\ R_{i2} \cdots R_{ip}\right]$, $i = 1, 2, \cdots, q$. Since r, R, and Σ are assumed to be known and the vector r is a random variable having expectation $E(r) = R\beta$, (Özkale and Nyquist, 2021) mentioned that the restriction in Eq. (3.30.) does not hold apart from the mean, as r is assumed to be known and all the expectations are conditional on r like that $E\left(\hat{\beta}|r\right)$. Özkale and Nyquist (2021) also stated that the distribution of r is assumed to be normal, as the ML estimator is asymptotically normally distributed. Then, under the assumption that the random vector y and r are independent (Özkale and Nyquist, 2021) obtained the mixed-GLM estimator by maximizing the log-likelihood function

$$\begin{aligned} l(\theta, \beta; y, r) \quad = \quad & \sum_{i=1}^{n} \left\{ \frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right\} - \frac{q}{2} \ln(2\pi) - \frac{q}{2} \ln(|a(\phi)\Sigma|) \\ & - \frac{1}{2a(\phi)} (r - R\beta)^T \Sigma^{-1} (r - R\beta). \end{aligned}$$

By using the notion of (Özkale and Nyquist, 2021), we combine the sample information and auxiliary information to estimate the parameter vector β in the context of Liu estimator which was originally proposed by (Liu, 1993) in linear regression. Liu estimator has also been proposed in GLMs by (Kurtoğlu and Özkale, 2016). They introduced a Liu estimator in GLMs by maximizing the objective function

$$l(\beta, d; y) = l(\beta; y) - \frac{1}{2a(\phi)} \left(\beta - d\hat{\beta}_{ML}\right)^T \left(\beta - d\hat{\beta}_{ML}\right)$$

where $l(\beta; y)$ is the log-likelihood function to obtain the ML estimator.

Thus, by combining the idea of (Kurtoğlu and Özkale, 2016) and (Özkale and

Nyquist, 2021), we define the log-likelihood function

$$l\left(\beta, d; y, r\right) = l\left(\theta, \beta; y, r\right) - \frac{1}{2a(\phi)}\left(\beta - d\hat{\beta}_{ML}\right)^{T}\left(\beta - d\hat{\beta}_{ML}\right)$$

to get the iterative stochastic restricted Liu estimator in GLMs. Differentiating the log-likelihood function $l\left(\beta, d; y, r\right)$ with respect to β_j we get the score function. For this we utilize from (Özkale and Nyquist, 2021) for the derivative of $l\left(\beta, d; y, r\right)$ with respect to β_j and we get

$$\begin{aligned} \frac{\partial l\left(\beta, d; y, r\right)}{\partial \beta_j} &= \sum_{i=1}^{n} \frac{y_i - \mu_i}{a\left(\phi\right) b''\left(\theta_i\right)} \frac{1}{g'\left(\mu_i\right)} x_{ij} + \frac{1}{a(\phi)} T_j - \frac{1}{a(\phi)} \sum_{j=1}^{p} \left(\beta_j - d\hat{\beta}_{MLj}\right) \\ &= \sum_{i=1}^{n} \frac{\left(y_i - \mu_i\right)}{a(\phi) w_{ii}} g'\left(\mu_i\right) x_{ij} + \frac{1}{a(\phi)} T_j - \frac{1}{a(\phi)} \sum_{j=1}^{p} \left(\beta_j - d\hat{\beta}_{MLj}\right) \end{aligned} \tag{3.31.}$$

where T_j are the elements of the $p-$vector $T = R^T \Sigma^{-1}(r - R\beta)$. The matrix form of Eq. (3.31.) can be written as

$$S\left(\beta, d\right) = \frac{1}{a(\phi)}[X^T W D\left(y - \mu\right) + R^T \Sigma^{-1}\left(r - R\beta\right) - \beta + d\hat{\beta}_{ML}].$$

where $W = diag\left(w_{ii}^{-1}\right)$ and $D = diag\left(g'\left(\mu_i\right)\right)$.

Now differentiating Eq. (3.31.) with respect to β_k we have

$$\begin{aligned} \frac{\partial^2}{\partial\beta_j\partial\beta_k} l(\beta,d;y,r) &= \sum_{i=1}^{n}(y_i-\mu_i)\frac{\partial}{\partial\beta_k}\frac{1}{a(\phi)w_{ii}}g^{'}(\mu_i)x_{ij}+\sum_{i=1}^{n}\frac{g^{'}(\mu_i)x_{ij}}{a(\phi)w_{ii}}(\frac{-1}{g^{'}(\mu_i)}x_{ik}) \\ &\quad -\frac{1}{a(\phi)}\delta_{jk}-\frac{1}{a(\phi)}u_{jk} \\ &= \sum_{i=1}^{n}(y_i-\mu_i)\frac{\partial}{\partial\beta_k}\frac{1}{a(\phi)w_{ii}}g^{'}(\mu_i)x_{ij}-\sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{a(\phi)w_{ii}}-\frac{1}{a(\phi)}\delta_{jk} \\ &\quad -\frac{1}{a(\phi)}u_{jk} \qquad (3.32.) \end{aligned}$$

where $\delta_{jk}=1$, if $j=k$ and 0 otherwise and u_{jk} are the elements in the $p\times p$ matrix $U=R^T\Sigma^{-1}R$.

By considering the expected value of the 2nd derivative in Eq. (3.32.) we obtain

$$Q_{jk}(\beta,d)=-E\left[\frac{\partial^2 l(\beta,d;y,r)}{\partial\beta_j\partial\beta_k}\right]=\frac{1}{a(\phi)}\left[\sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{w_{ii}}+\delta_{jk}+u_{jk}\right]$$

which can be written in matrix form as

$$Q(\beta,d)=\frac{1}{a(\phi)}\left[(X^TWX+I)+R^T\Sigma^{-1}R\right].$$

We know from the Fisher's scoring method that

$$\hat{\beta}_{c-l}(d)^{(t+1)}=\hat{\beta}_{c-l}(d)^{(t)}+\left\{[Q(\beta,d)]^{-1}[S(\beta,d)]\right\}_{\beta=\hat{\beta}_{c-l}(d)^{(t)}}.$$

Multiplying both sides by $[Q(\beta,d)]$ and substituting the values of $[Q(\beta,d)]$ and

$[S(\beta, d)]$, we obtain

$$\begin{aligned}
& \frac{1}{a(\phi)}\left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]_{\beta=\hat{\beta}_{c-l}(d)^{(t)}} \hat{\beta}_{c-l}(d)^{(t+1)} \\
= \; & \frac{1}{a(\phi)}\left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]_{\beta=\hat{\beta}_{c-l}(d)^{(t)}} \hat{\beta}_{c-l}(d)^{(t)} \\
& + \frac{1}{a(\phi)}[X^T W^{(t)} D\left(y-\mu^{(t)}\right) - \beta + d\hat{\beta}_{ML} + R^T \Sigma^{-1}(r - R\beta)]_{\beta=\hat{\beta}_{c-l}(d)^{(t)}}
\end{aligned}$$

where $W^{(t)}$ is the weight matrix evaluated at $\hat{\beta}_{c-l}(d)^{(t)}$. After simplifying we have

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} = & \left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]^{-1} [(X^T W^{(t)} X + I)\hat{\beta}_{c-l}(d)^{(t)} \\
& + X^T W^{(t)} D\left(y-\mu^{(t)}\right) - \hat{\beta}_{c-l}(d)^{(t)} + d\hat{\beta}_{ML} + R^T \Sigma^{-1} r].
\end{aligned}$$

If we further simplify for $\hat{\beta}_{c-l}(d)^{(t)}$ we obtain

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} = & \left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]^{-1} [X^T W^{(t)} z_{c-l}(d)^{(t)} \\
& + R^T \Sigma^{-1} r + d\hat{\beta}_{ML}].
\end{aligned}$$

where $z_{c-l}(d)^{(t)} = X\hat{\beta}_{c-l}(d)^{(t)} + D\left(y-\mu^{(t)}\right)$ is the working response evaluated at $\hat{\beta}_{c-l}(d)^{(t)}$.

On making use of inverse formula

$$\begin{aligned}
\left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]^{-1} = & \left(X^T W^{(t)} X + I\right)^{-1} - \left(X^T W^{(t)} X + I\right)^{-1} \\
& \times R^T \{\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T\}^{-1} \\
& \times R\left(X^T W^{(t)} X + I\right)^{-1},
\end{aligned}$$

we get

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} [(X^T W^{(t)} X + I)\hat{\beta}_{c-l}(d)^{(t)} \\
&\quad + X^T W^{(t)} D\left(y - \mu^{(t)}\right) - \hat{\beta}_{c-l}(d)^{(t)} + d\hat{\beta}_{ML} + R^T \Sigma^{-1} r] \\
&\quad - \left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} \\
&\quad \times R\left(X^T W^{(t)} X + I\right)^{-1} [(X^T W^{(t)} X + I)\hat{\beta}_{c-l}(d)^{(t)} \\
&\quad + X^T W^{(t)} D\left(y - \mu^{(t)}\right) - \hat{\beta}_{c-l}(d)^{(t)} + d\hat{\beta}_{ML} + R^T \Sigma^{-1} r].
\end{aligned} \tag{3.33.}$$

Since

$$\begin{aligned}
&[\left(X^T W^{(t)} X + I\right)^{-1} - \left(X^T W^{(t)} X + I\right)^{-1} R^T \\
&\times [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} R \left(X^T W^{(t)} X + I\right)^{-1}] R^T \Sigma^{-1} r \\
= &\left(X^T W^{(t)} X + I\right)^{-1} R^T \{I - [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} R \\
&\times \left(X^T W^{(t)} X + I\right)^{-1} R^T\} \Sigma^{-1} r \\
= &\left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} r
\end{aligned}$$

Eq. (3.33.) reduces to

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} [X^T W^{(t)} X \hat{\beta}_{c-l}(d)^{(t)} + X^T W^{(t)} D\left(y - \mu^{(t)}\right) \\
&\quad + d\hat{\beta}_{ML}] - \left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} \\
&\quad \times R\left(X^T W^{(t)} X + I\right)^{-1} [X^T W^{(t)} X \hat{\beta}_{c-l}(d)^{(t)} + X^T W^{(t)} D\left(y - \mu^{(t)}\right) \\
&\quad + d\hat{\beta}_{ML}] + \left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} r.
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} [X^T W^{(t)} X \hat{\beta}_{c-l}(d)^{(t)} + X^T W^{(t)} D\left(y - \mu^{(t)}\right) \\
&\quad + d\hat{\beta}_{ML}] - \left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} \\
&\quad \times \{R\left(X^T W^{(t)} X + I\right)^{-1} [X^T W^{(t)} X \hat{\beta}_{c-l}(d)^{(t)} + X^T W^{(t)} D\left(y - \mu^{(t)}\right) \\
&\quad + d\hat{\beta}_{ML}] - r\},
\end{aligned}$$

thus, we get an iterative stochastic restricted Liu (SRL) estimator in GLMs and called this estimator as SRL estimator in GLMs because the expression

$$\begin{aligned}
& \left(X^T W^{(t)} X + I\right)^{-1} \left[X^T W^{(t)} X \hat{\beta}_{c-l}(d)^{(t)} + X^T W^{(t)} D\left(y - \mu^{(t)}\right) + d\hat{\beta}_{ML}\right] \\
= & \left(X^T W^{(t)} X + I\right)^{-1} (X^T W^{(t)} z_{c-l}(d)^{(t)} + d\hat{\beta}_{ML})
\end{aligned}$$

has the closed form solution of the Liu estimator in GLMs which is introduced by (Kurtoğlu and Özkale, 2016) where $z_{c-l}(d)^{(t)} = X\hat{\beta}_{c-l}(d)^{(t)} + D\left(y - \mu^{(t)}\right)$ is the working response evaluated at $\hat{\beta}_{c-l}(d)^{(t)}$. This implies that

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} \left[X^T W^{(t)} z_{c-l}(d)^{(t)} + d\hat{\beta}_{ML}\right] - \left(X^T W^{(t)} X + I\right)^{-1} \\
&\quad \times R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} [R(X^T W^{(t)} X + I)^{-1} \\
&\quad \times \left(X^T W^{(t)} z_{c-l}(d)^{(t)} + d\hat{\beta}_{ML}\right) - r]. \qquad (3.34.)
\end{aligned}$$

If t approaches to infinity the successive estimates of $\hat{\beta}_{c-l}(d)^{(t+1)}$ might converge to $\hat{\beta}_{c-l}(d)$ then, we get the SRL estimator at convergence as

$$\begin{aligned}
\hat{\beta}_{c-l}(d) &= (X^T \hat{W}_{c-l} X + I)^{-1} [X^T \hat{W}_{c-l} \hat{z}_{c-l}(d) + d\hat{\beta}_{ML}] - (X^T \hat{W}_{c-l} X + I)^{-1} R^T \\
&\quad \times [\Sigma + R\left(X^T \hat{W}_{c-l} X + I\right)^{-1} R^T]^{-1} \{R\left(X^T \hat{W}_{c-l} X + I\right)^{-1} \\
&\quad \times \left(X^T \hat{W}_{c-l} \hat{z}_{c-l}(d) + d\hat{\beta}_{ML}\right) - r\}.
\end{aligned}$$

From Eq. (3.34.) we see that

$$\begin{aligned}
\hat{\beta}_{c-l}(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} [X^T W^{(t)} z_{c-l}(d)^{(t)} + d\hat{\beta}_{ML}] \\
&\quad - \left(X^T W^{(t)} X + I\right)^{-1} R^T [\Sigma + R\left(X^T W^{(t)} X + I\right)^{-1} R^T]^{-1} \\
&\quad \times [R(X^T W^{(t)} X + I)^{-1} \left(X^T W^{(t)} z_{c-l}(d)^{(t)} + d\hat{\beta}_{ML}\right) - r] \\
&= \left[(X^T W^{(t)} X + I) + R^T \Sigma^{-1} R\right]^{-1} [(X^T W^{(t)} X + I)\hat{\beta}_{c-l}(d)^{(t)} \\
&\quad + X^T W^{(t)} D\left(y - \mu^{(t)}\right) - \hat{\beta}_{c-l}(d)^{(t)} + d\hat{\beta}_{ML} + R^T \Sigma^{-1} r].
\end{aligned} \tag{3.35.}$$

A few special cases of the SRL estimator are also observed and discussed as under. If the working responses and weight matrices of the estimators are same then:

i) It is to be noticed that if there is no restriction in SRL estimator then, the Liu estimator in GLMs can be obtained, which was proposed by (Kurtoğlu and Özkale, 2016) as

$$\hat{\beta}(d)^{(t+1)} = \left(X^T W^{(t)} X + I\right)^{-1} \{X^T W^{(t)} z(d)^{(t)} + d\hat{\beta}_{ML}\}$$

where $z(d)^{(t)} = X\hat{\beta}(d)^{(t)} + D\left(y - \mu^{(t)}\right)$.

ii) If Σ tends to ∞ then, the SRL estimator approaches to the restricted Liu estimator introduced by (Kurtoğlu and Özkale, 2019b) which is

$$\begin{aligned}
\hat{\beta}_r(d)^{(t+1)} &= \left(X^T W^{(t)} X + I\right)^{-1} \left[X^T W^{(t)} z_r(d)^{(t)} + d\hat{\beta}_{ML}\right] - \left(X^T W^{(t)} X + I\right)^{-1} \\
&\quad \times R^T \{R\left(X^T W^{(t)} X + I\right)^{-1} R^T\}^{-1} \{R\left(X^T W^{(t)} X + I\right)^{-1} \\
&\quad \times \left[X^T W^{(t)} z_r(d)^{(t)} + d\hat{\beta}_{ML}\right] - r\}
\end{aligned}$$

where $z_r(d)^{(t)} = X\hat{\beta}_r(d)^{(t)} + D\left(y - \mu^{(t)}\right)$.

iii) As d goes to 1 and no restriction then, the SRL estimator approximately goes to the ML estimator:

$$\hat{\beta}^{(t+1)} = \left(X^T W^{(t)} X\right)^{-1} X^T W^{(t)} z^{(t)}$$

where $z^{(t)} = X\hat{\beta}^{(t)} + D\left(y - \mu^{(t)}\right)$.

3.3.1. FOA Stochastic Restricted Liu Estimator

The FOA form of the SRL estimator can be written from Eq. (3.35.) as

$$\begin{aligned} \hat{\beta}_{c-l}(d)^{(1)} &= \left[(X^T W^{(0)} X + I) + R^T \Sigma^{-1} R\right]^{-1} [(X^T W^{(0)} X + I)\hat{\beta}_{c-l}(d)^{(0)} \\ &\quad + X^T W^{(0)} D(y-\mu) - \hat{\beta}_{c-l}(d)^{(0)} + d\hat{\beta}^{(1)} + R^T \Sigma^{-1} r] \end{aligned}$$

where $\hat{\beta}^{(1)} = \left(X^T W^{(0)} X\right)^{-1} X^T W^{(0)} z^{(0)}$ is the FOA ML estimator with $z^{(0)} = X\beta^{(0)} + D\left(y - \mu^{(0)}\right)$ is the initial working response and $\beta^{(0)}$ can be used for both $\hat{\beta}_{c-l}(d)^{(0)}$ and $\hat{\beta}^{(1)}$ as an initial value of β.

The $\hat{\beta}_{c-l}(d)^{(1)}$ can be further simplified as

$$\hat{\beta}_{c-l}(d)^{(1)} = \left[(X^T W^{(0)} X + I) + R^T \Sigma^{-1} R\right]^{-1} [X^T W^{(0)} z^{(0)} + R^T \Sigma^{-1} r + d\hat{\beta}^{(1)}].$$

3.3.2. MSE of FOA Stochastic Restricted Liu Estimator

Since the weight matrices and the working responses of the estimators are changed at each iteration thus, we give the approximated matrix MSE of the FOA SRL estimator.

To find the bias and the variance of $\hat{\beta}_{c-l}(d)^{(1)}$ we utilize from $\hat{\beta}^{(1)} \sim$

$\left(\beta, a(\phi)\left(X^TWX\right)^{-1}\right)$ and $r \sim (R\beta, a(\phi)\Sigma)$.

The bias and variance of $\hat{\beta}_{c-l}(d)^{(1)}$ are obtained as

$$\begin{aligned}
Bias\left(\hat{\beta}_{c-l}(d)^{(1)}\right) &= E(\hat{\beta}_{c-l}(d)^{(1)}) - \beta \\
&= \left[(X^TW^{(0)}X + I) + R^T\Sigma^{-1}R\right]^{-1}[(X^TW^{(0)}X + R^T\Sigma^{-1}R)\beta \\
&\quad + d\beta] - \beta \\
&= \left[(X^TW^{(0)}X + I) + R^T\Sigma^{-1}R\right]^{-1}[(X^TW^{(0)}X + I + R^T\Sigma^{-1}R)\beta \\
&\quad + d\beta - \beta] - \beta \\
&= (d-1)\left[\left(X^TW^{(0)}X + I\right) + R^T\Sigma^{-1}R\right]^{-1}\beta.
\end{aligned}$$

where $E(z^{(0)}) = X\beta$ and

$$\begin{aligned}
var\left(\hat{\beta}_{c-l}(d)^{(1)}\right) &= a(\phi)\left[\left(X^TW^{(0)}X + I\right) + R^T\Sigma^{-1}R\right]^{-1}\{\left(X^TW^{(0)}X + dI\right) \\
&\quad \times \left(X^TW^{(0)}X\right)^{-1}\left(X^TW^{(0)}X + dI\right) + R^T\Sigma^{-1}R\} \\
&\quad \times \left[\left(X^TW^{(0)}X + I\right) + R^T\Sigma^{-1}R\right]^{-1}
\end{aligned}$$

Then, the matrix MSE of $\hat{\beta}_{c-l}(d)^{(1)}$ is obtained as

$$\begin{aligned}
MSE\left(\hat{\beta}_{c-l}(d)^{(1)}\right) &= \left[\left(X^TW^{(0)}X + I\right) + R^T\Sigma^{-1}R\right]^{-1}[a(\phi)\{\left(X^TW^{(0)}X + dI\right) \\
&\quad \times \left(X^TW^{(0)}X\right)^{-1}\left(X^TW^{(0)}X + dI\right) + R^T\Sigma^{-1}R\} \\
&\quad + (d-1)^2\beta\beta^T]\left[\left(X^TW^{(0)}X + I\right) + R^T\Sigma^{-1}R\right]^{-1}. \quad (3.36.)
\end{aligned}$$

3.3.3. Selection of the Tuning Parameter d through C_p Statistic

This section describes how to choose the tuning parameter d for the SRL estimator. The choice of the tuning parameter d is vital regarding the performance of the SRL estimator. There are numerous strategies available, that might be used for the selection of the parameter d. One of the fruitful strategy is to choose a criterion based on the C_p statistic presented by (Mallows, 1973). Many researchers have used the C_p statistic. For instance, (Liu, 1993) gave the C_p statistic for the Liu estimator in LRMs, (Smith and Marx, 1990) presented a C_p statistic for generalized ridge regression in GLMs and (Özkale, 2019) has also used the C_p statistic for the $r-d$ class estimator in GLMs. The C_p statistic for the FOA SRL estimator in GLMs can be defined as

$$C_p^{c-l(d)} = \frac{D_{c-l,r,d}}{D_{c-l,r,d=1}} - n + 2tr(H_{c-l}(d))$$

where $D_{c-l,r,d}$ is the deviance for the model with FOA SRL estimator, $D_{c-l,r,d=1}$ is an estimate of the scale parameter with $d=1$ and $H_{c-l}(d)$ is an approximated quasi projection matrix of the FOA SRL estimator.

$D_{c-l,r,d}$ is defined as $D_{c-l,r,d} = 2\sum_{i=1}^{n}\{y_i(\tilde{\theta}_i - \hat{\theta}_i) - [b(\tilde{\theta}) + b(\hat{\theta}_i)]\}$ where $\hat{\theta}_i$ denotes the estimated mean evaluated at x_i using $\hat{\beta}_{c-l}(d)^{(1)}$ and $\tilde{\theta}_i$ is an estimate of the natural parameter as a function of $\hat{\theta}_i$. $H_{c-l}(d)$ is defined by following (Özkale, 2015) for the particular case when $r=0$ and we give it as

$$H_{c-l}(d) = X\left[(X^TW^{(0)}X + I) + R^T\Sigma^{-1}R\right]^{-1}\left(X^TW^{(0)}X\right)^{-1}(X^TW^{(0)}X + dI)X^TW^{(0)}.$$

3.4. Stochastic Restricted OK Estimator in GLMs

In this section, we develop an iterative stochastic restricted OK estimator by applying stochastic restrictions on the parameters described in Section 3.3.

Thus, in order to develop an iterative stochastic restricted OK estimator in GLMs we consider the log-likelihood function given in Section 3.3., such as

$$\begin{aligned} l(\theta,\beta;y,r) &= \sum_{i=1}^{n}\left\{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i,\phi)\right\} - \frac{q}{2}\ln(2\pi) - \frac{q}{2}\ln|a(\phi)\Sigma| \\ &\quad - \frac{1}{2a(\phi)}(r-R\beta)^T\Sigma^{-1}(r-R\beta). \end{aligned} \tag{3.37.}$$

In addition, we also look for an estimator which has log-likelihood function given by Eq. (3.37.) in an equivalence class of estimators of β which are equal distance from $d\hat{\beta}_{ML}$ that is

$$p(\beta,d) = \frac{1}{2a(\phi)}\left(\beta - d\hat{\beta}_{ML}\right)^T\left(\beta - d\hat{\beta}_{ML}\right).$$

Thus, to find the stochastic restricted OK estimator in GLMs we combine the two log-likelihood functions and write the objective function as:

$$\begin{aligned} l(\theta,\beta,k,d;y,r) &= l(\theta,\beta;y,r) + p(\beta,d) \\ &= \sum_{i=1}^{n}\left\{\frac{\theta_i y_i - b(\theta_i)}{a(\phi)} + c(y_i,\phi)\right\} - \frac{k}{2a(\phi)}\sum_{i=1}^{p}(\beta_j - d\hat{\beta}_{MLj})^2 \\ &\quad - \frac{q}{2}\ln(2\pi) - \frac{q}{2}\ln|a(\phi)\Sigma| - \frac{1}{2a(\phi)}(r-R\beta)^T\Sigma^{-1}(r-R\beta) \end{aligned}$$

where the lagrange multiplier $\frac{k}{2} > 0$ and $0 < d < 1$ are the tuning parameters.

By taking the derivatives of $l(\theta,\beta,k,d;y,r)$ with respect to β_j and applying the chain rule we have:

$$\begin{aligned}\frac{\partial}{\partial\beta_j}l(\theta,\beta,k,d;y,r) &= \sum_{i=1}^{n}\frac{(y_i-\mu_i)}{a(\phi)b''(\theta_i)}\frac{1}{g'(\mu_i)}x_{ij}-\frac{k}{a(\phi)}\sum_{i=1}^{p}(\beta_j-d\hat{\beta}_{MLj})+\frac{1}{a(\phi)}T_j\\ &= \sum_{i=1}^{n}\frac{(y_i-\mu_i)}{a(\phi)var(y_i)}\frac{g'(\mu_i)}{(g'(\mu_i))^2}x_{ij}-\frac{k}{a(\phi)}\sum_{j=1}^{p}\beta_j+\frac{kd}{a(\phi)}\sum_{j=1}^{p}\hat{\beta}_{MLj}\\ &+ \frac{1}{a(\phi)}T_j\\ &= \sum_{i=1}^{n}\frac{(y_i-\mu_i)}{a(\phi)w_{ii}}g'(\mu_i)x_{ij}-\frac{k}{a(\phi)}\sum_{j=1}^{p}\beta_j+\frac{kd}{a(\phi)}\sum_{j=1}^{p}\hat{\beta}_{MLj}+\frac{1}{a(\phi)}T_j\end{aligned} \tag{3.38.}$$

where T_j are the elements of p vector $T=R^T\Sigma^{-1}(r-R\beta)$ (Özkale and Nyquist, 2021). In matrix notation Eq. (3.38.) can be written as

$$S(\beta,k,d)=\frac{1}{a(\phi)}[X^TWD(y-\mu)-k\beta+kd\hat{\beta}_{ML}+R^T\Sigma^{-1}(r-R\beta)]$$

where $W=diag(w_{ii}^{-1})$ and $D=diag(g'(\mu_i))$.

Now consider derivatives of Eq. (3.38.) with respect to β_k we get

$$\begin{aligned}\frac{\partial^2}{\partial\beta_j\partial\beta_k}l(\theta,\beta;y,r,k,d) &= \sum_{i=1}^{n}(y_i-\mu_i)\frac{\partial}{\partial\beta_k}\frac{1}{a(\phi)w_{ii}}g'(\mu_i)x_{ij}+\sum_{i=1}^{n}\frac{g'(\mu_i)x_{ij}}{a(\phi)w_{ii}}(\frac{-1}{g'(\mu_i)}x_{ik})\\ &\quad -\frac{k}{a(\phi)}\delta_{jk}-\frac{1}{a(\phi)}u_{jk}\\ &= \sum_{i=1}^{n}(y_i-\mu_i)\frac{\partial}{\partial\beta_k}\frac{1}{a(\phi)w_{ii}}g'(\mu_i)x_{ij}-\sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{a(\phi)w_{ii}}-\frac{1}{a(\phi)}k\delta_{jk}\\ &\quad -\frac{1}{a(\phi)}u_{jk}\end{aligned} \tag{3.39.}$$

since $\delta_{jk}=1$, if $j=k$, and 0 otherwise and u_{jk} are the components of $p\times p$ matrix $U=R^T\Sigma^{-1}R$.

Minus times the expected value of second derivative gives

$$Q_{jk}(\beta,k,d) = -E\left[\frac{\partial^2}{\partial\beta_j\partial\beta_k} l(\beta,y,r,k,d)\right] = \frac{1}{a(\phi)}\left[\sum_{i=1}^{n}\frac{x_{ij}x_{ik}}{w_{ii}} + kI_p + u_{jk}\right].$$

In matrix notation it might be written as

$$Q(\beta,k,d) = \frac{1}{a(\phi)}[(X^T WX + kI_p) + R^T\Sigma^{-1}R].$$

By following the technique of Fisher's scoring algorithm we have

$$\hat{\beta}^{(t+1)}_{kd-sr} = \hat{\beta}^{(t)}_{kd-sr} + \left\{[Q(\beta,k,d)]^{-1}[S(\beta,k,d)]\right\}_{\beta=\hat{\beta}^{(t)}_{kd-sr}}.$$

Premultiplying both sides by $Q(\beta, k, d)$ gives us

$$[Q(\beta,k,d)]_{\beta=\hat{\beta}^{(t)}_{kd-sr}}\hat{\beta}^{(t+1)}_{kd-sr} = [Q(\beta,k,d)]_{\beta=\hat{\beta}^{(t)}_{kd-sr}}\hat{\beta}^{(t)}_{kd-sr} + [S(\beta,k,d)]_{\beta=\hat{\beta}^{(t)}_{kd-sr}}.$$

Now replacing the values of $Q(\beta,k,d)$ and $S(\beta,k,d)$ we obtained

$$\begin{aligned}\frac{1}{a(\phi)}[(X^T\hat{W}^{(t)}_{kd-sr}X + kI_p) + R^T\Sigma^{-1}R]\hat{\beta}^{(t+1)}_{kd-sr} &= \frac{1}{a(\phi)}\{[(X^T\hat{W}^{(t)}_{kd-sr}X + kI_p) + R^T\Sigma^{-1}R]\hat{\beta}^{(t)}_{kd-sr}\\ &\quad + X^T\hat{W}^{(t)}_{kd-sr}D(y-\hat{\mu}^{(t)}_{kd-sr}) - k\hat{\beta}^{(t)}_{kd-sr}\\ &\quad + kd\hat{\beta}_{ML} + R^T\Sigma^{-1}(r - R\hat{\beta}^{(t)}_{kd-sr})\}_{\beta=\hat{\beta}^{(t)}_{kd-sr}}\end{aligned}$$

where t manifests the iteration step and $\hat{W}^{(t)}_{kd-sr}$ is a weight matrix computed at $\hat{\beta}^{(t)}_{kd-sr}$.

This leads to:

$$\begin{aligned}\hat{\beta}_{kd-sr}^{(t+1)} &= [(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)+R^T\Sigma^{-1}R]^{-1}\{X^T\hat{W}_{kd-sr}^{(t)}X\hat{\beta}_{kd-sr}^{(t)} \\ &\quad +X^T\hat{W}_{kd-sr}^{(t)}D(y-\hat{\mu}_{kd-sr}^{(t)})+kd\hat{\beta}_{ML}+R^T\Sigma^{-1}r\}.\end{aligned} \tag{3.40.}$$

By applying the inverse formula we get:

$$\begin{aligned}[(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)+R^T\Sigma^{-1}R]^{-1} &= (X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}-(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1} \\ &\quad \times R^T[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1} \\ &\quad \times R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}.\end{aligned}$$

Then, Eq. (3.40.) transformed to:

$$\begin{aligned}\hat{\beta}_{kd-sr}^{(t+1)} &= (X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}\{X^T\hat{W}_{kd-sr}^{(t)}X\hat{\beta}_{kd-sr}^{(t)}+X^T\hat{W}_{kd-sr}^{(t)}D(y-\hat{\mu}_{kd-sr}^{(t)}) \\ &\quad +kd\hat{\beta}_{ML}\}-(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1} \\ &\quad \times R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}\{X^T\hat{W}_{kd-sr}^{(t)}X\hat{\beta}_{kd-sr}^{(t)}+X^T\hat{W}_{kd-sr}^{(t)}D(y-\hat{\mu}_{kd-sr}^{(t)}) \\ &\quad +kd\hat{\beta}_{ML}\}+\{(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}-(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T \\ &\quad \times[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1}R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}\}R^T\Sigma^{-1}r.\end{aligned} \tag{3.41.}$$

Solving:

$$\begin{aligned}&\{(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}-(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T \\ &\times[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1}R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}\}R^T\Sigma^{-1} \\ = \;& (X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T\{I-[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1} \\ &\times R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T\}\Sigma^{-1} \\ = \;& (X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}_{kd-sr}^{(t)}X+kI_p)^{-1}R^T]^{-1}.\end{aligned}$$

Thus Eq. (3.41.) turns into:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-sr} &= (X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-sr}X\hat{\beta}^{(t)}_{kd-sr}+X^T\hat{W}^{(t)}_{kd-sr}D(y-\hat{\mu}^{(t)}_{kd-sr})\\ &\quad +kd\hat{\beta}_{ML}\}-(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T]^{-1}\\ &\quad \times R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\{X^T\hat{W}^{(t)}_{kd-sr}X\hat{\beta}^{(t)}_{kd-sr}+X^T\hat{W}^{(t)}_{kd-sr}D(y-\hat{\mu}^{(t)}_{kd-sr})\\ &\quad +kd\hat{\beta}_{ML}\}+(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T]^{-1}r.\end{aligned}$$

On further simplification we get

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-sr} &= (X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-sr}\{X\hat{\beta}^{(t)}_{kd-sr}+D(y-\hat{\mu}^{(t)}_{kd-sr})\}\\ &\quad +kd\hat{\beta}_{ML}]-(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T]^{-1}\\ &\quad \times R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\{R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\\ &\quad \times[X^T\hat{W}^{(t)}_{kd-sr}\{X\hat{\beta}^{(t)}_{kd-sr}+D(y-\hat{\mu}^{(t)}_{kd-sr})\}+kd\hat{\beta}_{ML}]-r\}.\end{aligned}$$

Thus, we obtained an iterative stochastic restricted OK estimator in GLMs as follows:

$$\begin{aligned}\hat{\beta}^{(t+1)}_{kd-sr} &= (X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}[X^T\hat{W}^{(t)}_{kd-sr}\hat{z}^{(t)}_{kd-sr}+kd\hat{\beta}_{ML}]-(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\\ &\quad \times R^T[\Sigma+R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}R^T]^{-1}\{R(X^T\hat{W}^{(t)}_{kd-sr}X+kI_p)^{-1}\\ &\quad \times[X^T\hat{W}^{(t)}_{kd-sr}\hat{z}^{(t)}_{kd-sr}+kd\hat{\beta}_{ML}]-r\}\end{aligned} \tag{3.42.}$$

where $\hat{z}^{(t)}_{kd-sr}=X\hat{\beta}^{(t)}_{kd-sr}+D(y-\hat{\mu}^{(t)}_{kd-sr})$ is the working response computed at $\hat{\beta}^{(t)}_{kd-sr}$. As t approaches to ∞ then, we acquired the asymptotic form of the iterative stochastic restricted OK estimator in GLMs as

$$\begin{aligned}\hat{\beta}_{kd-sr} &= (X^T\hat{W}_{kd-sr}X+kI_p)^{-1}[X^T\hat{W}_{kd-sr}\hat{z}_{kd-sr}+kd\hat{\beta}_{ML}]\\ &\quad -(X^T\hat{W}_{kd-sr}X+kI_p)^{-1}R^T[\Sigma+R(X^T\hat{W}_{kd-sr}X+kI_p)^{-1}R^T]^{-1}\\ &\quad \times\{R(X^T\hat{W}_{kd-sr}X+kI_p)^{-1}[X^T\hat{W}_{kd-sr}\hat{z}_{kd-sr}+kd\hat{\beta}_{ML}]-r\}\end{aligned} \tag{3.43.}$$

where $(X^T\hat{W}_{kd-sr}X+kI_p)^{-1}[X^T\hat{W}_{kd-sr}\hat{z}_{kd-sr}+kd\hat{\beta}_{ML}]$ is in the form of OK estimator established by (Özkale and Kaçıranlar, 2007), in the linear regression model with homoscedastic errors (also Özkale and Abbasi, 2022)[3]. The FOA form of the iterative stochastic restricted OK estimator is given as:

$$\begin{aligned}\hat{\beta}^{(1)}_{kd-sr} &= (X^TW^{(0)}X+kI_p)^{-1}[X^TW^{(0)}z^{(0)}+kd\hat{\beta}^{(1)}]-(X^TW^{(0)}X+kI_p)^{-1}\\ &\quad \times R^T[\Sigma+R(X^TW^{(0)}X+kI_p)^{-1}R^T]^{-1}\{R(X^TW^{(0)}X+kI_p)^{-1}\\ &\quad \times[X^TW^{(0)}z^{(0)}+kd\hat{\beta}^{(1)}]-r\}\end{aligned}$$

where $z^{(0)} = X\beta^{(0)} + D(y-\mu^{(0)})$ is the initial working response and $\hat{\beta}^{(1)} = (X^TW^{(0)}X)^{-1}X^TW^{(0)}z^{(0)}$ is the FOA ML estimator. Since $\beta^{(0)}$ is an initial value for both $\hat{\beta}^{(1)}_{kd-sr}$ and $\hat{\beta}^{(1)}$, $\hat{\beta}^{(1)}_{kd-sr}$ can be written as:

$$\begin{aligned}\hat{\beta}^{(1)}_{kd-sr} &= (X^TW^{(0)}X+kI_p)^{-1}(X^TW^{(0)}X+kdI_p)\hat{\beta}^{(1)}-(X^TW^{(0)}X+kI_p)^{-1}\\ &\quad \times R^T[\Sigma+R(X^TW^{(0)}X+kI_p)^{-1}R^T]^{-1}\{R(X^TW^{(0)}X+kI_p)^{-1}\\ &\quad \times(X^TW^{(0)}X+kdI_p)\hat{\beta}^{(1)}-r\}.\end{aligned}$$

[3] Since there are too many two-parameter estimators in the literature, (Guber, 2012) suggested abbreviating this estimator as OK estimator to distinguish the estimator defined by (Özkale and Kaçıranlar, 2007) from the other estimators in the literature, so that OK is the abbreviation of the authors' names who introduced it.

This implies that

$$\hat{\beta}^{(1)}_{kd-sr} = M^{(0)}_k S^{(0)}_{kd} \hat{\beta}^{(1)} + (S^{(0)}_k)^{-1} R^T [\Sigma + R(S^{(0)}_k)^{-1} R^T]^{-1} r$$

where $M^{(0)}_k = (S^{(0)}_k)^{-1} - (S^{(0)}_k)^{-1} R^T [\Sigma + R(S^{(0)}_k)^{-1} R^T]^{-1} R(S^{(0)}_k)^{-1}$, $S^{(0)}_k = X^T W^{(0)} X + kI_p$, $S^{(0)}_{kd} = X^T W^{(0)} X + kdI_p$.

We observed that newly developed stochastic restricted OK estimator is a manifold estimator, which contains the different estimators in GLMs under the particular conditions on R, d, and k if the working responses and weight matrices of the estimators are identical then, the following estimators are obtained:

i) If $k = 0$ and no restrictions then, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}^{(1)}$

ii) If $k = 1$ and no restrictions $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}^{(1)}_d$ (FOA Liu estimator by Kurtoğlu and Özkale, 2016)

iii) If $k = 1$ and Σ goes to 0, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}_r(d)^{(1)}$ (FOA restricted Liu estimator by Kurtoğlu and Özkale, 2019b)

iv) If $d = 0$, Σ goes to 0, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}_r(k)^{(1)}$ (FOA restricted ridge estimator by Kurtoğlu and Özkale, 2019a)

v) If $d = 0$ then, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}^{(1)}_{c-r}$ (FOA stochastic restricted ridge estimator by Özkale and Nyquist, 2021)

vi) If $k = 0$ then, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}^{(1)}_c$ (FOA mixed estimator by Özkale and Nyquist, 2021)

vii) If $k = 1$ then, $\hat{\beta}^{(1)}_{kd-sr} \to \hat{\beta}_{c-l}(d)^{(1)}$ (FOA stochastic restricted Liu estimator by Abbasi and Özkale, under review)

3.4.1. MSE of FOA Stochastic Restricted OK Estimator

This section gives the matrix MSE of the stochastic restricted OK estimator in its FOA form. Thus, by definition matrix MSE of $\hat{\beta}^{(1)}_{kd-sr}$ is defined as:

$$MSE\left(\hat{\beta}^{(1)}_{kd-sr}\right) = var(\hat{\beta}^{(1)}_{kd-sr}) + [Bias(\hat{\beta}^{(1)}_{kd-sr})][Bias(\hat{\beta}^{(1)}_{kd-sr})]^T.$$

To find the variance and bias of $\hat{\beta}^{(1)}_{kd-sr}$ we use that asymptotically $\hat{\beta}^{(1)} \sim N(\beta, a(\phi)(X^T W^{(0)} X)^{-1})$ and $r \sim N(R\beta, a(\phi)\Sigma)$. Then, we get:

$$\begin{aligned} var(\hat{\beta}^{(1)}_{kd-sr}) &= a(\phi)M_k^{(0)} S_{kd}^{(0)} (S^{(0)})^{-1} S_{kd}^{(0)} M_k^{(0)} + (S_k^{(0)})^{-1} R^T [\Sigma + R(S_k^{(0)})^{-1} R^T]^{-1} \\ &\quad \times \Sigma[\Sigma + R(S_k^{(0)})^{-1} R^T]^{-1} R(S_k^{(0)})^{-1} \end{aligned}$$

where $S^{(0)} = X^T W^{(0)} X$. The bias of $\hat{\beta}^{(1)}_{kd-sr}$ is found as:

$$\begin{aligned} Bias(\hat{\beta}^{(1)}_{kd-sr}) &= M_k^{(0)} S_{kd}^{(0)} \hat{\beta}^{(1)} + (S_k^{(0)})^{-1} R^T [\Sigma + R(S_k^{(0)})^{-1} R^T]^{-1} R\beta - \beta \\ &= k(d-1) M_k^{(0)} \beta. \end{aligned}$$

Thus, the matrix MSE of $\hat{\beta}^{(1)}_{kd-sr}$ is given by

$$\begin{aligned} MSE(\hat{\beta}^{(1)}_{kd-sr}) &= a(\phi)M_k^{(0)} S_{kd}^{(0)} (S^{(0)})^{-1} S_{kd}^{(0)} M_k^{(0)} + (S_k^{(0)})^{-1} R^T [\Sigma + R(S_k^{(0)})^{-1} \\ &\quad \times R^T]^{-1} \Sigma[\Sigma + R(S_k^{(0)})^{-1} R^T]^{-1} R(S_k^{(0)})^{-1} + k^2(d-1)^2 M_k^{(0)} \\ &\quad \times \beta\beta^T M_k^{(0)} \qquad (3.44.) \end{aligned}$$

3.4.2. Test Statistic for Testing Compatibility of Prior and Sample Information in GLMs

Here we demonstrate the test statistic for testing the compatibility of prior and sample information for the stochastic restricted OK estimator in GLMs. By following (Theil, 1963, Toutenburg, 1982, and Özkale and Nyquist, 2021), we devised a test for the hypothesis

H_0: Prior and sample information are in agreement

H_1:Prior and sample information are not in agreement

If the null hypothesis is true, then the difference is $\delta = r - R\hat{\beta}_{kd-sr}$ which might be near to 0. Since we have

$$\hat{\beta}_{kd-sr} = M_k S_{kd} \hat{\beta}_{ML} + (S_k)^{-1} R^T [\Sigma + R(S_k)^{-1} R^T]^{-1} r$$

we get

$$E(\delta) = k(1-d) R M_k \beta$$

and under the assumption that r and β are independent

$$\begin{aligned} var(\delta) &= \Sigma\left(\Sigma + R(S_k)^{-1}R^T\right)^{-1}\left[\Sigma + R(S_k)^{-1} S_{kd} var(\hat{\beta}_{ML}) S_{kd} (S_k)^{-1} R^T\right] \\ &\quad \times \left(\Sigma + R(S_k)^{-1}R^T\right)^{-1}\Sigma. \end{aligned}$$

As $var(\delta)$ is positive definite, we can write $var(\delta) = z^{1/2}z^{1/2} = z$. Thus, by following (Toutenburg, 1982) we find that: $\delta^T z^{-1/2} \sim \left(k(1-d)\beta^T M_k R^T z^{-1/2}, I_p\right)$

Moreover, under the normality condition it can be written as:

$$\delta^T z^{-1/2} \sim N\left(k(1-d)\beta^T M_k R^T z^{-1/2}, I_p\right).$$

Therefore, we define a test statistic

$$\begin{aligned} \gamma &= \delta^T z^{-1/2} z^{-1/2} \delta \\ &= \delta^T z^{-1} \delta. \end{aligned}$$

Since $\hat{\beta}_{kd-sr}$ is a biased estimator $\delta^T z^{-1}\delta$ follows a noncentral χ^2 distribution with p degrees of freedom and noncentrality parameter $\lambda = \frac{k^2(1-d)^2(RM_k\beta)^T z^{-1} RM_k\beta}{2}$.

If Σ is unknown we can use its estimator $\hat{\Sigma}$ for practical purpose and we reject H_0 if $\hat{\gamma} \geq \chi^2_{p,\lambda,(1-\alpha)}$, where α is the level of significance and p is the number of parameters.

4. NUMERICAL EXAMPLES

In this chapter, various applications of the proposed estimators to real data sets are provided to compare the performance of the newly developed estimators with already existing estimators. A SMSE criterion is used to assess the performance of the estimators. The MATLAB and R programming languages are used to obtain the results.

4.1. Swedish Football Data Set Poisson Response

In this section, we describe the Swedish football data set which was originally used by (Qasim et al., 2019). There are 242 observations in this data set and two response variables that are: the number of full-time home team goals (y_1) and the number of full-time away team goals (y_2), and six explanatory variables, which are pinnacle home win odds (x_1), pinnacle away win odds (x_2), maximum oddsportal maximum home win (x_3), oddsportal maximum away win (x_4), average oddsportal home win (x_5), and average oddsportal away win (x_6). Qasim et al. (2019) showed that each of the response variable follows a Poisson distribution. Therefore, the log link function is to be considered and the regression model is

$$\log(\hat{\mu}_i) = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\beta}_4 x_{4i} + \hat{\beta}_5 x_{5i} + \hat{\beta}_6 x_{6i}.$$

Our computations are carried out by using MATLAB programming language.

Before computations, we standardized the explanatory variables by unit length standardization to avoid any difficulty regarding comparisons of the regression estimates. As a result of standardization the intercept term is also vanished. Therefore, the standardization is also helpful in reducing an ill-conditioned situation which may arises because of the intercept term and the standardizing also increases the speed of convergence. The standardization is also very useful in order to achieve the desired convergence rapidly. The number of full-time home team goals (y_1) is considered as a response variable. We first obtained the ML estimator by an iterative procedure and used the convergence criterion 1×10^{-6} as a rule of thumb, which is close enough to zero. The iteration carries on until the absolute value of the difference of parameter estimates smaller than 1×10^{-6} is achieved. As an initial estimate the OLS is used; that is $\hat{\beta}_{ols} = \left(X^T X\right)^{-1} X^T y$ having values $\hat{\beta}_{ols} = [13.8992, 2.0119, -18.8602, -20.2691, 3.0540, 24.1594]^T$. After obtaining the ML estimator the collinearity diagnostics are applied by following (Özkale, 2021), who investigated the effects of centering and scaling the information matrix on the sensitivity of the collinearity diagnostics. She found that when the columns of $X^* = \hat{W}_{ML}^{1/2} X$ have been centered and scaled where $\hat{W}_{ML}$ is the weight matrix at final iteration of ML estimation, the collinearity in such a coordinate system is meaningful. Therefore, after getting $\hat{W}_{ML}$ we centered and scaled the columns of X^*. Following (Mackinnon and Puterman, 1989) and (Weissfeld and Sereika, 1991), the condition number $\varsigma_{\max} = \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2}$ is computed as $\kappa = (4623.714059)^{1/2} = 67.997897$, which is large enough to confirm that there exists a strong multicollinearity problem among the explanatory variables. Here, $\lambda_{\max}$ and $\lambda_{\min}$ denote respectively the maximum and minimum eigenvalues of $X^{*T} X^*$ in correlation form. Furthermore, the values of VIF are obtained as 357.9890, 90.7073, 596.7983, 205.6173, 432.4859, and 211.3393, which are all greater than 10 and revealed that there exists a severe problem of multicollinearity among all the explanatory variables.

Thus, to tackle the problem of multicollinearity we apply the $r-k$ class estimator to the Swedish football data set in the following section.

4.1.1. Application of $r-k$ Class Estimator to Swedish Football Data Set

In this section, we illustrate the application of $r-k$ class estimator as well as ML, ridge, and PCR estimators to a Swedish football data set. The $r-k$ class estimator is compared to the ML, ridge, and PCR estimators to evaluate which performs better when dealing with multicollinearity problems. The performance is evaluated by a SMSE criterion. By using the PTV method given in subsection 2.4., the number of PCs retained in the model is $r = 2$, which explained almost 95% of the total variation. There is no hard and fast rule for selecting the cut-off point for the PTV value so, it is arbitrarily chosen as 0.95[4].

The value of k is found by using the formula $k_1 = \frac{r}{\hat{\beta}_{ols}^T M_r M_r^T \hat{\beta}_{ols}} = 0.0325$ and another value of k is also chosen arbitrarily as $k_2 = 0.0001$. The results of the parameter estimates are shown in Table 4.1., for both the values of k. Whereas, SMSE values of the estimators given in Table 4.1., are attained in their FOA form because of comparative purpose.

Table 4.1. Results of Swedish football data estimated coefficients and the SMSE values for k_1 and k_2

	$\hat{\beta}_{ML}$	$\hat{\beta}_r$	$\hat{\beta}(k_1)$	$\hat{\beta}_r(k_1)$	$\hat{\beta}(k_2)$	$\hat{\beta}_r(k_2)$
x_1	44.5967	-2.3055	-1.0360	-2.2614	39.3589	-2.3054
x_2	2.27943	0.7011	0.7788	0.7154	2.0959	0.7011
x_3	-60.9131	-2.2586	-3.3571	-2.2154	-52.2539	-2.2585
x_4	-11.6586	0.5974	-0.7965	0.6144	-11.0773	0.5974
x_5	9.3973	-2.3201	-2.5623	-2.2762	6.0965	-2.3199
x_6	12.3219	0.79224	2.1481	0.8042	11.8368	0.7923
SMSE	2975.6742	1546.0270	1299.9149	1545.9282	2336.2812	1546.0267

Table 4.1., shows the results of estimated coefficients and SMSE values of the estimators. By making comparisons of the estimators on the basis of SMSE values, it is clear that all of the biased estimators have smaller SMSE values than the ML

[4]0.90 and 0.99 also gave the same number of PCs for this data set

estimator so, it is evident that all the biased estimators are superior to the ML estimator particularly when there is a problem of multicollinearity. Moreover, it is perceived that the ridge estimator has a smaller SMSE value as compared to all other estimators for k_1. However, it is important to note that the value of k also has a significant impact on the performance of estimators. As it can be seen that for k_2 the proposed $r-k$ class estimator outperforms all of its competitors in terms of the SMSE criterion because it attains the smallest SMSE value when compared to its counterparts.

Table 4.1., presents the performance of estimators only for particular values of k as discussed above. To make comparisons for other values of k Figs. 4.1., and 4.2., are provided. Fig. 4.1., displays the performance of PCR and $r-k$ class estimators on the basis of SMSE for different values of k. It is examined that the $r-k$ class estimator outperforms the PCR estimator for small values of k such as $k \leq 2.5$. However, when the values of k are increased that is when k is greater than approximately 2.5 then, the PCR estimator performs better than the $r-k$ class estimator. Similarly, Fig. 4.2., expresses the performance of the ridge estimator and the ML estimator on the basis of SMSE criterion for different values of k. It is to be noted that the SMSE values of the ridge estimator increases as the k value increases. Nevertheless, regardless of the values of k, it is clear that the ridge estimator dominated the ML estimator comprehensively. Thus, the graphical results confirm the tabular results.

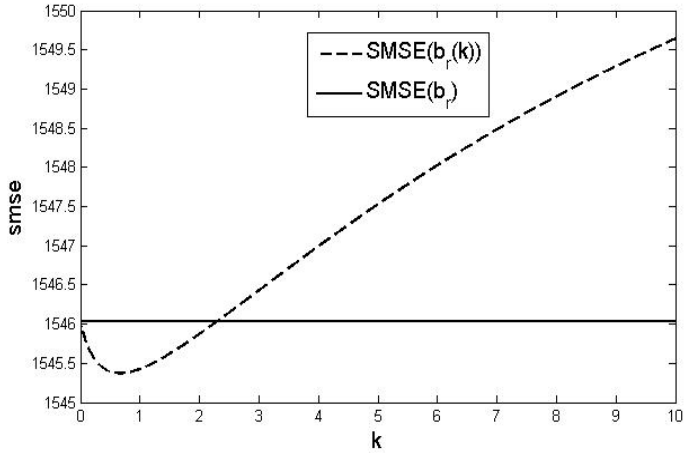


Figure 4.1. Graph of SMSE values against k for PCR and $r-k$ class estimators

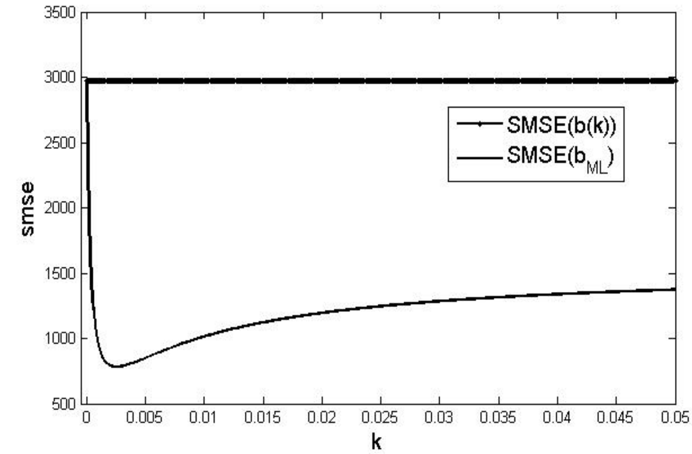


Figure 4.2. Graph of SMSE values against k for ML and ridge estimators

This numerical example indicates that there is always a prevailing value of k for which the proposed $r-k$ class estimator is superior to all of its counterparts and performs efficiently, particularly for small values of k in terms of SMSE criterion. However, for larger values of k results may be reversed.

4.1.2. Application of Restricted OK Estimator to Swedish Football Data Set

This section provides the application of the restricted OK estimator to a Swedish football data set as well as SMSE comparisons of the restricted OK estimator with ML, restricted ML, OK, Liu, restricted Liu, ridge, and restricted ridge estimators, respectively in GLMs. For comparisons, both the ML-based and FOA-based estimators are used. Before computing the results, explanatory variables are standardized through the unit length standardization technique and then, the intercept term is added to the model. The following log link function is used:

$$\log(\hat{\mu}_i) = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\beta}_4 x_{4i} + \hat{\beta}_5 x_{5i} + \hat{\beta}_6 x_{6i}.$$

Our computations are carried out by applying MATLAB and R programming languages. To acquire the results, iterative algorithm is applied by using the convergence criterion $||\tilde{\beta}^{(t+1)} - \tilde{\beta}^{(t)}|| < 1 \times 10^{-6}$, as far as the prescribed convergence criterion is satisfied this process continues. By considering the number of full-time home team goals (y_1) as a response variable first we acquired the ML estimator and considered the OLS estimator $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y_1$ as an initial estimate which comprises of the following values: $[1.4752, 13.8992, 2.0119, -18.8602, -20.2691, 3.0540, 24.1594]^T$. It is noticed that the ML estimator is obtained at the 5-th iteration.

After obtaining the ML estimator the collinearity diagnostics such as condition number and VIFs are performed to see the severity of multicollinearity after including the intercept term in the model. Since (Mackinnon and Puterman 1989, Weissfeld and Sereika, 1991, and Özkale, 2021) observed that collinearity is essential when the columns of $X^* = \hat{W}_{ML}^{1/2} X$ have been centered and scaled where $\hat{W}_{ML}$ is the weight matrix at the final iteration of ML estimator. Thus, the condition number is obtained as

$\zeta_{\max} = (\frac{\lambda_{\max}}{\lambda_{\min}})^{1/2} = (5.2952 \times 10^{+4})^{1/2} = 230.1127$ and VIFs are obtained as: VIFs: 148.3699, $3.1005 \times 10^{+3}$, 111.4931, $5.0295 \times 10^{+3}$, 672.3973, 367.5650, 710.5477 (by Özkale, 2021). The values of condition number and VIFs are sufficiently large to a certain nominal value and revealed that there exists a severe multicollinearity problem among the explanatory variables. To encounter the problem of multicollinearity we go on by biased estimation methods mentioned above.

The tuning parameters of the iterative OK estimator are optimized by using GA described in subsubsection 3.2.2.2., and its parameters are: number of chromosomes in the population, iterations (generations), elitism individuals, Crossover ratio and Mutation ratio are set to 100, 200, 10, 0.8, and 0.05, respectively. Our algorithms are carried out by using R programming language and GA package of (Scrucca, 2013) is used which contains already built-in commands in its library. The optimal $(k,\ d)$ pair obtained via GA by minimizing MSEP and MAE criteria are given as $(k_1, d_1) = (1.992158, 6.571125 \times 10^{-5})$ and $(k_2, d_2) = (1.992158, 6.571125 \times 10^{-5})$, respectively. As a third way to select (k, d) the method given in subsubsection 3.2.2.1., is used and gives $(k = k_{HM}, d) = (0.0083, 0.2522)$[5]. After the tuning parameters k and d are obtained for the iterative OK estimator by GA, they are fixed for the OK, ridge, and Liu estimators for both the restricted and unrestricted versions.

We imposed exact restrictions on the parameters to see its impact on the performance of the estimators. By considering the ML estimator at final iteration, we determined the restriction as: $\beta_0 + \beta_1 - 2\beta_6 = -0.0441$ which is approximately true $\hat{\beta}_{ML}$[6]. This restriction leads to R and r to be $R = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}$ and $r = [-0.0441]$. The estimated regression coefficients and the SMSE values of the

[5] Other (k, d) pairs are $(k_{AM}, d) = (5.6870, 0.8466)$ and $(k_{GM}, d) = (1.7808, 0.8448)$

[6] The method for obtaining optimum (k, d) given in subsubsection 3.2.2.1., is proposed when the restrictions are true. For that reason, we aim to use a restriction which is approximately true at least for $\hat{\beta}_{ML}$

estimators corresponding to the aforementioned optimum tuning parameter values are shown by Tables 4.2-4.3[7].

[7] Since (k_1,d_1) and (k_2,d_2) are same for Swedish football data, Table 4.10. is given for both (k_1,d_1) and (k_2,d_2)

Table 4.2. Estimates and SMSE when optimal $(k,d) = (0.0083, 0.2522)$ are obtained via MSE

FOA	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.7956	16.5551	1.8395	-22.8978	-17.6340	4.0531	20.9067	453.3039
Rest. ML	0.7952	29.4692	2.6001	−31.2725	−12.9404	−1.0437	15.1543	932.3543
Ridge	0.8015	6.1900	1.1127	-6.7961	-12.1253	-1.9813	15.7382	444.2457
Rest. Ridge	0.8032	15.9586	2.2901	−11.8069	−6.3316	−7.3345	8.3880	665.4275
Liu	0.8139	3.4604	1.3848	-6.6027	-3.7336	0.1433	6.7067	865.7525
Rest. Liu	0.8272	5.2938	2.9870	−7.4990	−1.9842	−0.8004	3.0825	724.9798
OK	0.8000	8.8038	1.2960	-10.8566	-13.5145	-0.4595	17.0416	347.1706
Rest. OK	0.8017	18.5403	2.4732	−15.8663	−7.7221	−5.8115	9.6930	600.1318
Iterative	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.3240	22.2255	1.5541	-27.9149	-10.2011	2.0604	11.2968	-
Rest. ML	0.3240	22.2255	1.5541	−27.9149	−10.2011	2.0604	11.2968	-
Ridge	0.3307	3.1136	0.8662	-4.3620	-3.5882	-2.6763	4.8956	-
Rest. Ridge	0.3309	4.8262	1.6381	−5.1793	−2.1719	−3.7123	2.6006	-
Liu	0.3434	4.9240	0.9311	-7.7700	-2.1007	-0.2221	3.4446	-
Rest. Liu	0.3475	5.2343	1.1822	−7.9018	−1.8459	−0.3609	2.8130	-
OK	0.3298	7.9283	1.0330	-10.2432	-5.2541	-1.5051	6.5231	-
Rest. OK	0.3299	9.2163	1.6086	−10.8686	−4.1814	−2.2787	4.7951	-
ML-based	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.3240	22.2255	1.5541	-27.9149	-10.2011	2.0604	11.2968	1572.7010
Rest. ML	0.3240	22.2255	1.5541	−27.9149	−10.2011	2.0604	11.2968	1244.3428
Ridge	0.3348	3.0626	0.8306	-4.0543	-3.5749	-2.7788	4.9609	1136.1243
Rest. Ridge	0.3347	4.8217	1.5944	−4.9487	−2.0840	−3.8141	2.6003	1074.5967
Liu	0.3480	4.9544	0.9387	-7.7299	-2.0905	-0.1861	3.4547	940.3761
Rest. Liu	0.3531	3.8285	1.1244	−6.6029	−3.8400	0.1771	5.1058	924.8459
OK	0.3321	7.8951	1.0130	-10.0714	-5.2459	-1.5585	6.5586	2781.8486
Rest. OK	0.3320	9.2106	1.5842	−10.7403	−4.1310	−2.3327	4.7933	772.1671

Table 4.3. Estimates and SMSE when optimal $(k,d) = (1.992158, 6.571125 \times 10^{-5})$ are obtained via GA by minimizing MSEP and MAE

FOA	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.7956	16.5551	1.8395	-22.8978	-17.6340	4.0531	20.9067	453.3039
Rest. ML	0.7952	29.4692	2.6001	−31.2725	−12.9404	−1.0437	15.1543	932.353
Ridge	0.8276	-0.9282	1.2748	-0.9978	1.1322	-1.0474	1.6311	1518.8218
Rest. Ridge	0.8363	−0.2593	1.8486	−1.2941	1.7332	−1.3586	0.3106	1574.7693
Liu	0.8200	-0.9542	1.2315	-1.1092	0.9526	-1.1749	1.9195	1493.5833
Rest. Liu	0.8258	−0.1622	1.9236	−1.4964	1.7084	−1.5825	0.3538	1554.4181
OK	0.8276	-0.9270	1.2748	-0.9992	1.1309	-1.0471	1.6324	1518.6224
Rest. OK	0.8363	−0.2578	1.8489	−1.2957	1.7322	−1.3584	0.3113	1574.6030
Iterative	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.3240	22.2255	1.5541	-27.9149	-10.2011	2.0604	11.2968	-
Rest. ML	0.3240	22.2255	1.5541	−27.9149	−10.2011	2.0604	11.2968	-
Ridge	0.3560	-0.7619	0.6833	-0.7965	0.6330	-0.8109	0.7234	-
Rest. Ridge	0.3631	−0.4099	0.9274	−0.9284	0.8790	−0.9491	−0.0014	-
Liu	0.3478	-0.9243	0.7171	-0.9990	0.6251	-1.0097	0.7923	-
Rest. Liu	0.3531	−0.5089	1.0537	−1.1773	0.9667	−1.1973	−0.0559	-
OK	0.3560	-0.7604	0.6834	-0.7983	0.6323	-0.8107	0.7241	-
Rest. OK	0.3631	−0.4084	0.9274	−0.9302	0.8783	−0.9489	−0.0006	-
ML-based	β_0	β_1	β_2	β_3	β_4	β_5	β_6	SMSE
ML	0.3240	22.2255	1.5541	-27.9149	-10.2011	2.0604	11.2968	1572.7010
Rest. ML	0.3240	22.2255	1.5541	−27.9149	−10.2011	2.0604	11.2968	1244.3428
Ridge	0.3669	-0.7001	0.6934	-0.7216	0.6454	-0.7409	0.7359	1503.0390
Rest. Ridge	0.3757	−0.3493	0.9330	−0.8397	0.8871	−0.8656	0.0353	1501.6367
Liu	0.3561	-0.8682	0.7312	-0.9250	0.6439	-0.9434	0.8108	1499.4278
Rest. Liu	0.3621	−0.4545	1.0649	−1.0953	0.9819	−1.1234	−0.0233	1497.6607
OK	0.3669	-0.6986	0.6934	-0.7234	0.6446	-0.7408	0.7366	1506.3268
Rest. OK	0.3757	−0.3478	0.9331	−0.8415	0.8864	−0.8654	0.0360	1501.4396

From Table 4.2., it is clear that the restricted FOA and ML-based OK estimator dominated its restricted FOA and ML-based counterparts as it achieves the smallest SMSE values for (k,d). Thereafter, FOA OK estimator acquires the smallest SMSE value among unrestricted FOA ML, FOA ridge, and FOA Liu estimators for (k,d), however, this changes in ML-based comparison and unrestricted ML-based Liu estimator has the smallest SMSE value. It is worthy to mention here that each of the restricted ML-based estimators acquires a smaller SMSE value than its corresponding unrestricted ML-based estimator which means that the restricted estimators perform efficiently as compared to unrestricted ones. But this case is not valid for the FOA estimators under the restriction considered. It is also seen that the coefficients of the ML-based and iteratively calculated estimators are different from each other.

The results in Table 4.3., do not confirm the results and comments in Table 4.2., such that biased estimation methods as well as the restricted estimators yield worse results than the ML method. The ML-based restricted estimators give a smaller SMSE value than their corresponding unrestricted estimators however it is not true for the FOA estimators.

Tables 4.2., and 4.3., show the performance of estimators only for particular values of tuning parameters (k,d) therefore, Figs. 4.3., and 4.4., are provided for the remaining values of the tuning parameters. Fig. 4.3., is set up to review the performance of estimators for various values of d and when k is fixed at its value which is obtained by minimizing MSE.

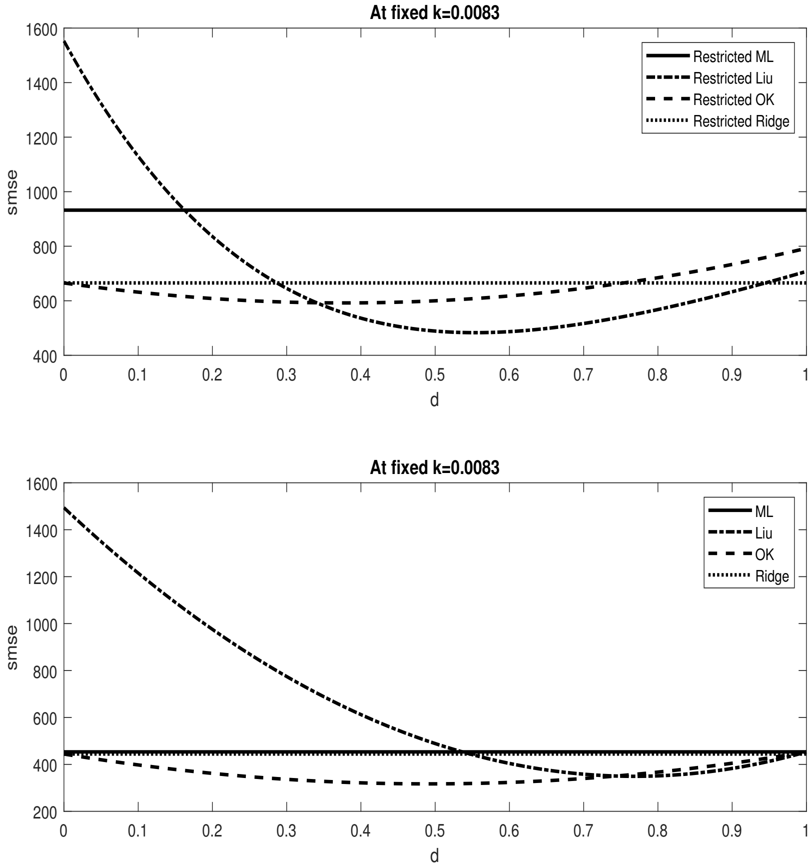


Figure 4.3. Graph of the SMSE values of the FOA estimators against d when k is fixed at 0.0083 (Swedish Football Data)

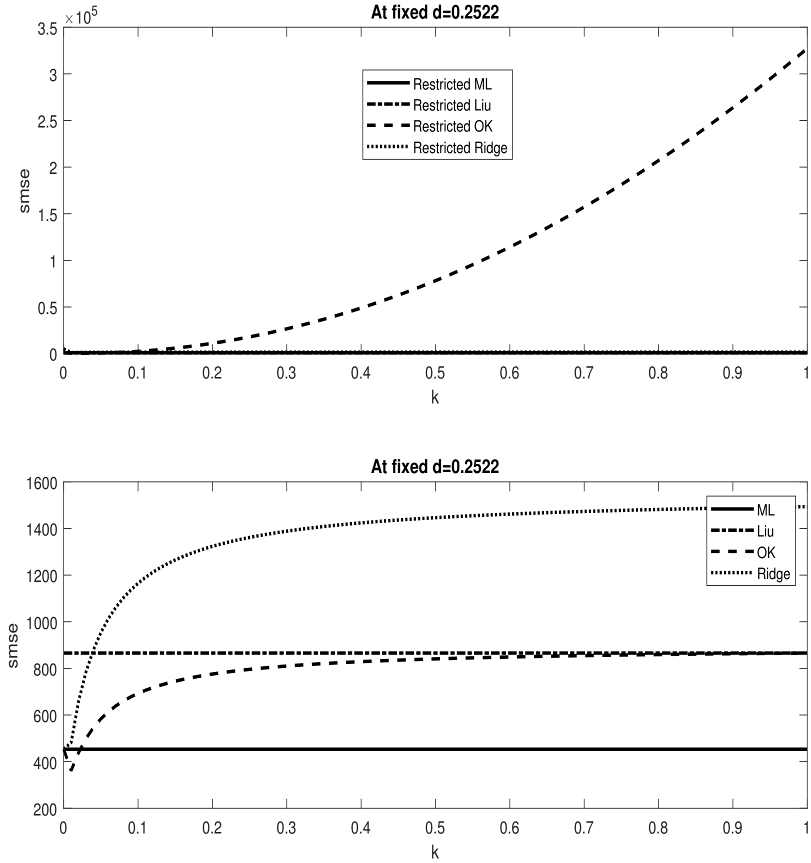


Figure 4.4. Graph of the SMSE values of the FOA estimators against k when d is fixed at 0.2522 (Swedish Football Data)

From Fig. 4.3., it is perceived that the restricted FOA OK estimator performs efficiently as compared to its counterparts when the values of d fall below approximately 0.34. However, when the values of d lie in the range of 0.34 to 0.95 then, the restricted Liu estimator performs better than the other estimators. The SMSE values of the restricted FOA OK estimator decrease and then increase as d increases like the restricted FOA Liu estimator. Furthermore, it is seen from Fig. 4.3., that the FOA OK estimator is better than its counterparts for $d < 0.74$ and $k = 0.0083$ while the FOA Liu estimator is better than its counterparts for $d > 0.74$ and $k = 0.0083$. Although the FOA OK estimator outperforms the FOA ML and ridge estimators for all d values the FOA Liu estimator is only good for d values greater than 0.55.

Similarly, Fig. 4.4., is established to figure out the performance of the estimators for various values of k when d is fixed at its value obtained via MSE. When the graph is examined in detail (not given here) it is observed that the restricted FOA Liu estimator dominates the restricted FOA ML estimator as in Table 4.2. Additionally, the restricted FOA ridge estimator stabilizes at small k values. The restricted FOA ridge estimator has larger SMSE values than the restricted FOA ML and Liu estimators for $d = 0.2522$ and k values lie in the interval of $(0,1)$. The SMSE values of the restricted FOA OK estimator decrease as k decreases and the restricted FOA OK estimator has a smaller SMSE value in the range of $0.01 < k < 0.05$ than the other estimators. On the other hand, Fig. 4.4., shows that the SMSE values of the FOA OK and ridge estimators increase as k increases, and then, the SMSE values of the FOA OK estimator are always less than the FOA ridge estimator. The FOA OK estimator is the best for $d = 0.2522$ and k values fall in the range of $(0,0.02)$ for Fig. 4.4. As k increases the FOA OK estimator gives results close to the FOA Liu estimator.

4.1.3. Application of Stochastic Restricted Liu Estimator to Swedish Football Data Set

In this section, we demonstrate the application of the SRL estimator on the Swedish football data set. The Liu and mixed estimators in GLMs proposed by (Kurtoğlu and Özkale, 2016) and (Özkale and Nyquist, 2021), respectively are also considered along with the ML estimator for the purpose of comparisons with the SRL estimator. This data set has a serious multicollinearity problem as stated before.

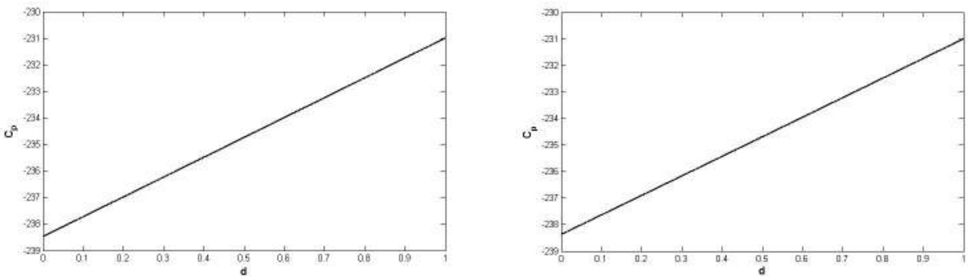
Thus, in order to deal with the multicollinearity problem we apply the SRL estimator estimator by applying stochastic restrictions on the parameters.

In order to see the effects of restrictions on the efficiency of the estimators two different types of restrictions are considered by following (Özkale and Nyquist, 2021) as:

(i) $\beta_2+\beta_4+\beta_5+G=0.02$, $G\sim(0,0.005^2)$, whereas in order to impose the stochastic restrictions on the parameters ML estimator is used and the principle of Chebyshev inequality is followed hence resulted as: $P(|\beta_2+\beta_4+\beta_5-0.02|<3\sigma)\geq 1-\frac{1}{3^2}=0.8888$. This implies that $\beta_2+\beta_4+\beta_5\in[0.005,0.035]$ with probability 0.8888 for $\sigma=0.005$, $r=0.02$ and the R matrix is $R=[0\ 1\ 0\ 1\ 1\ 0]$ while 0.005 is chosen to be the value of an estimate of $var(G)=\Sigma=\sigma^2 I$ corresponding to 0.8888 probability.

(ii) $\beta_2+\beta_3+\beta_5+4\beta_6+G=0.05$, $G\sim(0,0.004^2)$, $P(|\beta_2+\beta_3+\beta_5+4\beta_6-0.05|<3\sigma)\geq 0.8888$ and $\beta_2+\beta_3+\beta_5+4\beta_6\in[0.038,0.062]$ with probability 0.8888 for $\sigma=0.004$, $r=0.05$ and the R matrix is $R=[0\ 1\ 1\ 0\ 1\ 4]$ as well as 0.004 is selected to be the value of an estimate of $var(G)=\Sigma=\sigma^2 I$ correspond to 0.8888 probability. Furthermore, the value of the tuning parameter d is calculated by minimizing the $C_p^{c-l(d)}$ statistic. The deviance for the Poisson regression is evaluated as $D_{c-l,r,d}=2\sum\{y\log(y/\hat{\mu})-(y-\hat{\mu})\}$, where $\hat{\mu}$ is found by applying $\hat{\beta}_{c-l}(d)^{(1)}$. Fig. 4.5., is given to find the optimum tuning parameter value which results in minimum $C_p^{c-l(d)}$ statistic for restrictions (i) and (ii). Fig. 4.5., shows that $C_p^{c-l(d)}$ is an increasing function of d for this data set[8]. Therefore, we arbitrarily selected optimum d value as $\hat{d}_{C_p}=0.001$ for both restrictions. We also used the d estimator proposed by (Kurtoğlu and Özkale, 2016) for the Liu estimator in GLMs which gives $\hat{d}=0.5077$.

[8]The data set affects the tendency of the C_p statistic (Özkale, 2015)



(a) $C_p^{c-l(d)}$ versus d for restriction (i) (b) $C_p^{c-l(d)}$ versus d for restriction (ii)

Figure 4.5. $C_p^{c-l(d)}$ values of the FOA SRL estimator versus d (Swedish football data set for y_1) under restrictions (i) and (ii)

The results of parameter estimate along with the SMSE for the FOA estimators and reduction rate (RR) corresponding to restrictions (i) and (ii) are shown by Table 4.4., and Table 4.5., respectively. In order to measure the efficiency of the SRL estimator over its counterparts RR is used. The RR is evaluated corresponding to the SRL estimator by using the following formula:

$$RR = \frac{SMSE\left(\tilde{\beta}\right) - SMSE\left(\hat{\beta}_{SRL}\right)}{SMSE\left(\tilde{\beta}\right)}.100$$

where $\tilde{\beta}$ is any estimator for making comparison with the SRL estimator. All the estimates in Tables 4.4., and 4.5., are obtained iteratively by using the convergence criterion $||\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}|| < 1 \times 10^{-6}$ but to compare the performance of the estimators the SMSE values are obtained for FOA estimates. [9]

[9]Columns 9 and 10 of Tables 4.4., and 4.5., show the RR for ($\hat{d}_{C_p} = 0.001$) and ($\hat{d} = 0.5077$), respectively. This is because we do the comparison for same d value

Table 4.4. Results of Swedish football data set for y_1 with restriction (i)

Estimators	β_1	β_2	β_3	β_4	β_5	β_6	SMSE	RR%	RR%
ML	44.5968	2.2794	-60.9131	-11.6586	9.3973	12.322	1973.6305	99.95	58.98
Mixed	44.5964	2.2795	-60.9145	-11.6585	9.399	12.3219	1228.3878	99.92	34.10
Liu ($\hat{d}_{C_p}$)	-1.2765	0.8952	-1.4542	0.7830	-1.3939	0.9886	1.1510	16.30	–
Liu ($\hat{d}$)	22.0179	1.6054	-31.5774	-5.5191	4.1104	6.7468	890.3192		9.07
SRL($\hat{d}_{C_p}$)	-1.2766	0.8956	-1.4542	0.7835	-1.3930	0.9882	0.9633	–	
SRL($\hat{d}$)	22.0178	1.6058	-31.5775	-5.5187	4.1109	6.7465	809.4980		–

Table 4.5. Results of Swedish football data set for y_1 with restriction (ii)

Estimators	β_1	β_2	β_3	β_4	β_5	β_6	SMSE	RR%	RR%
ML	44.5968	2.2794	-60.9131	-11.6586	9.3973	12.3220	1973.6305	99.95	57.27
Mixed	44.5964	2.2795	-60.9145	-11.6585	9.3990	12.3219	1687.0440	99.94	50.01
Liu ($\hat{d}_{C_p}$)	-1.2766	0.8952	-1.4542	0.7830	-1.3939	0.9887	1.1510	7.99	–
Liu ($\hat{d}$)	22.0179	1.6055	-31.5775	-5.5191	4.1104	6.7468	890.3192		5.28
SRL($\hat{d}_{C_p}$)	-1.2766	0.8952	-1.4539	0.7826	-1.3936	0.9896	1.0591	–	
SRL($\hat{d}$)	22.0179	1.6054	-31.5770	-5.5198	4.1109	6.7484	843.3094		–

It is obvious from Table 4.4., that the SRL estimator contains the smallest SMSE values relative to its competitors, which signifies that SRL estimator performs efficiently over its counterparts. The RR also testifies the efficiency of the SRL estimator, which indicates that it is about 99.95%, 99.92%, and 16.30% more efficient against multicollinearity than ML, mixed, and Liu estimators in case of restriction (i) and for $\hat{d}_{C_p} = 0.001$. While for $\hat{d} = 0.5077$ it is about 58.98%, 34.10%, and 9.07% effective to confront the issue of multicollinearity as compared to its counterparts.

Table 4.5., shows that in case of restriction (ii) SRL estimator performs dominantly for both the values of tuning parameter under consideration as it acquires the smallest SMSE values by comparisons to its contestants. Moreover, RR demonstrates that it is 99.95%, 99.94%, and 7.99%, worthwhile to tackle the multicollinearity as compared to its participants for $\hat{d}_{C_p} = 0.001$ whereas for $\hat{d} = 0.5077$ it is 57.27%, 50.01%, and 5.28% more valuable to cope multicollinearity problem relative to ML,

mixed, and Liu estimators, respectively. It can be seen that types of restrictions affect the performance of the estimators.

Tables 4.4., and 4.5., show the performance of the estimators only for the specific values of the tuning parameters that are $\hat{d}_{C_p}$ and $\hat{d}$, respectively. Whereas, Figs. 4.6., and 4.7., are set up corresponding to restrictions (i) and (ii) to monitor the performance for other values of d. Fig. 4.6., indicates that under restriction (i) SRL estimator performs efficiently as compared to its opponents particularly for small to moderate values of d. Although, the behavior of Liu estimator looks slightly similar to the SRL estimator only for small range of values of d. Besides, as the value of d becomes larger than approximately 0.20 then, the SRL estimator outperforms the Liu estimator. As long as the mixed estimator is concerned its SMSE values become low by comparisons with SRL estimator when the value of d approaches nearly above 0.85 approximately, which refers that its performance is better for larger values of d under restriction (i). Fig. 4.7., explains that under restriction (ii) SRL outperforms all of its counterparts precisely regardless of the values of tuning parameter d even if the Liu estimator reacts somewhat like to SRL estimator but only for small range of d values. Further, as the value of d approaches near to 1 then mixed estimator seems to be slightly better than the SRL estimator.

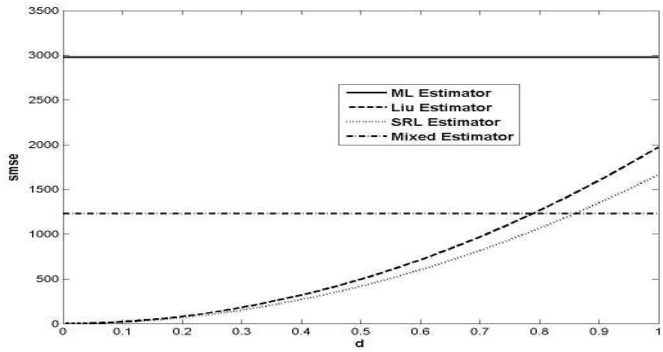


Figure 4.6. SMSE values of the FOA estimators versus d under restriction (i) (Swedish football data set for y_1)

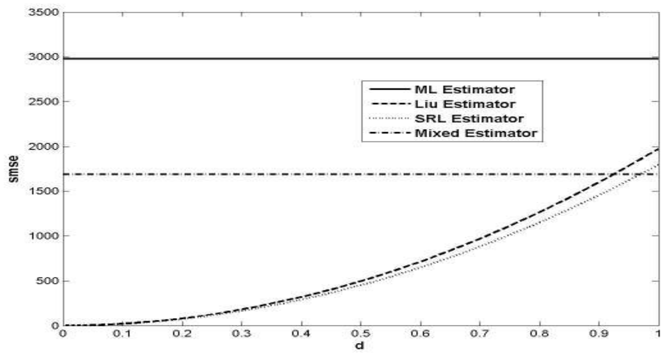


Figure 4.7. SMSE values of the FOA estimators versus d under restriction (ii) (Swedish football data set for y_1)

As a second analysis, the number of full-time away team goals (y_2) is treated as a response variable in order to perceive the effects of multicollinearity on the estimators. For detection of multicollinearity the same collinearity diagnostics are performed as in the case of response variable (y_1). The initial value of β is obtained as $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y_2 = [-2.6669, -1.5939, -20.8366, 5.1596, 29.4340, -5.7847]$ for all the estimators considered. Hence, the condition number and VIFs are given as follows, $\varsigma_{\max} = (3858.9)^{1/2} = 62.1200$, VIFs: 165.1293, 188.6727, 464.8536, 191.4244, 293.5432, 315.1214. Consequently, the condition number and VIFs values are significantly large to a prescribed level and certify that multicollinearity is existing among the explanatory variables. The stochastic restrictions in this case are assumed as: (iii) $\beta_1 + 2\beta_3 + \beta_4 + \beta_5 + G = 0.02$, $G \sim (0, 0.005^2)$. While, implementing the restrictions on parameters the same rule is followed as in the case of restrictions (i) and (ii), respectively. Therefore, resulted as $P(|\beta_1 + 2\beta_3 + \beta_4 + \beta_5 - 0.02| < 3\sigma) \geq 1 - \frac{1}{3^2} = 0.8888$ which leads to $\beta_1 + 2\beta_3 + \beta_4 + \beta_5 \in [0.005, 0.035]$ having 0.8888 probability. Further, at 0.8888 probability $\sigma = 0.005$ is assumed to be value of an estimate of $var(G) = \Sigma = \sigma^2 I$ along with $r = 0.02$ and R matrix is $R = [1\ 0\ 2\ 1\ 1\ 0]$. Fig. 4.8., is established to find out the optimum value of tuning parameter d at which $C_p^{c-l(d)}$ statistic is minimum that is $\hat{d}_{C_p} = 0.001$ which is arbitrarily chosen because $C_p^{c-l(d)}$ is an increasing function of d. Moreover, we also assume the value $\hat{d} = 0.8429$, which is found by using the same formula as mentioned in previous restrictions (i) and (ii), respectively. The findings obtained are reported in Table 4.6., for $\hat{d}_{C_p} = 0.001$ and $\hat{d} = 0.8429$.

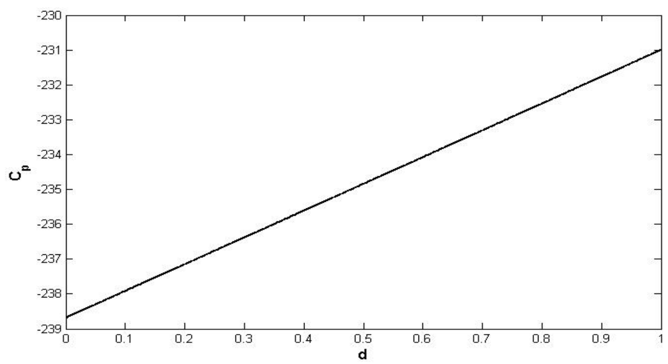


Figure 4.8. $C_p^{c-l(d)}$ values of the FOA SRL estimator versus d (Swedish football data set for y_2) under restriction (iii)

Table 4.6. Results of Swedish football data set for y_2 with restriction (iii)

Estimators	β_1	β_2	β_3	β_4	β_5	β_6	SMSE	RR%	RR%
ML	-0.3083	-3.8030	-11.0644	8.4449	14.0087	-9.9722	1225.1581	99.94	44.40
Mixed	-0.3089	-3.8034	-11.0628	8.4469	14.0075	-9.9738	741.9192	99.89	8.19
Liu ($\hat{d}_{C_p}$)	0.8418	-1.3474	0.8473	-1.2687	1.0461	-1.3762	0.9467	16.07	–
Liu ($\hat{d}$)	0.2695	-2.5904	-5.1864	3.6604	7.6280	-5.7336	880.9246		22.68
SRL($\hat{d}_{C_p}$)	0.8417	-1.3474	0.8480	-1.2679	1.0461	-1.3762	0.7945	–	
SRL($\hat{d}$)	0.2694	-2.5904	-5.1843	3.6626	7.6280	-5.7336	681.1686		–

From, Table 4.6., it is apparent that the enhancement of the SRL estimator is superior in terms of SMSE criterion over its counterparts as it consists of the smallest SMSE values. Moreover, RR justifies that the SRL estimator is relatively 99.94%, 99.89%, and 16.07% more efficient to confront the multicollinearity than the ML, Liu, and mixed estimators for $\hat{d}_{C_p} = 0.001$. Further, for $\hat{d} = 0.8429$ RR represents that it is 44.40%, 8.19%, and 22.68% more useful as compared to its rivals to circumvent the issue of multicollinearity. Table 4.6., expresses the performance of estimators only for special values of d whereas Fig. 4.9., is build up to review the performance of estimators for the remaining values of d in the interval $[0,1]$. Fig. 4.9., manifests that SRL estimator again performs better than its competitors although the behavior of the Liu estimator seems to be slightly similar to SRL estimator for those d values fall in the range 0 to 0.20. But, as the value of d goes above almost 0.20 then, the SRL outperforms the Liu estimator. Also as the value of d reaches approximately near to 0.90 then,the mixed estimator seems to be marginally better than the SRL estimator.

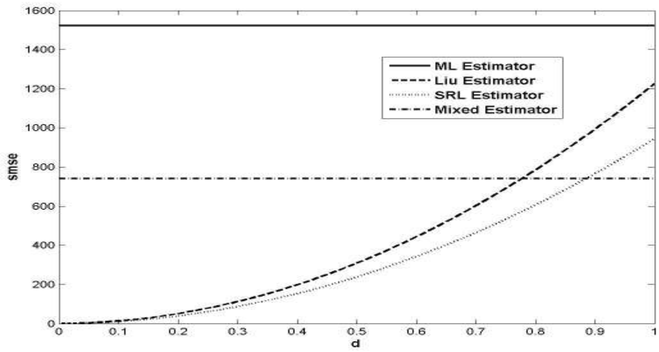


Figure 4.9. SMSE values of the FOA estimators versus d under restriction (iii) (Swedish football data set for y_2)

4.1.4. Application of Stochastic Restricted OK Estimator to Swedish Football Data Set

In this section, we illustrate the application of stochastic restricted OK to a Swedish football data set to address the problem of multicollinearity. Moreover, ML, mixed, and stochastic restricted ridge estimators are also considered owing to the comparisons with stochastic restricted OK estimator. The comparison is made in terms of SMSE criterion. Prior to computing the results by means of unit length standardization technique the explanatory variables are standardized. Hence, a log link function is utilized having the following regression model:

$$\log(\hat{\mu}_i) = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\beta}_4 x_{4i} + \hat{\beta}_5 x_{5i} + \hat{\beta}_6 x_{6i}.$$

An iterative technique is adopted to achieve the results by using the convergence criterion $||\tilde{\beta}^{(t+1)} - \tilde{\beta}^{(t)}|| < 1 \times 10^{-6}$. This iterative procedure continues until the required convergence criterion is fulfilled. In addition, the technique of GA explained in subsubsection 3.2.2.2., is utilized to find the values of tuning parameters (k,d). The algorithms are figured out by using the R programming language and GA package of (Scrucca, 2013) which consists of already built-in commands in its library is applied and the seed value is considered to be 100. The parameters of GA are considered as number of chromosomes in the population 100, iterations (generations) 100, elitism individuals 10, Crossover ratio 0.8, and Mutation ratio 0.05. Initially, the ML estimator is obtained by following OLS estimator $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y$ as a premier estimate having the values: $[13.8992, 2.0119, -18.8602, -20.2691, 3.0540, 24.1594]^T$. As there is severe multicollinearity in this data set, therefore, we impose stochastic restrictions on the parameters to diminish the multicollinearity problem. Three types of restrictions are examined, the first two of which are the same as those used in the application of SRL estimator. For the implementation of restrictions the ML estimator is used and rule of Chebyshev inequality is applied and the restrictions are obtained as:

i) $\beta_{2+}\beta_4 + \beta_5 + G = 0.02$, $G \sim (0,\ 0.005^2)$. By following the rule of Chebyshev inequality we obtained $P(|\beta_{2+}\beta_4 + \beta_5 - 0.02| < 3\sigma) \geq 1 - \frac{1}{3^2} = 0.8888$ which specifies that $\beta_{2+}\beta_4 + \beta_5 \in [0.005,\ 0.035]$ having 0.8888 probability with $r = 0.02$, $R = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$ and $\sigma = 0.005$ is selected to be the estimated value of $var(G) = \Sigma = \sigma^2 I$ at this probability [10].

ii) $\beta_2 + \beta_3 + \beta_5 + 4\beta_6 + G = 0.05$, $G \sim (0,\ 0.004^2)$ and $P(|\beta_{2+}\beta_3 + \beta_5 + 4\beta_6 - 0.05| < 3\sigma) \geq 1 - \frac{1}{3^2} = 0.8888$, which determined that $\beta_{2+}\beta_3 + \beta_5 + 4\beta_6 \in [0.038,\ 0.062]$ at 0.8888 probability besides, $r = 0.05$, $R = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 4 \end{bmatrix}$ and $\sigma = 0.004$ that is considered to be an estimated value of $var(G) = \Sigma = \sigma^2 I$.

[10]This stochastic restriction is created by considering the ML estimator given in Table 4.8., since we do not have prior information on the data. Then, the value of σ is determined such that the ML estimates of second, fourth and fifth coefficients fall within the specified interval

iii) $\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + G = 0$, which resembles the homogeneous type of restriction and $G \sim (0,\ 0.000001^2)$. In this particular restriction $r = 0$, $R = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$ and the estimated value of $var(G) = \Sigma = \sigma^2 I$ is assumed to be $\sigma = 0.000001$ which is near to 0 but not exactly equal to 0 (if it is equal to zero it would be a exact restriction not stochastic) and this signifies that stochastic restrictions are assumed to hold on the parameters.

To apply the restrictions, we test the null hypothesis

H_0: Prior and sample information are in agreement

The test results are shown in Table 4.7. [11]

Table 4.7. Results of the test statistic when $\lambda = 0.000000, \alpha = 0.10, p = 6$

Restrictions	Test Statistic ($\hat{\gamma}$)	Noncentrality parameter (λ)	$\chi^2_{p,\lambda,1-\alpha}$
(i)	0.0000135	$1.997892e^{-07}$	0.01087
(ii)	0.00000132	$8.37747e^{-08}$	
(iii)	$6.361851e^{-13}$	$4.761127e^{-16}$	

From Table 4.7., it is clear that the calculated value of the test statistic $\hat{\gamma}$ is significantly below the critical value for all the three restrictions imposed therefore, we accept the null hypothesis and conclude that prior and sample information are in agreement. Thus, the stochastic restrictions are appropriate in this numerical illustration.

To estimate the values of tuning parameters (k,d) of the stochastic restricted OK estimator the technique of GA described in subsubsection 3.2.2.2 is applied. Thus, the values of tuning parameters (k,d) corresponding to restrictions (i), (ii), and (iii) are computed as: under restriction (i) $(k_1, d_1) = (0.03396, 0.11029)$, $(k_2, d_2) = (0.03512, 0.09615)$ are pairs correspond to minimizer of MSEP and MAE, respectively. Under

[11] we show only $\lambda = 0.000000$, because the same critical value appears corresponding to all three values of λ if it is round up to six digits

restriction (ii) $(k_1, d_1) = (0.09731, 0.03043)$, $(k_2, d_2) = (0.99019, 0.12752)$ are pairs correspond to minimizer of MSEP and MAE, respectively. Under restriction (iii) $(k_1, d_1) = (0.94064, 0.01077)$, $(k_2, d_2) = (0.01067, 0.88014)$ are pairs correspond to minimizer of MSEP and MAE, respectively. For the restrictions (i) and (ii) the CN and VIF are used defined in subsubsection 3.2.2.2., for the restricted OK estimator and for the restriction (iii) the CN and VIF are used defined for the stochastic restricted OK estimator in the particular case when $r = 0$ given in subsubsection 3.2.2.2. The results of the iterative parameter estimates along with their SMSE values in the FOA form are shown in Tables 4.8., to 4.10.

Table 4.8. Iterative estimates and the SMSE values of the estimators under restriction (i)

	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
β_1	44.5968	-4907.3961	-1.7518	-1.2423	-1.3146	-0.8149
β_2	2.2794	98.4972	1.3126	1.1729	0.8265	0.5910
β_3	-60.9131	8508.1254	-3.9929	-4.5012	-1.3873	-2.1348
β_4	-11.6586	11069.5849	-0.1299	0.0691	0.7287	1.0086
β_5	9.3973	-2401.2431	-1.1630	-1.2222	-1.5347	-1.5775
β_6	12.3220	-10370.2436	1.0123	0.9154	1.0434	0.7768
SMSE	1973.6305	1228.3863	1167.2903	980.8183	1347.8550	1046.3430

Table 4.9. Iterative estimates and the SMSE values of the estimators under restriction (ii)

	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
β_1	44.5968	-5965.3158	-0.7123	-0.4635	-1.3110	-0.8702
β_2	2.2794	225.6443	0.9709	0.8594	0.9268	0.8132
β_3	-60.9131	8123.4527	-3.5911	-4.0824	-1.5345	-2.1375
β_4	-11.6586	10639.8317	-0.2543	-0.23176	0.9848	1.1932
β_5	9.39730	-947.3284	-2.7633	-2.5208	-1.5482	-1.1737
β_6	12.3220	-10054.4364	1.3584	1.4485	0.5516	0.6370
SMSE	1973.63050	1687.0440	857.9861	730.6889	1009.0030	798.0855

Table 4.10. Iterative estimates and the SMSE values of the estimators under restriction (iii)

	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
β_1	44.5968	-4812.2260	-1.0249	-0.9834	-1.0313	-0.9902
β_2	2.2794	-282.0278	1.0832	1.0772	1.0929	1.0870
β_3	-60.9131	6318.1983	-1.1052	-1.1660	-1.1175	-1.1778
β_4	-11.6586	11925.9288	0.9618	0.9974	0.9632	0.9985
β_5	9.3973	-521.0430	-1.1081	-1.0866	-1.1185	-1.0972
β_6	12.3220	-10869.5781	1.1932	1.1614	1.2112	1.1797
SMSE	1973.6305	1870.5260	1528.6220	1496.1860	1527.5050	1495.39100

From the tabular results, it is evident that all of the biased estimators perform better than the ML estimator to encountering the problem of multicollinearity in terms of the SMSE criterion. Anyhow, the performance of the stochastic restricted OK estimator dominated over its counterparts comprehensively as it gains the smallest SMSE values then, the stochastic restricted ridge estimator performs better than the mixed estimator. It is obvious that the performance of stochastic restricted OK estimator is superior to its counterparts regardless of the type of restrictions imposed and the values of tuning parameters either they are obtained by minimizing MSEP or MAE as it acquires the smallest SMSE values. Tables 4.8.-4.10., show the performance of the estimators for particular values of tuning parameters k and d. To view the performance of the estimators for further values of k and d, Figs. 4.10., and 4.11., are developed [12].

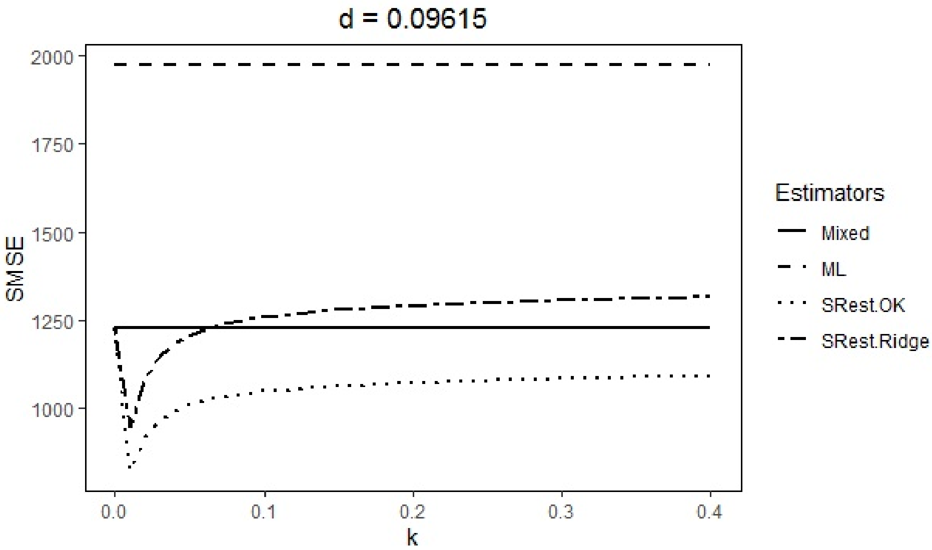


Figure 4.10. Graph of the SMSE values of the FOA estimators against k when d is fixed at minimizer of MSEP under restriction (i)

[12]We provide Figs. 4.10. and 4.11. only for restriction (i) because the same trend was observed for restriction (ii) and (iii)

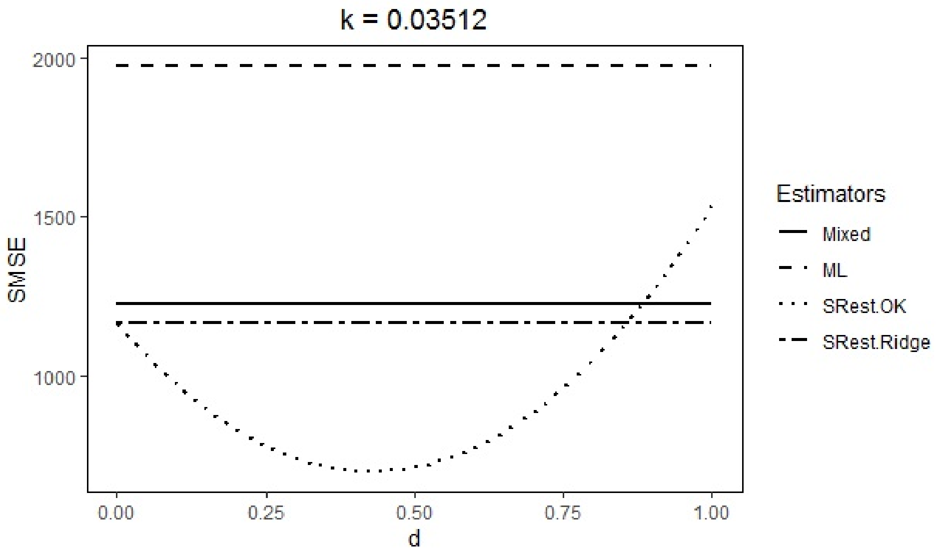


Figure 4.11. Graph of the SMSE values of the FOA estimators against d when k is fixed at minimizer of MAE under restriction (i)

From Fig. 4.10., it is clear that all of the biased estimators outperform the ML estimator in terms of lower SMSE values and reacted better in confronting the problem of multicollinearity. Whereas, enhanced performance of the stochastic restricted OK estimator is observed when comparing it to its counterparts for all the values of k under consideration when the value of d is fixed at 0.09615. As the value of k increases, the more efficient performance of the stochastic restricted OK estimator is noticed. When the value of k falls roughly below 0.1 then, stochastic restricted ridge estimator performs better than the mixed estimator.

Similarly, Fig. 4.11., shows the superiority of the biased estimators over ML estimator in terms of lesser SMSE values. However, the stochastic restricted OK estimator performs better than its counterparts when the value of d lies roughly in the range 0 to 0.80. As the value of d gets approximately larger than 0.80 the performance of stochastic restricted ridge and mixed estimators seem to be better than the stochastic restricted OK estimator.

Moreover, Figs. 4.12., 4.13., and 4.14., are also established under restrictions (i), (ii), and (iii) respectively to see the combine effects of k and d on the SMSE of $\hat{\beta}^{(1)}_{kd-sr}$.

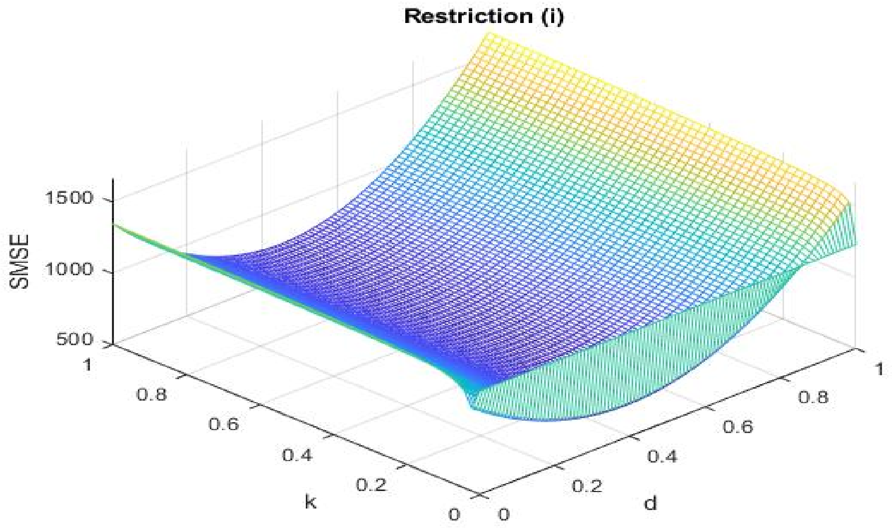


Figure 4.12. The SMSE values of $\hat{\beta}^{(1)}_{kd-sr}$ against k and d under restriction (i)

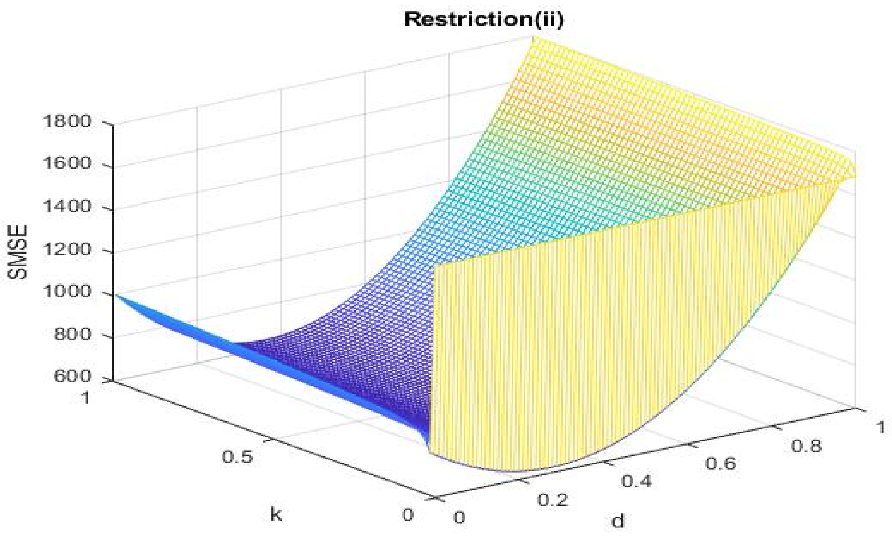


Figure 4.13. The SMSE values of $\hat{\beta}^{(1)}_{kd-sr}$ against k and d under restriction (ii)

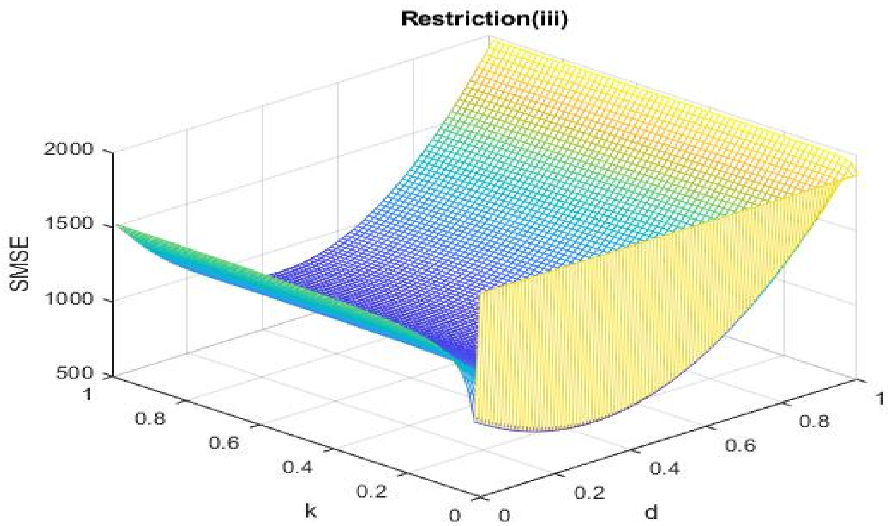


Figure 4.14. The SMSE values of $\hat{\beta}^{(1)}_{kd-sr}$ against k and d under restriction (iii)

Fig. 4.12., shows that when the value of d falls approximately between 0.4-0.60 and for all values of k then, $\hat{\beta}^{(1)}_{kd-sr}$ acquires smaller SMSE values. Similarly, Fig. 4.13., manifests that when the value of d falls around 0.20-0.40 and for all the values of k, $\hat{\beta}^{(1)}_{kd-sr}$ achieves small SMSE values. Likewise, Fig. 4.14., demonstrates that when the value of d falls in the range 0.20-0.40 and the value of k approximately lies in the range 0.001-0.20 then, the SMSE values of $\hat{\beta}^{(1)}_{kd-sr}$ are small.

This numerical example illustrates that to address the problem of multi-collinearity in GLMs, the performance of the stochastic restricted OK estimator is best among all other estimators for the appropriate selected values of (k, d) pairs.

4.2. Weather Data Set Gamma Response

In this section, we explain the weather data set which was originally used by (Chatterjee and Hadi, 1988). This data set is pertaining to weather factors and nitrogen dioxide concentrations in parts per hundred million (p.p.h.m) for 26 days in September 1984 at a monitoring station in the San Francisco Bay area. This data set contains four explanatory variables that are: mean wind speed in miles per hour m.p.h. (x_1), maximum temperature (x_2) in ^0F, insolation (x_3) langleys per day and stability factor (x_4) in ^0F as well as the response variable nitrogen dioxide concentrations (y). Chatterjee and Hadi (1988) showed that this data set follows a gamma distribution. Moreover, this data set has also been used by (Kurtoğlu and Özkale, 2019b)[13]. Before computation, we standardized the explanatory variables through unit length standardization scheme and then, included the intercept term in the model. Thus, a reciprocal link function $\hat{\mu}_i = (\sum_{i=1}^{p} \hat{\beta}_i x_{ij})^{-1}$ is utilized to estimate the GLM model. The results are found by using MATLAB programming language and R programming language is used only for optimal (k,d) pairs found via GA by using GA package of (Scrucca, 2013) which contains already built-in commands in its library.

For the estimation of parameters an iterative technique is applied by considering the convergence criterion $||\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}|| < 1 \times 10^{-6}$ until the prerequisite convergence criterion is fulfilled this iterative procedure continues. In this iterative process the initial values are treated as $\hat{\mu}_i^{(0)} = (\frac{y+\bar{y}}{2})$, then, we defined a weight matrix $\hat{W}^{(0)} = diag(\hat{\mu}_i^{(0)})^2$ besides, the initial working response for the i-th observation is $\hat{z}_i^{(0)} = \frac{1}{\hat{\mu}_i^{(0)}} - \hat{w}_{ii}^{(0)}(y_i - \hat{\mu}_i^{(0)})$. The Pearson estimate of $a(\phi)$ is found as $\hat{\phi} = 0.0757$. We used the OLS estimator $\hat{\beta}_{OLS} = (X^TX)^{-1}X^Ty$ as an initial estimate for all the estimators that contains the values: [6.269231, −5.622560, 7.502692, −2.000585, −0.356589]. It is observed that the ML estimator is obtained in the 4-th iteration. Then, the condition number and VIFs are computed in order to see the problem of multicollinearity. Since the collinearity is meaningful when the columns of $X^* = \hat{W}_{ML}^{1/2}X$ have been centered and scaled where $\hat{W}_{ML}$ is a weight matrix acquired in the final

[13]Kurtoğlu and Özkale (2019b) used this data set as raw data with constant term.

iteration of ML estimator by (Mackinnon and Puterman 1989, Weissfeld and Sereika 1991, Özkale 2021). Thus, the condition number and VIFs values are found as: $\varsigma_{\max} = \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} = (6.1930 \times 10^{16})^{1/2} = 2.4886 \times 10^{8}$ and VIFs [14]: 0.1918, 3.3096, 5.4907, 1.0525, and 1.6419. It is clear that the values of condition number and VIFs[15] are extremely large and indicate that there is a severe problem of multicollinearity in this data set. Therefore, to tackle the problem of multicollinearity we go ahead with biased estimation methods and place the restrictions on the parameters in the following section.

4.2.1. Application of Restricted OK Estimator to Weather Data Set

In this section, we demonstrate the application of the restricted OK estimator to a weather data set which follows a gamma distribution. The other biased estimators such as restricted ML, ridge, restricted ridge, Liu, restricted Liu, and OK estimators in GLMs are also considered along with the ML estimator to make comparisons with the restricted OK estimator. Both ML-based and FOA estimators are utilized in the comparisons. As discussed above there is a multicollinearity problem in this data and we impose the restrictions on the parameters to address the problem of multicollinearity. The restriction is $\beta_2 + \beta_3 + \beta_4 = 0.1245$ which leads to $R = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \end{bmatrix}$ and $r = 0.1245$. This restriction is obtained by considering the $\hat{\beta}_{ML}$ estimator given in Table 4.13., where $r = R\beta$ is approximately true.

Since there are ML-based and FOA estimators in the literature, the results of ML-based and FOA estimators have also been reported in order to be able to compare with these estimators. In computing ML-based and FOA estimators the restrictions,

[14] When finding the VIF values from the centered and scaled $X^{*T}X^{*}$, close to singular error exists which yields into negative VIF values in Matlab and cannot computed in R. Again, this is due to the problem of multicollinearity. Therefore, pinv (Moore-Penrose pseudoinverse) is used in calculating VIF values for ML estimator. For this reason, VIF values are smaller than 10.

[15] Due to the previous note

initial values, and (k,d) pairs used in the calculation of the iterative estimator are used.

The values of the tuning parameters k and d are found as $k = k_{HM} = 1.0925$ and $d = 0.9920$ by using the procedure adopted in subsection 4.1.2. [16]. Additionally, the tuning parameters k and d obtained via GA by minimizing MSEP and MAE are respectively given as $(k_2,d_2) = (1.997748,0.0008013207)$ and $(k_3,d_3) = (1.997065,0.001288869)$. The parameter used in this GA are given as: number of chromosomes in the population 100, iterations (generations) 200, elitism individuals 10, Crossover ratio 0.8, and Mutation ratio 0.05. The results obtained are reported in Tables, 4.11.,- 4.13.

[16]Other (k,d) computations are as $(k_{AM},d) = (7.3323, 0.9987)$ and $(k_{GM},d) = (0.3248, 0.9732)$

Table 4.11. Estimates and SMSE when optimal $(k,d) = (1.0925, 0.9920)$ are obtained via MSE

FOA	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1736	0.1812	-0.1621	0.0153	0.0019	0.018952
Rest. ML	0.1736	0.2178	-0.1175	0.0241	-0.0195	0.036357
Ridge	0.1730	0.1778	-0.1578	0.0157	-0.0013	0.172554
Rest. Ridge	0.1731	0.2138	-0.1145	0.0252	-0.0215	0.058634
Liu	0.1736	0.1811	-0.1621	0.0153	0.0019	0.018930
Rest. Liu	0.1736	0.2177	-0.1180	0.0248	-0.0187	0.035749
OK	0.1736	0.1811	-0.1611	0.0153	0.0019	0.018929
Rest. OK	0.1736	0.2177	-0.1181	0.0249	-0.0186	0.035471
Iterative	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
Rest. ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
Ridge	0.1808	0.1991	-0.1203	0.0304	-0.0219	-
Rest. Ridge	0.1810	0.2065	-0.1131	0.0310	-0.0250	-
Liu	0.1819	0.2098	-0.1161	0.0306	-0.0234	-
Rest. Liu	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
OK	0.1819	0.2098	-0.1161	0.0306	-0.0234	-
Rest. OK	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
ML-based	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	0.019709
Rest. ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	0.007862
Ridge	0.1808	0.1990	-0.1204	0.0304	-0.0219	0.015529
Rest. Ridge	0.1810	0.2065	-0.1130	0.0310	-0.0250	0.007214
Liu	0.1819	0.2098	-0.1161	0.0306	-0.0234	0.019675
Rest. Liu	0.1818	0.2047	-0.1217	0.0307	-0.0199	0.007862
OK	0.1819	0.2098	-0.1161	0.0306	-0.0234	0.026193
Rest. OK	0.1819	0.2099	-0.1160	0.0306	-0.0234	0.007863

Table 4.12. Estimates and SMSE when optimal (k_1, d_1) are obtained via GA by minimizing MSEP

FOA	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1736	0.1812	-0.1621	0.0153	0.0019	0.018952
Rest. ML	0.1736	0.2178	-0.1175	0.0241	-0.0195	0.036357
Ridge	0.1725	0.1749	-0.1547	0.0160	-0.0036	0.447044
Rest. Ridge	0.1726	0.2106	-0.1120	0.0259	-0.0230	0.149558
Liu	0.1730	0.1781	-0.1582	0.0157	-0.0011	0.149921
Rest. Liu	0.1731	0.2141	-0.1147	0.0251	-0.0213	0.104882
OK	0.1725	0.1749	-0.1547	0.0160	-0.0036	0.446353
Rest. OK	0.1726	0.2106	-0.1120	0.0259	-0.0230	0.149335
Iterative	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
Rest. ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
Ridge	0.1800	0.1918	-0.1224	0.0301	-0.0212	-
Rest. Ridge	0.1802	0.2038	-0.1107	0.0314	-0.0261	-
Liu	0.1809	0.1999	-0.1200	0.0304	-0.0220	-
Rest. Liu	0.1810	0.2068	-0.1133	0.0310	-0.0248	-
OK	0.1800	0.1918	-0.1224	0.0301	-0.0212	-
Rest. OK	0.1802	0.2038	-0.1107	0.0314	-0.0261	-
ML-based	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	0.019709
Rest. ML	0.1819	0.2099	−0.1160	0.0306	−0.0234	0.007862
Ridge	0.1800	0.1916	-0.1225	0.0301	-0.0212	0.013379
Rest. Ridge	0.1802	0.2038	−0.1107	0.0313	−0.0260	0.006781
Liu	0.1809	0.1999	-0.1201	0.0304	-0.0220	0.015805
Rest. Liu	0.1810	0.2068	−0.1133	0.0310	−0.0248	0.007264
OK	0.1800	0.1916	-0.1225	0.0301	-0.0212	0.018126
Rest. OK	0.1802	0.2039	−0.1107	0.0313	−0.0260	0.006781

Table 4.13. Estimates and SMSE when optimal (k_2, d_2) are obtained via GA by minimizing MAE

FOA	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1736	0.1812	-0.1621	0.0153	0.0019	0.018952
Rest. ML	0.1736	0.2178	−0.1175	0.0241	−0.0195	0.036357
Ridge	0.1725	0.1749	-0.1547	0.0160	-0.0036	0.446808
Rest. Ridge	0.1726	0.2106	−0.1120	0.0259	−0.0230	0.149468
Liu	0.1730	0.1781	-0.1582	0.0157	-0.0011	0.149792
Rest. Liu	0.1731	0.2141	−0.1147	0.0251	−0.0213	0.104772
OK	0.1725	0.1749	-0.1547	0.0160	-0.0036	0.445697
Rest. OK	0.1726	0.2106	−0.1120	0.0259	−0.0230	0.149110
Iterative	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	-
Rest. ML	0.1819	0.2099	−0.1160	0.0306	−0.0234	-
Ridge	0.1800	0.1918	-0.1224	0.0301	-0.0212	-
Rest. Ridge	0.1802	0.2038	−0.1107	0.0314	−0.0261	-
Liu	0.1809	0.1999	-0.1200	0.0304	-0.0220	-
Rest. Liu	0.1810	0.2068	−0.1133	0.0310	−0.0248	-
OK	0.1800	0.1918	-0.1224	0.0301	-0.0212	-
Rest. OK	0.1802	0.2038	−0.1107	0.0314	−0.0261	-
ML-based	β_1	β_2	β_3	β_4	β_5	SMSE
ML	0.1819	0.2099	-0.1160	0.0306	-0.0234	0.019709
Rest. ML	0.1819	0.2099	−0.1160	0.0306	−0.0234	0.007862
Ridge	0.1800	0.1916	-0.1225	0.0301	-0.0212	0.013380
Rest. Ridge	0.1802	0.2039	−0.1107	0.0313	−0.0260	0.006781
Liu	0.1809	0.1999	-0.1201	0.0304	-0.0220	0.015807
Rest. Liu	0.1810	0.2068	−0.1133	0.0310	−0.0248	0.007264
OK	0.1800	0.1916	-0.1225	0.0301	-0.0212	0.018131
Rest. OK	0.1802	0.2039	−0.1107	0.0313	−0.0260	0.006782

Table 4.11., presents that while the restricted FOA OK estimator gives the smallest SMSE value but the restricted ML-based OK estimator does not and the restricted ridge seems better as ML-based estimator. Considering the FOA estimators, all of the restricted estimators give a larger SMSE than the unrestricted ones except the FOA restricted ridge estimator. However, all the ML-based restricted estimators give smaller SMSE value than their corresponding unrestricted ML-based estimators. FOA OK estimator gives the smallest SMSE value among the restricted and unrestricted FOA estimators and the ML-based ridge estimator gives the smallest SMSE among the unrestricted ML-based estimators. Considering Tables 4.12., and 4.13., it is seen that both the restricted FOA and ML-based restricted estimators give smaller SMSE than their corresponding unrestricted estimators except the restricted FOA ML estimator. Besides, the restricted FOA ML estimator dominates the other restricted FOA estimators and the restricted ML-based ridge estimator dominates the other restricted ML-based estimators. When unrestricted estimators are compared to each other, it is apparent that the FOA ML estimator and ML-based ridge estimators have smaller SMSE values than the other unrestricted estimators.

To see the performance for the remaining values of k and d Figs., 4.15., and 4.16., are given.

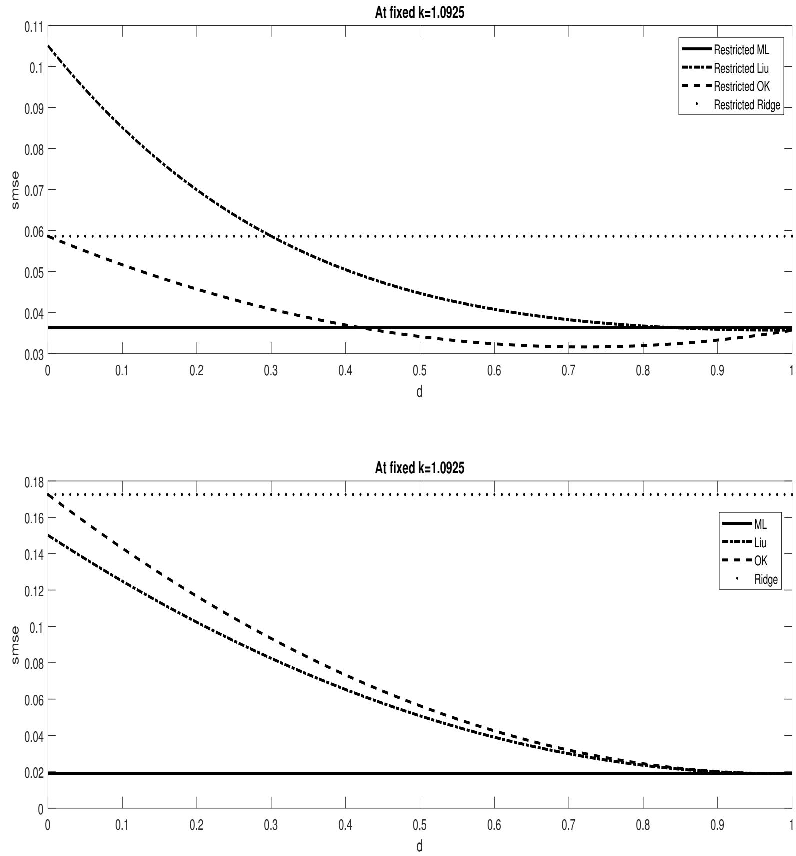


Figure 4.15. Graph of the SMSE values of the FOA estimators against d when k is fixed at 1.0925 (Weather Data)

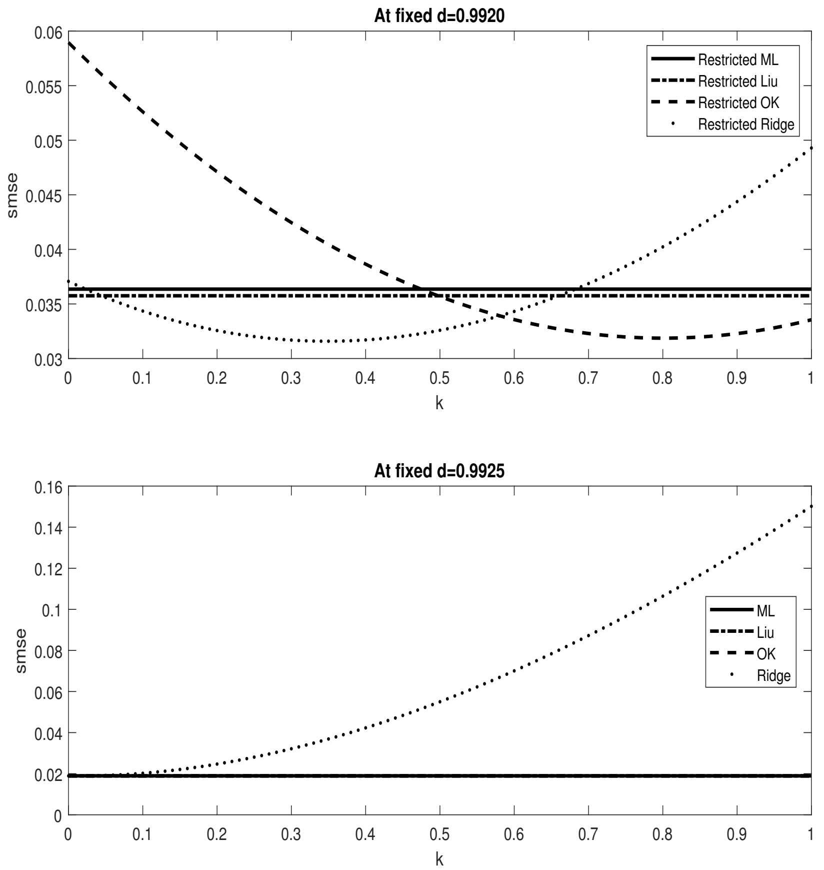


Figure 4.16. Graph of the SMSE values of the FOA estimators against k when d is fixed at 0.9920 (Weather Data)

Fig. 4.15., shows that as d increases the SMSE values of FOA Liu, FOA OK, restricted FOA Liu, and restricted FOA OK estimators decreases and then increase slightly. This is most evident in the restricted FOA OK estimator. The restricted FOA OK estimator is superior to the other restricted FOA estimators for $0.42 < d < 1$ and $k = 1.0925$. Fig. 4.16., presents that as k increases the SMSE values of restricted FOA ridge and restricted FOA OK estimators first decrease and then increase. The restricted

FOA ridge is better than the other restricted FOA estimators for $0.0001 < k < 0.57$ and $d = 0.9920$ while the restricted FOA OK is better than the other restricted FOA estimators for $0.57 < k < 1$ and $d = 0.9920$.

4.3. Reaction Data Set Gamma Response

In this section, we illustrate the reaction data set which is taken from (Huet et al., 2006). This data set contains 24 observations and three explanatory variables, which are partial pressure of hydrogen (x_1), partial pressure of n-pentane (x_2), and partial pressure of iso-pentane (x_3), as well as the response variable reaction rate (y). Moreover, this data set has also been used by (Amin et al., 2019) and they explained that the response variable follows a gamma distribution by implementing the goodness of fit tests like; Anderson-darling, Cramer-Von Mises, and pearson χ^2 test. Before computing the results, explanatory variables are standardized and then intercept term is added to the model. For the computations of estimators, we used $\hat{\mu}_i^{(0)} = (\frac{y+\bar{y}}{2})$ as an initial estimate. The reciprocal link function, $\hat{\mu}_i = \left(\sum_{j=1}^{p} \hat{\beta}_i x_{ij}\right)^{-1}$ is to be considered. After that we define a weight matrix $W^{(0)} = diag(\hat{\mu}_i^{(0)})^2$ along with the initial working response variable z having the i-th observation $z_i^{(0)} = 1/\hat{\mu}_i^{(0)} - \left(W^{(0)}\right)^{-1}(y_i - \hat{\mu}_i^{(0)})$. Then, the ML estimator is computed. After getting the ML estimator the same collinearity diagnostics explained in Section 4.2., are applied in order to detect the problem of multicollinearity. Thus, we obtained the VIF values as $1.62070e+16$, $7.0041e+14$, $3.9239e+15$, $3.3081e+15$, while the condition number is calculated as $\varsigma_{\max} = \left(\frac{\lambda_{\max}}{\lambda_{\min}}\right)^{1/2} = (6.9625e+16)^{1/2} = 5522321.2271$. Since the values of VIF and condition number are very large to the certain prescribed value and revealed that there exists a strong multicollinearity problem in this data set. Therefore, to handle the problem of multicollinearity we apply the SRL estimator on reaction data set by imposing the stochastic restrictions to the parameters in the following section.

4.3.1. Application of Stochastic Restricted Liu Estimator to Reaction Data Set

This section demonstrate the application of SRL estimator to a reaction data set. In addition, ML, mixed, and Liu estimators are also considered to make comparisons with SRL estimator.

As it is indicated that there is a severe multicollinearity problem in this data set, therefore, we impose the stochastic restriction on the parameters to deal with multicollinearity. For the implementation of restrictions the ML estimator is used given in Table 4.14., which gives: $\beta_1+\beta_2+0.5\beta_3+G=0.1653$, $G\sim(0, 0.01^2)$. By following the rule of Chebyshev inequality we have $P(|\beta_1+\beta_2+0.5\beta_3-0.1653|<3\sigma)\geq 1-\frac{1}{3^2}=0.8888$, which refers that $\beta_1+\beta_2+0.5\beta_3\in[0.1353, 0.1953]$ with 0.8888 probability along with $\sigma=0.01$ is considered to be an estimated value of $var(G)=\Sigma=\sigma^2 I$ at 0.8888 probability also $r=0.1653$ and R matrix is $R=[0, 1, 1, 0.5]$. The Pearson estimate of $a(\phi)$ is obtained as $\hat{\phi}=0.1854$.

In addition, the optimum value of tuning parameter d is found by minimizing the $C_p^{c-l(d)}$ statistic and deviance for the gamma regression is calculated as $D_{c-l,r,d}=2\sum\{-\log(y/\hat{\mu})-(y-\hat{\mu})/\hat{\mu}\}$, where $\hat{\mu}$ is found by using $\hat{\beta}_{c-l}(d)^{(1)}$. Fig. 4.17., is made to examine the optimum value of d and it is arbitrarily selected as $\hat{d}_{C_p}=0.001$, because the $C_p^{c-l(d)}$ statistic is an increasing function of d values. We also used the tuning parameter d given by (Kurtoğlu and Özkale, 2016) that is computed as $\hat{d}=0.8229$.

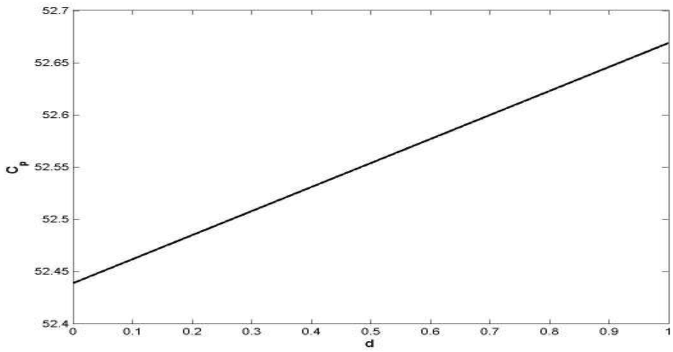


Figure 4.17. $C_p^{c-l(d)}$ values of the FOA SRL estimator versus d (Reaction data set for the gamma regression)

The results of the parameters estimates along with SMSE for the FOA estimators and RR are shown in Table 4.14. The RR is computed by using the same formula given in subsection 4.1.3.

Table 4.14. Results of reaction data with SMSE and RR for the gamma regression.

Estimators	β_0	β_1	β_2	β_3	SMSE	RR%	RR%
ML	0.355234	0.186071	-0.312965	0.584397	0.0363	89.44	88.96
Mixed	0.355234	0.186069	-0.312966	0.584396	0.0228	83.20	82.44
Liu ($\hat{d}_{C_p}$)	0.343153	0.186889	-0.303918	0.521844	0.0064	39.88	–
Liu ($\hat{d}$)	0.353000	0.186184	-0.311284	0.573047	0.0068		41.46
SRL($\hat{d}_{C_p}$)	0.349040	0.252700	-0.254692	0.558960	0.0038	–	
SRL($\hat{d}$)	0.359401	0.250880	-0.263611	0.611262	0.0040		–

Table 4.14., explains that the performance of the SRL estimator is best among all other estimators in terms of the SMSE criterion. As it acquires the smallest SMSE values for both the values of the tuning parameter. RR depicts that it is about 89.44%, 83.20%, and 39.88% extra beneficial as compared to its rival to reduce the problem of multicollinearity for $\hat{d}_{C_p} = 0.001$. Likewise, for $\hat{d} = 0.8229$ SRL estimator seems to be 88.96%, 82.84%, and 41.46% more effective than its rivals to address the issue of

multicollinearity. Table 4.14., shows the performance of the estimators only for two values of *d* while Fig. 4.18., is provided to examine the performance for the rest of the *d* values.

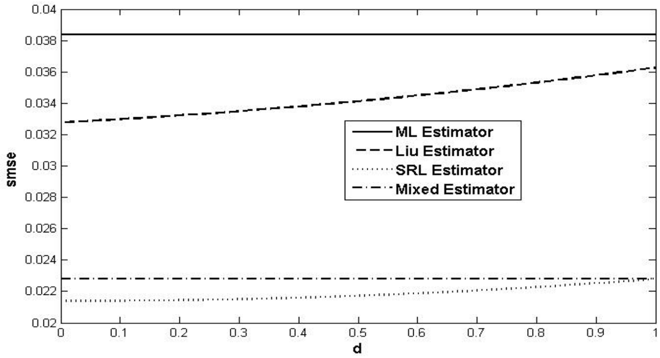


Figure 4.18. SMSE values of the FOA estimators versus *d* for gamma regression, (reaction data set)

Fig. 4.18., signifies that the SRL estimator dominated over its counterparts comprehensively regardless of the values of tuning parameter *d*. As the *d* value increases then the SMSE values of the FOA SRL estimator become close to the SMSE values of the FOA mixed estimator. For *d* values larger than 0.9990 the FOA mixed estimator seems slightly better than the FOA SRL estimator in terms of SMSE criterion.

In the light of tabular and graphical results, it is evident that the SRL estimator performs effectively in contrast to its counterparts in terms of SMSE criterion against multicollinearity problems. However, the values of tuning parameter *d*, Σ, and the type of restriction also affected the performance of the estimators.

Table 4.15. Values of the tuning parameters k and d corresponding to better performance of the estimators among the restricted and unrestricted estimators for numerical examples

Estimators	Data Sets	Restrictions	Better estimator compared to	(k, d) obtained by MSE	(k,d) obtained by GA
FOA					
ML	Swedish Football data	-	FOA restricted and unrestricted	-	$(1.992158, 6.571125 \times 10^{-5})$
	Weather data	-	FOA restricted and unrestricted		$(1.997748, 0.0008013207)$, $(1.997065, 0.001288869)$
Rest.ML	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	FOA restricted and unrestricted estimators except ML		$(1.992158, 6.571125 \times 10^{-5})$
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	FOA restricted and unrestricted estimators except ML	-	$(1.997748, 0.0008013207)$, $(1.997065, 0.001288869)$
OK	Swedish Football data	-	FOA restricted and unrestricted for 1st pair and among unrestricted for 2nd pair	$(0.0083, 0.2522)$, $(0.0083, d<0.74)$	-
	Weather data	-	FOA restricted and unrestricted estimators		$(1.0925, 0.9920)$
Rest.OK	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	FOA restricted estimators	$(0.0083, 0.2522)$, $(0.0083, d<0.34)$,$(0.01<k<0.05, 0.2522)$	-
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	FOA restricted estimators	$(1.0925, 0.9920)$, $(k=1.0925, 0.42<d<1)$, $(0.57<k<1, 0.9920)$	-
Ridge	Swedish Football data	-	ML, PCR, $r-k$ for $k=0.0325$ and $(k \leq 0.05)$ for ML	$(k \leq 0.05, k=0.0325)$	-
	Weather data	-	-	-	-
Rest.Ridge	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	-	-	-
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	FOA restricted estimators	$(0.0001<k<0.57, 0.2522)$	-
Liu	Swedish Football data	-	FOA unrestricted estimators	$(k=0.0083, d>0.74)$	-
	Weather data	-	-	-	-
Rest.Liu	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	FOA restricted estimators	$(k=0.0083, 0.34<d<0.95)$	-
	Weather data	-	-	-	-
SRest.Liu	Swedish Football data	$\beta_2+\beta_4+\beta_5+G=0.02$	ML, Liu and Mixed	$(d=0.001, 0.5077, 0.20<d<0.85)$	-
	Swedish Football data	$\beta_2+\beta_3+\beta_5+4\beta_6+G=0.05$	ML, Liu and Mixed	$(d=0.001, 0.5077, 0.20<d<0.99)$	-
	Swedish Football data	$\beta_1+2\beta_3+\beta_4+\beta_5+G=0.02$	ML, Liu and Mixed	$(d=0.001, 0.8429, 0.10<d<0.85)$	-
	Reaction data	$\beta_1+\beta_2+0.5\beta_3+G=0.1653$	ML, Liu and Mixed	$(d=0.001, 0.8229, 0<d<0.999)$	-
SRest.OK	Swedish Football data	$\beta_2+\beta_4+\beta_5+G=0.02$	ML, Mixed and SRest.ridge	-	$(0.03396, 0.11029)$, $(0.03512, 0.09615)$, $(k<0.4, 0.09615)$, $(0.03512, d<0.80)$
	Swedish Football data	$\beta_2+\beta_3+\beta_5+4\beta_6+G=0.05$	ML, Mixed and SRest.ridge		$(0.09731, 0.03043)$, $(0.99019, 0.12752)$
	Swedish Football data	$\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6+G=0$	ML, Mixed and SRest.ridge	-	$(0.94064, 0.01077)$, $(0.01067, 0.88014)$
SRest.Ridge	Swedish Football data	$\beta_2+\beta_4+\beta_5+G=0.02$	ML, Mixed, SRest.OK for 1st pair and better than Mixed for 2nd pair	-	$(k=0.03512, d>0.80, k<0.1, d=0.09615)$
	Swedish Football data	$\beta_2+\beta_3+\beta_5+4\beta_6+G=0.05$	ML and Mixed	-	$(0.09731, 0.03043)$, $(0.99019, 0.12752)$
	Swedish Football data	$\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\beta_6+G=0$	ML and Mixed	-	$(0.94064, 0.01077)$, $(0.01067, 0.88014)$
Mixed	Swedish Football data	$\beta_2+\beta_4+\beta_5+G=0.02$	ML, Liu and SRest.Liu	$(d>0.85)$	-
	Swedish Football data	$\beta_2+\beta_3+\beta_5+4\beta_6+G=0.05$	ML, Liu and SRest.Liu	$(d>0.99)$	-
	Swedish Football data	$\beta_1+2\beta_3+\beta_4+\beta_5+G=0.02$	ML, Liu and SRest.Liu	$(d>0.85)$	-
	Reaction data	$\beta_1+\beta_2+0.5\beta_3+G=0.1653$	ML, Liu and SRest.Liu	$(d>0.9990)$	-
$r-k$	Swedish Football data	-	ML, ridge, PCR for $k=0.0001$ and $(k \leq 2.5)$ for PCR	$(k \leq 2.5, k=0.0001)$	-
ML-Based					
ML	Swedish Football data	-	-	-	-
	Weather data	-	-	-	-
Rest.ML	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	Restricted and unrestricted ML-based estimators	-	$(1.992158, 6.571125 \times 10^{-5})$
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	-	-	-
OK	Swedish Football data	-	-	-	-
	Weather data	-	-	-	-
Rest.OK	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	ML-based restricted and unrestricted	$(0.0083, 0.2522)$	-
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	-	-	-
Ridge	Swedish Football data	-	-	-	-
	Weather data	-	Unrestricted ML-based estimators	$(1.0925, 0.9920)$	$(1.997748, 0.0008013207)$, $(1.997065, 0.001288869)$
Rest.Ridge	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	-	-	-
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	ML-based restricted and unrestricted estimators	$(1.0925, 0.9920)$	$(1.997748, 0.0008013207)$, $(1.997065, 0.001288869)$
Liu	Swedish Football data	-	ML-based unrestricted estimators	$(0.0083, 0.2522)$	$(1.992159, 6.571125 \times 10^{-5})$
	Weather data	-	-	-	-
Rest.Liu	Swedish Football data	$\beta_0+\beta_1-2\beta_6=-0.0441$	-	-	-
	Weather data	$\beta_2+\beta_3+\beta_4=0.1245$	-	-	-

5. MONTE CARLO SIMULATION STUDIES

This chapter gives the application of the proposed estimators given in Section 3, through the Monte Carlo simulations experiments. The Monte Carlo simulation experiments are performed when the response belongs to binomial, Poisson, gamma, and negative binomial distributions, respectively. Comparisons between the newly developed estimators and existing estimators are also done in order to assess their performance. The performance of the estimators is evaluated using either the expected mean square error (EMSE) or the prediction mean square error (PMSE) criteria which are calculated by using the formulas below:

$$EMSE(\tilde{\beta}) = \frac{1}{MCN} \sum_{s=1}^{MCN} \left(\tilde{\beta}_{(s)} - \beta\right)^T \left(\tilde{\beta}_{(s)} - \beta\right)$$

$$PMSE(\tilde{\mu}) = \frac{1}{MCN} \sum_{s=1}^{MCN} \left(\tilde{\mu}_{(s)} - \mu\right)^T \left(\tilde{\mu}_{(s)} - \mu\right)$$

where the subscript (s) denotes the s-th replication of the experiment under consideration, while $\tilde{\beta}_{(s)}$ and $\tilde{\mu}_{(s)}$ are the estimates of β and μ, respectively. MCN represents the number of replications in the Monte Carlo simulation and is replicated up to 1000 times. It is obvious that degrees of multicollinearity, sample size, number of explanatory variables, selection of the tuning parameters (k,d), as well as the distribution of response variable significantly affect the EMSE or PMSE of the estimators. As a result, simulation experiments are carried out with various choices of these parameters. Our computations are carried out by employing MATLAB and R programming languages. The explanatory variables are generated throughout the simulation process using the equation given by (Mcdonald and Galarneau, 1975) which is:

$$x_{ij} = (1-\gamma^2)^{1/2}\nu_{ij} + \gamma\nu_{i,p+1}, i = 1, \cdots, n, \ j = 1, \cdots, p$$

where ν_{ij} are independent standard normal pseudo-random numbers and γ is specified in such a way that correlation among any of two explanatory variables is considered to be γ^2. After generating the explanatory variables they are assumed to be fixed during the simulation process.

5.1. Simulation Study for Poisson Response

In this section, simulation experiments are performed when the response variable belongs to the Poisson distribution for the estimators under consideration.

5.1.1. The $r-k$ Class Estimator

This section provides an application of the $r-k$ class estimator via Monte Carlo simulation study having a response variable from Poisson distribution. Apart from the ML estimator, ridge, and PCR estimators are also taken into account when comparing the performance with the $r-k$ class estimator. The performance is evaluated by the EMSE criterion mentioned above. Comparison is provided under different conditions such as various degrees of multicollinearity, choices about the number of explanatory variables as well as various sample sizes. The simulation results are done by using the MATLAB programming language. The remaining steps in the Monte Carlo simulation experiment technique are as follows:

1. To perceive the impact of the sample sizes on the performance of the estimators we selected $n = 200,\ 400,\ 600$, and 800, whereas the number of explanatory variables

are to be selected as $p = 8$, 10, and 12.

2. The explanatory variables are standardized through unit length standardization approach so that $X^T X$ is a matrix of correlation.

3. To examine the influence of the correlation in the simulation study we considered the values $\gamma^2 = 0.90$, 0.95, and 0.99.

4. For every set of explanatory variables, parameter vector β is selected as an eigenvector corresponding to the substantially large eigenvalue of $X^T X$ matrix.

5. By using the pseudo-random numbers the response variable is generated from a Poisson distribution that is $y_i \sim P(\mu_i)$ having the log link function, which might be expressed as $\mu_i = \exp\left(\beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_p x_{ip}\right)$, where $i = 1, 2, ..., n$ and $\beta_j,\ j = 1, ..., p$.

6. The initial values are set to be $\mu_i^{(0)} = \exp(X\hat{\beta}_{ols})$, where $\hat{\beta}_{ols}$ is the OLS estimator. And the initial working response for the i-th observation is considered as $z_i^{(0)} = X\hat{\beta}_{ols} + (W^{(0)})^{-1}\left(y_i - \mu_i^{(0)}\right)$ for the Poisson distribution, whereas $W^{(0)}$ is a weight matrix and calculated as diagonal elements of $\left(\mu_i^{(0)}\right)$.

7. The ML estimator is acquired by using an iterative algorithm along with the weight matrix $\hat{W}$ which remains to be fixed at the final iteration during the process of finding the iterative estimators. The value 1×10^{-6} is to be used as a convergence criterion for achieving the ML estimator and the iteration is carried on until the desired convergence criterion is met that is $||\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}|| \leq 1 \times 10^{-6}$.

8. For selecting the number of PCs we used the method of PTV. An arbitrarily selected cut-off point value is chosen as 0.90.

9. The value of k for the FOA ridge estimator is found by using $k = \frac{r}{\hat{\beta}_{ols}^T M_r M_r^T \hat{\beta}_{ols}}$.

Reduction rate (RR) is also computed with respect to the $r-k$ class estimator that is

$$RR = \frac{EMSE(\tilde{\beta}) - EMSE(\hat{\beta}_{rk})}{EMSE(\tilde{\beta})}$$

where $\tilde{\beta}$ is any estimator to be compared with the $r-k$ class estimator. RR is used to quantify the strength of the $r-k$ class estimator over its competitors.

The simulated results of the estimators are shown in Tables A1-A3 given in Appendix-A. Tables A1-A3 reveal the performance of ML, ridge, PCR, and $r-k$ class estimators in terms of their EMSE values. The findings obtained are described as follows:

Tables A1-A3, demonstrate that the ML estimator $\hat{\beta}_{ML}$ have the largest EMSE values for all the sample sizes, as well as for all levels of multicollinearity, and the number of explanatory variables when compared to the other biased estimators. It is to be noted that as the levels of multicollinearity increased the EMSE values of the ML estimator also increased. This dramatic change in the EMSE values of ML estimator can also be observed as the number of explanatory variables increased. This shows that ML estimator is positively related with levels of multicollinearity and the number of explanatory variables. However, the role of the sample size is not considered to be a significant factor in affecting the performance of the ML estimator. Although, a small amount of difference in the EMSE values of the ML estimator is observed by a change in sample size. This difference is to be seen either in the increasing or decreasing form. Likewise, the RR of the ML estimator is observed to be almost 90% at the lowest level of multicollinearity which is 0.90 and it increased as the levels of multicollinearity, the number of explanatory variables, and sample size are increased. This increment almost reached a value of 100% RR as the level of multicollinearity moved from 0.90 to 0.99. Therefore, the RR of the ML estimator is positively related with the levels of multicollinearity, number of explanatory variables, and sample size. Moreover, the RR of the ML estimator is seen to be highest as compared to ridge and PCR estimators in all aspects.

Similarly, the ridge estimator is also affected by changing the levels of multicollinearity and the number of explanatory variables. It is to be perceived that as the levels of multicollinearity increased, the EMSE values of the ridge estimator increased as well. This increment can also be noticed as the number of explanatory variables increased. In so far as a ridge estimator is to be concerned it reacted almost very similar to the ML estimator against multicollinearity regardless of the number of explanatory variables and the sample size as well. That is the EMSE values of $\hat{\beta}(k)$ are seen to be increased as the number of explanatory variables and the levels of multicollinearity increased. Thus, EMSE of $\hat{\beta}(k)$ is observed to be positively related with the levels of multicollinearity and the number of explanatory variables. Whereas, the effect of sample size is to be monitored either in the increasing or decreasing form as in the case of ML estimator. As long as the RR of the ridge estimator is concerned it is a bit low at lowest level of multicollinearity and then increased gradually as the number of explanatory variables and levels of multicollinearity are increased. It is to be noticed that the RR is approximately 95% in the presence of high multicollinearity and a large number of explanatory variables. Nevertheless, the sample size is not considered to be as an important factor in affecting the RR of ridge estimator. Although, a minor change is to be observed in RR either in the increasing or decreasing form by change of a sample size.

For the PCR estimator, it is conceived that at the lowest level of multicollinearity the EMSE values of the PCR estimator are high. Furthermore, as the number of explanatory variables and sample size increased the EMSE of the PCR estimator increased as well at the aforementioned level of multicollinearity. So, the number of explanatory variables and the sample size are positively related to the EMSE of the PCR estimator at that point of multicollinearity. As the levels of multicollinearity and the number of explanatory variables increased the EMSE values of the PCR estimator decreased quickly. Therefore, an inverse relationship is investigated between the EMSE values of the PCR estimator with high levels of multicollinearity and a large number of explanatory variables. Similarly, the RR for the PCR estimator seemed to be high at the lowest level of multicollinearity and it increased as the sample size and the explanatory variables are increased at this level and shows almost 77% level of RR. But, as the levels of multicollinearity and number of explanatory variables increased the RR of PCR estimator becomes very low which means that it did not differ very much from $r-k$ class estimator in the presence of high multicollinearity and a large number of explanatory variables. The role of the sample size is again not considered to be significant in affecting the RR although, a small change either in the increasing or decreasing form is to be observed in RR.

Here, we discuss the behavior of the $r-k$ class estimator under different scenarios. From the results, it is explored that at the lowest level of multicollinearity the EMSE values of $\hat{\beta}_{rk}$ are high and seem to be gradually increased as the number of explanatory variables increases. Thus, $\hat{\beta}_{rk}$ is positively related to the number of explanatory variables at that particular point of multicollinearity. Nevertheless, as the levels of multicollinearity and number of explanatory variables increased the EMSE values of $\hat{\beta}_{rk}$ decreased rapidly which displays that the high levels of multicollinearity and a large number of explanatory variables are inversely related to the EMSE of $\hat{\beta}_{rk}$. Besides this, no serious effect of the sample size is realized on the behavior of $\hat{\beta}_{rk}$.

Now, we make comparisons of the $r-k$ class estimator against its counterparts which are $\hat{\beta}_{ML}$, $\hat{\beta}(k)$, and $\hat{\beta}_r$ respectively in terms of the EMSE criterion. It is apparent that the $r-k$ class estimator performed efficiently as compared to all of its competitors regardless of the level of multicollinearity, the number of explanatory variables, and the sample size. The $\hat{\beta}_{rk}$ contains the smallest EMSE values as compared to its counterparts. As a result, the $r-k$ class estimator outperforms all of its competitors.

5.1.2. Restricted OK Estimator

In this section, we examine the performance of the restricted FOA OK estimator with the FOA ML, restricted FOA ML, FOA ridge, restricted FOA ridge, FOA Liu, restricted FOA Liu, and FOA OK estimators. Furthermore, the performance of these estimators in their iterative form is investigated. The EMSE and PMSE criteria described in Section 5 are used to assess the performance: The following is a detailed procedure for conducting the Monte Carlo simulation study:

1. In order to monitor the impact of sample size on the behavior of estimators it is chosen as: $n = 25, 50, 100, 200, 500, 900$, and the number of explanatory variables is $p = 4$.
2. A constant term is introduced to the matrix of explanatory variables after the explanatory variables have been generated.
3. The correlation values are chosen as: $\gamma^2 = 0.75, 0.85, 0.95, 0.99$ to perceive the effects of multicollinearity on the performance of estimators.
4. The parameter vector β is figured out as $\beta_i = (\|\beta\| / \sum u_i^2) \times u_i$ where u_i is sampled from a uniform distribution on the interval [-1,1] which followed from (Delaney and Chaterjee, 1986) and (Özkale and Altuner, 2021) and $\|\beta\|$ is considered as 3 in this study which shows the effect of signal to noise ratio.
5. To analyze the effects of restrictions on the behavior of the estimators, we consider the restriction $R = j_{1\times p}$ and $r = \sqrt{p} * \kappa$ where κ is arbitrarily chosen as 0.5. This restriction may be useful in situations to identify whether the sum of the explanatory

variables [17] has as much effect as $\sqrt{p} * \kappa$ on the linear predictor and κ shows the degree of the effect.

6. The response variable is generated from Poisson distribution such as $y_i \sim P(\mu_i)$, which contains the log link function and evaluated as $\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})$, where $i = 1, \cdots, n$ and $j = 1, \cdots, p$.

7. The OLS estimator $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y$ is used as an initial estimate for all the estimators such as $\hat{\mu}^{(0)} = \exp(X\hat{\beta}_{OLS})$ along with the initial working response $\hat{z}^{(0)} = X\hat{\beta}_{OLS} + (\hat{W}^{(0)})^{-1}(y - \hat{\mu}^{(0)})$ where $\hat{W}^{(0)} = diag(\hat{\mu}_i^{(0)})$ is a weight matrix with the i-th element of $\hat{\mu}^{(0)}$. All the estimators are computed in the FOA and iterative form.

8. The tuning parameter values are found by using the procedure given in subsubsections 3.2.2.1., and 3.2.2.2, such that the first (k,d) pair is obtained by minimizing MSE, and the second and third (k,d) pairs are found by GA via MSEP and MAE which are respectively denoted as (k_1,d_1), (k_2,d_2) and (k_3,d_3).

The simulation results are reported in Tables A-11-A-16 given in Appendix A [18] and the findings obtained from the simulation results are outlined as follows:

i) For (k_1,d_1) the EMSE values of the iterative unrestricted estimators increase as γ^2 increases. However, for the FOA estimators as γ^2 increases there is not any systematic pattern is seen in the EMSE of the estimators.

ii) When the iterative restricted estimators are compared among themselves it is seen that while the sample size is large (200, 500, 900), $\hat{\beta}_{kd-R}^{(t+1)}$ for (k_1,d_1) is better than other iterative restricted estimators in the sense of EMSE. When the iterative unrestricted estimators are examined among themselves it is seen that $\hat{\beta}_{kd}^{(t+1)}$ for (k_1,d_1)

[17] The sum of elements of β chosen in item 4 is zero

[18] Bold text is used to show the minimum values corresponding to unrestricted estimators and bold italic text is used to show the minimum values corresponding to restricted estimators. PMSE values are shown in parenthesis.

is better than other iterative unrestricted estimators when the sample size is 900 in terms of EMSE, however for the other sample sizes the performance of the estimators changes as the sample size and degree of correlation change so no definite arguments stand regarding which estimator performs better. For the FOA estimators when the sample size is relatively small less than 500 both $\hat{\beta}(k)^{(1)}$ and $\hat{\beta}_R(k)^{(1)}$ outperform the other corresponding unrestricted and restricted FOA estimators for (k_1, d_1). When the restricted FOA estimators are taken into account, $\hat{\beta}^{(1)}_{kd-R}$ for (k_1, d_1) performs efficiently as compared to other restricted FOA estimators when $n = 900$, 500 (for $\gamma^2 = 0.85, 0.95, 0.99$), 200 (for $\gamma^2 = 0.95, 0.99$).

iii) Restricted iterative and FOA estimators have larger EMSE values as compared to the unrestricted iterative and FOA estimators for (k_1, d_1).

iv) The iterative restricted estimators have larger PMSE values than their corresponding iterative unrestricted estimators for (k_1, d_1). However, no significant difference is seen for the FOA estimators.

v) $\hat{\beta}_R(k)^{(t+1)}$ for (k_1, d_1) has the smallest PMSE values among the iterative restricted estimators (except $n = 900$, $\gamma^2 = 0.85$ and $n = 100$, $\gamma^2 = 0.95$), while there is no obvious arguments about the iterative unrestricted estimators. Because the PMSE values of the estimators change as the sample size and degree of correlation change, and this change is not a systematic way so it is difficult to say which estimator stands alone as a better estimator in terms of the PMSE criterion. On the other hand, $\hat{\beta}_R(k)^{(1)}$ and $\hat{\beta}(k)^{(1)}$ for (k_1, d_1) have smallest PMSE values among the corresponding restricted and unrestricted FOA estimators.

vi) As γ^2 increases the PMSE values of iterative restricted estimators increases. On the contrary, there are no obvious arguments about the iterative unrestricted estimators and the restricted and unrestricted FOA estimators.

vii) When sample size is small (25, 50, 100) and large (900), $\hat{\beta}(d)^{(1)}$ and $\hat{\beta}_R(d)^{(1)}$ for both (k_2,d_2) and (k_3,d_3) often dominate the other unrestricted and restricted FOA estimators in the sense of EMSE. Although there is no clear rule regarding the EMSE of the iterative estimators for both (k_2,d_2) and (k_3,d_3), there are many cases where $\hat{\beta}_R(k)^{(t+1)}$ and $\hat{\beta}_R(d)^{(t+1)}$ are better than other restricted estimators. Similarly, there are many cases where $\hat{\beta}^{(t+1)}$, $\hat{\beta}(k)^{(t+1)}$, and $\hat{\beta}(d)^{(t+1)}$ are better than other unrestricted estimators.

viii) $\hat{\beta}(d)^{(1)}$ and $\hat{\beta}_R(d)^{(1)}$ for both (k_2,d_2) and (k_3,d_3) generally give the best PMSE among unrestricted and restricted estimators, respectively. $\hat{\beta}(d)^{(t+1)}$ is generally the best in terms of PMSE among the unrestricted estimators, while $\hat{\beta}_R(d)^{(t+1)}$ and $\hat{\beta}_R(k)^{(t+1)}$ for both (k_2,d_2) and (k_3,d_3) are generally give better results among the restricted estimators in having small PMSE value. Furthermore, no clear rules exists about when other estimators are good.

As a result of the comparisons made by considering (k_1,d_1) when the sample size is large and the correlation is small $\hat{\beta}^{(t+1)}_{kd-R}$ and $\hat{\beta}^{(1)}_{kd-R}$ are better than the other restricted estimators according to the EMSE criterion. On the other hand, when the sample size is small $\hat{\beta}(k)^{(1)}$ and $\hat{\beta}_R(k)^{(1)}$ perform efficiently as compared to their corresponding unrestricted and restricted FOA estimators while this case is not observed for the iterative estimators. Furthermore, contrary to this result for (k_1,d_1) no generalizations can be made for (k_2,d_2) and (k_3,d_3).

5.2. Simulation Study for Gamma Response

In this section, simulation studies are conducted when the response variable belongs to a gamma distribution.

5.2.1. The $r-k$ Class Estimator

This section provides the comparisons of the $r-k$ class estimator with its counterparts that are; ML, ridge, and PCR estimators, respectively. The response variable is originated from a gamma distribution like; $y \sim gamma(\nu = \mu^2/var, u = var/\mu)$, where $var = \mu^2$ and $\mu = \exp(X\beta)$ as well as the initial fitted values are set to be as: $\mu_i^{(0)} = \frac{(y+\bar{y})}{2}$ to compute IRLS and the weight matrix $W^{(0)}$ is computed as $\hat{W}^{(0)} = diag\left(\mu_i^{(0)}\right)^2$. In addition, the initial working response z for the i-th observation is considered as: $z_i^{(0)} = \frac{1}{\mu_i^{(0)}} - \left(W^{(0)}\right)^{-1}\left(y_i - \mu_i^{(0)}\right)$. The cut-off point value which is arbitrarily chosen as 0.80 for selecting the number of PCs. Other aspects in this experimental set up are considered to be the same as in the case of the Poisson distribution explained in subsection 5.1.1., such as the number of explanatory variables, sample size, choice of the degree of correlation, and selection of the parameter vector β.

The results are reported in Tables A4-A6 given in Appendix-A and findings are outlined as:

These results show that the ML estimator has larger EMSE values as compared to other biased estimators for all the aspects that are considered. The values of RR for the ML estimator are also higher as compared to ridge and PCR estimators, which means that it differs substantially from the $r-k$ class estimator. The ridge estimator also contains large EMSE values as compared to PCR and $r-k$ class estimators. The RR values of the ridge estimator are also, higher than the PCR estimator. Furthermore, it is evident that the $r-k$ class estimator consists of the smallest EMSE values as compared to its counterparts for all the different scenarios which are considered in the simulation study. Although, the RR values of the PCR estimator are seemed to be small, which means that the EMSE values of the PCR estimator are not much deviated from the EMSE values of the $r-k$ class estimator. Nevertheless, the $r-k$

class estimator performed efficiently as compared to its counterparts according to the EMSE criterion at all levels of multicollinearity, number of explanatory variables as well as for each sample size.

5.2.2. Restricted OK Estimator

In this section, we compare the performance of the estimators when the response variable comes from a gamma distribution. The same strategy as mentioned in subsection 5.1.2, is followed to compare the performance of the restricted OK estimator with its counterparts, which are the ML, restricted ML, ridge, restricted ridge, Liu, restricted Liu, and OK estimators in the FOA and iterative forms. The only difference is that the response variable is generated from a gamma distribution rather than a Poisson that is $y \sim gamma(\nu = \mu^2/var, u = var/\mu)$, where, $var = \mu^2$ and μ is a vector with elements $\mu_i = \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})$ and the initial weight matrix $\hat{W}_0$ for the i-th element of $\hat{\mu}^{(0)}$ is computed as $\hat{W}^{(0)} = diag\left(\mu_i^{(0)}\right)^2$.

Tables A17-A22, Appendix A demonstrate the results which are summarized as follows:

i) It is found that the estimators $\hat{\beta}(k)^{(t+1)}$ and $\hat{\beta}_R(k)^{(t+1)}$ outperform the unrestricted and restricted iterative estimators in terms of EMSE criterion for (k_1, d_1). When the sample size is small and correlation is 0.99 and the sample size is large (500, 900) and in all correlations, $\hat{\beta}(k)^{(1)}$ and $\hat{\beta}_R(k)^{(1)}$ are the best within the class of unrestricted and restricted FOA estimators according to EMSE for (k_1, d_1). $\hat{\beta}^{(1)}_{kd-R}$ for (k_1, d_1) is best in the class of restricted estimators according to EMSE when sample size is small and correlation is small at the same time.

ii) Increase of γ^2 does not cause systematic increase or decrease in EMSE values of all the estimators for (k_1, d_1).

iii) In most of the cases, $\hat{\beta}_R(k)^{(1)}/\hat{\beta}(k)^{(1)}$ are the best among the restricted/unrestricted

FOA estimators in terms of PMSE when (k_1, d_1) is used. However, a general result cannot be obtained for the iterative estimators considering (k_1, d_1) when PMSE is taken into account.

iv) In many cases, the FOA and iterative unrestricted/restricted ridge and Liu estimators for (k_2, d_2) and (k_3, d_3) dominate the other unrestricted/restricted estimators in achieving a small EMSE value.

v) For the gamma response GLM, EMSE and PMSE are minimal at the same values of estimators for (k_2, d_2) and (k_3, d_3). Therefore, the results obtained based on EMSE are also valid for PMSE for (k_2, d_2) and (k_3, d_3).

As a result of comparisons done via gamma response, we get the results for the restricted estimators that $\hat{\beta}_R(k)^{(1)}$ for (k_1, d_1) is the best particularly when the sample size and the degree of correlation are simultaneously small then, $\hat{\beta}^{(1)}_{kd-R}$ for (k_1, d_1) is the best in having small EMSE value. While $\hat{\beta}(k)^{(1)}$ seems to be better for (k_1, d_1) among the unrestricted estimators in having smaller EMSE or PMSE values. On the other hand, when (k_2, d_2) and (k_3, d_3) are considered it is seen that $\hat{\beta}(k)^{(1)}/\hat{\beta}(d)^{(1)}$ or $\hat{\beta}(k)^{(t+1)}/\hat{\beta}(d)^{(t+1)}$ are the best among the unrestricted estimators and $\hat{\beta}_R(k)^{(1)}/\hat{\beta}_R(d)^{(1)}$ or $\hat{\beta}_R(k)^{(t+1)}/\hat{\beta}_R(d)^{(t+1)}$ dominated the other restricted estimators in the sense of EMSE.

5.2.3. Stochastic Restricted OK Estimator

In this section, the performance of the stochastic restricted OK estimator is compared with its competitors which are, the ML estimator, mixed, and stochastic restricted ridge estimators, respectively. R programming language is used to compute the results and the seed value is considered to be 1122. The rest steps of performing the Monte Carlo simulation are outlined below:

1. We choose the sample size n =100, 250, 350, 500, 700, and 900 to see its impact

on the performance of the estimators and the number of explanatory variables used $p = 4$.

3. The values of multicollinearity degree γ^2 are assumed to be $\gamma^2 = 0.75$, 0.85, 0.90 and 0.95.

4. The parameter vector β is determined as a normalized eigenvector corresponding to the largest eigenvalue of $X^T X$ matrix in such a way that $\beta^T \beta = 1$.

5. By following (Varathan and Wijekoon, 2018) and (Zuo and Li, 2018) the stochastic restrictions are considered as:

$$R = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, r = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

6. The values of tuning parameters k and d are obtained by using the technique of GA adopted in subsection 4.1.4.

7. The response variable is generated as $y_i \sim gamma\left(\mu_i^2/v, v/\mu_i\right)$ where $\mu_i = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_p x_{ip})$ and v is equal to μ_i^2.

8. OLS estimator $\hat{\beta}_{ols} = \beta^{(0)} = \left(X^T X\right)^{-1} X^T y$ is treated as a primitive estimate and the initial weight matrix $W^{(0)}$ which is figured out to be $W^{(0)} = diag(\mu_i^{(0)})^2$.

The results of the estimators in their FOA form are shown in Table A-8 as given in Appendix-A and described as follows:

i) It is observed that the performance of the stochastic restricted OK estimator dominated thoroughly over its counterparts as it acquired the smallest EMSE values as compared to its counterparts regardless of all the aspects considered in the simulation study. While the effects of degrees of multicollinearity and sample sizes on the behavior of the estimators are described below:

ii) A positive relationship is investigated between the degrees of multicollinearity and EMSE of ML and stochastic restricted OK estimators, because as the degrees of multi-

collinearity increases EMSE values of both the estimators increases. Multicollinearity did not have persistent effects on the behavior of mixed and stochastic restricted ridge estimators though, the EMSE values of these estimators are found in the increased or decreased form as the degrees of multicollinearity increased.
iii) It is seen that as the sample size goes up to 350 then, the EMSE of the mixed and stochastic restricted OK estimator decreased continuously. While the sample size did not have a systematic effect on the EMSE of ML and stochastic restricted ridge estimators.

5.3. Simulation Study for Negative Binomial Response

In this section, a simulation study is conducted when the response variable is generated from a negative binomial distribution.

5.3.1. Stochastic Restricted OK Estimator

This section evaluates the performance of the stochastic restricted OK estimator and its counterparts that are ML, mixed, and stochastic restricted ridge estimators, respectively. In this Monte Carlo simulation experiment the same experimental set up is considered as in the case of gamma response variable explained in subsection 5.2.3. The only difference is that the response variable is generated from a negative binomial distribution instead of gamma that is; $y_i \sim NB\left(\mu_i, \mu_i + \alpha\mu_i^2\right)$ we choose $\alpha = 1,2$ by (Mansson, 2012) along with the initial weight matrix $W^{(0)}$ which is computed as $W^{(0)} = diag(\mu_i^{(0)} + \alpha(\mu_i^{(0)})^2)$.
The results are summarized in Tables A9-A10 as shown in Appendix-A and elaborated as:
i) All of the biased estimators outperform the ML estimator as they acquired smaller EMSE values than the ML estimator. The performance of the stochastic restricted

OK estimator dominated its counterparts for every aspect mentioned in the simulation study. Then, the stochastic restricted ridge estimator performs better than the mixed estimator.

ii) By raising the degrees of multicollinearity, a small amount of increment is observed in the EMSE values of the stochastic restricted OK estimator and the EMSE values of the ML estimator are also increased for both the values of α. Besides, the EMSE values of the mixed estimator increased as the degrees of multicollinearity increased except when $\alpha = 2$ and $n = 500$ at higher degrees of multicollinearity. While an increasing or decreasing trend is observed in the EMSE values of the stochastic restricted ridge estimator by increasing the degrees of multicollinearity for both values of α.

iii) As the sample size increases from 100 to 500 then, the EMSE of the mixed, stochastic restricted OK, and stochastic restricted ridge estimators decrease when $\alpha = 1$ and gradually increases as the sample size moves from 700 to 900. However, when $\alpha = 2$ the sample size has no obvious effect on the EMSE values of these estimators.

iv) There appears to be a small decrease in the EMSE values of the stochastic restricted OK estimator for all the sample sizes except $n = 500$ as α changes from 1 to 2 and the EMSE of mixed and stochastic restricted ridge estimators also decrease when the value of α changes from 1 to 2 except the sample sizes $n = 500$ and 700. There is no obvious effect of the value of α on the EMSE of the ML estimator.

5.4. Simulation Study for Binomial Response

In this section, simulation experiments are performed when a response variable comes from a binomial distribution.

5.4.1. Stochastic Restricted OK Estimator

This section illustrate the performance of the stochastic restricted OK estimator and its counterparts which are the ML, mixed, and stochastic restricted ridge estimators. The Monte Carlo simulation experimental settings are the same as we used in the case of gamma and negative binomial response variables. The main distinction is that the response variable in this experiment is derived from a binomial distribution as; $y_i \sim Bern(\mu_i)$ containing the link function $\mu_i = \frac{\exp(\beta_1 x_{i1}+\cdots+\beta_p x_{ip})}{1+\exp(\beta_1 x_{i1}+\cdots+\beta_p x_{ip})}$, $i = 1,\ldots,n$, along with the initial weight matrix $W^{(0)}$ which is figured out to be diagonal elements of $\mu_i^{(0)}(1-\mu_i^{(0)})$.

The results are reported in Table A-7 presented in Appendix-A and described as follows:

i) It is found that the performance of biased estimators is better than the ML estimator in terms of EMSE criterion, as they acquired the smaller EMSE values than the ML estimator.

ii) The performance of the stochastic restricted OK estimator is superior when compared to its counterparts regardless of whether the tuning parameter values are obtained by minimizing MSEP or MAE. Thereafter, the performance of stochastic restricted ridge estimator is better than the mixed estimator.

iii) The performance of stochastic restricted OK estimator prevails over its counterparts for all the degrees of multicollinearity as well as for all the sample sizes considered.

iv) By increasing the degrees of multicollinearity, a small increment is observed in the EMSE values of the stochastic restricted OK estimator, while the EMSE values of ML estimator are highly increased. Whereas, the mingling effects of multicollinearity are seen on the EMSE values of mixed and stochastic restricted ridge estimators that are either in the increasing or decreasing mode by raising the degrees of multicollinearity.

v) As the sample size goes up to 500 then, the EMSE values of the stochastic restricted

OK estimator decreased continuously. However, sample size has no significant effect on other estimators either in the increasing or decreasing trend.

5.4.2. Stochastic Restricted Liu Estimator

This section includes a simulation analysis to assess the performance of the stochastic restricted Liu estimator, ML, mixed, and Liu estimators, respectively. The EMSE is the performance evaluation criterion. The MATLAB programming language is used to carry out our calculations. The following is the rest of the procedure for obtaining Monte Carlo simulation results:

1. To monitor the effect of sample sizes on the behavior of an estimator they are chosen as $n = 300$, 600, and 900 whereas the number of explanatory variables are chosen as $p = 12$.
2. The explanatory variables are standardized through unit length standardization approach, hence $X^T X$ is a matrix of correlation.
3. The values of γ^2 are selected as $\gamma^2 = 0.9, 0.95, 0.99$ to see the effects of multi-collinearity.
4. Parameter vector β is specified as an eigenvector corresponding to the largest eigenvalue of $X^T X$ matrix.
5. To impose the stochastic restrictions, we first randomly generate the R matrix of size $q \times p$ from standard normal distribution where $q = 10$ and $p = 12$ with q less than or equal to p, to avoid any complications regarding computations of results. Then, the random vector G is generated from normal distribution with $(0, \sigma^2 I)$ where the values of σ are arbitrarily chosen as 0.1, 1 and 5. After R and ϕ are generated, then r values are obtained by imposing the equality $r = R\beta + G$ where β is defined as in step 4.
6. The values of tuning parameter d are chosen arbitrarily from 0.1 to 1 with an increment of 0.1.

7. By applying pseudo-random numbers the response variable is produced from a Bernoulli distribution as $y_i \sim Bern(\mu_i)$ having the link function $\mu_i = \frac{\exp(\beta_1 x_{i1}+\cdots+\beta_p x_{ip})}{1+\exp(\beta_1 x_{i1}+\cdots+\beta_p x_{ip})}$, where, $i = 1,\ldots,n$ and β_j, $j = 1,\ldots,p$ as given in step 4.

8. OLS estimator $\hat{\beta}_{ols} = \beta^{(0)} = \left(X^T X\right)^{-1} X^T y$ is considered as an initial estimate for all the estimators and the initial working response for the i-th observation is treated as $z_i^{(0)} = X\beta^{(0)} + (W^{(0)})^{-1}(y - \mu_i^{(0)})$ where $W^{(0)}$ is a weight matrix computed as diagonal elements of $\mu_i^{(0)}(1-\mu_i^{(0)})$.

9. The ML estimator is obtained by an iterative algorithm as well as the weight matrix $\hat{W}$ in the final iteration which is settled to be fixed during the iteration process. The value of convergence criterion is assumed as 1×10^{-6} and the iteration process continues until $||\hat{\beta}^{(t+1)} - \hat{\beta}^{(t)}|| < 1 \times 10^{-6}$ can be achieved.

Figures 5.1.-5.9., summarize the simulation results which are described as follows:

From Figs. 5.1-5.9., it is apparent that the SRL estimator contains the smallest EMSE values and outperforms all other estimators, particularly when the value of d lies in the range from 0.1 to 0.5 and $\sigma = 0.1$, and 1 at all levels of multicollinearity. However, for a large value of σ that is when $\sigma = 5$ then, the SRL estimator again dominates over its competitors precisely especially when the value of d lies in the interval from 0.1 to 0.7 approximately and levels of multicollinearity are $\gamma^2 = 0.90$ and 0.95. Also at a high level of multicollinearity namely $\gamma^2 = 0.99$ it performs efficiently for a large value of $\sigma = 5$ but, for a small range of values of d i.e., when the value of d lies in the range of 0.1 to 0.5 approximately. It is also noted that the behavior of SRL and Liu estimators is slightly similar only for some small values of d but, the SRL estimator still contains smaller EMSE values as compared to the Liu estimator and outperforms it. Figs. 5.1.-5.9., demonstrate that as the value of d becomes greater than approximately 0.5 then, the mixed estimator performs better than all other estimators for $\sigma = 0.1$ and 1 at all levels of multicollinearity. However, for $\sigma = 5$, it performs better only for those values of d which are approximately greater than or equal to 0.7 and $\gamma^2 = 0.90$ and 0.95, respectively. Also at a high level of multicollinearity viz $\gamma^2 = 0.99$ and $\sigma = 5$ mixed estimator performs better for those values of d which are greater than 0.5.

These simulation results indicate that the SRL estimator performs efficiently as compared to the mixed estimator for small to moderate values of d that is when the value of d lies in the range from 0.1 to 0.5. However, it outperforms the ML and Liu estimator regardless of the d values. Furthermore, it is found that sample size is not an important factor in affecting the performance of an estimator in the simulation study. The most important factors to determine the performance of an estimator in this simulation study are the values of d, σ, and γ^2, respectively. Nevertheless, the types of restrictions are also a significant component to determine the performance of the estimators as we noticed in the empirical application of the SRL estimator.

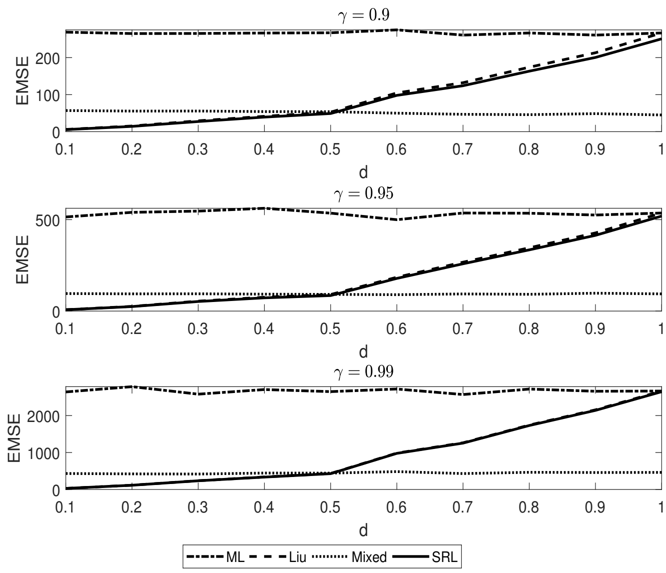


Figure 5.1. Graph of estimated MSE values against d when $n = 300$, $p = 12$ and $\sigma = 0.1$

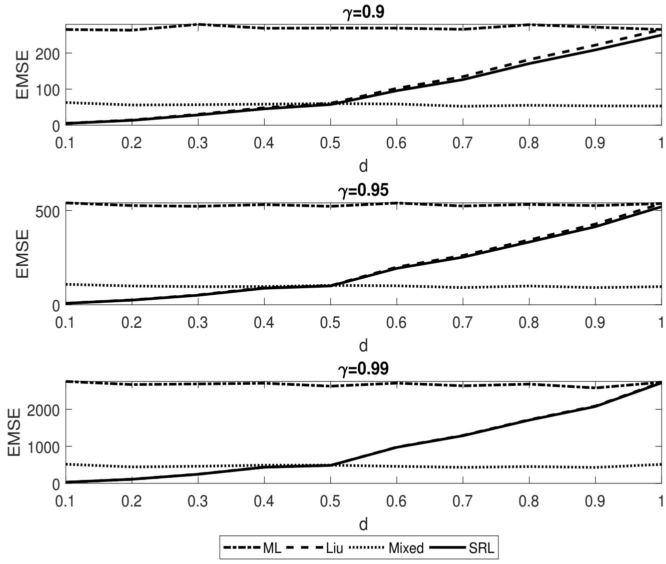


Figure 5.2. Graph of estimated MSE values against d when $n = 300$, $p = 12$ and $\sigma = 1$

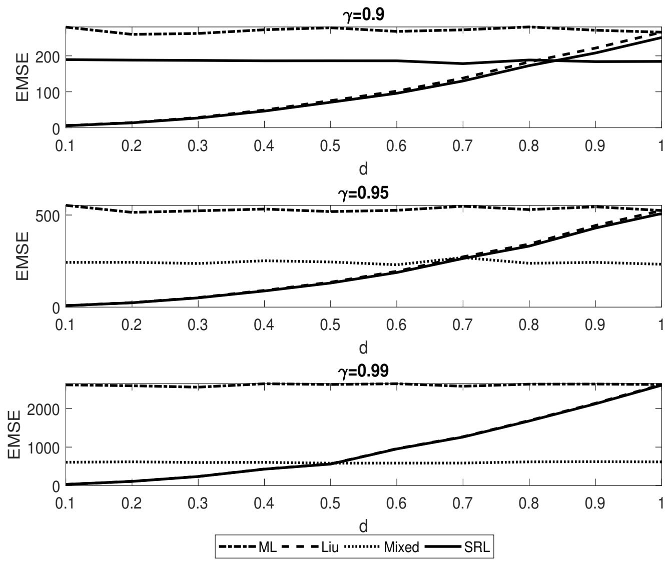


Figure 5.3. Graph of estimated MSE values against d when $n = 300$, $p = 12$ and $\sigma = 5$

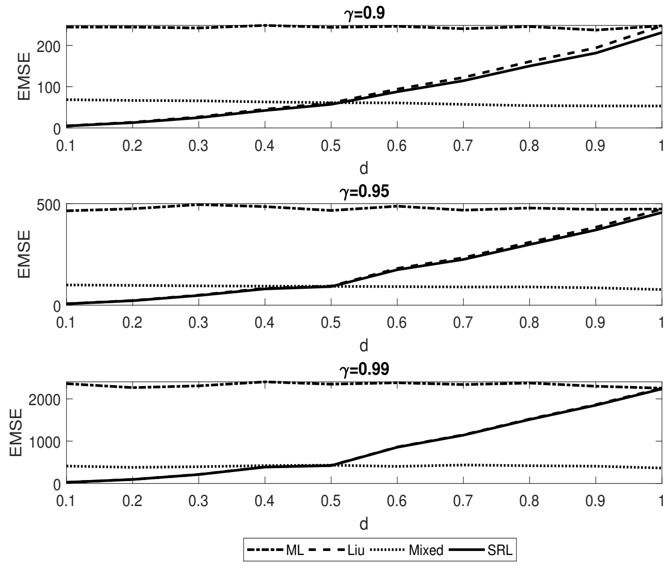


Figure 5.4. Graph of estimated MSE values against d when $n = 600$, $p = 12$ and $\sigma = 0.1$

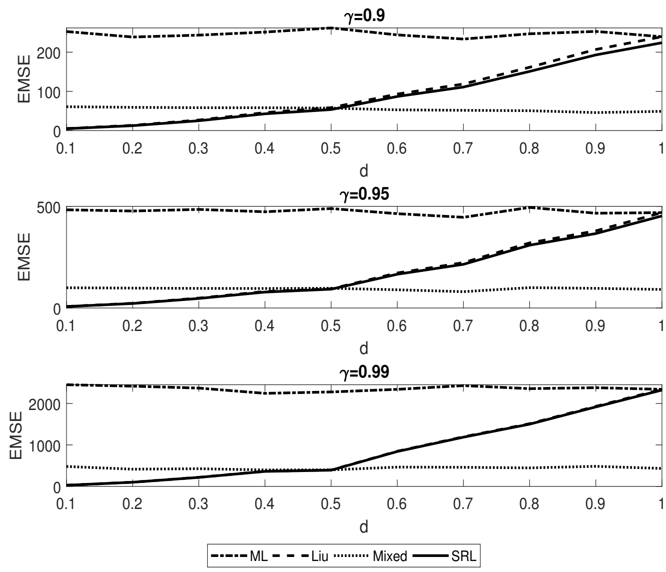


Figure 5.5. Graph of estimated MSE values against d when n=600, $p = 12$ and $\sigma = 1$

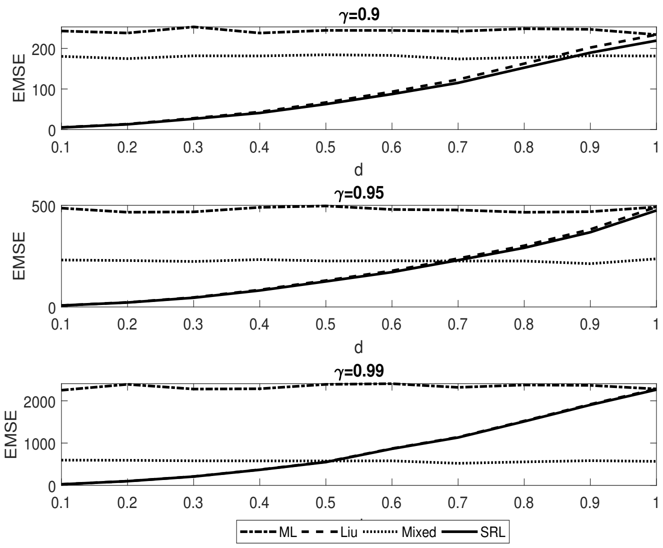


Figure 5.6. Graph of estimated MSE values against d when $n = 600$, $p = 12$ and $\sigma = 5$

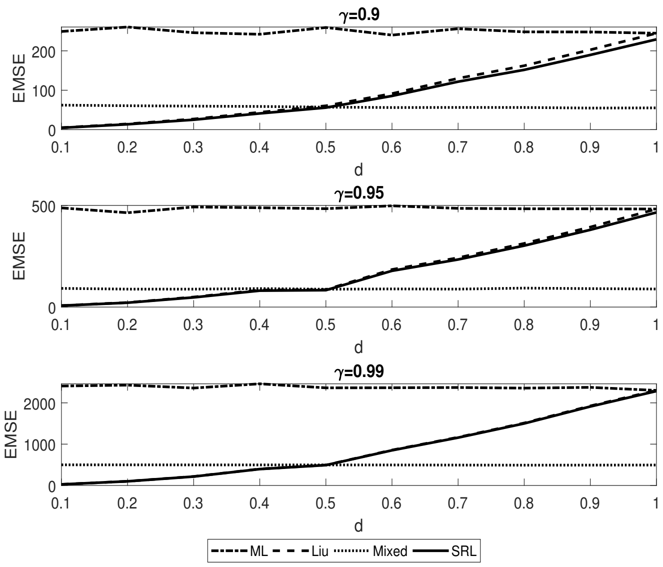


Figure 5.7. Graph of estimated MSE values against d when $n = 900$, $p = 12$ and $\sigma = 0.1$

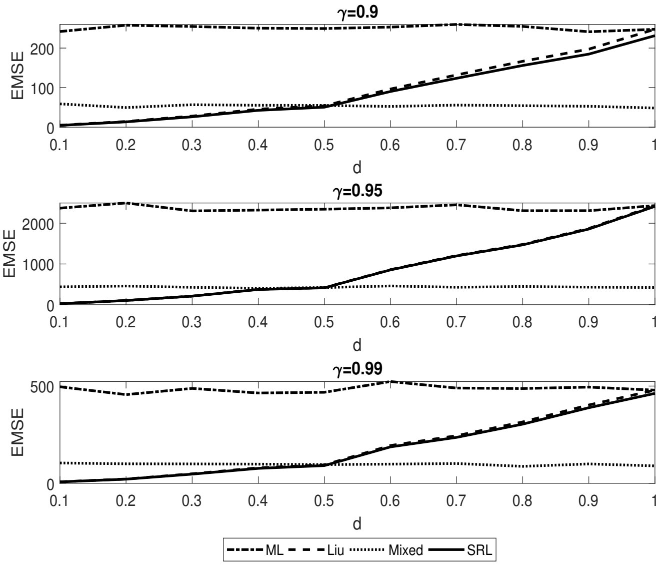


Figure 5.8. Graph of estimated MSE values against d when $n = 900$, $p = 12$ and $\sigma = 1$

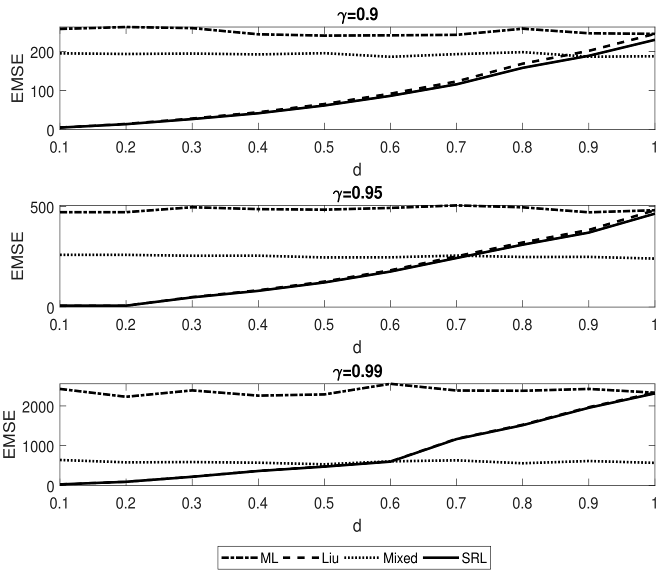


Figure 5.9. Graph of estimated MSE values against d when $n = 900$, $p = 12$ and $\sigma = 5$

6. CONCLUSION

Multicollinearity is a well-known phenomenon that affects not only normal linear models but also GLMs. In both normal linear models and GLMs, the problem of multicollinearity has a significant impact on parameter estimation. The ML approach is often used to estimate the parameters in GLMs however it is seriously affected by the multicollinearity problem.

Four alternative iterative restricted biased estimators to the ML estimator have been proposed in this dissertation to address the problem of multicollinearity such as; the $r-k$ class estimator, stochastic restricted Liu estimator, restricted OK estimator, and stochastic restricted OK estimator, respectively. Exact and stochastic linear restrictions on the parameters are imposed in order to develop these estimators. Furthermore, Fisher's scoring algorithm is used to obtain all of the estimators in an iterative process. The FOA form of the estimators is also given for comparison purposes along with the iterative form. The properties of the estimators as well as the technique for choosing tuning parameters k and d are also discussed. Numerical illustrations and simulation studies are used to assess the performance of various estimators. The SMSE, EMSE, and PMSE are the performance evaluation criteria.

The $r-k$ class estimator a hybrid of ridge and PCR estimators was first obtained and compared to the ML, ridge, and PCR estimators. The estimators are compared in their FOA form. A numerical example with a response variable from a Poisson distribution as well as two simulation studies with responses from Poisson and gamma distributions are used to evaluate the estimators' performance. The results revealed that the performance of the $r-k$ class estimator is better than its counterparts for the appropriate value of ridge parameter k in terms of the SMSE criterion when numerical illustration is considered. However, regardless of the value of k it outperforms its competitors in simulation studies in terms of the EMSE criterion. The number of explanatory variables, the value of ridge parameter k, and the degrees of multicollinearity are all found to influence the estimators' performance.

Then, by applying stochastic restrictions on the parameters the SRL estimator is presented. The SRL estimator is compared with ML, Liu, and mixed estimators in the FOA form. The performance of the estimators is evaluated using two numerical examples with a response variable from the Poisson and gamma distributions, as well as a simulation study with a response from the binomial distribution. According to the results of the empirical applications, the SRL estimator outperforms all other estimators when the values of tuning parameter d lie below 0.85 approximately under the Poisson response. However, the SRL estimator dominated over its counterparts regardless of the value of tuning parameter d when gamma response is considered. Simulation results revealed that the performance of the SRL estimator is better than the mixed estimator for small to moderate values of d that is 0.1 to 0.5 approximately, however it outperforms the ML and Liu estimator regardless of the value of d. Liu estimator behaves slightly similar to the SRL estimator when d values are in the range of $0.1 < d < 0.20$, although the SRL estimator still has lower EMSE values. The important factors which affected the performance of the estimators are the value of tuning parameter d, degree of correlation, the value of dispersion parameter Σ, and the restriction under consideration.

Thereafter, a versatile two parameter restricted OK estimator is proposed when exact linear restrictions are assumed to be hold on the parameters. Under certain conditions, this estimator contains a number of different estimators. This estimator is compared with ML, restricted ML, ridge, restricted ridge, Liu, restricted Liu, and OK estimator. For comparisons, the FOA form and ML-based or iterative versions of these estimators are evaluated. The performance evaluation criteria are; SMSE, EMSE, and

PMSE, respectively. To determine the performance of estimators with a response variable from the Poisson and gamma distributions two empirical applications and simulation experiments are conducted. The simulation studies and the numerical examples summarize that the most important factors which affected the performance of the estimators are the values of tuning parameters (k,d) and whether the FOA and ML-based or iterative estimators are used under the chosen restriction. The results of the numerical illustrations revealed that the restricted OK estimator performs better than other restricted estimators in the FOA form when Poisson response is concerned, however, it outperforms restricted as well as unrestricted estimators when the ML-based version is taken into account and (k,d) are obtained by MSE. However, when it comes to numerical application of gamma response, the superiority of one estimator over another is determined by the values of tuning parameters achieved as well as the use of FOA or ML-based versions. Under the configurations in the simulation studies, the restricted OK estimator in the FOA form as well as the iterative or ML-based version for (k_1,d_1) under the Poisson response dominates over the other estimators along with small sample size and a large degree of correlation. However, this case is not valid for (k_2,d_2) and (k_3,d_3). On the other hand, for the gamma response, GLM restricted FOA OK estimator for (k_1,d_1) outperforms the other estimators when sample size and degree of correlation are small, while the iterative restricted ridge estimator for (k_1,d_1) outperforms the other estimators. The superiority of one estimator over another is determined by whether the FOA, or iterative version is used, as well as the tuning parameter values utilized. Furthermore, the impact of restrictions appears to be significant in influencing the performance of estimators over one another, both in simulation studies and numerical examples.

Finally, we constructed a stochastic restricted OK estimator by imposing stochastic restrictions on the parameters. This is a manifold estimator that consists of many other estimators under specific conditions. The estimator's performance is compared to that of ML, mixed, and stochastic restricted ridge estimators. To examine the performance of the estimators an empirical application with a Poisson response and three simulation experiments with responses from binomial, negative binomial, and gamma distributions are performed. The stochastic restricted OK estimator outperforms its counterparts for all values of k and when the d values are ≤ 0.85 according to the results of the empirical application. However, when simulation results are taken into account it surpasses all of its competitors regardless of the values of (k,d) or all of the factors considered.

Based on the results of numerical applications and simulation studies it is obvious that multicollinearity has a significant impact on GLM estimators, notably the ML estimator. As a result, the MSE of the ML estimator increases, resulting in decreased precision and efficiency of the ML estimator. In addition, the performance of our proposed estimators is better than ML and other estimators considered under certain circumstances as indicated in the empirical applications and simulation studies. Thus, in order to combat the multicollinearity our proposed estimators can serve as a better alternative to ML and other estimators mentioned in this dissertation as they are less affected by the problem of multicollinearity.

In future studies, it is recommended that inequality restrictions be applied to these estimators to investigate how they affect the estimators' performance. Other GLMs, such as inverse Gaussian, generalized Poisson, and zero-inflated Poisson can also be performed with these estimators to see how well they work. In addition, qualitative explanatory variables may be employed in future studies instead of quantitative or continuous explanatory variables.

REFERENCES

Abbasi, A., Özkale, M. R., 2021. The $r-k$ class estimator in generalized linear models applicable with simulation and empirical study using a Poisson and Gamma responses. Hacetepe Journal of Mathematics and Statistics, 50(2), 594-611.

Abbasi, A., Özkale, M. R., 2021. Iterative Stochastic Restricted Liu Estimator In Generalized Linear Models: Exemplifying Through Poisson, Gamma, and Binomial Distributions. Submitted

Abbasi, A., Özkale, M. R., 2021. An Iterative Stochastic Restricted OK Estimator In Generalized Linear Models and Selecting the Tuning Parameters Via Genetic Algorithm. Submitted.

Abdeslam, A., El Bouanani, F., Ben-Azza, H., 2014. Four parallel decoding schemas of product block codes. Transactions on Networks and Communications, 2, 49-69.

Agresti, A., 2003. Categorical data analysis. John Wiley & Sons.

Agresti, A., 2015. Foundations of linear and generalized linear models. John Wiley & Sons.

Ameen, M., Qasim, M., Amanullah, M., 2019. Performance of Asar and Genç and Huang and Yang two-parameter estimation methods for the gamma regression model. Iranian Journal of Science and Technology, Transactions A: Science, 43(6), 2951-2963.

Asar, Y., Arashi, M., Wu, J., 2017. Restricted ridge estimator in the logistic regression model. Communications in Statistics-Simulation and Computation, 46(8), 6538-6544.

Asar, Y., Erişoğlu, M., Arashi, M., 2017. Developing a restricted two-parameter Liu-type estimator. A comparison of restricted estimators in the binary logistic regression model. Communications in Statistics-Theory and Methods, 46(14), 6864-6873.

Asar, Y., Genç, A., 2017. Two-parameter ridge estimator in the binary logistic regression.Communications in Statistics-Simulation and Computation, 46(49), 7088-7099.

Asar, Y., Genç, A., 2018. A New Two-Parameter Estimator for the Poisson Regression Model. Iranian Journal of Science and Technology, Transactions A: Science, 42(2), 793-803.

Belsley, D. A., Kuh, E., Welsch, R. E., 1980. Regression diagnostics: Identifying influential data and sources of collinearity. John Wiley & Sons.

Breslow, N. E., 1996. Generalized linear models: checking assumptions and strengthening conclusions. Statistica Applicata, 8(1), 23-41.

Cordeiro, G. M., McCullagh, P., 1991. Bias correction in generalized linear models. Journal of the Royal Statistical Society: Series B (Methodological), 53(3), 629-643.

Das, P., 2019. Econometrics in theory and practice. Springer, 10, 978-981.

Deroche, C. B., 2015. Diagnostics and model selection for generalized linear models and generalized estimating equations (Doctoral dissertation, University of South Carolina).

Du, J., Which estimator of the dispersion parameter for the Gamma family generalized linear models is to chosen? Dalarna University, D-level Master's thesis.

Fallah, R., Arashi, M., Tabatabaey, S. M. M., 2017. On the ridge regression estimator with sub-space restriction. Communications in Statistics-Theory and Methods, 46(23), 11854-11865.

Fahrmeir, L., Tutz, G., 1994. Multivariate statistical modelling based on generalized linear models. New York: Springer-Verlag.

Farrar, D. E., Glauber, R. R., 1967. Multicollinearity in regression analysis: the problem revisited. The Review of Economic and Statistics, 92-107.

Fomby, T. B., Hill, R. C., 1978. Deletion criteria for principal components regression analysis. American Journal of Agricultural Economics, 60(3), 524-527.

Fox, J., 2015. Applied regression analysis and generalized linear models. Sage Publications.

Frisch, R., 1934. Statistical confluence analysis by means of complete regression systems", Institute of Economics, Olso University, Publication No. 5.

Gill, J., 2001. Generalized Linear Models: A unified approach. Sage University Papers Series on Quantitative Application in the Social Sciences.

Groß, J., 2003. Restricted ridge estimation. Statistics & probability letters, 65(1), 57-64.

Gruber, M. H., 2012. Liu and ridge estimators-a comparison. Communications in Statistics-Theory and Methods,, 41(20), 3739-3749.

Gunst, R. F., 1983. Regresion analysis with multicollinear predictor variables: definition, derection, and effects. Communications in Statistics-Theory and Methods, 12(19), 2217-2260.

Hamed, R., Hefnawy, A. E., Farag, A., 2013. Selection of the ridge parameter using mathematical programming. Communications in Statistics-Simulation and Computation, 42(6), 1409-1432.

Hardin, J. W., Hilbe, J. M., 2012. Generalized linear models and extensions. Stata press.

Holland, J. H., 1975. Adaptation in natural and artificial systems: an introductory analysis with application to biology, control and artificial intelligence. MIT press.

Hoerl, A. E., Kennard, R. W., 1970. Ridge regression: applications to nonorthogonal problems. Technometrics, 12(1), 69-82.

Hoerl, A. E., Kannard, R. W., Baldwin, K. F., 1975. Ridge regression: some simulations. Communications in Statistics-Theory and Methods, 4(2), 105-123.

Huang, J., Yang, H., 2014. A two-parameter estimator in the negative binomial regression model. Journal of Statistical Computation and Simulation, 84(1), 124-134.

Huang, C. C. L., Jou, Y. J., Cho, H. J., 2016. A new multicollinearity diagnostic for generalized linear models. Journal of Applied Statistics, 43(11), 2029-2043.

Huet, S., Bouvier, A., Poursat, M. A., Jolivet, E., Bouvier, A. M., 2004. Statistical tools for nonlinear regression: a practical guide with S-PLUS and R examples (p. 233). New York: Springer.

Iquebal, M. A., Prajneshu, Ghosh, H., 2012. Genetic algorithm optimization technique for linear regression models with heteroscedastic errors. Indian Journal of Agricultural Sciences, 82(5), 422-425.

Johnson, S. R., Reimer, S. C., Rothrock, T. P., 1973. Principal components and the problem of multicollinearity. Metroeconomica, 25(3), 306-317.

Kejian, L., 1993. A new class of biased estimate in linear regression. Communications in Statistics-Theory and Methods, 22(2), 393-402.

Khuri, A. I., 2009. Linear model methodology. Chapman and Hall/CRC.

Kibria, B. G., 2003. Performance of some new ridge regression estimators. Communications in Statistics-Simulation and Computation, 32(2), 419-435.

Kurtoğlu, F., Özkale, M. R., 2016. Liu estimation in generalized linear models: application on Gamma distributed response variable. Statistical Papers, 57(4), 911-928.

Kurtoğlu, F., Özkale, M. R., 2019a. Restricted ridge estimator in generalized linear models: Monte Carlo simulation studies on Poisson and binomial distributed responses. Communications in Statistics-Simulation and Computation, 48(4), 1191-1218.

Kurtoğlu, F., Özkale, M. R., 2019b. Restricted Liu estimator in generalized linear models: Monte Carlo simulation studies on Gamma and Poisson distributed responses. Hacetepe Journal of Mathematics and Statistics, 48(4), 1250-1276.

Lamari, Y., Freskura, B., Abdessamad, A., Eichberg, S., de Bonviller, S., 2020. Predicting Spatial Crime Occurrences through an Efficient Ensemble-Learning Model. ISPRS International Journal of Geo-Information, 9(11), 645.

Le Cessie, S., Van Houwelingen, J. C., 1992. Ridge estimators in logistic regression. Journal of the Royal Statistical Society: Series C (Applied Statistics), 41(1), 191-201.

Lee, A. H., Silvapulle, M. J., 1988. Ridge estimation in logistic regression. Communication in Statistics-Simulation and Computation, 17(4), 1231-1257.

Lesaffre, E., Marx, B. D., 1993. Collinearity in generalized linear regression. Communications in Statistics-Theory and Methods, 22(7), 1933-1952.

Li, Y., Yang, H., 2010. A new stochastic mixed ridge estimator in linear regression model. Statistical Papers, 51(2), 315-323.

Lindsey, J. K., 2000. Applying generalized linear models. Springer Science & Business Media.

Mackinnon, M. J., Puterman, M. L., 1989. Collinearity in generalized linear models. Communications in statistics-theory and methods, 18(9), 3463-3472.

Mallawaarachchi, V., 2017. How to define a Fitness Function in a Genetic Algorithm? Towards Data Science. Pridobljeno s https://towardsdatascience. com/howto-define-a-fitness-function-in-a-genetic-algorithm-be572b9ea3b4.

Mallows, C. L., 1973. Some remarks of Cp. Technometrics, 15, 661-675.

Månsson, K., 2012. On ridge estimator for the negative binomial regression model. Economic Modeling, 29(2), 178-184.

Månsson, K., Shukur, G., 2011. A Poisson ridge regression estimator. Economic Modeling, 28(4), 1475-1481.

Månsson, K., Kibria, B. G., Shukur, G., 2012. On Liu estimators for the logit regression model. Economic Modeling, 29(4), 1483-1488.

McCullagh, P., Nelder, J. A., 1989. Generalized linear models. Chapman and Hall. London, UK.

McDonald, G. C., Galarneau, D.I., 1975. A Monte Carlo evaluation of ridge-type estimators. Journal of the American Statistical Association, 70(350), 407-416.

Montgomery, D. C., Peck, E. A., Vining, G. G., 2012. Introduction to linear regression analysis. John Wiley & Sons.

Myers, R. H., Montgomery, D. C., Vining, G. G., Robinson, T. J., 2010. Generalized linear models with applications in engineering and the sciences. John Wiley & Sons.

Ndabashinze, B., Şiray, G. U., 2020. Comparing ordinary ridge and generalized ridge regression results obtained using genetic algorithms for ridge param- eter selection. Communications in Statistics-Simulation and Computation, 1-1.

Nelder, J. A., Wedderburn, R. W. M., 1972. Generalized linear models. Journal of the Royal Statistical Society: Series A (General), 135(3), 370-384.

Nyquist, H., 1991. Restricted estimation of generalized linear models. Journal of the Royal Statistical Society: Series C (Applied Statistics), 40(1), 133-141.

Özkale, M. R., Kaçıranlar, S., 2007. The restricted and unrestricted two-parameter estimators. Communication in Statistics-Theory and Methods, 36(15), 2707-2725.

Özkale, M. R., 2014. The relative efficiency of the restricted estimators in linear regression models. Journal of Applied Statistics, 41(5), 998-1027.

Özkale, M. R., 2015. Predictive Performance of linear regression models. Statistical Papers, 41(5), 998-1027.

Özkale, M. R., 2016. Iterative algorithms of biased estimation methods in binary logistic regression. Statistical Papers, 57(4), 991-1016.

Özkale, M. R., Arıcan, E., (2016). A new biased estimator in logistic regression model. Statistics, 50(2), 233-253.

Özkale, M. R., 2019. The $r-d$ class estimator in generalized linear models: applications on gamma, Poisson and binomial distributed responses. Journal of Statistical Computation and Simulation, 89(4), 615-640.

Özkale, M. R., 2021. The red indicator and corrected VIFs in generalized linear models. Communications in Statistics-Simulation and Computation, 50(12), 4144-4170.

Özkale, M. R., Nyquist, H., 2021. The stochastic restricted ridge estimator in generalized linear models. Statistical Papers, 62(3), 1421-1460.

Özkale, M. R., Altuner, H., 2021. Bootstrap selection of ridge regularization parameter: a comparative study via a simulation study. Communications in Statistics-Simulation and Computation, 1-19.

Özkale, M. R., Abbasi, A., 2022. Iterative restricted OK estimator in generalized linear models and the selection of tuning parameters via MSE and genetic algorithm. Statistical Papers, 1-62.

Praga-Alejo, R. J., Torres-Trevio, L. M., Pia-Monarrez, M. R., 2008. Optimal determination of k constant of ridge regression using a simple genetic algorithm. In 2008 Electronics, Robotics and Automotive Mechanics Conference (CERMA'08) (pp. 39-44). IEEE.

Qasim, M., Kibria, B. M. G., Mansson, K., Sjölander, P., 2020. A new Poisson Liu regression estimator: method and application. Journal of Applied Statistics, 47(12), 2258-2271.

Rao, C. R., Toutenburg, H., 1995. Linear models Least Squares and Alternatives (pp. 3-18). Springer, New York, NY.

Segerstedt, B., 1992. On ordinary ridge regression in generalized linear models. Communications in Statistics-Theory and Methods, 21(8), 2227-2246.

Scrucca, L., 2013. GA: a package for genetic algorithms in R. Journal of Statistical Software, 53, 1-37.

Schaefer R. L., Roi, L. D., Wolfe, R. A., 1984. A ridge logistic estimator. Communication in Statistics-Theory and Methods, 13(1), 99-113.

Schaefer, R. L., 1986. Alternative estimators in logistic regression when the data are collinear. Journal of Statistical Computation and Simulation, 25(1-2), 75-91.

Shalabh, S., 2013. Econometric Theory. http//home.iitk.ac.in Shalab.

Siriwardene, N. R., Perera B. J. C., 2006. Selection of genetic algorithm operators for urban drainage model parameter optimization. Mathematical and Computer Modelling. 44(5-6):415-429

Smith, E. P., Marx, B. D., 1990. Ill-conditioned information matrices, generalized linear models and estimation of the effects of acid rain. Environmetrics, 1(1), 57-71.

Tekeli, E., Kaçıranlar, S., Özbay, N., 2019. Optimal determination of the parameters of some biased estimators using genetic algorithm. Journal of Statistical Computation and Simulation, 89(18), 3331-53.

Theil, H., 1963. On the use of incomplete prior information in regression analysis. Journal of the American Statistical Association, 58(302), 401-414.

Theil, H., Golberger, A. S., 1961. On pure and mixed statistical estimation in economics. International Economic Review, 2(1), 65-78.

Thiart, C. 1990. Collinearity and consequences for estimation: a study and simulation (Doctoral dissertation, University of Cape Town).

Toutenburg, H., 1982. Prior information in linear models. John Wiley & Sons.

Varathan, N., Wijekoon, P., 2018. Liu-Type logistic estimator under stochastic linear restrictions. Ceylon Journal of Science, 47(1), 21-34.

Varathan, N., Wijekoon, P., 2015. Stochastic restricted maximum likelihood estimator in logistic regression model. Open Journal of Statistics, 5(7), 837-851.

Varathan, N., Wijekoon, P., 2019. Logistic Liu Estimator under stochastic linear restrictions. Statistical Papers, 60(3), 945-962.

Weissfeld, L. A., Sereika, S. M., 1991. A multicollinearity diagnostic for generalized linear models. Communication in Statistics-Theory and Methods, 20(4), 1183-1198.

Zhong Z., Yang, H., 2007. Ridge estimation to the restricted linear model. Communication in Statistics-Theory and Methods, 36(11), 2099-2115.

Zuo, W., Li, Y., 2018. A new stochastic restricted Liu estimator for the logistic regression model. Open Journal of Statistics, 8(1), 25-37.

CURRICULUM VITAE

Atif Abbasi was born in Azad Jammu & Kashmir Muzaffarabad, Pakistan. He completed his Primary and Middle school education at Government Middle School Dangali in 1998 and Secondary School education at Higher Secondary School Barsala, Muzaffarabad in 2000. He received his F.Sc. and B.Sc. degrees from Post Graduate College (Boys) Muzaffarabad in 2002 and 2005. He received his M.Sc. in Statistics from The University of Azad Jammu & Kashmir (UAJ&K) Muzaffarabad, Department of Statistics in 2008. In 2015, he earned his M.Phil in Statistics from Riphah International University Islamabad, Pakistan. In addition, from 2010 to 2013, he worked as Research Associate on a contract basis in the Department of Statistics at UAJ&K, Muzaffarabad. From December 2013 to September 2017 he has also worked as Research Associate on a permanent basis in the same University. Then, he was awarded study leave to pursue his PhD studies in Statistics from Turkey through the YÖK scholarship program. In 2018, he started his doctoral studies in the Department of Statistics at Çukurova University, Institute of Natural and Applied Sciences, Adana Turkey. Restricted Estimation Methods in Generalized Linear Models was the topic of his PhD dissertation. In the year 2022, he completed his doctoral studies.

APPENDICES

A. Appendix-A

A.1. Tables of Simulations Results

Table A1. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p=8$ and Poisson response

n	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	37.525551600	6.634948370	14.816844830	3.620857272
	RR	0.903509552	0.454274989	0.755625620	—
	0.95	74.898961730	9.025821874	1.134143020	1.126853379
	RR	0.984955020	0.875152269	0.006427444	—
	0.99	357.633127400	12.956143160	1.125907322	1.122861521
	RR	0.996860298	0.913333659	0.002705197	—
400	0.90	36.757518980	6.475868688	14.669363070	3.652012230
	RR	0.900645845	0.436058326	0.751044936	—
	0.95	75.525567000	9.015851375	1.143600032	1.133592781
	RR	0.984990595	0.874266696	0.008750657	—
	0.99	358.991520900	13.734417470	1.124515272	1.120107231
	RR	0.99687985	0.918445232	0.003919948	—
600	0.90	37.379177570	6.320171343	16.199015530	3.360099991
	RR	0.910107707	0.468353023	0.792573815	—
	0.95	74.889655430	8.413768773	1.142149944	1.140927777
	RR	0.984765215	0.864397536	0.001070059	—
	0.99	373.446567200	13.571149240	1.132983179	1.128861009
	RR	0.996977182	0.916819056	0.003638333	—
800	0.90	36.430436340	6.212917176	15.0342819700	3.378795924
	RR	0.907253487	0.456165948	0.7752605720	—
	0.95	69.644517380	7.633279444	1.137273251	1.124261111
	RR	0.983857148	0.852715845	0.011441525	—
	0.99	344.454171700	12.227834970	1.120663201	1.119735478
	RR	0.996749247	0.908427331	0.000827834	—

Table A2. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p=10$ and Poisson response

n	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	46.450003420	8.207749747	18.066915660	4.316719839
	RR	0.907067395	0.474067805	0.761070460	—
	0.95	88.607949890	10.529122080	1.113766959	1.107497599
	RR	0.987501149	0.894815770	0.005628969	—
	0.99	428.426512200	16.199528900	1.104810291	1.102194135
	RR	0.997427344	0.931961346	0.002367968	—
400	0.90	49.314254300	9.062197277	15.740118510	4.047476215
	RR	0.917924822	0.553367016	0.742856052	—
	0.95	94.107980950	11.422366900	1.110830353	1.104596150
	RR	0.988262460	0.903295336	0.005612202	—
	0.99	459.319592000	17.9122536100	1.099700932	1.097529787
	RR	0.997610531	0.938727431	0.001974305	—
600	0.90	51.333409800	8.852738585	18.092092880	4.196684320
	RR	0.918246531	0.525945076	0.768037653	—
	0.95	99.445072440	11.649305330	1.106013423	1.100245662
	RR	0.988936147	0.905552681	0.0052149110	—
	0.99	480.176492300	17.885007240	1.100365920	1.097872610
	RR	0.997713606	0.938614919	0.002265891	—
800	0.90	47.878570820	8.2520536760	21.898757710	4.482281784
	RR	0.906382298	0.456828329	0.795317988	—
	0.95	94.975872200	11.138258900	1.108479909	1.104095402
	RR	0.988374991	0.900873609	0.003955424	—
	0.99	468.697664700	17.957672400	1.098806610	1.096708442
	RR	0.997660094	0.938928141	0.001909497	—

Table A3. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p = 12$ and Poisson response

	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	61.515596600	11.760815610	20.006249900	5.743763706
	RR	0.906629147	0.511618590	0.712901532	—
	0.95	118.990357100	15.455487480	1.094932081	1.09167392
	RR	0.990825526	0.929366581	0.002975674	—
	0.99	597.959631000	25.211374200	1.088032985	1.086625068
	RR	0.998182779	0.956899411	0.001294002	—
400	0.90	58.598507160	11.383937040	19.626232310	5.331225031
	RR	0.909021146	0.531688816	0.728362278	—
	0.95	112.529613400	14.439576700	1.090225373	1.086905414
	RR	0.990341161	0.924727335	0.003045205	—
	0.99	551.727960700	23.056078060	1.082692108	1.081275818
	RR	0.998040201	0.953102353	0.001308119	—
600	0.90	56.206296330	11.507211350	26.659439460	6.593502267
	RR	0.882691038	0.427011283	0.752676635	—
	0.95	112.478571500	15.6504103500	1.097124998	1.093592122
	RR	0.990277329	0.930123741	0.003220122	—
	0.99	537.647396700	24.231441960	1.076318516	1.075055926
	RR	0.998000444	0.955633844	0.001173064	—
800	0.90	57.145665640	10.889680990	27.232295920	6.190831589
	RR	0.891665772	0.431495597	0.772665823	—
	0.95	111.711436400	13.8972465400	1.096952756	1.092815095
	RR	0.990217518	0.921364632	0.003771959	—
	0.99	554.175703600	23.293587300	1.077799657	1.076469580
	RR	0.998057530	0.953786870	0.001234067	—

Table A4. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p=8$ and Gamma response

n	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	6.470187	5.26132	1.694170	1.692702
	RR	0.738384	0.678277	0.000867	—
	0.95	7.037896	5.259769	2.509046	2.505413
	RR	0.644011	0.523665	0.001448	—
	0.99	8.937555	4.320989	1.856880	1.854712
	RR	0.792481	0.570767	0.00116	—
400	0.90	4.840330	4.533189	3.360391	3.357198
	RR	0.306411	0.259418	0.000950	—
	0.95	4.816353	4.414173	3.690117	3.684973
	RR	0.234904	0.165195	0.001394	—
	0.99	10.250065	6.677664	2.659756	2.655296
	RR	0.740948	0.602362	0.001677	—
600	0.90	4.557798	4.390382	2.563674	2.563325
	RR	0.437596	0.416150	0.000136	—
	0.95	5.223324	4.923769	4.841758	4.837445
	RR	0.073876	0.017532	0.000891	—
	0.99	8.053276	7.299779	6.891482	6.879726
	RR	0.145723	0.057543	0.001706	—
800	0.90	4.119247	4.029864	3.217761	3.217077
	RR	0.219013	0.201691	0.000212	—
	0.95	5.379205	5.146413	3.689900	3.688674
	RR	0.314272	0.283253	0.000332	—
	0.99	11.442052	9.059892	3.385485	3.383178
	RR	0.704321	0.626576	0.000681	—

Table A5. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p=10$ and Gamma response

n	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	6.725315	4.879247	2.012484	2.011344
	RR	0.700929	0.587776	0.000566	—
	0.95	10.996022	6.556518	1.919557	1.916631
	RR	0.825698	0.707675	0.001524	—
	0.99	18.696256	6.604226	1.355227	1.355139
	RR	0.927518	0.794807	0.000065	—
400	0.90	4.660879	4.350649	2.347800	2.347003
	RR	0.496446	0.460540	0.000340	—
	0.95	5.345370	4.782703	3.523911	3.520888
	RR	0.341320	0.263829	0.000858	—
	0.99	8.580232	5.563526	3.445426	3.439806
	RR	0.599101	0.381722	0.001631	—
600	0.90	4.650520	4.458327	2.596699	2.596115
	RR	0.441758	0.417693	0.000225	—
	0.95	5.316252	4.933863	3.463860	3.461015
	RR	0.348975	0.298518	0.000821	—
	0.99	10.645504	7.741610	2.947297	2.943354
	RR	0.723512	0.619801	0.001338	—
800	0.90	4.655214	4.508783	3.186016	3.185615
	RR	0.315689	0.293465	0.000126	—
	0.95	5.402928	5.113568	3.857560	3.855624
	RR	0.286383	0.246001	0.000502	—
	0.99	8.918563	7.141924	2.508333	2.507115
	RR	0.718888	0.648958	0.000486	—

Table A6. RR, EMSE values of the ML, ridge, PCR and $r-k$ class estimators when $p=12$ and Gamma response

n	γ^2	ML	Ridge	PCR	$r-k$
200	0.90	6.844589	4.885149	2.333045	2.329927
	RR	0.659596	0.523059	0.001336	—
	0.95	11.680592	6.570061	1.269806	1.268581
	RR	0.891394	0.806915	0.000965	—
	0.99	29.221371	7.414583	1.335053	1.334850
	RR	0.954319	0.819970	0.000153	—
400	0.90	5.943753	5.274643	2.075884	2.075177
	RR	0.650864	0.606575	0.000341	—
	0.95	5.772376	4.909312	3.577512	3.572440
	RR	0.381115	0.272314	0.001418	—
	0.99	8.982236	5.243724	2.418714	2.415251
	RR	0.731108	0.539402	0.001432	—
600	0.90	5.147963	4.841649	2.218392	2.217934
	RR	0.569163	0.541905	0.000207	—
	0.95	5.543631	5.061856	2.627509	2.626005
	RR	0.526302	0.481217	0.000572	—
	0.99	9.198872	6.611784	2.207942	2.206136
	RR	0.760173	0.666333	0.000818	—
800	0.90	4.689527	4.502824	2.599765	2.599365
	RR	0.445708	0.422726	0.000154	—
	0.95	4.892099	4.624351	3.759408	3.758130
	RR	0.231796	0.187317	0.000340	—
	0.99	9.470270	7.376626	2.329244	2.327379
	RR	0.754244	0.684493	0.000801	—

Table A7. EMSE values of the estimators when optimal (k,d) are obtained via GA for the Binomial response variable

n	γ^2	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
100	0.75	48.5449	3.7745	3.6996	1.5590	3.6923	1.6633
	0.85	79.1186	3.9690	3.7942	1.8147	3.9680	1.8222
	0.90	117.1301	4.0824	3.8768	2.0473	4.0812	2.0639
	0.95	230.7551	4.2300	4.0992	2.4781	4.1765	2.4914
250	0.75	83.0668	7.8919	7.8884	1.6492	7.8881	1.6792
	0.85	131.6176	7.7115	7.7078	1.9097	7.7065	1.9100
	0.90	192.0188	7.6387	7.6205	2.1886	7.4869	2.2346
	0.95	372.1139	7.5891	7.2644	2.8602	7.2996	2.8416
350	0.75	14.6385	3.3095	3.3057	1.1034	3.3086	1.1033
	0.85	23.6282	3.3839	3.3824	1.1763	3.3829	1.1762
	0.90	34.8888	3.4296	3.4284	1.2715	3.4275	1.2717
	0.95	68.7600	3.4894	3.4477	1.5551	3.4482	1.5547
500	0.75	57.2466	4.5279	4.5255	1.2091	4.5264	1.2090
	0.85	93.8814	3.8315	3.8325	1.3158	3.8316	1.3159
	0.90	140.3530	3.5091	3.5069	1.4595	3.5077	1.4593
	0.95	280.7951	3.1689	3.0935	1.9019	3.0928	1.9009
700	0.75	72.8829	3.5283	3.5275	1.1435	3.5268	1.1435
	0.85	121.0312	3.19990	3.1984	1.2276	3.1974	1.2278
	0.90	181.7504	3.0586	3.0579	1.3403	3.0581	1.3401
	0.95	365.0050	2.9440	2.9221	1.6831	2.9221	1.6821
900	0.75	60.2647	4.8392	4.6899	1.1403	2.8064	1.0791
	0.85	120.4783	2.6842	2.3689	1.2163	2.6773	1.1502
	0.90	180.4573	4.3647	4.3018	1.3105	4.5207	1.2962
	0.95	360.5381	4.4044	4.3787	1.5887	2.5567	1.5389

Table A8. EMSE values of the estimators when optimal (k, d) are obtained via GA for the Gamma response variable

n	γ^2	ML	Mixed	SR.Ridge (k_1, d_1)	SR.OK (k_1, d_1)	SR.Ridge (k_2, d_2)	SR.OK (k_2, d_2)
100	0.75	17.3608	11.3117	10.4216	1.4697	9.7701	1.4696
	0.85	25.2906	10.3622	10.3497	1.5727	10.3188	1.5722
	0.90	35.5947	10.2442	10.1754	1.7146	10.0721	1.7190
	0.95	67.3479	9.9939	9.7946	2.1171	9.7468	2.1164
250	0.75	10.2025	7.3488	7.2396	1.1934	5.2248	1.1932
	0.85	13.4270	7.3709	7.2144	1.1987	5.3160	1.1976
	0.90	17.5547	7.3861	7.2841	1.2083	5.3319	1.2065
	0.95	30.1163	7.3941	7.1881	1.2504	5.3486	1.2432
350	0.75	20.6971	9.8363	9.6987	1.2461	7.0747	1.2454
	0.85	28.9407	9.7591	9.6294	1.2691	7.2350	1.2678
	0.90	39.3025	9.7234	9.4827	1.3006	7.3903	1.2949
	0.95	70.4619	9.7236	9.3243	1.3974	7.3973	1.3898
500	0.75	9.3140	6.3280	4.7698	1.1180	4.6284	1.1181
	0.85	12.3885	6.3080	4.7773	1.1262	4.6847	1.1259
	0.90	16.2954	6.3054	4.7914	1.1385	4.7254	1.1381
	0.95	28.1076	6.3194	5.0111	1.1834	4.7273	1.1806
700	0.75	18.8680	5.6981	5.6960	1.0881	4.0354	1.0879
	0.85	29.3557	5.4911	5.4902	1.1068	3.9808	1.1067
	0.90	42.5081	5.3452	5.3444	1.1316	3.9721	1.1316
	0.95	82.0637	5.1849	5.1166	1.2134	3.9484	1.1666
900	0.75	13.4647	3.1049	2.9644	1.0168	2.2705	1.0168
	0.85	21.7419	3.1195	2.9820	1.0303	2.3045	1.0302
	0.90	32.0493	3.1259	2.9953	1.0458	2.3287	1.0457
	0.95	62.9300	3.1519	2.9967	1.0944	2.3430	1.0913

Table A9. EMSE values of the estimators when optimal (k,d) are obtained via GA for the Negative Binomial response variable with $\alpha = 1$

n	γ^2	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
100	0.75	4.6159	3.5411	2.9125	1.1834	2.9311	1.1835
	0.85	5.7995	3.6149	2.9171	1.1848	2.9088	1.1849
	0.90	7.4505	3.6973	2.9341	1.1889	2.9278	1.1897
	0.95	12.6420	3.8252	2.9556	1.1891	2.9535	1.1898
250	0.75	10.1694	3.1177	2.5838	1.1415	2.5961	1.1411
	0.85	15.4427	3.1985	2.5909	1.1558	2.5913	1.1553
	0.90	22.1163	3.2924	2.6018	1.1728	2.6043	1.1720
	0.95	42.2362	3.4482	2.6548	1.2061	2.6438	1.2077
350	0.75	10.8358	3.0458	2.3004	1.0846	2.2912	1.0847
	0.85	17.1900	3.1270	2.3015	1.1147	2.2966	1.1149
	0.90	25.2219	3.1786	2.2879	1.1538	2.2887	1.1533
	0.95	49.4997	3.2356	2.2916	1.2617	2.2919	1.2665
500	0.75	5.9908	1.6391	1.2152	0.9789	1.2113	0.9589
	0.85	9.7018	1.7232	1.2571	0.9915	1.4617	0.9757
	0.90	14.3691	1.7777	1.2998	1.0067	1.2954	1.0070
	0.95	28.4340	1.8723	1.3612	1.0492	1.3557	1.0492
700	0.75	5.4447	2.4473	1.9141	1.0140	1.9122	1.0140
	0.85	8.4134	2.5322	1.9346	1.0204	1.9208	1.0205
	0.90	12.1763	2.5928	2.0200	1.0290	1.9263	1.0303
	0.95	23.5587	2.6782	1.9456	1.0618	1.9444	1.0616
900	0.75	3.4783	3.3247	2.7155	1.0414	2.7118	1.0415
	0.85	4.0803	3.5733	2.7170	1.0417	2.6993	1.0418
	0.90	5.1065	3.6391	2.7117	1.0420	2.7142	1.0419
	0.95	10.9135	4.2107	2.7146	1.0621	3.1826	1.0802

Table A10. EMSE values of the estimators when optimal (k,d) are obtained via GA for the Negative Binomial response variable with $\alpha = 2$

n	γ^2	ML	Mixed	SR.Ridge (k_1,d_1)	SR.OK (k_1,d_1)	SR.Ridge (k_2,d_2)	SR.OK (k_2,d_2)
100	0.75	5.1132	2.1149	1.7838	1.0712	1.7784	1.0713
	0.85	6.0632	2.3248	1.8983	1.0842	1.8917	1.0844
	0.90	8.5492	2.4650	1.9737	1.0979	1.9711	1.0982
	0.95	16.1224	2.6538	2.0881	1.1246	2.0846	1.1251
250	0.75	2.4063	2.2467	1.9346	1.0181	1.9275	1.0181
	0.85	2.9682	2.6120	2.0196	1.0227	2.0145	1.0228
	0.90	3.8613	2.7437	2.0724	1.0285	2.0734	1.0285
	0.95	6.5230	2.9096	2.1497	1.0460	2.1409	1.0459
350	0.75	6.4362	2.3222	1.6473	1.0183	1.6347	1.0184
	0.85	10.2008	2.4763	1.7330	1.0366	1.7292	1.0367
	0.90	14.8879	2.5146	1.7860	1.0565	1.7855	1.0562
	0.95	28.9472	2.5168	1.8733	1.1070	1.8788	1.1066
500	0.75	6.3928	3.0672	1.8726	0.9987	1.8878	0.9987
	0.85	10.2153	3.0914	1.8507	1.0117	1.8560	1.0116
	0.90	15.0031	2.9783	1.8031	1.0284	1.8080	1.0284
	0.95	29.3866	2.6896	1.7563	1.0780	1.7582	1.0789
700	0.75	5.6979	3.1820	2.3402	1.0074	2.3429	1.0074
	0.85	8.8856	3.2803	2.3199	1.0149	2.3195	1.0149
	0.90	12.9191	3.3074	2.2795	1.0261	2.2709	1.0263
	0.95	25.1059	3.3765	2.1869	1.0637	2.1871	1.0637
900	0.75	4.1899	1.8412	2.3989	0.9580	1.2881	0.9674
	0.85	6.8167	1.9644	1.3674	0.9758	1.3708	0.9757
	0.90	10.0729	2.0133	1.4048	0.9838	1.3622	1.0526
	0.95	19.8003	2.0627	1.5767	0.9951	1.4414	1.0560

Table A11. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained via GA by minimizing MSEP for Poisson response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	8.3548 ($4.4666\times10^{+4}$)	9.4415 ($4.5045\times10^{+4}$)	8.2409 ($4.4220\times10^{+4}$)	9.4026 ($4.4581\times10^{+4}$)	**8.2378** (**4.4208**$\times10^{+4}$)	***9.4014*** (***4.4567***$\times10^{+4}$)	8.2431 ($4.4229\times10^{+4}$)	9.40343 ($4.4589\times10^{+4}$)
	0.85	10.9089 ($1.2228\times10^{+4}$)	12.1354 ($1.2268\times10^{+4}$)	10.4999 ($1.1927\times10^{+4}$)	11.9674 ($1.1947\times10^{+4}$)	**10.4968** (**1.1925**$\times10^{+4}$)	***11.9656*** (***1.1944***$\times10^{+4}$)	10.5123 ($1.1936\times10^{+4}$)	11.9722 ($1.1956\times10^{+4}$)
	0.95	15.3082 ($1.4093\times10^{+3}$)	16.2553 ($1.3965\times10^{+3}$)	13.1415 ($1.2913\times10^{+3}$)	14.9219 ($1.2518\times10^{+3}$)	**13.0634** (**1.2871**$\times10^{+3}$)	***14.8692*** (***1.2463***$\times10^{+3}$)	13.1662 ($1.2926\times10^{+3}$)	14.9371 ($1.2534\times10^{+3}$)
	0.99	65.2098 ($7.8288\times10^{+2}$)	59.8493 ($7.9968\times10^{+2}$)	35.1225 ($6.0392\times10^{+2}$)	40.2003 ($5.6809\times10^{+2}$)	**30.3847** (**5.8013**$\times10^{+2}$)	***36.2808*** (***5.2845***$\times10^{+2}$)	35.7800 ($6.0726\times10^{+2}$)	40.6801 ($5.7306\times10^{+2}$)
50	0.75	3.8162 ($1.6611\times10^{+4}$)	3.5842 ($1.6349\times10^{+4}$)	3.7828 ($1.6486\times10^{+4}$)	3.5708 ($1.6248\times10^{+4}$)	**3.7824** ($1.6484\times10^{+4}$)	***3.5706*** ($1.6490\times10^{+4}$)	3.7838 (**1.6251**$\times10^{+4}$)	3.5712 (***1.6247***$\times10^{+4}$)
	0.85	3.6479 ($6.0139\times10^{+3}$)	3.0166 ($5.8932\times10^{+3}$)	3.5549 ($5.9371\times10^{+3}$)	2.9834 ($5.8342\times10^{+3}$)	**3.5518** (**5.9345**$\times10^{+3}$)	***2.9822*** (***5.8323***$\times10^{+3}$)	3.5565 ($5.9383\times10^{+3}$)	2.9840 ($5.8353\times10^{+3}$)
	0.95	4.9291 ($2.8904\times10^{+3}$)	2.9113 ($2.8237\times10^{+3}$)	**4.3541** (**2.8191**$\times10^{+3}$)	***2.7293*** (***2.7778***$\times10^{+3}$)	4.3667 ($2.8207\times10^{+3}$)	2.7365 ($2.7794\times10^{+3}$)	4.3809 ($2.8224\times10^{+3}$)	2.7409 ($2.7805\times10^{+3}$)
	0.99	28.0073 ($2.7037\times10^{+3}$)	14.5864 ($2.6381\times10^{+3}$)	16.7937 ($2.5108\times10^{+3}$)	9.8190 ($2.5137\times10^{+3}$)	**16.6992** (**2.5091**$\times10^{+3}$)	***9.8160*** (***2.5134***$\times10^{+3}$)	17.1098 ($2.5159\times10^{+3}$)	10.0017 ($2.5180\times10^{+3}$)
100	0.75	3.6889 ($7.2923\times10^{+4}$)	3.5912 ($6.9481\times10^{+4}$)	3.6772 ($7.2664\times10^{+4}$)	3.5832 ($6.9276\times10^{+4}$)	**3.6770** ($7.2660\times10^{+4}$)	***3.5830*** ($7.2671\times10^{+4}$)	3.6776 (**6.9272**$\times10^{+4}$)	3.5834 ($6.9281\times10^{+4}$)
	0.85	3.3267 ($1.7925\times10^{+4}$)	3.1868 ($1.6732\times10^{+4}$)	3.2959 ($1.7793\times10^{+4}$)	3.1685 ($1.6629\times10^{+4}$)	**3.2958** (**1.7793**$\times10^{+4}$)	***3.1685*** (***1.6629***$\times10^{+4}$)	3.2969 ($1.7798\times10^{+4}$)	3.1692 ($1.6633\times10^{+4}$)
	0.95	3.5021 ($1.5708\times10^{+3}$)	2.9878 ($1.7527\times10^{+3}$)	3.2899 ($1.5262\times10^{+3}$)	2.8569 ($1.7128\times10^{+3}$)	**3.2859** ($1.5253\times10^{+3}$)	***2.8549*** ($1.7122\times10^{+3}$)	3.2948 ($1.5272\times10^{+3}$)	2.8604 ($1.7139\times10^{+3}$)
	0.99	5.5057 ($4.1144\times10^{+2}$)	4.1525 ($6.4412\times10^{+2}$)	**3.7527** (**3.8451**$\times10^{+2}$)	***3.0343*** (***6.1501***$\times10^{+2}$)	3.7672 ($3.8471\times10^{+2}$)	3.0533 ($3.8575\times10^{+2}$)	3.8371 ($6.1542\times10^{+2}$)	3.0973 ($6.1657\times10^{+2}$)
200	0.75	**0.4033** (231.9036)	***1.0257*** (613.2740)	0.4054 (228.8268)	1.0288 (***609.6148***)	0.4054 (**228**.7965)	1.0287 (609.6185)	0.4053 (228.9211)	1.0286 (609.7661)
	0.85	**0.4159** (101.6331)	***1.1460*** (511.2616)	0.4189 (99.9652)	1.1542 (508.4187)	0.4189 (**99.9391**)	1.1542 (***508.4161***)	0.4187 (100.0148)	1.1538 (508.5449)
	0.95	**0.6673** (51.1861)	***1.4035*** (509.2843)	0.6711 (50.2492)	1.4364 (506.8212)	0.6714 (**50.2114**)	1.4375 (***506.7415***)	0.6711 (50.2595)	1.4357 (506.8717)
	0.99	1.4954 (39.4517)	***1.5741*** (522.0365)	1.3379 (38.4605)	1.6488 (519.5329)	**1.288421** (**38.0502**)	1.7024 (***518.3708***)	1.3395 (38.4711)	1.6473 (519.5924)
500	0.75	23.6214440 ($2.7658\times10^{+17}$)	33.400590 ($2.5720\times10^{+17}$)	23.6214440 ($2.7658\times10^{+17}$)	33.400590 ($2.5720\times10^{+17}$)	23.6214440 ($2.7658\times10^{+17}$)	33.400590 ($2.5720\times10^{+17}$)	23.6214 ($2.7658\times10^{+17}$)	33.4006 ($2.5720\times10^{+17}$)
	0.85	22.212130 ($2.3226\times10^{+17}$)	32.224760 ($2.1854\times10^{+17}$)	22.212130 ($2.3226\times10^{+17}$)	32.224760 ($2.1854\times10^{+17}$)	22.212130 ($2.3226\times10^{+17}$)	32.224760 ($2.1854\times10^{+17}$)	22.212130 ($2.3226\times10^{+17}$)	32.224760 ($2.1854\times10^{+17}$)
	0.95	26.224870 ($4.0334\times10^{+17}$)	33.10266 ($3.7963\times10^{+17}$)	26.224870 ($4.0334\times10^{+17}$)	33.10266 ($3.7963\times10^{+17}$)	26.224870 ($4.0334\times10^{+17}$)	***33.10265*** ($3.7963\times10^{+17}$)	26.224870 ($4.0334\times10^{+17}$)	33.1027 ($3.7963\times10^{+17}$)
	0.99	39.169690 ($1.1114\times10^{+18}$)	32.3337 ($1.0579\times10^{+18}$)	39.169690 ($1.1114\times10^{+18}$)	32.3337 ($1.0579\times10^{+18}$)	**39.169680** ($1.1114\times10^{+18}$)	32.3337 ($1.0579\times10^{+18}$)	39.169690 ($1.1114\times10^{+18}$)	32.3337 ($1.0579\times10^{+18}$)
900	0.75	3.7924 ($6.4861\times10^{+6}$)	4.1115 ($5.1591\times10^{+6}$)	3.79167 ($6.4853\times10^{+6}$)	***4.111190*** ($5.1585\times10^{+6}$)	**3.791667** (**6.4852**$\times10^{+6}$)	4.111191 (***5.1585***$\times10^{+6}$)	3.7917 ($6.4853\times10^{+6}$)	4.1112 ($5.1585\times10^{+6}$)
	0.85	2.6750 ($3.9010\times10^{+5}$)	2.7999 ($3.0476\times10^{+5}$)	2.6725 ($3.8990\times10^{+5}$)	2.798783 ($3.0464\times10^{+5}$)	**2.6724** (**3.8990**$\times10^{+5}$)	***2.798745*** (***3.0463***$\times10^{+5}$)	2.6725 ($3.8991\times10^{+5}$)	2.7988 ($3.0464\times10^{+5}$)
	0.95	1.5806 ($3.8492\times10^{+4}$)	2.0784 ($3.3507\times10^{+4}$)	1.57184 ($3.8438\times10^{+4}$)	2.07397 ($3.3468\times10^{+4}$)	**1.57164** (**3.8437**$\times10^{+4}$)	***2.073894*** (***3.3467***$\times10^{+4}$)	1.5721 ($3.8440\times10^{+4}$)	2.0741 ($3.3469\times10^{+4}$)
	0.99	1.4578 ($1.2840\times10^{+4}$)	2.3358 ($1.3680\times10^{+4}$)	1.4221 ($1.2810\times10^{+4}$)	2.3073 ($1.3655\times10^{+4}$)	**1.42058** (**1.2809**$\times10^{+4}$)	***2.306012*** (***1.3654***$\times10^{+4}$)	1.4226 ($1.2811\times10^{+4}$)	2.3076 ($1.3656\times10^{+4}$)

Table A12. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained via GA by minimizing MSEP for Poisson response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	0.8949	0.8644	0.7023	0.8514	**0.6990**	***0.8510***	0.7051	0.8513
		(74.6112)	(137.4901)	(63.3029)	(126.7304)	(**63.0662**)	(***126.5749***)	(63.4906)	(127.0173)
	0.85	1.7836	1.3135	1.2601	1.2036	**1.2590**	***1.203253***	1.2722	1.2062
		(67.3494)	(128.3608)	(53.0605)	(***113.8489***)	(**53.0173**)	(113.9832)	(53.4353)	(114.4547)
	0.95	3.6167	2.4888	2.0839	1.9623	**2.0547**	***1.9513***	2.0961	1.9672
		(39.2574)	(93.2453)	(26.3735)	(78.1439)	(**26.0999**)	(***77.8268***)	(26.4905)	(78.3636)
	0.99	18.3565	9.1638	2.9386	2.78378	**2.4137**	***2.4774***	3.1077	2.8878
		(30.5432)	(85.9155)	(11.5532)	(61.5872)	(**10.7769**)	(***60.2384***)	(11.7895)	(62.1846)
50	0.75	0.0578	***0.6008***	0.0397	0.6160	**0.0396**	0.6156	0.0401	0.6150
		(21.0960)	(144.2576)	(18.4453)	(***141.2339***)	(**18.4241**)	(141.2951)	(18.5177)	(141.4106)
	0.85	0.1223	***0.6109***	0.0652	0.6399	**0.0639**	0.6403	0.0658	0.6388
		(16.1699)	(138.8017)	(13.4936)	(135.6349)	(**13.4241**)	(***135.5892***)	(13.5281)	(135.7339)
	0.95	0.5185	***0.5825***	**0.1209**	0.6768	0.1273	0.6675	0.1322	0.665195
		(10.0991)	(133.8669)	(**6.4758**)	(***129.9416***)	(6.5370)	(130.2599)	(6.5854)	(130.3382)
	0.99	6.2826	0.7030	**0.3351**	0.5474	0.3503	0.5328	0.3940	*0.5242*
		(19.9069)	(144.3472)	(**9.0853**)	(***132.8631***)	(9.1099)	(133.2301)	(9.2099)	(133.4817)
100	0.75	0.0567	***0.5448***	0.05363	0.55159	**0.05359**	0.55154	0.05369	0.55124
		(18.2253)	(278.5314)	(17.0624)	(275.5351)	(**17.0432**)	(***275.5202***)	(17.0881)	(275.6486)
	0.85	0.1013	***0.6497***	0.09260	0.6598	**0.092159**	0.6595	0.0928	0.6591
		(10.6043)	(247.7561)	(8.6113)	(***244.1418***)	(**8.6082**)	(244.1674)	(8.6666)	(244.2852)
	0.95	0.4543	0.8667	0.3201	0.8313	**0.3186**	*0.8304*	0.3224	0.8313
		(9.2895)	(233.7580)	(6.1049)	(228.0863)	(**6.0668**)	(***228.0251***)	(6.1616)	(228.2299)
	0.99	1.5052	2.0309	**0.5501**	***1.2134***	0.5549	1.2140	0.5633	1.2295
		(4.7558)	(223.5042)	(**1.6260**)	(***216.7049***)	(1.6454)	(216.7762)	(1.6784)	(216.9224)
200	0.75	**0.0546**	***1.0232***	0.05941	1.0329	0.05946	1.0327	0.05924	1.0323
		(**6.6129**)	(440.6551)	(7.0596)	(***439.4446***)	(7.0651)	(439.4948)	(7.0432)	(439.5404)
	0.85	**0.1337**	***1.1373***	0.1363	1.1539	0.1364	1.1537	0.1361	1.15292
		(**6.5916**)	(493.0939)	(6.6059)	(***491.6778***)	(6.6081)	(491.7399)	(6.6022)	(491.7990)
	0.95	0.4788	***1.3659***	0.4611	1.4134	0.4613	1.4147	**0.4609**	1.4121
		(6.1589)	(548.7239)	(5.9293)	(***547.6648***)	(**5.9285**)	(547.6700)	(5.9296)	(547.7185)
	0.99	1.6040	***1.5028***	1.1236	1.6066	**1.0762**	1.6736	1.1262	1.6030
		(5.1287)	(567.8079)	(4.3837)	(***567.2539***)	(**4.2710**)	(567.3121)	(4.3882)	(567.3116)
500	0.75	0.002405	0.494285	0.002405	0.494287	0.002424	***0.494233***	0.002405	0.494285
		(**22.3229**)	(***6750.9799***)	(22.3275)	(6751.0155)	(22.6042)	(6753.9249)	(22.3249)	(6751.0001)
	0.85	**0.009361**	0.525491	0.009364	0.525496	0.009469	***0.525450***	0.009362	0.525491
		(**69.9013**)	(***7302.5483***)	(69.9159)	(7302.5958)	(70.5632)	(7305.5965)	(69.9065)	(7302.5723)
	0.95	0.004214	0.532681	**0.004213**	0.532691	0.004228	***0.532503***	0.004214	0.532680
		(**128.8128**)	(***7807.1400***)	(128.8179)	(7807.1749)	(129.1670)	(7810.4586)	(128.8145)	(7807.1604)
	0.99	**0.11453**	0.94234	0.11458	0.94222	0.11698	***0.928769***	0.11455	0.94217
		(**68.7669**)	(***7707.1843***)	(68.8022)	(7707.2791)	(70.1159)	(7712.4908)	(68.7826)	(7707.2471)
900	0.75	0.00481	***0.59016***	0.004518	0.59111	**0.004516**	0.59108	0.004527	0.59105
		(43.5038)	(2810.5376)	(42.0677)	(***2808.7904***)	(**42.0604**)	(2808.8192)	(42.1122)	(2808.8818)
	0.85	0.01208	***0.56228***	0.01086	0.56411	**0.01082**	0.56415	0.01088	0.56406
		(13.3647)	(2499.2135)	(12.1515)	(2496.9709)	(**12.1082**)	(***2496.9099***)	(12.1686)	(2497.0232)
	0.95	**0.0422**	*0.7051*	0.0437	0.7115	0.0438	0.7116	0.0437	0.7113
		(12.4758)	(2191.4018)	(12.2936)	(2189.4525)	(**12.2916**)	(***2189.4305***)	(12.2969)	(2189.5189)
	0.99	**0.2658**	***0.9867***	0.2871	1.0091	0.2884	1.0099	0.2869	1.0086
		(**10.6533**)	(2104.5704)	(10.8146)	(2102.6201)	(10.8325)	(***2102.5445***)	(10.8114)	(2102.6471)

Table A13. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained via GA by minimizing MAE for Poisson response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	8.3548	9.4415	**8.2390**	***9.4019***	8.2393	***9.4019***	8.2426	9.4030
		($4.4666 \times 10^{+4}$)	($4.5045 \times 10^{+4}$)	(***4.4213***$\times 10^{+4}$)	(**4.4573**$\times 10^{+4}$)	($4.4214 \times 10^{+4}$)	(**4.4573**$\times 10^{+4}$)	($4.4227 \times 10^{+4}$)	($4.4586 \times 10^{+4}$)
	0.85	10.9089	12.1354	10.9666	11.9674	**10.4968**	***11.9657***	10.5105	11.9715
		($1.2228 \times 10^{+4}$)	($1.2268 \times 10^{+4}$)	($1.1926 \times 10^{+4}$)	($1.1945 \times 10^{+4}$)	(**1.1925**$\times 10^{+4}$)	(***1.1944***$\times 10^{+4}$)	($1.1935 \times 10^{+4}$)	($1.1955 \times 10^{+4}$)
	0.95	15.3082	16.2553	13.1455	14.9247	**13.1231**	***14.9056***	13.2272	14.9743
		($1.4093 \times 10^{+3}$)	($1.3965 \times 10^{+3}$)	($1.2915 \times 10^{+3}$)	($1.2521 \times 10^{+3}$)	(**1.2902**$\times 10^{+3}$)	(***1.2501***$\times 10^{+3}$)	($1.2959 \times 10^{+3}$)	($1.2574 \times 10^{+3}$)
	0.99	65.2098	59.8493	49.5086	50.4611	**30.5525**	***36.4105***	49.9783	50.7621
		($7.8288 \times 10^{+2}$)	($7.9968 \times 10^{+2}$)	($6.8449 \times 10^{+2}$)	($6.8201 \times 10^{+2}$)	(**5.8091**$\times 10^{+2}$)	(***5.2971***$\times 10^{+2}$)	($6.8728 \times 10^{+2}$)	($6.8559 \times 10^{+2}$)
50	0.75	3.8163	3.5842	**3.7824**	***3.570603***	3.7825	3.570638	3.7835	3.5711
		($1.6611 \times 10^{+4}$)	($1.6349 \times 10^{+4}$)	(**1.6484**$\times 10^{+4}$)	(*1.6247*$\times 10^{+4}$)	($1.6485 \times 10^{+4}$)	(***1.6247***$\times 10^{+4}$)	($1.6488 \times 10^{+4}$)	($1.6250 \times 10^{+4}$)
	0.85	3.6479	3.0166	3.5545	2.9832	**3.5515**	***2.9822***	3.5556	2.9837
		($6.0139 \times 10^{+3}$)	($5.8932 \times 10^{+3}$)	($5.9366 \times 10^{+3}$)	($5.8339 \times 10^{+3}$)	(**5.9341**$\times 10^{+3}$)	(***5.8320***$\times 10^{+3}$)	($5.9376 \times 10^{+3}$)	($5.8346 \times 10^{+3}$)
	0.95	4.9291	2.9113	4.3668	2.7335	**4.3491**	***2.7288***	4.3762	2.7375
		($2.8904 \times 10^{+3}$)	($2.8237 \times 10^{+3}$)	($2.8207 \times 10^{+3}$)	($2.7789 \times 10^{+3}$)	(**2.8185**$\times 10^{+3}$)	(***2.7776***$\times 10^{+3}$)	($2.8219 \times 10^{+3}$)	($2.7798 \times 10^{+3}$)
	0.99	28.0073	14.5864	19.9062	11.2016	**16.8526**	***9.9049***	20.2597	11.4006
		($2.7037 \times 10^{+3}$)	($2.6381 \times 10^{+3}$)	($2.5626 \times 10^{+3}$)	($2.5485 \times 10^{+3}$)	(**2.5116**$\times 10^{+3}$)	(***2.5155***$\times 10^{+3}$)	($2.5685 \times 10^{+3}$)	($2.5533 \times 10^{+3}$)
100	0.75	3.6889	3.5912	3.6772	3.5832	**3.6769**	***3.5829***	3.6774	3.5833
		($7.2923 \times 10^{+4}$)	($6.9481 \times 10^{+4}$)	($7.2663 \times 10^{+4}$)	($6.9275 \times 10^{+4}$)	(**7.2656**$\times 10^{+4}$)	(***6.9269***$\times 10^{+4}$)	($7.2666 \times 10^{+4}$)	($6.9277 \times 10^{+4}$)
	0.85	3.3267	3.1868	3.2962	3.1686	**3.2958**	***3.1685***	3.2971	3.1693
		($1.7925 \times 10^{+4}$)	($1.6732 \times 10^{+4}$)	($1.7794 \times 10^{+4}$)	($1.6630 \times 10^{+4}$)	(**1.7793**$\times 10^{+4}$)	(***1.6629***$\times 10^{+4}$)	($1.7798 \times 10^{+4}$)	($1.6633 \times 10^{+4}$)
	0.95	3.5021	2.9878	3.2904	2.8572	**3.2876**	***2.8561***	3.2968	2.8618
		($1.5708 \times 10^{+3}$)	($1.7527 \times 10^{+3}$)	($1.5263 \times 10^{+3}$)	($1.7129 \times 10^{+3}$)	(**1.5257**$\times 10^{+3}$)	(***1.7125***$\times 10^{+3}$)	($1.5276 \times 10^{+3}$)	($1.7143 \times 10^{+3}$)
	0.99	5.5057	4.1525	3.7525	***3.0342***	**3.7481**	3.0389	3.8185	3.0833
		($4.1144 \times 10^{+2}$)	($6.4412 \times 10^{+2}$)	($3.8451 \times 10^{+2}$)	(***6.1500***$\times 10^{+2}$)	(**3.8442**$\times 10^{+2}$)	($6.1506 \times 10^{+2}$)	($3.8547 \times 10^{+2}$)	($6.1622 \times 10^{+2}$)
200	0.75	**0.4033**	***1.0256***	0.4054	1.0288	0.4054	1.0289	0.4053	1.0287
		(231.9036)	(613.2740)	(228.8429)	(609.6339)	(**228.7393**)	(*609.5265*)	(228.8820)	(609.6965)
	0.85	**0.4159**	***1.1460***	0.4189	1.1543	0.4189	1.1542	0.4188	1.1539
		(101.6331)	(511.2616)	(99.9501)	(***508.3919***)	(**99.9329**)	(508.4009)	(99.9942)	(508.5048)
	0.95	**0.6673**	***1.4035***	0.6711	1.4361	0.6714	1.4375	0.6710	1.4355
		(51.1861)	(509.2843)	(50.2542)	(506.8357)	(**50.2128**)	(***506.7462***)	(50.2657)	(506.8904)
	0.99	1.4954	***1.5741***	1.3419	1.6486	**1.2885**	1.7024	1.3435	1.6472
		(39.4517)	(522.0365)	(38.4797)	(519.5768)	(**38.0503**)	(***518.3715***)	(38.4901)	(519.6352)
500	0.75	23.6214	33.4006	23.6214	33.4006	23.6214	33.4006	23.6214	33.4006
		($2.7658 \times 10^{+17}$)	($2.5720 \times 10^{+17}$)	($2.7658 \times 10^{+17}$)	($2.5720 \times 10^{+17}$)	($2.7658 \times 10^{+17}$)	($2.5720 \times 10^{+17}$)	($2.7658 \times 10^{+17}$)	($2.5720 \times 10^{+17}$)
	0.85	22.2121	32.2248	22.2121	32.2248	22.2121	32.2248	22.2121	32.2248
		($2.3226 \times 10^{+17}$)	($2.1854 \times 10^{+17}$)	($2.3226 \times 10^{+17}$)	($2.1854 \times 10^{+17}$)	($2.3226 \times 10^{+17}$)	($2.1854 \times 10^{+17}$)	($2.3226 \times 10^{+17}$)	($2.1854 \times 10^{+17}$)
	0.95	26.2249	33.1027	26.2249	33.1027	26.2249	33.1027	26.2249	33.1027
		($4.0334 \times 10^{+17}$)	($3.7963 \times 10^{+17}$)	($4.0334 \times 10^{+17}$)	($3.7963 \times 10^{+17}$)	($4.0334 \times 10^{+17}$)	($3.7963 \times 10^{+17}$)	($4.0334 \times 10^{+17}$)	($3.7963 \times 10^{+17}$)
	0.99	39.1697	32.3337	39.1697	32.3337	39.1697	32.3337	39.1697	32.3337
		($1.1114 \times 10^{+18}$)	($1.0579 \times 10^{+18}$)	($1.1114 \times 10^{+18}$)	($1.0579 \times 10^{+18}$)	($1.1114 \times 10^{+18}$)	($1.0579 \times 10^{+18}$)	($1.1114 \times 10^{+18}$)	($/1.0579 \times 10^{+18}$)
900	0.75	3.7924	4.1115	2.3462	2.33869	**2.3461**	***2.33868***	2.3462	2.33870
		($6.4861 \times 10^{+6}$)	($5.1591 \times 10^{+6}$)	($6.4853 \times 10^{+6}$)	($5.1585 \times 10^{+6}$)	(**6.4852**$\times 10^{+6}$)	($5.1585 \times 10^{+6}$)	($6.4853 \times 10^{+6}$)	($5.1585 \times 10^{+6}$)
	0.85	2.6750	2.7999	2.6725	2.798771	**2.6725**	***2.798770***	2.6726	2.79881
		($3.9010 \times 10^{+5}$)	($3.0476 \times 10^{+5}$)	($3.0464 \times 10^{+5}$)	($3.8990 \times 10^{+5}$)	(**3.8990**$\times 10^{+5}$)	(***3.0463***$\times 10^{+5}$)	($3.8991 \times 10^{+5}$)	($3.0464 \times 10^{+5}$)
	0.95	1.5806	2.0784	1.5718	2.073960	**1.5717**	***2.073943***	1.5721	2.0741
		($3.8492 \times 10^{+4}$)	($3.3507 \times 10^{+4}$)	($3.8438 \times 10^{+4}$)	($3.3467 \times 10^{+4}$)	(**3.8438**$\times 10^{+4}$)	(***3.3467***$\times 10^{+4}$)	($3.8440 \times 10^{+4}$)	($3.3469 \times 10^{+4}$)
	0.99	1.4578	2.3358	1.4219	2.3073	**1.4206**	***2.3060***	1.4225	2.3075
		($1.2840 \times 10^{+4}$)	($1.3680 \times 10^{+4}$)	($1.2810 \times 10^{+4}$)	($1.3655 \times 10^{+4}$)	(**1.2809**$\times 10^{+4}$)	(***1.3654***$\times 10^{+4}$)	($1.2811 \times 10^{+4}$)	($1.3656 \times 10^{+4}$)

Table A14. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained via GA by minimizing MAE for Poisson response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	0.8949	0.8644	**0.7001**	0.8513	0.7009	***0.8509***	0.7047	0.8511
		(74.6112)	(137.4901)	(**63.1457**)	(***126.5656***)	(63.1913)	(126.7673)	(63.4583)	(127.0445)
	0.85	1.7836	1.3135	**1.2586**	***1.203273***	1.2591	1.203345	1.2707	1.2059
		(67.3494)	(128.3608)	(**53.0105**)	(***113.7915***)	(53.0182)	(113.9829)	(53.3872)	(114.3991)
	0.95	3.6167	2.4888	2.0859	1.9630	**2.0829**	***1.962551***	2.1262	1.9791
		(39.2574)	(93.2453)	(26.3895)	(***78.1653***)	(**26.3759**)	(78.3532)	(26.7780)	(78.8965)
	0.99	18.3566	9.1638	6.5909	4.6517	**2.4459**	***2.497598***	6.8521	4.7952
		(30.5432)	(85.9155)	(16.3152)	(69.4571)	(**10.8252**)	(***60.3728***)	(16.6417)	(70.0974)
50	0.75	0.0578	***0.6008***	**0.0396**	0.6161	0.0396	0.6156	0.0400	0.6152
		(21.0960)	(144.2576)	(18.4194)	(***141.2014***)	(**18.4248**)	(141.2969)	(18.4933)	(141.3817)
	0.85	0.1223	***0.6109***	0.0649	0.6401	**0.0638**	0.6406	0.0654	0.6393
		(16.1699)	(138.8017)	(13.4810)	(135.6177)	(**13.4151**)	(***135.5626***)	(13.5068)	(135.6916)
	0.95	0.5185	***0.5825***	0.1253	0.6746	**0.1200**	0.6750	0.1292	0.6704
		(10.0991)	(133.8669)	(6.5191)	(130.0168)	(**6.4659**)	(***129.9990***)	(6.5574)	(130.1551)
	0.99	6.2826	0.7031	0.8322	0.4873	**0.3784**	0.5228	0.9655	***0.4815***
		(19.9069)	(144.3472)	(10.1391)	(135.0854)	(**9.1686**)	(***133.5324***)	(10.3939)	(135.8958)
100	0.75	0.0567	0.5449	0.0536	0.5516	**0.0536**	0.5517	0.0537	***0.5514***
		(18.2253)	(278.5314)	(17.0574)	(275.5226)	(**17.0278**)	(***275.4556***)	(17.0681)	(275.5745)
	0.85	0.1013	***0.6497***	0.0926	0.6597	**0.0925**	0.6595	0.0928	0.6591
		(10.6043)	(247.7561)	(8.6223)	(***244.1631***)	(**8.6065**)	(244.1636)	(8.6758)	(244.3019)
	0.95	0.4543	0.8667	0.3203	0.8314	**0.3193**	***0.8304***	0.3233	0.8314
		(9.2895)	(233.7580)	(6.1098)	(228.0968)	(**6.0856**)	(***228.0721***)	(6.1849)	(228.2858)
	0.99	1.5052	2.0309	**0.5499**	1.2134	0.5523	***1.2107***	0.5601	1.2258
		(4.7558)	(223.5042)	(**1.6259**)	(***216.7046***)	(1.6341)	(216.7279)	(1.6658)	(216.8729)
200	0.75	**0.0547**	***1.0232***	0.0594	1.0329	0.0596	1.0331	0.0593	1.0326
		(**6.6129**)	(440.6551)	(7.0571)	(439.4509)	(7.0757)	(***439.4365***)	(7.0508)	(439.4904)
	0.85	**0.1338**	***1.1373***	0.1364	1.1542	0.1364	1.1538	0.1362	1.1532
		(**6.5916**)	(493.0939)	(6.6078)	(***491.6651***)	(6.6083)	(491.7254)	(6.6040)	(491.7736)
	0.95	0.4789	***1.3659***	0.4609	1.4131	0.4614	1.4147	**0.4609**	1.4118
		(6.1589)	(548.7239)	(**5.9289**)	(***547.6695***)	(5.9291)	(547.6752)	(5.9298)	(547.7279)
	0.99	1.6040	***1.5027***	1.1277	1.6062	**1.0763**	1.6736	1.1306	1.6028
		(5.1287)	(567.8079)	(4.3884)	(***567.2736***)	(**4.2711**)	(567.3128)	(4.3934)	(567.3307)
500	0.75	**0.0024**	***0.4942***	0.0025	0.4945	0.0025	0.4945	0.0025	0.4945
		(**22.3229**)	(*6750.9799*)	(22.9675)	(6755.9856)	(22.9776)	(6756.0912)	(22.9529)	(6755.9028)
	0.85	**0.0094**	***0.5254***	0.0097	0.5261	0.0097	0.5261	0.0097	0.5260
		(**9.9013**)	(***7302.5483***)	(71.7152)	(7308.3788)	(71.7539)	(7308.5270)	(71.6831)	(7308.3003)
	0.95	**0.0042**	***0.5326***	0.0043	0.5343	0.0043	0.5342	0.0043	0.5342
		(**128.8128**)	(***7807.1400***)	(129.8935)	(7812.6862)	(129.8809)	(7812.7303)	(129.8360)	(7812.5061)
	0.99	**0.1145**	0.9423	0.1201	0.9344	0.1199	***0.9329***	0.1195	0.9335
		(**68.7669**)	(***7707.1843***)	(71.6878)	(7714.6650)	(71.6084)	(7714.7955)	(71.3741)	(7714.1826)
900	0.75	0.0048	***0.5901***	0.00452	0.5911	**0.00451**	0.5911	0.0045	0.5911
		(43.5038)	(2810.5376)	(42.0747)	(2808.7991)	(**42.0334**)	(***2808.7632***)	(42.0931)	(2808.8365)
	0.85	0.0121	***0.5622***	0.0109	0.5641	**0.0108**	0.5641	0.0109	0.5640
		(13.3647)	(2499.2135)	(12.1392)	(***2496.9475***)	(**12.1290**)	(2496.9731)	(12.1765)	(2497.0611)
	0.95	**0.0423**	***0.7051***	0.0438	0.7115	0.0438	0.7115	0.0437	0.7112
		(12.4758)	(2191.4018)	(12.2934)	(***2189.4442***)	(**12.2923**)	(2189.4572)	(12.2974)	(2189.5365)
	0.99	**0.2658**	***0.9866***	0.2873	1.0092	0.2885	1.0099	0.2870	1.0087
		(10.6533)	(2104.5704)	(10.8165)	(2102.6139)	(10.8326)	(***2102.5445***)	(**10.8134**)	(2102.6409)

Table A15. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained by MSE Poisson response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	23.1596	23.8761	**22.9498**	***23.6414***	23.1362	23.8447	23.0965	23.7945
		$(1.2804 \times 10^{+11})$	$(1.2805 \times 10^{+11})$	$(1.2804 \times 10^{+11})$	$(1.2805 \times 10^{+11})$	$(1.2804 \times 10^{+11})$	$(1.2805 \times 10^{+11})$	$(1.2804 \times 10^{+11})$	$(1.2805 \times 10^{+11})$
	0.85	**1.8506**	2.4582	1.9675	2.5669	1.8626	2.4280	1.8770	***2.4186***
		$(2.8127 \times 10^{+3})$	$(2.8496 \times 10^{+3})$	$(\mathbf{2.7529} \times 10^{+3})$	$(\boldsymbol{2.7807} \times 10^{+3})$	$(2.8054 \times 10^{+3})$	$(2.8428 \times 10^{+3})$	$(2.7956 \times 10^{+3})$	$(2.8318 \times 10^{+3})$
	0.95	3.0475	3.3930	**2.4750**	***3.2024***	2.8289	3.2426	2.7621	3.2127
		(103.1005)	(147.9965)	(**84.9373**)	(***121.3511***)	(99.8243)	(146.4904)	(94.5070)	(141.5373)
	0.99	27.6106	25.4649	**12.6354**	***13.1363***	23.1215	21.8655	21.4840	20.4238
		$(9.1009 \times 10^{+3})$	$(8.4703 \times 10^{+3})$	$(\mathbf{8.4972} \times 10^{+3})$	$(\boldsymbol{7.8328} \times 10^{+3})$	$(8.9561 \times 10^{+3})$	$(8.2986 \times 10^{+3})$	$(8.8228 \times 10^{+3})$	$(8.1245 \times 10^{+3})$
50	0.75	**1.9282**	2.4878	1.9913	2.5184	1.9315	2.4808	1.9344	***2.4752***
		(326.9882)	(367.8967)	(**318.4131**)	(***355.3449***)	(326.5183)	(367.9316)	(326.0980)	(367.7597)
	0.85	12.4686	10.4771	**12.0421**	***10.3611***	12.4616	10.4774	12.4179	10.4822
		$(\mathbf{6.3471} \times 10^{+3})$	$(6.0829 \times 10^{+3})$	$(6.3942 \times 10^{+3})$	$(\boldsymbol{6.0596} \times 10^{+3})$	$(6.4363 \times 10^{+3})$	$(6.0827 \times 10^{+3})$	$(6.4316 \times 10^{+3})$	$(6.0809 \times 10^{+3})$
	0.95	1.5833	1.4960	**1.4447**	***1.4448***	1.5640	1.5179	1.5699	1.5346
		(47.1426)	(153.60404)	(**41.8719**)	(***142.1514***)	(46.2699)	(152.2779)	(46.3054)	(152.4184)
	0.99	6.5557	9.6350	**3.4844**	***4.8961***	5.1530	6.1492	5.0047	6.0635
		(22.8108)	(134.0435)	(**20.2422**)	(***127.8409***)	(21.5652)	(133.5622)	(21.4613)	(135.9269)
100	0.75	8.6446	8.2570	**8.6010**	***8.2362***	8.6433	8.2568	8.6365	8.2559
		$(1.4003 \times 10^{+8})$	$(1.2967 \times 10^{+8})$	$(\mathbf{1.3991} \times 10^{+8})$	$(\boldsymbol{1.2962} \times 10^{+8})$	$(1.4002 \times 10^{+8})$	$(1.2967 \times 10^{+8})$	$(1.4000 \times 10^{+8})$	$(1.2966 \times 10^{+8})$
	0.85	4.8610	4.5104	**4.7952**	***4.4816***	4.8556	4.5097	4.8501	4.5096
		$(6.9141 \times 10^{+5})$	$(6.0419 \times 10^{+5})$	$(\mathbf{6.8930} \times 10^{+5})$	$(\boldsymbol{6.0314} \times 10^{+5})$	$(6.9121 \times 10^{+5})$	$(6.0410 \times 10^{+5})$	$(6.9101 \times 10^{+5})$	$(6.0401 \times 10^{+5})$
	0.95	33.3149	28.0779	**33.2580**	***28.0558***	33.3093	28.0761	33.2891	28.0698
		$(2.4650 \times 10^{+9})$	$(2.2634 \times 10^{+9})$	$(\mathbf{2.4643} \times 10^{+9})$	$(\boldsymbol{2.2631} \times 10^{+9})$	$(2.4650 \times 10^{+9})$	$(2.2634 \times 10^{+9})$	$(2.4647 \times 10^{+9})$	$(2.2653 \times 10^{+9})$
	0.99	3.9782	4.1070	**2.7053**	***3.6202***	3.5676	3.8734	3.3539	3.7299
		(323.2629)	(572.0573)	(**284.5876**)	(***523.7959***)	(314.9502)	(571.5454)	(303.1493)	(569.8653)
200	0.75	5.5562	5.2278	**5.5496**	***5.2255***	5.5561	5.2279	5.5560	5.2279
		$(3.5423 \times 10^{+6})$	$(3.3627 \times 10^{+6})$	$(\mathbf{3.5368} \times 10^{+6})$	$(\boldsymbol{3.3590} \times 10^{+6})$	$(3.5423 \times 10^{+6})$	$(3.3627 \times 10^{+6})$	$(3.5422 \times 10^{+6})$	$(3.3627 \times 10^{+6})$
	0.85	1.9735	1.8022	**1.9417**	***1.7896***	1.9730	1.8026	1.9722	1.8031
		$(3.4205 \times 10^{+4})$	$(3.2589 \times 10^{+4})$	$(\mathbf{3.3752} \times 10^{+4})$	$(\boldsymbol{3.2247} \times 10^{+4})$	$(3.4199 \times 10^{+4})$	$(3.2599 \times 10^{+4})$	$(3.4188 \times 10^{+4})$	$(3.2616 \times 10^{+4})$
	0.95	**1.0462**	2.5259	1.1824	2.6983	1.0494	2.4832	1.0623	***2.3434***
		(85.4167)	(544.4987)	(**78.9317**)	(***529.9843***)	(85.1877)	(547.1995)	(84.4136)	(558.3936)
	0.99	**1.8150**	3.9989	2.0803	4.6544	1.8296	3.5146	1.8482	***3.4773***
		(47.1379)	(506.1389)	(**44.5115**)	(***495.6190***)	(46.7202)	(513.8217)	(46.7163)	(514.0110)
500	0.75	2.5370	2.5058	**2.5351**	***2.5048***	2.5370	2.5059	2.5370	5.5059
		$(4.7506 \times 10^{+5})$	$(4.3471 \times 10^{+5})$	$(\mathbf{4.7443} \times 10^{+5})$	$(\boldsymbol{4.3432} \times 10^{+5})$	$(4.7506 \times 10^{+5})$	$(4.3472 \times 10^{+5})$	$(4.7506 \times 10^{+5})$	$(4.3472 \times 10^{+5})$
	0.85	**1.6396**	2.5505	1.6461	2.5555	1.6399	2.5486	1.6400	***2.5482***
		$(0.3563 \times 10^{+3})$	$(1.2159 \times 10^{+3})$	$(\mathbf{0.3557} \times 10^{+3})$	$(\boldsymbol{1.2133} \times 10^{+3})$	$(0.3563 \times 10^{+3})$	$(1.2171 \times 10^{+3})$	$(0.3563 \times 10^{+3})$	$(1.2174 \times 10^{+3})$
	0.95	**1.1139**	2.5748	1.1620	2.6270	1.1149	2.5611	1.1189	***2.5205***
		$(0.2533 \times 10^{+3})$	$(1.3831 \times 10^{+3})$	$(\mathbf{0.2491} \times 10^{+3})$	$(\boldsymbol{1.3703} \times 10^{+3})$	$(0.2531 \times 10^{+3})$	$(1.3866 \times 10^{+3})$	$(0.2528 \times 10^{+3})$	$(1.3969 \times 10^{+3})$
	0.99	**1.0237**	3.3524	1.1097	3.4368	1.0332	3.1359	1.0401	***3.0444***
		$(0.1023 \times 10^{+3})$	$(1.4555 \times 10^{+3})$	$(\mathbf{0.1009} \times 10^{+3})$	$(\boldsymbol{1.4505} \times 10^{+3})$	$(0.1021 \times 10^{+3})$	$(1.4647 \times 10^{+3})$	$(0.1020 \times 10^{+3})$	$(1.4707 \times 10^{+3})$
900	0.75	1.0607	1.2634	1.0328	1.2636	**1.0307**	1.2633	1.0307	***1.2632***
		$(3.0837 \times 10^{+4})$	$(3.1763 \times 10^{+4})$	$(\mathbf{3.0744} \times 10^{+4})$	$(\boldsymbol{3.1670} \times 10^{+4})$	$(3.0836 \times 10^{+4})$	$(3.1766 \times 10^{+4})$	$(3.0836 \times 10^{+4})$	$(3.1770 \times 10^{+4})$
	0.85	0.9913	2.0647	1.0186	2.0901	**0.9915**	2.0629	0.9923	***2.0539***
		$(0.3701 \times 10^{+3})$	$(2.5696 \times 10^{+3})$	$(\mathbf{0.3668} \times 10^{+3})$	$(\boldsymbol{2.5530} \times 10^{+3})$	$(0.3701 \times 10^{+3})$	$(2.5712 \times 10^{+3})$	$(0.3699 \times 10^{+3})$	$(2.5794 \times 10^{+3})$
	0.95	0.0940	3.5718	0.1203	3.6623	0.0940	3.5108	**0.0939**	***3.4587***
		$(0.0090 \times 10^{+3})$	$(3.3315 \times 10^{+3})$	$(0.0103 \times 10^{+3})$	$(\boldsymbol{3.3215} \times 10^{+3})$	$(0.0090 \times 10^{+3})$	$(3.3431 \times 10^{+3})$	$(\mathbf{0.0089} \times 10^{+3})$	$(3.3543 \times 10^{+3})$
	0.99	**0.6814**	2.2840	0.8819	2.5562	0.6854	2.2141	0.6978	***2.0321***
		$(0.1621 \times 10^{+3})$	$(2.8532 \times 10^{+3})$	$(\mathbf{0.1596} \times 10^{+3})$	$(\boldsymbol{0.8442} \times 10^{+3})$	$(0.1620 \times 10^{+3})$	$(2.8594 \times 10^{+3})$	$(0.1618 \times 10^{+3})$	$(2.8763 \times 10^{+3})$

Table A16. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained by MSE for Poisson response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	0.3558	***0.6483***	0.3588	0.7779	**0.3545**	0.6498	0.3828	0.6759
		(25.0682)	(98.5990)	(26.1504)	(***97.7590***)	(**24.8871**)	(99.7758)	(25.6415)	(100.4658)
	0.85	2.3281	2.3917	**1.3908**	2.4662	1.9090	2.1628	1.9318	***2.1277***
		(8.0595)	(54.9552)	(**7.3833**)	(***49.0311***)	(7.6364)	(60.5812)	(7.6385)	(62.2585)
	0.95	10.1329	3.7087	**2.0297**	***2.9813***	5.9594	4.5240	5.9774	4.4051
		(5.8389)	(72.2275)	(**3.9696**)	(***63.0613***)	(4.7289)	(82.7340)	(4.5525)	(85.7387)
	0.99	6.6161	5.7780	**2.6441**	***3.5508***	4.4608	4.4377	4.3807	4.3523
		(12.6629)	(67.6560)	(**8.0275**)	(***61.3199***)	(10.1798)	(68.2210)	(9.8945)	(68.3792)
50	0.75	0.5310	2.4471	**0.4940**	2.5287	0.5143	2.3175	0.5067	***2.2473***
		(11.6749)	(140.4673)	(**9.9830**)	(***131.0026***)	(11.4885)	(148.7566)	(11.3615)	(162.4514)
	0.85	0.1629	1.1571	0.3772	1.4429	0.1618	1.1285	**0.1586**	***0.9849***
		(23.5088)	(159.2415)	(37.5934)	(***157.8098***)	(23.3197)	(160.3397)	(**22.5762**)	(171.4932)
	0.95	1.3427	***0.9899***	**1.1703**	1.0749	1.2837	1.0537	1.2984	1.0750
		(6.1593)	(127.9580)	(**5.1058**)	(***118.3274***)	(5.9142)	(128.3328)	(5.9464)	(128.2985)
	0.99	14.5253	12.5982	**2.2654**	5.0393	7.8369	7.4318	7.8811	***1.3742***
		(5.8989)	(172.8037)	(**3.0497**)	(***160.2040***)	(4.2903)	(181.9657)	(4.3832)	(186.4917)
100	0.75	0.0763	0.8699	0.1270	1.0071	**0.0762**	0.8614	0.0768	***0.8221***
		(30.9338)	(421.5839)	(39.3035)	(***413.0584***)	(30.8170)	(422.9022)	(**30.4007**)	(429.9826)
	0.85	0.1479	0.8855	**0.1386**	0.9498	0.1467	0.8762	0.1449	***0.8654***
		(19.6125)	(359.6621)	(19.6575)	(***350.2502***)	(19.4428)	(361.5243)	(**19.3061**)	(364.0841)
	0.95	0.1896	0.8950	0.2201	1.0609	**0.1820**	***0.8986***	0.1875	0.9074
		(40.1703)	(***468.5702***)	(47.1127)	(473.1134)	(**39.1901**)	(468.7864)	(40.3728)	(470.0672)
	0.99	6.0502	5.8520	1.9652	***4.6315***	3.9670	5.1059	**3.8435**	4.7337
		(14.9625)	(424.6825)	(**11.1442**)	(***418.5808***)	(12.4865)	(438.3449)	(11.9800)	(463.0178)
200	0.75	0.0397	0.8753	**0.0384**	0.8907	0.0396	0.8690	0.0396	***0.8645***
		(23.2980)	(614.4115)	(**22.8812**)	(***611.7764***)	(23.2778)	(616.1912)	(23.2471)	(617.4408)
	0.85	0.0691	0.8973	0.0759	0.9494	**0.0690**	0.8864	0.0691	***0.8699***
		(15.6044)	(487.2846)	(16.5544)	(***479.0240***)	(15.5875)	(489.7291)	(**15.5667**)	(493.2165)
	0.95	0.5604	2.7443	0.7295	2.9096	0.5474	2.6102	**0.5340**	***2.2313***
		(8.7660)	(671.5844)	(12.1773)	(***661.1939***)	(8.6352)	(677.4129)	(**8.4487**)	(707.2889)
	0.99	2.0112	4.9118	2.3247	5.5950	**1.9108**	3.7112	1.9341	***3.6642***
		(4.6818)	(560.7849)	(4.9369)	(***552.6999***)	(4.5204)	(574.8568)	(**4.5456**)	(575.9478)
500	0.75	0.0174	1.0857	**0.0172**	1.0916	0.0174	1.0822	0.0174	***1.0804***
		$(0.0251 \times 10^{+3})$	$(1.6454 \times 10^{+3})$	$(\mathbf{0.0248} \times 10^{+3})$	$(\boldsymbol{1.6414} \times 10^{+3})$	$(0.0251 \times 10^{+3})$	$(1.6484 \times 10^{+3})$	$(0.0251 \times 10^{+3})$	$(1.6501 \times 10^{+3})$
	0.85	0.0735	2.3998	**0.0712**	2.4066	0.0733	2.3839	0.0734	***2.3800***
		$(0.0165 \times 10^{+3})$	$(1.7060 \times 10^{+3})$	$(\mathbf{0.0159} \times 10^{+3})$	$(1.7005 \times 10^{+3})$	$(0.0164 \times 10^{+3})$	$(1.7163 \times 10^{+3})$	$(0.0164 \times 10^{+3})$	$(1.7188 \times 10^{+3})$
	0.95	0.1796	3.1834	0.2522	3.2017	**0.1787**	3.1157	0.1812	***2.9363***
		$(0.0119 \times 10^{+3})$	$(1.7819 \times 10^{+3})$	$(0.0144 \times 10^{+3})$	$(\boldsymbol{1.7736} \times 10^{+3})$	$(0.0119 \times 10^{+3})$	$(1.7903 \times 10^{+3})$	$(0.0119 \times 10^{+3})$	$(1.8197 \times 10^{+3})$
	0.99	0.8815	4.5551	0.8657	4.4278	**0.8538**	3.9772	0.8548	***3.7593***
		$(1.0237 \times 10^{+3})$	$(1.7416 \times 10^{+3})$	$(\mathbf{0.0052} \times 10^{+3})$	$(\boldsymbol{1.7396} \times 10^{+3})$	$(0.0055 \times 10^{+3})$	$(1.7552 \times 10^{+3})$	$(0.0055 \times 10^{+3})$	$(1.7668 \times 10^{+3})$
900	0.75	**0.0204**	1.9393	0.0217	1.9446	**0.0204**	1.9353	**0.0204**	***1.9313***
		$(0.0521 \times 10^{+3})$	$(3.5636 \times 10^{+3})$	$(0.0547 \times 10^{+3})$	$(\boldsymbol{3.5534} \times 10^{+3})$	$(0.0521 \times 10^{+3})$	$(3.5756 \times 10^{+3})$	$(0.0522 \times 10^{+3})$	$(3.5878 \times 10^{+3})$
	0.85	0.0384	1.9933	0.0482	2.0311	0.0384	1.9841	**0.0379**	***1.9403***
		$(0.0111 \times 10^{+3})$	$(3.6383 \times 10^{+3})$	$(0.0120 \times 10^{+3})$	$(3.6201 \times 10^{+3})$	$(0.0111 \times 10^{+3})$	$(3.6472 \times 10^{+3})$	$(\mathbf{0.0110} \times 10^{+3})$	$(3.6944 \times 10^{+3})$
	0.95	**0.9431**	2.9200	0.9929	2.9804	0.9437	2.9057	0.9443	***2.8934***
		$(0.3759 \times 10^{+3})$	$(2.6493 \times 10^{+3})$	$(\mathbf{0.3709} \times 10^{+3})$	$(\boldsymbol{2.6318} \times 10^{+3})$	$(0.3758 \times 10^{+3})$	$(2.6542 \times 10^{+3})$	$(0.3758 \times 10^{+3})$	$(2.6586 \times 10^{+3})$
	0.99	0.4907	2.7980	0.7976	2.9670	0.4842	2.5773	**0.4809**	***2.1427***
		$(0.0050 \times 10^{+3})$	$(\boldsymbol{3.3810} \times 10^{+3})$	$(0.0069 \times 10^{+3})$	$(3.3862 \times 10^{+3})$	$(0.0050 \times 10^{+3})$	$(3.3917 \times 10^{+3})$	$(0.0050 \times 10^{+3})$	$(3.4252 \times 10^{+3})$

Table A17. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained via GA by minimizing MSEP for Gamma response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	17.9888	18.1947	**17.9170**	***18.1504***	17.9234	18.1543	17.9245	18.1550
		(7730.9470)	(7718.3280)	(7727.0090)	(***7712.8460***)	(7713.3460)	(7727.3100)	(**7713.2590**)	(7727.3720)
	0.85	15.0631	13.6155	14.8012	13.4847	**14.7988**	***13.4840***	14.8082	13.4888
		(655.6418)	(680.5180)	(650.7882)	(675.0309)	(**650.6919**)	(***674.9450***)	(650.8722)	(675.1506)
	0.95	12.4368	9.0809	**10.9280**	***8.4152***	11.2146	8.6160	11.2765	8.6416
		(73.4909)	(105.0058)	(**68.2530**)	(***99.5475***)	(69.2369)	(101.2052)	(69.4621)	(101.4142)
	0.99	12.6731	6.8412	10.8979	6.4049	**5.5154**	***4.6155***	10.9634	6.4349
		(31.3634)	(68.0404)	(30.6515)	(67.3383)	(**28.2126**)	(***64.8151***)	(30.6785)	(67.4054)
50	0.75	10.8089	12.3488	**10.7726**	***12.3123***	10.7884	12.3257	10.7890	12.3264
		(1681.3500)	(1603.9800)	(**1678.2840**)	(***1599.7900***)	(1679.6730)	(1601.3880)	(1679.7250)	(1601.4680)
	0.85	11.4235	12.4665	**11.2864**	***12.3392***	11.3352	12.3813	11.3369	12.3829
		(752.8328)	(716.6898)	(**748.3756**)	(***710.9133***)	(749.9900)	(712.8124)	(750.0432)	(712.8844)
	0.95	14.6291	14.8125	**13.5764**	***13.8420***	13.8050	14.0504	13.8440	14.0867
		(368.1667)	(365.5289)	(**360.5653**)	(***355.5503***)	(362.2196)	(357.6973)	(362.5063)	(358.0788)
	0.99	28.7023	27.6024	27.8515	26.8321	**17.6528**	***17.3635***	27.9375	26.9122
		(289.5643)	(303.3613)	(288.6228)	(302.1216)	(**276.8689**)	(***285.1183***)	(288.7234)	(302.2590)
100	0.75	7.0937	5.6175	7.0472	***5.5968***	**7.0469**	5.5969	7.0488	5.5978
		(739.5130)	(820.3079)	(736.8929)	(***818.0202***)	(**736.8887**)	(818.0859)	(737.0063)	(818.1776)
	0.85	7.0513	5.1282	6.9511	***5.0858***	**6.9503**	5.0862	6.9547	5.0881
		(411.7009)	(524.3464)	(408.8039)	(***521.8563***)	(**408.7998**)	(521.9229)	(408.9285)	(522.0304)
	0.95	7.1160	4.5156	**6.7372**	***4.3747***	6.7650	4.3975	6.7743	4.4006
		(186.6831)	(342.1177)	(**183.4979**)	(***339.4986***)	(183.7378)	(339.9967)	(183.8170)	(340.0528)
	0.99	7.9597	3.8406	**6.0562**	***3.3657***	6.1229	3.4556	6.2389	3.4810
		(116.7167)	(291.5884)	(**113.6544**)	(***289.4448***)	(113.7491)	(290.0110)	(113.9515)	(290.1203)
200	0.75	15.3255	15.0654	15.3240	15.0646	**15.3235**	***15.0645***	15.3240	15.0646
		(29484.3400)	(29691.9000)	(29481.4500)	(29688.5900)	(**29481.4200**)	(***29688.5600***)	(29481.5200)	(29688.6900)
	0.85	13.8748	13.1006	13.8670	13.0961	**13.8669**	***13.0960***	13.8672	13.0962
		(6935.4980)	(7248.8050)	(6932.4580)	(7245.3490)	(**6932.4260**)	(***7245.3200***)	(6932.5520)	(7245.4630)
	0.95	12.1200	10.6186	12.0661	10.5846	**12.0652**	***10.5842***	12.0673	10.5855
		(1899.8620)	(2217.2760)	(1896.5570)	(2213.6520)	(**1896.5020**)	(***2213.6050***)	(1896.6300)	(2213.7470)
	0.99	10.1021	8.2615	9.8154	8.1003	**9.8034**	***8.0945***	9.8189	8.1032
		(1190.0330)	(1530.2470)	(1186.9830)	(1526.9480)	(**1186.8580**)	(***1526.8240***)	(1187.0200)	(1527.0020)
500	0.75	4.4694	3.4807	**4.4653**	***3.4788***	4.4678	3.4808	4.4679	3.4807
		(34007.0400)	(34127.4400)	(**34005.3200**)	(***34126.4100***)	(34006.3700)	(34127.8600)	(34006.4200)	(34127.8300)
	0.85	4.3436	3.3440	**4.3375**	***3.3409***	4.3407	3.3435	4.3409	3.3435
		(35405.5600)	(35516.6500)	(**35403.9100**)	(***35515.7600***)	(35404.7800)	(35516.9700)	(35404.8300)	(35516.9500)
	0.95	4.0312	3.2421	**4.0184**	***3.2332***	4.0218	3.2363	4.0224	3.2367
		(37245.7100)	(37305.1100)	(**37244.3600**)	(***37304.3500***)	(37244.7100)	(37304.8700)	(37244.7800)	(37304.8900)
	0.99	3.7775	3.3016	**3.7160**	***3.2516***	3.7221	3.2559	3.7240	3.2575
		(38440.0600)	(38447.4000)	(**38439.2600**)	(***38446.5700***)	(38439.3400)	(38446.6400)	(38439.3600)	(38446.6700)
900	0.75	8.2299	6.9022	**8.2270**	***6.9002***	8.2271	6.9010	8.2271	6.9011
		(10350.3100)	(11034.9700)	(**10347.8300**)	(***11032.7200***)	(10347.8400)	(11032.7600)	(10347.9000)	(11032.8100)
	0.85	7.8131	5.8105	7.8053	5.8073	**7.8051**	***5.8072***	7.8054	5.8074
		(5436.0560)	(6329.2670)	(5433.1580)	(6326.7770)	(**5433.0840**)	(***6326.7400***)	(5433.2060)	(6326.8430)
	0.95	7.2957	4.5934	7.2631	***4.5816***	**7.2630**	4.5819	7.2641	4.5823
		(2627.4840)	(4031.8910)	(2624.3320)	(***4029.4270***)	(**2624.3240**)	(4029.5150)	(2624.4330)	(4029.5970)
	0.99	6.8114	3.6589	**6.6434**	***3.6164***	6.6479	3.6209	6.6515	3.6217
		(1819.6790)	(3531.3490)	(**1816.6320**)	(***3529.5970***)	(1816.7130)	(3529.8450)	(1816.7790)	(3529.8780)

Table A18. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained via GA by minimizing MSEP for Gamma response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	11.9951	7.4501	**10.1865**	***7.0337***	10.3224	7.0954	10.3441	7.1010
		(170.7891)	(220.0739)	(**153.1231**)	(***207.1051***)	(154.4451)	(208.8750)	(154.6842)	(209.0509)
	0.85	11.7867	6.9730	9.0697	***6.3587***	**9.0597**	6.3599	9.1285	6.3799
		(97.2824)	(150.1952)	(84.9850)	(***140.1717***)	(**84.8501**)	(140.2194)	(85.2010)	(140.5370)
	0.95	11.4882	6.5803	**5.9930**	***5.2473***	6.9802	5.7455	7.0946	5.7813
		(48.8360)	(103.4048)	(**43.0983**)	(***97.7288***)	(44.2699)	(100.6117)	(44.4215)	(100.7148)
	0.99	13.7588	7.6452	7.1932	6.5738	**3.2938**	***3.8887***	7.4087	6.5824
		(35.2670)	(89.9996)	(34.5637)	(***89.7655***)	(**33.7579**)	(90.8419)	(34.6044)	(89.8632)
50	0.75	12.6408	9.7006	**10.7753**	***8.8252***	11.6143	9.3349	11.6390	9.3448
		(399.6054)	(494.4707)	(**381.5204**)	(***474.4040***)	(389.5676)	(486.3870)	(389.8344)	(486.6079)
	0.85	14.3981	10.2043	**11.0755**	***8.8543***	12.1801	9.5022	12.2099	9.5130
		(320.8657)	(432.2142)	(**303.7596**)	(***411.6184***)	(309.7875)	(421.6973)	(309.9334)	(421.8622)
	0.95	15.4411	13.0724	**10.5728**	***9.2154***	11.9988	10.1561	12.1868	10.2505
		(273.5242)	(394.1230)	(**248.2727**)	(***372.0959***)	(250.8076)	(377.3906)	(251.1749)	(377.9443)
	0.99	28.4388	26.5725	25.1736	23.7139	**8.3468**	***9.2404***	25.6507	24.0077
		(258.8695)	(390.6796)	(256.0782)	(387.3908)	(**239.0383**)	(***372.1941***)	(256.2554)	(387.7436)
100	0.75	8.0427	4.2324	**7.5102**	***4.1707***	7.5272	4.1774	7.5465	4.1796
		(591.7564)	(776.8677)	(**583.3936**)	(***772.5340***)	(583.6644)	(773.3648)	(583.9756)	(773.5010)
	0.85	8.1511	3.9518	**7.3127**	***3.8651***	7.3257	3.8735	7.3575	3.8767
		(371.2352)	(588.0413)	(**363.4569**)	(***584.5174***)	(363.5658)	(585.1398)	(363.8729)	(585.2579)
	0.95	8.4883	3.4002	**6.3309**	***3.2325***	6.4967	3.2951	6.5326	3.2973
		(202.3108)	(454.0343)	(**196.7926**)	(***452.0858***)	(197.3283)	(453.3398)	(197.4509)	(453.3472)
	0.99	9.6333	2.4887	**4.3660**	***2.2950***	4.5716	2.4506	4.7727	2.4540
		(145.5359)	(412.3310)	(**142.9489**)	(***413.3249***)	(143.2671)	(414.4543)	(143.2887)	(414.3354)
200	0.75	8.0501	6.1378	7.9560	***6.1117***	**7.9552**	6.1120	7.9586	6.1129
		(709.7314)	(1090.2871)	(700.0023)	(***1082.9962***)	(**699.9164**)	(1083.0559)	(700.2653)	(1083.3152)
	0.85	7.9411	5.9158	7.7756	***5.8723***	**7.7740**	5.8729	7.7807	5.8747
		(379.3958)	(780.7428)	(371.1661)	(***774.4089***)	(**371.0884**)	(774.4847)	(371.4182)	(774.7369)
	0.95	7.4528	5.5175	6.9868	***5.4019***	**6.9807**	5.4024	6.9972	5.4066
		(151.7326)	(576.1429)	(145.4917)	(571.1987)	(145.4072)	(571.2100)	(145.6282)	(571.3903)
	0.99	6.3148	4.9230	4.9205	4.5506	**4.8852**	***4.5424***	4.9367	4.5588
		(87.1924)	(523.6791)	(83.5048)	(520.6283)	(**83.4118**)	(***520.5572***)	(83.5472)	(520.6880)
500	0.75	9.4341	4.4162	**9.3431**	***4.4073***	9.4251	4.4224	9.4284	4.4219
		(34961.0300)	(36616.4200)	(**34941.2800**)	(***36608.1200***)	(34950.0800)	(36624.7600)	(34950.5200)	(36624.1700)
	0.85	8.8876	4.2063	**8.8717**	***4.1940***	9.0891	4.2102	9.0643	4.2100
		(36404.3300)	(38068.0800)	(**36372.2600**)	(***38060.6900***)	(36361.9400)	(38074.2300)	(36366.1400)	(38073.8600)
	0.95	8.4585	4.2516	**8.1592**	***4.2131***	8.2566	4.2327	8.2689	4.2339
		(38177.4000)	(39843.4800)	(**38161.9700**)	(***39836.0200***)	(38162.5800)	(39843.2100)	(38163.6200)	(39843.2300)
	0.99	8.8051	5.1927	**7.4812**	***4.8673***	7.6167	4.9006	7.6434	4.9090
		(39218.4600)	(40893.7000)	(**39206.2900**)	(***40882.2800***)	(39205.9800)	(40885.7700)	(39206.4400)	(40885.9900)
900	0.75	9.3637	4.8125	**9.3024**	***4.8055***	9.3028	4.8058	9.3042	4.8060
		(7154.2830)	(9545.1920)	(**7142.2540**)	(***9538.3820***)	(7142.3410)	(9538.8040)	(7142.6260)	(9538.9570)
	0.85	9.1572	4.3861	9.0603	***4.3764***	**9.0579**	4.3765	9.0619	4.3769
		(4683.2110)	(7412.3060)	(4672.7460)	(***7407.2210***)	(**4672.4850**)	(7407.3010)	(4672.9200)	(7407.5040)
	0.95	8.7716	3.8134	8.5045	***3.7926***	**8.5041**	3.7940	8.5130	3.7947
		(2840.7450)	(5958.7490)	(2831.5020)	(***5955.7110***)	(**2831.4870**)	(5956.1000)	(2831.7950)	(5956.1890)
	0.99	8.3226	3.0336	**7.2449**	***2.9897***	7.2758	3.0007	7.2954	3.0014
		(2236.5750)	(5522.2620)	(**2228.9760**)	(***5521.8170***)	(2229.1940)	(5522.3290)	(2229.3390)	(5522.3260)

Table A19. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained via GA by minimizing MAE for Gamma response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	17.9888	18.1947	17.9815	18.1913	**17.9332**	***18.1603***	17.9843	18.1930
		(7730.9470)	(7718.3280)	(7730.7370)	(7718.0870)	(**7727.8490**)	(***7714.0040***)	(7730.8030)	(7718.1750)
	0.85	15.0631	13.6155	15.0425	13.6075	**14.8865**	***13.5357***	15.0494	13.6115
		(655.6418)	(680.5180)	(655.4045)	(680.2792)	(**652.3681**)	(***677.0824***)	(655.4859)	(680.3949)
	0.95	12.4368	9.0809	**10.9276**	***8.4146***	11.2144	8.6159	11.2761	8.6409
		(73.4909)	(105.0058)	(**68.2489**)	(***99.5421***)	(69.2357)	(101.2044)	(69.4577)	(101.4083)
	0.99	12.6731	6.8412	9.2401	5.9283	**5.5206**	***4.6187***	9.3563	5.9867
		(31.3634)	(68.0404)	(29.9619)	(66.6110)	(**28.2150**)	(***64.8232***)	(30.0109)	(66.7385)
50	0.75	10.8089	12.3488	**10.7751**	***12.3149***	10.7899	12.3268	10.7918	12.3290
		(1681.3500)	(1603.9800)	(**1678.5060**)	(***1600.0960***)	(1679.7460)	(1601.4730)	(1679.9080)	(1601.7290)
	0.85	11.4235	12.4665	11.4101	12.4557	**11.3251**	***12.3723***	11.4149	12.4597
		(752.8328)	(716.6898)	(752.4686)	(716.2636)	(**749.6215**)	(***712.3827***)	(752.5950)	(716.4182)
	0.95	14.6291	14.8125	14.5251	14.7213	**13.6983**	***13.9525***	14.5463	14.7402
		(368.1667)	(365.5289)	(367.4436)	(364.6385)	(**361.4389**)	(***356.6841***)	(367.5891)	(364.8307)
	0.99	28.7023	27.6024	27.3495	26.3672	**17.7889**	***17.4940***	27.4885	26.4974
		(289.5643)	(303.3613)	(288.1391)	(301.3683)	(**277.0348**)	(***285.3707***)	(288.2880)	(301.5834)
100	0.75	7.0937	5.6175	**7.0478**	***5.5971***	7.0703	5.6130	7.0715	5.6132
		(739.5130)	(820.3079)	(**736.9126**)	(***818.0457***)	(738.1904)	(819.8245)	(738.2596)	(819.8499)
	0.85	7.0513	5.1282	**6.9494**	***5.0851***	6.9870	5.1121	6.9887	5.1125
		(411.7009)	(524.3464)	(**408.7591**)	(***521.8227***)	(409.8478)	(523.4875)	(409.8992)	(523.5144)
	0.95	7.1160	4.5156	6.9547	4.4559	**6.8048**	***4.4261***	6.9874	4.4790
		(186.6831)	(342.1177)	(185.3422)	(341.0251)	(**184.0671**)	(***340.5680***)	(185.6145)	(341.4951)
	0.99	7.9597	3.8406	7.8607	3.8160	**6.0521**	***3.4106***	7.8678	3.8202
		(116.7167)	(291.5884)	(116.5569)	(291.4731)	(**113.6337**)	(***289.7470***)	(116.5686)	(291.4975)
200	0.75	15.3255	15.0654	**15.3240**	***15.0646***	15.3251	15.0652	15.3251	15.0652
		(29484.3400)	(29691.9000)	(**29481.4600**)	(29688.6100)	(29483.4600)	(29690.9500)	(29483.5000)	(29690.9900)
	0.85	13.8748	13.1006	**13.8684**	***13.0969***	13.8709	13.0988	13.8718	13.0992
		(6935.4980)	(7248.8050)	(**6933.0860**)	(***7246.1130***)	(6933.9860)	(7247.2070)	(6934.3680)	(7247.6550)
	0.95	12.1200	10.6186	12.1163	10.6162	**12.0746**	***10.5916***	12.1170	10.6168
		(1899.8620)	(2217.2760)	(1899.6340)	(2217.0270)	(**1897.0710**)	(***2214.3340***)	(1899.6770)	(2217.0820)
	0.99	10.1021	8.2615	10.0826	8.2504	**9.8171**	***8.1054***	10.0837	8.2513
		(1190.0330)	(1530.2470)	(1189.8210)	(1530.0170)	(**1186.9980**)	(1527.0210)	(1189.8330)	(1530.0340)
500	0.75	4.4694	3.4809	**4.4652**	***3.4788***	4.4677	3.4807	4.4678	3.4808
		(34007.0400)	(34127.8000)	(**34005.2800**)	(***34126.3900***)	(34006.3200)	(34127.4400)	(34006.3600)	(34127.7800)
	0.85	4.3436	3.3440	**4.3375**	***3.3409***	4.3407	3.3434	4.3408	3.3435
		(35405.5600)	(35516.6500)	(**35403.9000**)	(***35515.7600***)	(35404.7600)	(35516.9500)	(35404.8100)	(35516.9300)
	0.95	4.0312	3.2421	**4.0182**	***3.2331***	4.0218	3.2363	4.0222	3.2366
		(37245.7100)	(37305.1100)	(**37244.3400**)	(***37304.3400***)	(37244.7100)	(37304.8700)	(37244.7600)	(37304.8800)
	0.99	3.7775	3.3016	3.7726	3.2976	**3.7215**	***3.2554***	3.7732	3.2981
		(38440.0600)	(38447.4000)	(38440.0000)	(38447.3300)	(**38439.3300**)	(***38446.6300***)	(38440.0000)	(38447.3400)
900	0.75	8.2299	6.9022	**8.2271**	***6.9010***	8.2290	6.9021	8.2290	6.9021
		(10350.3100)	(11034.9700)	(**10347.8600**)	(***11032.7600***)	(10349.5500)	(11034.8300)	(10349.5800)	(11034.8400)
	0.85	7.8131	5.8105	**7.8060**	***5.8076***	7.8092	5.8098	7.8097	5.8099
		(5436.0560)	(6329.2670)	(**5433.4390**)	(***6327.0160***)	(5434.6080)	(6328.8440)	(5434.8030)	(6328.8980)
	0.95	7.2957	4.5934	7.2934	4.5926	**7.2696**	***4.5865***	7.2939	4.5929
		(2627.4840)	(4031.8910)	(2627.2560)	(4031.7120)	(**2624.9630**)	(***4030.5980***)	(2627.3080)	(4031.8000)
	0.99	6.8114	3.6589	6.8007	3.6562	**6.6652**	***3.6322***	6.8023	3.6572
		(1819.6790)	(3531.3490)	(1819.4850)	(3531.2370)	(**1817.0280**)	(***3530.4450***)	(1819.5140)	(3531.2930)

Table A20. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained via GA by minimizing MAE for Gamma response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	11.9951	7.4501	11.8197	7.4261	**10.5496**	***7.1942***	11.8640	7.4408
		(170.78912)	(220.0739)	(170.2459)	(219.4009)	(**156.8277**)	(***211.7920***)	(170.5239)	(219.7781)
	0.85	11.7867	6.9730	11.5096	6.9392	**9.8963**	***6.7365***	11.5822	6.9614
		(97.2824)	(150.1952)	(96.8382)	(149.6414)	(**88.9635**)	(***146.3421***)	(97.1033)	(150.0332)
	0.95	11.4882	6.5803	**5.9916**	***5.2467***	6.9799	5.7450	7.0933	5.7804
		(48.8360)	(103.4048)	(**43.0970**)	(***97.7264***)	(44.2691)	(100.6123)	(44.4193)	(100.7102)
	0.99	13.7588	7.6452	5.1638	5.6733	**3.2988**	***3.8914***	5.4052	5.7065
		(35.2670)	(89.9996)	(34.2215)	(90.8517)	(**33.7595**)	(***89.6944***)	(34.2691)	(89.8609)
50	0.75	12.6408	9.7006	**10.8833**	***8.8806***	11.6336	9.3471	11.7178	9.3789
		(399.6054)	(494.4707)	(**382.6248**)	(***475.6856***)	(389.8053)	(486.7218)	(390.6555)	(487.4250)
	0.85	14.3981	10.2043	14.0540	10.0962	**11.9213**	***9.3449***	14.1671	10.1452
		(320.8657)	(432.2142)	(319.2982)	(430.5336)	(**308.3631**)	(***419.3132***)	(319.8132)	(431.3578)
	0.95	15.4411	13.0724	14.7283	12.5798	**11.3595**	***9.7329***	14.9412	12.6834
		(273.5242)	(394.1230)	(269.6394)	(391.2938)	(**249.6900**)	(***375.0438***)	(269.9796)	(391.9328)
	0.99	28.4388	26.5725	23.1542	21.9593	**8.5387**	***9.4085***	24.0050	22.4508
		(258.8695)	(390.6796)	(253.9304)	(385.3070)	(**239.1211**)	(***372.3432***)	(254.1712)	(385.8869)
100	0.75	8.0427	4.2324	**7.5155**	***4.1713***	7.7734	4.2513	7.7864	4.2504
		(591.7564)	(776.8677)	(**583.4795**)	(***772.5796***)	(587.5768)	(779.1395)	(587.7812)	(779.0218)
	0.85	8.1511	3.9518	**7.3043**	***3.8640***	7.6121	3.9602	7.6268	3.9596
		(371.2352)	(588.0413)	(**363.3732**)	(***584.4928***)	(366.2885)	(589.3830)	(366.4255)	(589.3172)
	0.95	8.4883	3.4002	7.4483	***3.3260***	**6.6850**	3.3591	7.6506	3.3822
		(202.3108)	(454.0343)	(199.3602)	(***453.1912***)	(**197.9517**)	(454.4675)	(199.9698)	(454.2059)
	0.99	9.6333	2.4887	8.8826	2.4689	**4.4195**	***2.3871***	8.9366	2.4765
		(145.5359)	(416.3610)	(145.4399)	(414.0530)	(**143.1509**)	(***412.3344***)	(145.4582)	(415.3812)
200	0.75	8.0501	6.1378	**7.9564**	***6.1119***	8.0214	6.1427	8.0226	6.1425
		(709.7314)	(1090.2871)	**(700.0537**)	(***1083.0378***)	(706.7757)	(1091.3735)	(706.8947)	(1091.3297)
	0.85	7.9411	5.9158	**7.8076**	***5.8814***	7.8590	5.9117	7.8785	5.9135
		(379.3958)	(780.7428)	(**372.8682**)	(***775.7747***)	(375.2798)	(779.9392)	(376.3248)	(780.2560)
	0.95	7.4528	5.5175	7.4181	5.5092	**7.0607**	***5.4387***	7.4247	5.5120
		(151.7326)	(576.1429)	(151.2655)	(575.7876)	(**146.4607)**	(***572.6923***)	(151.3538)	(575.9057)
	0.99	6.3148	4.9230	6.1832	4.8930	**4.9465**	***4.5736***	6.1908	4.8961
		(87.1924)	(523.6791)	(86.8381)	(523.4242)	(**83.5672**)	(***520.7775***)	(86.8582)	(523.4487)
500	0.75	9.4341	4.4162	**9.3408**	***4.4071***	9.4216	4.4218	9.4242	4.4215
		(34961.0300)	(36616.4200)	(**34940.9900**)	(***36607.9600***)	(34949.7000)	(36624.0900)	(34950.0400)	(36623.6900)
	0.85	8.8876	4.2063	**8.8723**	***4.1939***	9.0854	4.2099	9.0633	4.2097
		(36404.3300)	(38068.0800)	(**36371.9200**)	(***38060.6300***)	(36372.0600)	(38073.9800)	(36365.8100)	(38073.6700)
	0.95	8.4585	4.2516	**8.1546**	***4.2124***	8.2565	4.2326	8.2660	4.2336
		(38177.4000)	(39843.4800)	(**38161.7100**)	(***39835.9000***)	(38162.5800)	(39843.2100)	(38163.3700)	(39843.2200)
	0.99	8.8051	5.1927	8.6581	5.1648	**7.6034**	***4.8972***	8.6757	5.1682
		(39218.4600)	(40893.7000)	(39217.7100)	(40892.7100)	(**39205.9600**)	(***40885.4600***)	(39217.6900)	(40892.9900)
900	0.75	9.3637	4.8125	**9.3033**	***4.8056***	9.3451	4.8170	9.3458	4.8168
		(7154.2830)	(9545.1920)	(**7142.4320**)	(***9538.4840***)	(7150.6320)	(9551.6210)	(7150.7710)	(9551.3740)
	0.85	9.1572	4.3861	**9.0696**	***4.3774***	9.1087	4.3895	9.1152	4.3890
		(4683.2110)	(7412.3060)	(**4673.7450**)	(***7407.7060***)	(4677.9810)	(7416.2730)	(4678.6730)	(7415.7410)
	0.95	8.7716	3.8134	8.7517	3.8119	**8.5579**	***3.8072***	8.7562	3.8129
		(2840.7450)	(5958.7490)	(2840.0560)	(5959.1460)	(**2833.3550**)	(***5958.5240***)	(2840.2140)	(5958.7760)
	0.99	8.3226	3.0336	8.2420	3.0308	**7.3833**	***3.0256***	8.2541	3.0331
		(2236.5750)	(5522.2620)	(2236.0270)	(5523.4220)	(**2229.9650**)	(***5522.2280***)	(2236.1100)	(5522.3310)

Table A21. EMSE and PMSE values of the FOA estimators when (k,d) values are obtained by MSE Gamma response variable

n	γ^2	$\hat{\beta}^{(1)}$	$\hat{\beta}_R^{(1)}$	$\hat{\beta}(k)^{(1)}$	$\hat{\beta}_R(k)^{(1)}$	$\hat{\beta}(d)^{(1)}$	$\hat{\beta}_R(d)^{(1)}$	$\hat{\beta}_{kd}^{(1)}$	$\hat{\beta}_{kd-R}^{(1)}$
25	0.75	**5.8872**	9.1020	6.0079	9.1320	5.8909	9.0511	5.8910	***9.0453***
		$(1.4431 \times 10^{+9})$	$(1.4613 \times 10^{+9})$	$(\mathbf{1.4430} \times 10^{+9})$	$(\boldsymbol{1.4606} \times 10^{+9})$	$(1.4431 \times 10^{+9})$	$(1.4612 \times 10^{+9})$	$(1.4431 \times 10^{+9})$	$(\boldsymbol{1.4606} \times 10^{+9})$
	0.85	7.0211	7.9241	**7.0181**	7.9478	7.0199	7.9202	7.0207	***7.9187***
		$(1.4806 \times 10^{+7})$	$(1.4827 \times 10^{+7})$	$(\mathbf{1.4804} \times 10^{+7})$	$(\boldsymbol{1.4825} \times 10^{+7})$	$(1.4806 \times 10^{+7})$	$(1.4827 \times 10^{+7})$	$(1.4806 \times 10^{+7})$	$(1.4827 \times 10^{+7})$
	0.95	**4.5365**	7.1184	5.0251	7.3829	4.5415	6.8053	4.5366	***6.7782***
		$(8.8714 \times 10^{+3})$	$(9.0753 \times 10^{+3})$	$(\mathbf{8.6529} \times 10^{+3})$	$(\boldsymbol{8.6463} \times 10^{+3})$	$(8.8579 \times 10^{+3})$	$(8.9968 \times 10^{+3})$	$(8.8404 \times 10^{+3})$	$(8.8775 \times 10^{+3})$
	0.99	161.5864	158.2638	**123.3115**	***121.7599***	155.1821	152.7774	151.5802	148.5528
		$(2.6171 \times 10^{+14})$	$(2.6170 \times 10^{+14})$	$(2.6171 \times 10^{+14})$	$(2.6170 \times 10^{+14})$	$(2.6171 \times 10^{+14})$	$(2.6170 \times 10^{+14})$	$(2.6171 \times 10^{+14})$	$(2.6170 \times 10^{+14})$
50	0.75	**3.3573**	4.4807	3.5118	4.5867	3.3598	4.4592	3.3604	***4.4558***
		$(5.7189 \times 10^{+5})$	$(6.8372 \times 10^{+5})$	$(\mathbf{5.7153} \times 10^{+5})$	$(\boldsymbol{6.8309} \times 10^{+5})$	$(5.7188 \times 10^{+5})$	$(6.8340 \times 10^{+5})$	$(5.7188 \times 10^{+5})$	$(6.8345 \times 10^{+5})$
	0.85	**3.5410**	5.1554	3.8382	5.3172	3.5457	5.0586	3.5465	***5.0516***
		$(1.9187 \times 10^{+5})$	$(2.1423 \times 10^{+5})$	$(\mathbf{1.9166} \times 10^{+5})$	$(\boldsymbol{2.1398} \times 10^{+5})$	$(1.9187 \times 10^{+5})$	$(2.1417 \times 10^{+5})$	$(1.9187 \times 10^{+5})$	$(2.1416 \times 10^{+5})$
	0.95	**3.7155**	4.3607	4.0482	4.6687	3.7313	4.3608	3.7280	***4.3470***
		$(1.9453 \times 10^{+3})$	$(1.9792 \times 10^{+3})$	$(\mathbf{1.8553} \times 10^{+3})$	$(\boldsymbol{1.7869} \times 10^{+3})$	$(1.9425 \times 10^{+3})$	$(1.9656 \times 10^{+3})$	$(1.9350 \times 10^{+3})$	$(1.8798 \times 10^{+3})$
	0.99	7.7777	6.2367	**5.1756**	***4.6975***	7.4224	6.1814	7.4660	6.2044
		$(\mathbf{9.8925} \times 10^{+4})$	$(9.8590 \times 10^{+4})$	$(9.8992 \times 10^{+4})$	$(9.8622 \times 10^{+4})$	$(9.8931 \times 10^{+4})$	$(\boldsymbol{9.8511} \times 10^{+4})$	$(9.8931 \times 10^{+4})$	$(9.8517 \times 10^{+4})$
100	0.75	**3.7985**	5.1213	3.8447	5.1350	3.7985	5.1196	3.7986	***5.1140***
		$(1.0051 \times 10^{+8})$	$(1.1330 \times 10^{+8})$	$(\mathbf{1.0049} \times 10^{+8})$	$(\boldsymbol{1.1327} \times 10^{+8})$	$(1.0051 \times 10^{+8})$	$(1.1329 \times 10^{+8})$	$(1.0051 \times 10^{+8})$	$(1.1329 \times 10^{+8})$
	0.85	**4.3324**	13.8063	4.4622	13.7833	4.3333	13.7773	4.3346	***13.7454***
		$(1.1541 \times 10^{+19})$	$(1.1560 \times 10^{+19})$	$(1.1541 \times 10^{+19})$	$(1.1560 \times 10^{+19})$	$(1.1541 \times 10^{+19})$	$(1.1560 \times 10^{+19})$	$(1.1541 \times 10^{+19})$	$(1.1560 \times 10^{+19})$
	0.95	4.3610	8.6800	4.3767	8.6384	**4.3568**	8.5920	4.3581	***8.5777***
		$(4.4036 \times 10^{+10})$	$(4.5083 \times 10^{+10})$	$(4.4036 \times 10^{+10})$	$(4.5082 \times 10^{+10})$	$(4.4036 \times 10^{+10})$	$(4.5083 \times 10^{+10})$	$(4.4036 \times 10^{+10})$	$(4.5083 \times 10^{+10})$
	0.99	18.6909	19.6411	**16.2514**	***16.7409***	18.9547	19.4995	18.9889	19.4232
		$(0.8541 \times 10^{+7})$	$(\boldsymbol{0.0554} \times 10^{+7})$	$(1.2399 \times 10^{+7})$	$(0.4639 \times 10^{+7})$	$(0.4358 \times 10^{+7})$	$(1.006 \times 10^{+7})$	$(\mathbf{0.0354} \times 10^{+7})$	$(0.0718 \times 10^{+7})$
200	0.75	**4.8868**	4.7253	4.9786	4.7777	4.8869	4.7249	4.8872	***4.7220***
		$(1.9367 \times 10^{+9})$	$(2.1464 \times 10^{+9})$	$(\mathbf{1.9366} \times 10^{+9})$	$(2.1463 \times 10^{+9})$	$(1.9367 \times 10^{+9})$	$(2.1463 \times 10^{+9})$	$(1.9367 \times 10^{+9})$	$(2.1463 \times 10^{+9})$
	0.85	**1.3576**	2.3446	1.4296	2.4010	1.3577	2.3344	1.3578	***2.3319***
		$(1.6870 \times 10^{+4})$	$(1.7398 \times 10^{+4})$	$(\mathbf{1.6850} \times 10^{+4})$	$(\boldsymbol{1.7259} \times 10^{+4})$	$(1.6870 \times 10^{+4})$	$(1.7401 \times 10^{+4})$	$(1.6869 \times 10^{+4})$	$(1.7403 \times 10^{+4})$
	0.95	**2.4123**	3.6823	2.5737	3.8542	2.4138	3.6410	2.4140	***3.6404***
		$(1.4977 \times 10^{+4})$	$(1.6082 \times 10^{+4})$	$(\mathbf{1.4971} \times 10^{+4})$	$(\boldsymbol{1.6062} \times 10^{+4})$	$(1.4977 \times 10^{+4})$	$(1.6086 \times 10^{+4})$	$(1.4977 \times 10^{+4})$	$(1.6086 \times 10^{+4})$
	0.99	25.9028	18.1283	**24.4529**	***17.3559***	25.8922	18.2227	25.8838	18.2654
		$(6.3791 \times 10^{+5})$	$(5.5117 \times 10^{+5})$	$(\mathbf{6.3362} \times 10^{+5})$	$(\boldsymbol{5.4876} \times 10^{+5})$	$(6.3788 \times 10^{+5})$	$(5.5142 \times 10^{+5})$	$(6.3789 \times 10^{+5})$	$(5.5150 \times 10^{+5})$
500	0.75	**7.8521**	8.7036	7.8627	***8.6995***	7.8521	8.7037	**7.8521**	8.7052
		$(1.3662 \times 10^{+20})$	$(1.3704 \times 10^{+20})$	$(1.3662 \times 10^{+20})$	$(1.3704 \times 10^{+20})$	$(1.3662 \times 10^{+20})$	$(1.3704 \times 10^{+20})$	$(1.3662 \times 10^{+20})$	$(1.3704 \times 10^{+20})$
	0.85	74.9174	70.3490	**74.9147**	***70.3481***	74.9174	70.3490	74.9174	70.3491
		$(9.7285 \times 10^{+17})$	$(9.4299 \times 10^{+17})$	$(9.7285 \times 10^{+17})$	$(9.4299 \times 10^{+17})$	$(9.7285 \times 10^{+17})$	$(9.4299 \times 10^{+17})$	$(9.7285 \times 10^{+17})$	$(9.4299 \times 10^{+17})$
	0.95	14.6901	11.6463	**14.5741**	***11.5863***	14.6900	11.6488	14.6898	11.6538
		$(1.1021 \times 10^{+7})$	$(0.9819 \times 10^{+7})$	$(\mathbf{1.0997} \times 10^{+7})$	$(\boldsymbol{0.9804} \times 10^{+7})$	$(1.1021 \times 10^{+7})$	$(0.9819 \times 10^{+7})$	$(1.1021 \times 10^{+7})$	$(0.9820 \times 10^{+7})$
	0.99	7.5614	4.3834	**6.9533**	***4.1579***	7.5589	4.4241	7.5560	4.4439
		$(1.1402 \times 10^{+6})$	$(1.1326 \times 10^{+6})$	$(\mathbf{1.1398} \times 10^{+6})$	$(\boldsymbol{1.1323} \times 10^{+6})$	$(1.1402 \times 10^{+6})$	$(1.1326 \times 10^{+6})$	$(1.1402 \times 10^{+6})$	$(1.1326 \times 10^{+6})$
900	0.75	**4.1144**	4.5120	4.1241	***4.5087***	**4.1144**	4.5120	**4.1144**	4.5118
		$(2.3396 \times 10^{+10})$	$(3.1952 \times 10^{+10})$	$(2.3396 \times 10^{+10})$	$(\boldsymbol{3.1951} \times 10^{+10})$	$(2.3396 \times 10^{+10})$	$(3.1952 \times 10^{+10})$	$(2.3396 \times 10^{+10})$	$(3.1952 \times 10^{+10})$
	0.85	47.3313	45.4758	**47.3300**	***45.4754***	47.3313	45.4758	47.3313	45.4759
		$(2.6931 \times 10^{+18})$	$(2.6796 \times 10^{+18})$	$(2.6931 \times 10^{+18})$	$(2.6796 \times 10^{+18})$	$(2.6931 \times 10^{+18})$	$(2.6796 \times 10^{+18})$	$(2.6931 \times 10^{+18})$	$(2.6796 \times 10^{+18})$
	0.95	**4.9850**	3.9441	4.9986	***3.9335***	**4.9850**	3.9436	4.9851	3.9416
		$(4.8786 \times 10^{+10})$	$(4.8800 \times 10^{+10})$	$(4.8786 \times 10^{+10})$	$(4.8800 \times 10^{+10})$	$(4.8786 \times 10^{+10})$	$(4.8800 \times 10^{+10})$	$(4.8786 \times 10^{+10})$	$(4.8800 \times 10^{+10})$
	0.99	49.7921	56.5380	**49.6807**	***56.4264***	49.7918	56.5374	49.7909	56.5351
		$(9.1127 \times 10^{+10})$	$(8.3462 \times 10^{+10})$	$(\mathbf{9.1096} \times 10^{+10})$	$(\boldsymbol{8.3426} \times 10^{+10})$	$(9.1127 \times 10^{+10})$	$(\boldsymbol{8.3426} \times 10^{+10})$	$(9.1126 \times 10^{+10})$	$(8.3459 \times 10^{+10})$

Table A22. EMSE and PMSE values of the iterative estimators when (k,d) values are obtained by MSE for Gamma response variable

n	γ^2	$\hat{\beta}^{(t+1)}$	$\hat{\beta}_R^{(t+1)}$	$\hat{\beta}(k)^{(t+1)}$	$\hat{\beta}_R(k)^{(t+1)}$	$\hat{\beta}(d)^{(t+1)}$	$\hat{\beta}_R(d)^{(t+1)}$	$\hat{\beta}_{kd}^{(t+1)}$	$\hat{\beta}_{kd-R}^{(t+1)}$
25	0.75	63.1681	49.5074	**38.2752**	***35.4480***	65.8099	56.4268	65.4425	54.2879
		$(0.0647\times10^{+6})$	$(0.3731\times10^{+6})$	$(1.4113\times10^{+6})$	$(0.4159\times10^{+6})$	$(0.0865\times10^{+6})$	(***0.1821***$\times10^{+6}$)	(**0.0622**$\times10^{+6}$)	$(0.2577\times10^{+6})$
	0.85	145.5952	92.8090	**61.7773**	***53.3469***	147.7372	127.8379	149.5842	116.8219
		(**0.0082**$\times10^{+6}$)	$(0.2627\times10^{+6})$	$(3.9933\times10^{+6})$	(***0.0230***$\times10^{+6}$)	$(1.2959\times10^{+6})$	$(0.2304\times10^{+6})$	$(0.1307\times10^{+6})$	$(0.0252\times10^{+6})$
	0.95	119.7964	89.9383	**33.6588**	***31.3969***	110.7359	107.3187	113.4759	106.3352
		(**0.0865**$\times10^{+5}$)	$(0.9795\times10^{+5})$	$(0.1710\times10^{+5})$	$(0.7073\times10^{+5})$	$(1.2056\times10^{+5})$	$(0.1327\times10^{+5})$	$(0.5391\times10^{+5})$	(***0.1297***$\times10^{+5}$)
	0.99	23.4968	29.7050	**13.0373**	***14.1648***	21.8749	23.2003	21.7786	23.0914
		$(0.0012\times10^{+8})$	$(3.4666\times10^{+8})$	(**0.0005**$\times10^{+8}$)	(***0.0009***$\times10^{+8}$)	$(0.0019\times10^{+8})$	$(0.0459\times10^{+8})$	$(0.0022\times10^{+8})$	$(0.0010\times10^{+8})$
50	0.75	67.5411	47.9731	**47.4688**	***40.3775***	71.1201	52.7856	70.1636	51.7539
		$(0.0002\times10^{+9})$	$(0.0002\times10^{+9})$	(**0.0001**$\times10^{+9}$)	$(0.0003\times10^{+9})$	$(0.0010\times10^{+9})$	(***0.0001***$\times10^{+9}$)	$(2.8495\times10^{+9})$	$(0.0010\times10^{+9})$
	0.85	50.0900	37.3674	**35.5182**	***31.7747***	52.2464	40.1731	51.8200	40.0517
		$(0.5159\times10^{+6})$	$(0.2760\times10^{+6})$	(**0.0543**$\times10^{+6}$)	(***0.0737***$\times10^{+6}$)	$(1.3893\times10^{+6})$	$(0.4732\times10^{+6})$	$(0.3131\times10^{+6})$	$(1.4273\times10^{+6})$
	0.95	84.5640	63.9574	**41.1864**	***38.7477***	85.0312	75.6503	85.4580	72.8189
		$(4.2352\times10^{+5})$	(***0.4461***$\times10^{+5}$)	(**0.2809**$\times10^{+5}$)	$(0.7933\times10^{+5})$	$(0.8633\times10^{+5})$	$(0.4519\times10^{+5})$	$(1.0854\times10^{+5})$	$(2.3734\times10^{+5})$
	0.99	45.6728	44.5738	**20.4061**	***20.6414***	44.1190	42.7269	44.1204	42.6742
		$(1.0009\times10^{+5})$	$(1.0137\times10^{+5})$	(**0.9921**$\times10^{+5}$)	(***0.9922***$\times10^{+5}$)	$(1.0725\times10^{+5})$	$(1.0250\times10^{+5})$	$(1.1896\times10^{+5})$	$(6.5467\times10^{+5})$
100	0.75	25.5397	23.6203	**23.6064**	***22.7746***	25.6984	23.6845	25.8214	23.8040
		(**0.0122**$\times10^{+8}$)	$(0.0117\times10^{+8})$	$(0.0223\times10^{+8})$	$(0.0033\times10^{+8})$	$(0.1106\times10^{+8})$	(***0.0027***$\times10^{+8}$)	$(1.0035\times10^{+8})$	$(0.0701\times10^{+8})$
	0.85	52.6464	30.4325	**38.1625**	***27.6538***	53.7246	31.6934	53.6724	32.6544
		$(0.0175\times10^{+7})$	$(0.0212\times10^{+7})$	$(1.0780\times10^{+7})$	(***0.0119***$\times10^{+7}$)	$(0.0091\times10^{+7})$	$(0.0546\times10^{+7})$	(**0.0074**$\times10^{+7}$)	$(0.0295\times10^{+7})$
	0.95	61.6642	39.7624	**42.8064**	***33.5328***	62.5298	42.4186	62.1560	41.9015
		$(4.6546\times10^{+6})$	$(0.2353\times10^{+6})$	(**0.0644**$\times10^{+6}$)	$(0.4150\times10^{+6})$	$(0.1640\times10^{+6})$	(***0.0469***$\times10^{+6}$)	$(3.8538\times10^{+6})$	$(1.0394\times10^{+6})$
	0.99	18.6909	19.6411	**16.2514**	***16.7409***	18.9547	19.4995	$18.9889/0.0354\times10^{+7}$	19.4232
		$(0.8541\times10^{+7})$	(***0.0554***$\times10^{+7}$)	$(1.2399\times10^{+7})$	$(0.4639\times10^{+7})$	$(0.4358\times10^{+7})$	$(1.006\times10^{+7})$	(**0.0354**$\times10^{+7}$)	$(0.0718\times10^{+7})$
200	0.75	49.1105	34.9816	**41.3632**	***33.0203***	49.6659	35.1938	50.2062	35.9575
		$(1.2828\times10^{+6})$	$(1.1183\times10^{+6})$	$(0.1206\times10^{+6})$	(***0.1016***$\times10^{+6}$)	(**0.0746**$\times10^{+6}$)	$(1.2220\times10^{+6})$	$(0.5523\times10^{+6})$	$(0.3387\times10^{+6})$
	0.85	20.4463	21.7630	**20.0273**	***21.5049***	20.4780	21.7628	20.4998	21.7506
		$(1.1599\times10^{+6})$	$(5.8951\times10^{+6})$	(**0.9945**$\times10^{+6}$)	$(8.1605\times10^{+6})$	$(1.7677\times10^{+6})$	$(8.7474\times10^{+6})$	$(1.0193\times10^{+6})$	(***2.6409***$\times10^{+6}$)
	0.95	68.3371	44.9242	**53.5836**	***40.0949***	70.9368	47.4366	70.4452	47.2142
		$(0.4582\times10^{+6})$	$(0.7471\times10^{+6})$	$(0.2583\times10^{+6})$	(***0.1996***$\times10^{+6}$)	(**0.2562**$\times10^{+6}$)	$(4.6077\times10^{+6})$	$(0.3527\times10^{+6})$	$(7.0859\times10^{+6})$
	0.99	18.1706	19.7277	**16.8880**	***18.2702***	18.3830	19.5997	18.4269	19.5399
		$(0.0010\times10^{+11})$	(***0.0001***$\times10^{+11}$)	$(0.0002\times10^{+11})$	$(9.9022\times10^{+11})$	$(0.0006\times10^{+11})$	$(0.0002\times10^{+11})$	(**0.0001**$\times10^{+11}$)	(***0.0001***$\times10^{+11}$)
500	0.75	31.3465	26.2578	**29.4466**	***25.9218***	31.3948	26.2676	31.6368	26.3684
		$(5.3046\times10^{+8})$	$(0.0514\times10^{+8})$	$(1.0925\times10^{+8})$	(***0.0057***$\times10^{+8}$)	$(0.0077\times10^{+8})$	$(1.5536\times10^{+8})$	(**0.0026**$\times10^{+8}$)	$(2.5237\times10^{+8})$
	0.85	12.2646	14.9651	**12.2021**	***14.9026***	12.2654	14.9629	12.2743	14.9441
		$(1.7545\times10^{+8})$	$(0.6095\times10^{+8})$	(**0.4232**$\times10^{+8}$)	$(1.3387\times10^{+8})$	$(0.5032\times10^{+8})$	$(2.2025\times10^{+8})$	$(0.6045\times10^{+8})$	(***0.0738***$\times10^{+8}$)
	0.95	17.0357	17.8687	**16.8533**	***17.7012***	17.0467	17.8656	17.0644	17.8619
		$(0.0509\times10^{+9})$	$(0.2712\times10^{+9})$	$(0.0721\times10^{+9})$	$(9.3616\times10^{+9})$	(**0.0293**$\times10^{+9}$)	(***0.0183***$\times10^{+9}$)	$(0.0705\times10^{+9})$	$(0.1956\times10^{+9})$
	0.99	20.4700	21.5202	**19.6015**	***20.6032***	20.5888	21.4762	20.6227	21.4446
		$(0.3504\times10^{+9})$	$(0.0348\times10^{+9})$	(**0.0075**$\times10^{+9}$)	$(0.0431\times10^{+9})$	$(0.0365\times10^{+9})$	$(0.0309\times10^{+9})$	$(1.0252\times10^{+9})$	(***0.0242***$\times10^{+9}$)
900	0.75	27.2238	23.7809	**26.2580**	***23.6529***	27.2388	23.7836	27.3558	23.8091
		$(5.4911\times10^{+7})$	$(0.2792\times10^{+7})$	$(0.9659\times10^{+7})$	(***0.1390***$\times10^{+7}$)	$(0.4116\times10^{+7})$	$(0.6686\times10^{+7})$	(**0.2319**$\times10^{+7}$)	$(0.2073\times10^{+7})$
	0.85	13.1704	16.6888	**13.1359**	***16.6468***	13.1710	16.6858	13.1753	16.6702
		$(0.4980\times10^{+10})$	$(0.0063\times10^{+10})$	(**0.0020**$\times10^{+10}$)	(***0.0007***$\times10^{+10}$)	$(0.0287\times10^{+10})$	$(0.0025\times10^{+10})$	$(1.9146\times10^{+10})$	$(0.0054\times10^{+10})$
	0.95	48.4817	31.7265	**45.1920**	***30.8069***	48.6620	31.8120	48.9599	31.9958
		$(4.8736\times10^{+10})$	$(4.8714\times10^{+10})$	$(4.8965\times10^{+10})$	$(4.8714\times10^{+10})$	$(4.8708\times10^{+10})$	(***4.8707***$\times10^{+10}$)	(**4.8706**$\times10^{+10}$)	$(4.8722\times10^{+10})$
	0.99	13.7328	14.3618	**13.6321**	***14.2515***	13.7472	14.3616	13.7639	14.3726
		$(0.1223\times10^{+9})$	$(0.0545\times10^{+9})$	(**0.1617**$\times10^{+9}$)	$(5.4424\times10^{+9})$	$(0.1857\times10^{+9})$	(***0.0352***$\times10^{+9}$)	$(1.0209\times10^{+9})$	$(0.6664\times10^{+9})$