Hacettepe Journal of
Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 52 (4) (2023), 1046 – 1065
DOI : 10.15672/hujms.1122207

Research Article

James-Stein type estimators in beta regression model: simulation and application

Solmaz Seifollahi*, Hossein Bevrani

Department of Statistics, University of Tabriz, Tabriz, Iran

Abstract

Recently, the beta regression model has been used in several fields of science to model data in the form of rate or proportion. In this paper, some novel and improved methods to estimate parameters in the beta regression model are proposed. We consider a sub-space on the regression coefficients of the beta regression model and combine the unrestricted and restricted estimators then we present Stein-type and preliminary estimators. We develop the expressions for the proposed estimators' asymptotic biases and their quadratic risks. Numerical studies through Monte Carlo simulations are used to evaluate the performance of the proposed estimators in terms of their simulated relative efficiency. The results show that the proposed estimators outperform the unrestricted estimator when the restrictions hold. Finally, an empirical application is given to show how useful the proposed estimators are in the practical area.

Mathematics Subject Classification (2020). 62J05, 62J07

Keywords. Beta regression model, James-Stein type estimator, maximum likelihood estimator, preliminary test estimator, restricted estimator

1. Introduction

It is common knowledge that linear regression models can be used to represent a relationship between a response variable and a few predictors. However, for the bounded response variables, this approach is not appropriate. The beta regression model, proposed by [13], is a suitable model where the response variable is in the form of rates or proportions. Beta regression models have become increasingly popular in analyzing biological, chemical, or environmental data in the form of percentages and rates, such as poverty rates, migration rates, unemployment rates, the percentage of income spent on food, and the proportion of crude oil converted to gasoline.

The use and interpretation of the models depend on the quality of the parameters' estimation. If the model parameters are poorly calibrated, the model output will not be relevant to make predictions or data assimilation. The maximum likelihood method (MLE) is a well-known parameters estimation technique in regression models. Ferrari and Cribari-Neto [13] explained the MLE of parameters in beta regression models by considering the precision parameter as a constant value. Lately, some authors such as

*Corresponding Author.
Email addresses: s.seifollahi@tabrizu.ac.ir (S. Seifollahi), bevrani@tabrizu.ac.ir (H. Bevrani)
Received: 27.05.2022; Accepted: 24.01.2023

Ferrari and Pinheiro [14], Simas et al. [21] and Espinheira et al. [10, 11] tried to improve the MLE in beta regression models. In this paper, the MLE of the parameters is called the unrestricted estimator. Recently, the multicollinearity issue has been considered for beta regression models and for that many estimators have been introduced such as the ridge estimator [3, 18], modified ridge-type estimator [5], Liu estimator [17], Liu-type estimator [7], two-parameters estimator [1] and Dawoud-Kibria estimator [2].

In the data analysis process, the practitioners may have some information about the value of the parameters or any combination of parameters. This information can be obtained through model selection techniques such as the Akaike information criterion (AIC), corrected AIC [6], Bayesian information criterion (BIC), adjusted R^2 [19], or it can be obtained based on past data or expert opinions. The use of this prior information may increase the efficiency of the estimators, especially when the sample information is somewhat limited. Consider that the prior information can be written as follows:

$$H_0 : \boldsymbol{R\beta} = r, \tag{1.1}$$

in which $\boldsymbol{R}$ is a known $(q \times p)$ matrix of full rank $(rank(\boldsymbol{R}) = q)$ and r is a $(q \times 1)$ length-vector of constants. The primary aim of this study is to introduce the model's parameter estimator, $\boldsymbol{\beta}$, under the sub-space defined in (1.1). This estimator emerging from the model by considering restrictions in (1.1) is called the restricted estimator.

In practice, the prior information in (1.1) is uncertain. The doubt on this prior information can be removed by testing hypothesis $H_0 : \boldsymbol{R\beta} = r$ against the alternative $H_1 : \boldsymbol{R\beta} \neq r$. Accordingly, we will study the preliminary test estimator (PTE) which selects between unrestricted and restricted estimators based on the rejection or acceptance of H_0. The preliminary test estimator was first proposed by Bancroft [9] and then has been developed by several authors. In the sequel, our goal is to combine unrestricted and restricted estimators optimally. The shrinkage estimators, such as stein-type estimators, are well-known methods in this context. [20], [4], [8], and [15] are some of the most important sources on stein-type estimation and its application. Stein-type estimators shrink the unrestricted estimator toward the restricted estimator based on a function of sample test statistics for restrictions in (1.1). Therefore, we will define James-Stein and positive James-Stein estimators in the beta regression models, which outperform unrestricted estimators in most cases.

This paper unfolded as follows: In section 2, the parameters estimation problem in the beta regression models using the MLE is considered. Then, by considering restrictions in (1.1), the restricted estimator is proposed. In Section 3, the James-Stein and positive James-Stein estimators and also the preliminary test estimator are proposed in beta regression models. Then asymptotic characteristics of proposed estimators such as asymptotic bias and the quadratic risk functions will be obtained. Section 4 offers the numerical study of evaluating the performance of the proposed estimators using Monte Carlo simulation in the situation when (1.1) is valid and also it is considered to be false. Finally, the empirical data is analyzed in Section 5 and the conclusion is given in Section 6. The proof of the theoretical results is given in Appendix.

2. Beta regression model: unrestricted and restricted estimators

Let $y_1, y_2, \cdots, y_n$ be independent observations of response variable which follows a beta distribution with parameters $(\mu\tau, (1-\mu)\tau)$; e.g.

$$f(y_i|\mu,\tau) = \frac{\Gamma(\tau)}{\Gamma(\mu\tau)\Gamma((1-\mu)\tau)} y_i^{\mu\tau-1}(1-y_i)^{(1-\mu)\tau-1}, \quad 0 < \mu, y_i < 1, \tag{2.1}$$

where the precision parameter, $\tau > 0$, is considered constant over observations. Then the beta regression model is defined by

$$g(\mu_i) = x_i^T\boldsymbol{\beta} = \eta_i, \quad i = 1, 2, \cdots, n, \tag{2.2}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_p)^T \in \mathbb{R}^p$, $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T$ is the ith observation of the independent variables and the link function $g(.)$ is a continuous and double differentiable function from $(0, 1)$ into $\mathbb{R}$. Another form of the beta regression model with varing precision parameter in the literature was employed first by [22] and then it was used by [21]. However, in this paper, τ is considered a known and constant value.

Although, many link functions are considered for beta regression model, we prefer to use the logit link function, $g(\mu) = log(\frac{\mu}{1-\mu})$, so

$$\mu_i = \frac{\exp\{x_i^T \boldsymbol{\beta}\}}{1 + \exp\{x_i^T \boldsymbol{\beta}\}}, \qquad i = 1, 2, \cdots, n, \tag{2.3}$$

The log-likelihood function of the beta regression model is given by

$$\mathfrak{L}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \mathfrak{L}_i(\boldsymbol{\beta}), \tag{2.4}$$

where

$$\begin{aligned} \mathfrak{L}_i(\boldsymbol{\beta}) = \log\left(\Gamma(\tau)\right) - \log\left(\Gamma(\mu_i \tau)\right) - \log\left(\Gamma((1-\mu_i)\tau)\right) - log(\frac{y_i}{1-y_i}) \\ + \mu_i \tau \log(y_i) + (1-\mu_i)\tau \log(1-y_i). \end{aligned}$$

Since this function is nonlinear in $\boldsymbol{\beta}$, we use the iterative reweighted least-squares algorithm to derive the MLE of $\boldsymbol{\beta}$. Let $y^* = (y_1^*, y_2^*, \cdots, y_n^*)^T$, $\mu^* = (\mu_1^*, \mu_2^*, \cdots, \mu_n^*)$, $y_i^* = logit(y_i)$ and $\mu_i^* = \psi(\mu_i \tau) - \psi((1-\mu_i)\tau)$ such that $\psi(.)$ denotes the digamma function. Thus, in the beta regression model [3, 18], the MLE will be

$$\hat{\boldsymbol{\beta}}_{MLE} = (\boldsymbol{X}^T \hat{V} \boldsymbol{X})^{-1} \boldsymbol{X}^T \hat{V} Z, \tag{2.5}$$

where

$$\begin{aligned} \hat{V} &= diag(v_1, \cdots, v_n), \\ v_i &= \tau \Big\{ \psi'(\hat{\mu}_i \tau) - \psi'((1-\hat{\mu}_i)\tau) \Big\} \frac{1}{\{g'(\hat{\mu}_i)\}^2}, \\ \hat{Z} &= \boldsymbol{X}^T \hat{\boldsymbol{\beta}} + \hat{V}^{-1} \frac{(y^* - \mu^*)}{\{g'(\hat{\mu})\}^2}. \end{aligned}$$

The value of $\hat{V}$ and $\hat{Z}$ are evaluated at the final iteration. The MLE is called in this paper the unrestricted estimator and it is denoted by $\hat{\boldsymbol{\beta}}^{UN}$. Under regularity conditions [12],when n increases, we know that

$$\sqrt{n}(\hat{\boldsymbol{\beta}}^{UN} - \boldsymbol{\beta}) \xrightarrow{D} \mathcal{N}(\boldsymbol{0}, \frac{1}{\tau}(\boldsymbol{X}^T \hat{V} \boldsymbol{X})^{-1}), \tag{2.6}$$

where $\xrightarrow{D}$ denotes the convergence in distribution.

Now, we consider the restriction in (1.1) for beta regression model. The restricted estimation, $\hat{\boldsymbol{\beta}}^{RE}$, is obtained by maximizing the log-likelihood function in (2.4) under sub-space restriction $\boldsymbol{R\beta} - r = 0$. By using Lagrange multipliers, $\hat{\boldsymbol{\beta}}^{RE}$ is derived as follows:

$$\hat{\boldsymbol{\beta}}^{RE} = \hat{\boldsymbol{\beta}}^{UN} - \mathcal{C}_x \boldsymbol{R}^T \Big[\boldsymbol{R} \mathcal{C}_x \boldsymbol{R}^T \Big]^{-1} (\boldsymbol{R} \hat{\boldsymbol{\beta}}^{UN} - r), \tag{2.7}$$

where $\mathcal{C}_x = (\boldsymbol{X}^T \hat{V} \boldsymbol{X})^{-1}$.

It is possible that this sub-space may not be true, as mentioned before we can check it by testing hypothesis $H_0 : \boldsymbol{R\beta} = r$ against the alternative $H_1 : \boldsymbol{R\beta} \neq r$. Let define

$\mathfrak{L}(\hat{\boldsymbol{\beta}}^{UN})$ and $\mathfrak{L}(\hat{\boldsymbol{\beta}}^{RE})$,respectively, as the value of the log-likelihood function at $\hat{\boldsymbol{\beta}}^{UN}$ and $\hat{\boldsymbol{\beta}}^{RE}$, then the test statistics will be

$$\begin{aligned}\mathfrak{F}_n &= 2\Big[\mathfrak{L}(\hat{\boldsymbol{\beta}}^{UN}) - \mathfrak{L}(\hat{\boldsymbol{\beta}}^{RE})\Big] \\ &= \tau(\boldsymbol{R}\hat{\boldsymbol{\beta}}^{UN} - r)^T\Big[\boldsymbol{R}\mathfrak{C}_x\boldsymbol{R}^T\Big]^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}}^{UN} - r) + o_p(1). \end{aligned} \tag{2.8}$$

When n increases, $\mathfrak{F}_n$ follows a chi-square distribution with q degrees of freedom.

3. Shrinkage estimators in beta regression models

In the last few decades, shrinkage procedures, specifically Stein-type estimates, have gained a considerable amount of attention when it comes to estimating the parameters of a model. In these methods, the unrestricted and restricted estimators are combined optimally to dominate the unrestricted estimator. Therefore, in this section, the James-Stein and positive James-Stein estimators for the beta regression models are constructed, which are superior to MLEs. In the sequel, preliminary test estimators for the beta regression models are provided. Finally, asymptotic theories for the proposed estimators are established.

3.1. James-Stein type estimators

The James-Stein estimator (JSE) of $\boldsymbol{\beta}$, denoted by $\hat{\boldsymbol{\beta}}^{JSE}$, which shrinks the unrestricted estimator towards the restricted estimator, is defined in beta regression models as follows:

$$\hat{\boldsymbol{\beta}}^{JSE} = \hat{\boldsymbol{\beta}}^{RE} + \{1 - c\mathfrak{F}_n^{-1}\}(\hat{\boldsymbol{\beta}}^{UN} - \hat{\boldsymbol{\beta}}^{RE}), \quad c > 0, \tag{3.1}$$

which can be rewritten as

$$\hat{\boldsymbol{\beta}}^{JSE} = \hat{\boldsymbol{\beta}}^{UN} - c\mathfrak{F}_n^{-1}(\hat{\boldsymbol{\beta}}^{UN} - \hat{\boldsymbol{\beta}}^{RE}), \quad c > 0, \tag{3.2}$$

where $c \in [0, 2(q-2))$ and often is set as $c = q - 2$. But when $0 \leq \mathfrak{F}_n < c$, consequently $1 - c\mathfrak{F}_n^{-1} < 0$, the James-Stein estimator will suffer from the over-shrinkage problem. To avoid this issue, we define a truncated form of $\hat{\boldsymbol{\beta}}^{JSE}$, called the positive James-Stein estimator [4]. Formally, the positive James-Stein estimator (PJSE), $\hat{\boldsymbol{\beta}}^{PJSE}$ is defined as follows:

$$\hat{\boldsymbol{\beta}}^{PJSE} = \hat{\boldsymbol{\beta}}^{RE} + \{1 - c\mathfrak{F}_n^{-1}\}^+(\hat{\boldsymbol{\beta}}^{UN} - \hat{\boldsymbol{\beta}}^{RE}), \quad c > 0, \tag{3.3}$$

where $z^+ = max(0, z)$. This estimator also can be rewritten as

$$\hat{\boldsymbol{\beta}}^{PJSE} = \hat{\boldsymbol{\beta}}^{JSE} - \{1 - c\mathfrak{F}_n^{-1}\}(\hat{\boldsymbol{\beta}}^{UN} - \hat{\boldsymbol{\beta}}^{RE})I_{\mathfrak{F}_n < c}, \quad c > 0. \tag{3.4}$$

3.2. Preliminary test estimator

As mentioned before, when the restrictions on parameters are suspicious, it may be reasonable to construct a preliminary test estimator. This estimator chooses only one value of the unrestricted estimator or the restricted estimator based on the test statistic for testing hypothesis $H_0 : \boldsymbol{R\beta} = r$ against $H_1 : \boldsymbol{R\beta} \neq r$.

Let $\mathfrak{F}_{n,\alpha}$ be the $(1-\alpha)$th percentile of the chi-Squared distribution with q degrees of freedom under H_0. Therefore, in beta regression models, we define the preliminary test estimator (PTE) as

$$\hat{\boldsymbol{\beta}}^{PTE} = \hat{\boldsymbol{\beta}}^{UE} - (\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE})I_{(\mathfrak{F}_n \leq \mathfrak{F}_{n,\alpha})}. \tag{3.5}$$

3.3. Asymptotic properties of the proposed estimators

Sometimes the subspace on parameters based on restrictions , $\boldsymbol{R\beta} = r$ may be wrong. In this case, we consider $\boldsymbol{R\beta} = r + \upsilon$. For fixed $\upsilon \neq 0$, when the test statistic converges ∞, the proposed estimators will be asymptotically equivalent in probability to the unrestricted estimator. So, due to obtain meaningful asymptotic, we consider the following consequence of local alternatives:

$$H_{(0n)} : \boldsymbol{R\beta} = r + \frac{\xi}{\sqrt{n}}, \tag{3.6}$$

where $\xi = (\xi_1, \xi_2, \cdots, \xi_q)^T \in \mathbb{R}^q$. The vector $\xi/\sqrt{n}$ measures how much local alternatives differ from $\boldsymbol{R\beta} = r$. It is noteworthy that $H_0 : \boldsymbol{R\beta} = r$ is a special case of $H_{(0n)}$ since $\xi = \mathbf{0}$ implies $\boldsymbol{R\beta} = r$.

Now, we provide the asymptotic properties of the proposed estimators, such as bias and quadratic risk function. To do so, we define the asymptotic distribution function of the estimator $\hat{\boldsymbol{\beta}}$ under (3.6) as

$$F(w) = \lim_{n\to\infty} p(\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \leq w | H_{(0n)}), \tag{3.7}$$

where $F(w)$ is a non-degenerated distribution function and the letter p is used for probability operation. Thus, the asymptotic distributional bias (ADB) will be defined as

$$ADB(\hat{\boldsymbol{\beta}}) = \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right] = \int \cdots \int w dF(w). \tag{3.8}$$

In addition to the asymptotic distributional bias, the asymptotic distributional quadratic risk (ADR) is defined as follows:

$$ADR(\hat{\boldsymbol{\beta}}) = tr(MSE(\hat{\boldsymbol{\beta}})), \tag{3.9}$$

where

$$MSE(\hat{\boldsymbol{\beta}}) = \lim_{n\to\infty} \mathbb{E}\left[\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T\right] = \int \cdots \int ww' dF(w). \tag{3.10}$$

In the following theorems, we determined the ADB and ADR of the proposed estimators which provide the main results of this section.

Theorem 3.1. *Under regularity conditions [12] and the sequence of alternatives* (3.6), *the ADB of the proposed estimators is given by*

(1) $ADB(\hat{\boldsymbol{\beta}}^{RE}) = -J\xi$,
(2) $ADB(\hat{\boldsymbol{\beta}}^{JSE}) = -cJ\xi\mathbb{E}[\chi^{-2}_{q+2}(\lambda)]$,
(3) $ADB(\hat{\boldsymbol{\beta}}^{PJSE}) = ADB(\hat{\boldsymbol{\beta}}^{JS}) + J\xi H_{q+2}(c; \lambda)$,
(4) $ADB(\hat{\boldsymbol{\beta}}^{PTE}) = -J\xi H_{q+2}(\mathfrak{F}_{n,\alpha}; \lambda)$,

where $J = \mathfrak{C}_x \boldsymbol{R}^T \left[\boldsymbol{R}\mathfrak{C}_x\boldsymbol{R}^T\right]^{-1}$, $H_{q+2}(.; \lambda)$ *and* $\mathbb{E}[\chi^{2j}_{q+2}(\lambda)]$ *stand, respectively, for the cumulative distribution function and* j*th order moment of a non-central chi-squared distribution with* $q+2$ *degrees of freedom and non-centrality parameter* λ.

Proof. See Appendix B. □

For the ADR of proposed estimators, we have the following theorem.

Theorem 3.2. *Under regularity conditions [12] and the sequence of alternatives* (3.6), *the ADR of the proposed estimators are as follows:*

$$
\begin{aligned}
ADR(\hat{\boldsymbol{\beta}}^{UE}) &= \tau tr(\mathcal{C}_x),\\
ADR(\hat{\boldsymbol{\beta}}^{RE}) &= \tau tr(\mathcal{C}_x - J_0) + J^T\xi^T\xi J,\\
ADR(\hat{\boldsymbol{\beta}}^{JSE}) &= \tau tr(\mathcal{C}_x) + 2c\xi^T J^T J\xi\mathbb{E}\Big[\chi^{-2}_{q+2}(\lambda)\Big]\\
&\quad + \tau c tr(J_0)\mathbb{E}\Big[\chi^{-4}_{q+2}(\lambda)\Big] - 2c\tau tr(J_0)\mathbb{E}\Big[\chi^{-2}_{q+2}(\lambda)\Big]\\
&\quad + c^2\xi^T J^T J\xi\mathbb{E}\Big[\chi^{-4}_{q+4}(\lambda)\Big] - 2c\xi^T J^T J\xi\mathbb{E}\Big[\chi^{-2}_{q+4}(\lambda)\Big],\\
ADR(\hat{\boldsymbol{\beta}}^{PJSE}) &= ADR(\hat{\boldsymbol{\beta}}^{JSE}) - 2\xi^T J^T J\xi\Big[H_{q+2}(c;\lambda) - H_{q+4}(c;\lambda)\Big]\\
&\quad + 3\tau tr(J_0)H_{q+2}(c;\lambda) - 2c\tau tr(J_0)\mathbb{E}\Big[\chi^{-2}_{q+2}(\lambda)I_{(\chi^{-2}_{q+2}(\lambda)<c)}\Big]\\
&\quad - 2c\xi^T J^T J\xi\mathbb{E}\Big[\chi^{-2}_{q+4}(\lambda)I_{(\chi^{-2}_{q+4}(\lambda)<c)}\Big] + 2\xi^T J^T J\xi H_{q+4}(c;\lambda),\\
ADR(\hat{\boldsymbol{\beta}}^{PTE}) &= \tau tr(\mathcal{C}_x) + \xi^T J^T J\xi(H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) - H_{q+4}(\mathfrak{F}_{n,\alpha};\lambda))\\
&\quad - \tau tr(J_0)H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda).
\end{aligned}
$$

where $J_0 = J\boldsymbol{R}\mathcal{C}_x^{-1}$.

Proof. See Appendix C. □

Obviously, the expressions for ADB and ADR of estimators are not in closed form. So, in the following section, we conduct a simulation study to compare all proposed estimators.

4. Simulation study

In this section, we examine the performance of the proposed estimators in the beta regression model via a simulation study by the estimated relative efficiency criterion. We display various scenarios of the performance of the proposed estimators by considering different values of sample size, the number of independent variables, and the precision parameter, which is illustrated in Table 1.

Table 1. Assumed values for the simulation study.

Title	Symbol	Values
Number of independent variables	p	4, 6, 8, 12
Value of dispersion	τ	1, 10, 100
Sample size	n	50, 100, 200
Number of replicates	m	2000

The n observations of the independent variables are generated from the standard normal distribution. The n observations of the response variable in the beta regression model, with logit link function, are generated from the beta distribution with parameters (μ_i, τ), where

$$\mu_i = \log(\frac{e^{\eta_i}}{1+e^{\eta_i}}), \qquad i = 1, 2, \cdots, n \tag{4.1}$$

and

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}, \qquad i = 1, 2, \cdots, n. \tag{4.2}$$

We chose β_0 as zero and $(\beta_1, \beta_2, \cdots, \beta_p) = (1, 1, \cdots, 1)/\sqrt{p}$ such that $\boldsymbol{\beta}^T\boldsymbol{\beta} = 1$. For the restrictions, we consider the following restrictions.

$$R_{\left(\frac{p}{2}\times p\right)} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}, \quad \text{and} \quad r_{\left(\frac{p}{2}\times 1\right)} = \frac{2}{\sqrt{p}}(\kappa, 1, \cdots, 1)^T. \quad (4.3)$$

The value of κ allows us to control whether the restrictions are valid or not. When $\kappa = 0$, it indicates that the restrictions are valid, but when $\kappa \neq 0$, it indicates that the distance between simulated parameters and candidate values is $\kappa/\sqrt{p}$. We use this parameter, κ, to study the behavior of the proposed estimators in different situations. To do this, various values of κ in the interval $[0, 1]$ have been considered in this study.

We use the unrestricted estimator as a benchmark and report the simulated relative efficiency (SRE) as a ratio of the simulated MSE of the proposed estimator to that of the unrestricted estimator. Therefore, the SRE of an estimator $\hat{\boldsymbol{\beta}}$ is defined as

$$SRE(\hat{\boldsymbol{\beta}}) = \frac{SMSE(\hat{\boldsymbol{\beta}}^{UN})}{SMSE(\hat{\boldsymbol{\beta}})}, \quad (4.4)$$

in which the $SMSE$ of an estimator is calculated by

$$SMSE(\hat{\boldsymbol{\beta}}) = \frac{1}{m}\sum_{l=1}^{m}(\hat{\boldsymbol{\beta}}_l - \boldsymbol{\beta})^T(\hat{\boldsymbol{\beta}}_l - \boldsymbol{\beta}), \quad (4.5)$$

where m represents the total number of replications which is set to be 2000 here and $\hat{\boldsymbol{\beta}}_l$ is the estimated value of $\boldsymbol{\beta}$ in the lth replication. The results are given in Tables 2–5 and plotted for easier comparison in Figures 1–4.

Table 2. The SRE of the suggested estimators for $p = 4$.

		$n = 50$				$n = 100$				$n = 200$			
τ	κ	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE
1.0	0.0	3.0068	0.6521	2.2140	3.0068	2.6809	0.5384	2.1222	2.6809	2.6823	0.5261	2.1503	2.6661
	0.2	1.8831	0.7456	1.6476	1.8707	1.3305	0.6807	1.3605	1.3245	0.8408	0.8650	1.1305	0.8312
	0.4	0.9239	0.7998	1.1393	0.9107	0.5386	0.7698	0.9669	0.5350	0.2794	0.9928	0.9938	0.3374
	0.6	0.5034	0.9047	0.9641	0.5072	0.2711	0.9257	0.9335	0.3086	0.1326	1.0055	1.0055	0.6037
	0.8	0.3083	0.9335	0.9388	0.3621	0.1601	0.9574	0.9574	0.3377	0.0765	1.0070	1.0070	1.0000
	1.0	0.2060	0.9482	0.9488	0.3604	0.1049	0.9720	0.9720	0.6805	0.0495	1.0069	1.0069	1.0000
10	0.0	2.5271	1.2717	1.3969	2.0250	2.8485	1.3018	1.4857	2.3413	2.7453	1.2282	1.4724	2.3450
	0.2	0.9089	1.0784	1.1006	0.8655	0.6321	1.0500	1.0577	0.7307	0.3439	1.0156	1.0158	0.7438
	0.4	0.3183	1.0077	1.0078	0.8104	0.1913	1.0066	1.0066	0.9679	0.0954	1.0016	1.0016	1.0000
	0.6	0.1534	1.0004	1.0004	0.9852	0.0886	1.0007	1.0007	1.0000	0.0433	0.9997	0.9997	1.0000
	0.8	0.0890	0.9985	0.9985	1.0000	0.0506	0.9991	0.9991	1.0000	0.0245	0.9993	0.9993	1.0000
	1.0	0.0579	0.9980	0.9980	1.0000	0.0326	0.9986	0.9986	1.0000	0.0158	0.9992	0.9992	1.0000
100	0.0	3.8002	1.3740	1.4813	2.3998	2.9568	1.2787	1.4345	2.1693	2.5501	1.1938	1.3956	2.0205
	0.2	0.2205	1.0182	1.0182	0.9471	0.0820	1.0024	1.0024	1.0000	0.0411	1.0013	1.0013	1.0000
	0.4	0.0577	1.0037	1.0037	1.0000	0.0209	0.9998	0.9998	1.0000	0.0104	1.0001	1.0001	1.0000
	0.6	0.0258	1.0014	1.0014	1.0000	0.0093	0.9995	0.9995	1.0000	0.0046	0.9999	0.9999	1.0000
	0.8	0.0146	1.0006	1.0006	1.0000	0.0053	0.9995	0.9995	1.0000	0.0026	0.9999	0.9999	1.0000
	1.0	0.0093	1.0003	1.0003	1.0000	0.0034	0.9996	0.9996	1.0000	0.0017	0.9999	0.9999	1.0000

Table 3. The SRE of the suggested estimators for $p = 6$.

		$n=50$				$n=100$				$n=200$			
τ	κ	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE
1.0	0.0	6.1177	0.5425	4.9999	6.1177	3.7743	0.4324	3.5587	3.7743	3.5169	0.3080	3.3577	3.5169
	0.2	3.7678	0.5421	3.3783	3.7678	1.7912	0.5061	1.8214	1.7912	1.0877	0.6509	1.2453	1.0824
	0.4	1.7683	0.5376	1.8347	1.7420	0.7067	0.7026	0.9765	0.7034	0.3583	0.8886	0.9038	0.3880
	0.6	0.9401	0.5508	1.1520	0.9235	0.3527	0.8276	0.8518	0.3823	0.1695	0.9551	0.9551	0.5666
	0.8	0.5681	0.5887	0.8652	0.5643	0.2075	0.8878	0.8881	0.3955	0.0976	0.9786	0.9786	1.0000
	1.0	0.3767	0.6362	0.7536	0.3922	0.1358	0.9204	0.9204	0.7248	0.0632	0.9889	0.9889	1.0000
10	0.0	4.7898	1.8219	2.2389	2.8666	4.1973	1.8193	2.3179	3.1318	4.0200	1.7388	2.3294	3.0633
	0.2	1.4594	1.3800	1.4565	1.1225	0.8811	1.2074	1.2413	0.8471	0.4305	1.0866	1.0904	0.7640
	0.4	0.4727	1.1251	1.1261	0.8522	0.2626	1.0509	1.0509	0.9350	0.1174	1.0187	1.0187	1.0000
	0.6	0.2222	1.0561	1.0561	0.9904	0.1211	1.0199	1.0199	1.0000	0.0531	1.0074	1.0074	1.0000
	0.8	0.1276	1.0315	1.0315	1.0000	0.0691	1.0095	1.0095	1.0000	0.0301	1.0036	1.0036	1.0000
	1.0	0.0824	1.0201	1.0201	1.0000	0.0445	1.0050	1.0050	1.0000	0.0193	1.0020	1.0020	1.0000
100	0.0	3.7828	1.7035	1.9291	2.2662	3.7866	1.7546	2.0271	2.5128	4.0062	1.8365	2.1303	2.7340
	0.2	0.2541	1.0430	1.0430	0.9605	0.1121	1.0166	1.0166	1.0000	0.0557	1.0105	1.0105	1.0000
	0.4	0.0670	1.0084	1.0084	1.0000	0.0287	1.0035	1.0035	1.0000	0.0141	1.0022	1.0022	1.0000
	0.6	0.0301	1.0027	1.0027	1.0000	0.0128	1.0013	1.0013	1.0000	0.0063	1.0008	1.0008	1.0000
	0.8	0.0170	1.0010	1.0010	1.0000	0.0072	1.0006	1.0006	1.0000	0.0035	1.0003	1.0003	1.0000
	1.0	0.0109	1.0003	1.0003	1.0000	0.0046	1.0003	1.0003	1.0000	0.0023	1.0002	1.0002	1.0000

Table 4. The SRE of the suggested estimators for $p = 8$.

		$n=50$				$n=100$				$n=200$			
τ	κ	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE
1.0	0.0	6.5220	0.3862	6.0861	6.4719	5.1415	0.3391	5.0155	5.1415	4.6390	0.3410	4.5639	4.6390
	0.2	3.9289	0.4325	3.7847	3.9114	2.2841	0.4391	2.3078	2.2841	1.4080	0.4491	1.4732	1.4061
	0.4	1.8478	0.5126	1.9041	1.8393	0.8691	0.5947	1.0478	0.8621	0.4614	0.6258	0.7398	0.4645
	0.6	0.9869	0.5897	1.1568	0.9737	0.4286	0.7266	0.7967	0.4399	0.2180	0.7553	0.7565	0.3382
	0.8	0.5985	0.6608	0.8749	0.5969	0.2509	0.8123	0.8160	0.3604	0.1255	0.8341	0.8341	0.7809
	1.0	0.3978	0.7180	0.7874	0.4253	0.1638	0.8666	0.8666	0.5825	0.0812	0.8818	0.8818	1.0000
10	0.0	4.8467	2.1134	2.6142	2.7575	5.2926	2.3516	3.0272	3.8051	5.0183	2.2783	3.0580	3.6135
	0.2	1.7303	1.5189	1.5969	1.2328	1.0425	1.3876	1.4004	0.9510	0.5301	1.1687	1.1724	0.8044
	0.4	0.6002	1.1704	1.1732	0.9013	0.3083	1.1057	1.1057	0.9931	0.1443	1.0388	1.0388	1.0000
	0.6	0.2881	1.0722	1.0722	0.9942	0.1420	1.0447	1.0447	1.0000	0.0652	1.0145	1.0145	1.0000
	0.8	0.1669	1.0365	1.0365	1.0000	0.0809	1.0234	1.0234	1.0000	0.0369	1.0066	1.0066	1.0000
	1.0	0.1083	1.0204	1.0204	1.0000	0.0521	1.0139	1.0139	1.0000	0.0237	1.0032	1.0032	1.0000
100	0.0	5.5196	2.1561	2.4670	2.4435	4.5111	2.0914	2.5244	2.7516	4.4693	2.1550	2.5932	2.8744
	0.2	0.3285	1.1084	1.1084	0.9935	0.1353	1.0390	1.0390	1.0000	0.0640	1.0195	1.0195	1.0000
	0.4	0.0862	1.0280	1.0280	1.0000	0.0346	1.0080	1.0080	1.0000	0.0162	1.0048	1.0048	1.0000
	0.6	0.0387	1.0124	1.0124	1.0000	0.0155	1.0028	1.0028	1.0000	0.0072	1.0021	1.0021	1.0000
	0.8	0.0218	1.0069	1.0069	1.0000	0.0087	1.0011	1.0011	1.0000	0.0041	1.0012	1.0012	1.0000
	1.0	0.0140	1.0044	1.0044	1.0000	0.0056	1.0004	1.0004	1.0000	0.0026	1.0007	1.0007	1.0000

Table 5. The SRE of the suggested estimators for $p = 12$.

		$n=50$				$n=100$				$n=200$			
τ	κ	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE	RE	JSE	PJSE	PTE
1	0.0	12.1344	0.5219	10.9643	11.5686	8.8704	0.2872	8.7380	8.8704	7.9636	0.2736	7.9360	7.9636
	0.2	7.0338	0.6274	6.5875	6.7385	3.7438	0.3734	3.7287	3.7438	2.0386	0.4385	2.0709	2.0386
	0.4	3.1758	0.8617	3.1414	3.0267	1.3878	0.5362	1.4669	1.3749	0.6424	0.7001	0.8729	0.6407
	0.6	1.6652	1.0816	1.8257	1.5698	0.6788	0.6772	0.8966	0.6754	0.3008	0.8426	0.8451	0.4514
	0.8	1.0009	1.1685	1.3431	0.9494	0.3961	0.7702	0.7949	0.4478	0.1726	0.9072	0.9072	0.9194
	1.0	0.6620	1.1670	1.1935	0.6978	0.2580	0.8295	0.8297	0.4874	0.1116	0.9396	0.9396	1.0000
10	0.0	8.1865	3.1488	3.6097	2.9794	7.9188	3.1847	4.2565	4.0340	7.1899	3.2489	4.5858	4.9873
	0.2	2.7256	2.0558	2.1297	1.4155	1.4519	1.6841	1.7312	1.0837	0.7251	1.3616	1.3641	0.9180
	0.4	0.9202	1.3941	1.3951	0.9706	0.4221	1.2110	1.2110	0.9832	0.1964	1.0955	1.0955	1.0000
	0.6	0.4383	1.1837	1.1837	0.9956	0.1935	1.0959	1.0959	1.0000	0.0887	1.0417	1.0417	1.0000
	0.8	0.2531	1.1018	1.1018	1.0000	0.1101	1.0538	1.0538	1.0000	0.0502	1.0229	1.0229	1.0000
	1.0	0.1640	1.0631	1.0631	1.0000	0.0708	1.0342	1.0342	1.0000	0.0322	1.0142	1.0142	1.0000
100	0.0	8.5843	2.8356	2.9918	2.0743	8.8429	3.3162	3.9046	3.3997	7.3468	3.2494	3.9494	3.8172
	0.2	0.3770	1.1651	1.1651	0.9931	0.1993	1.0970	1.0970	1.0000	0.0915	1.0436	1.0436	1.0000
	0.4	0.0974	1.0398	1.0398	1.0000	0.0507	1.0239	1.0239	1.0000	0.0231	1.0105	1.0105	1.0000
	0.6	0.0436	1.0163	1.0163	1.0000	0.0226	1.0103	1.0103	1.0000	0.0103	1.0044	1.0044	1.0000
	0.8	0.0246	1.0084	1.0084	1.0000	0.0127	1.0056	1.0056	1.0000	0.0058	1.0024	1.0024	1.0000
	1.0	0.0157	1.0049	1.0049	1.0000	0.0082	1.0035	1.0035	1.0000	0.0037	1.0014	1.0014	1.0000

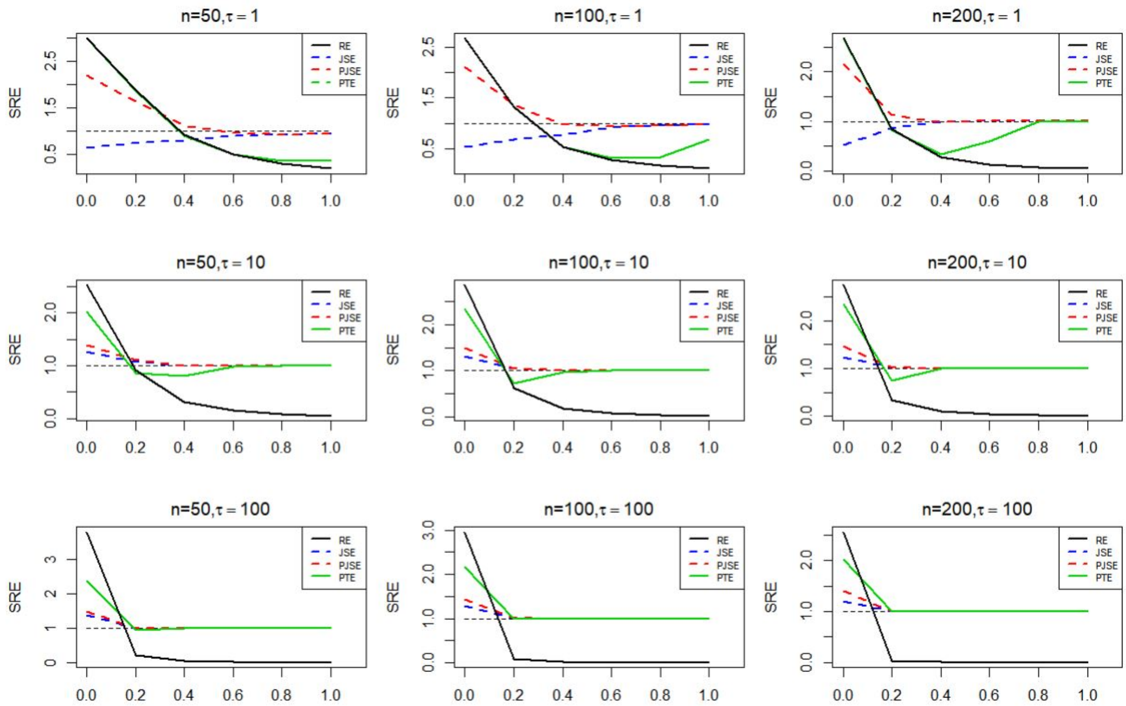

Figure 1. The SRE of proposed estimators for $p = 4$, $\tau = 1, 10, 100$ and $n = 50, 100, 200$.

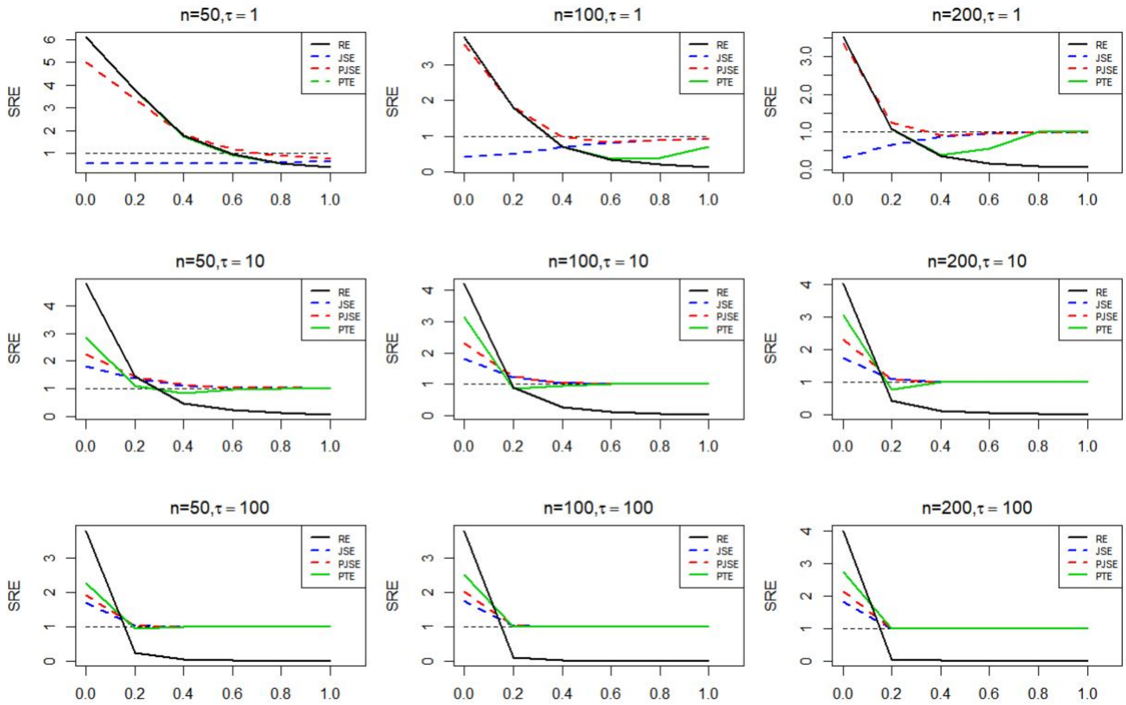

Figure 2. The SRE of proposed estimators for $p = 6$, $\tau = 1, 10, 100$ and $n = 50, 100, 200$.

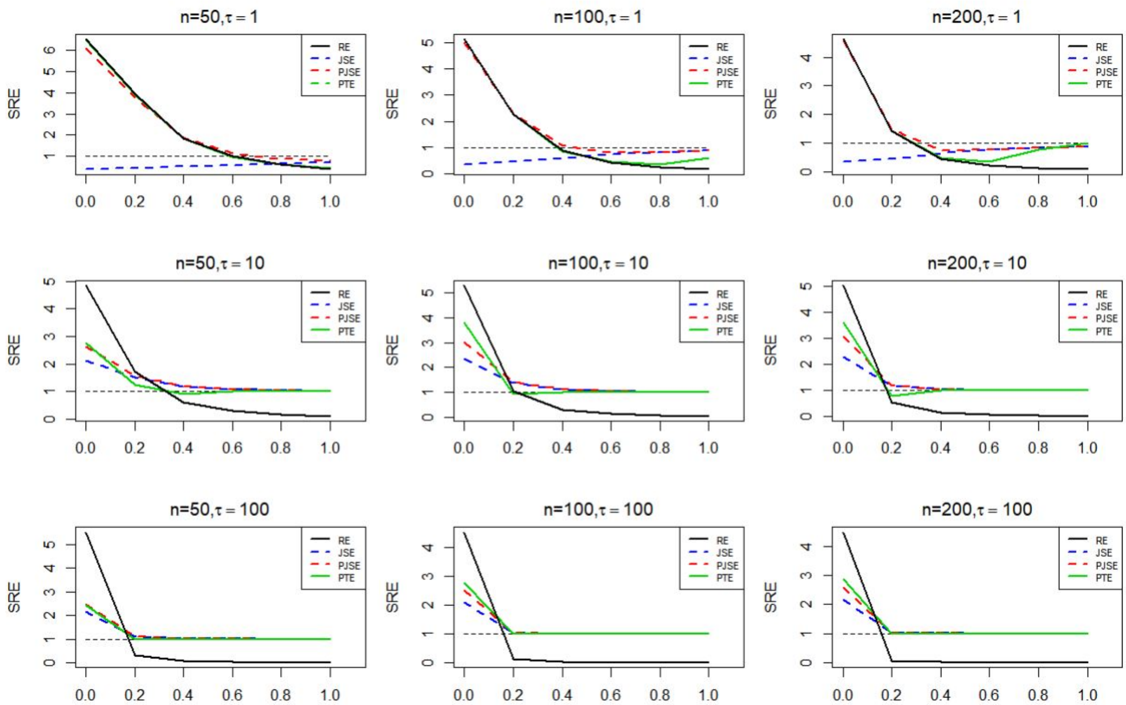

Figure 3. The SRE of proposed estimators for $p = 8$, $\tau = 1, 10, 100$ and $n = 50, 100, 200$.

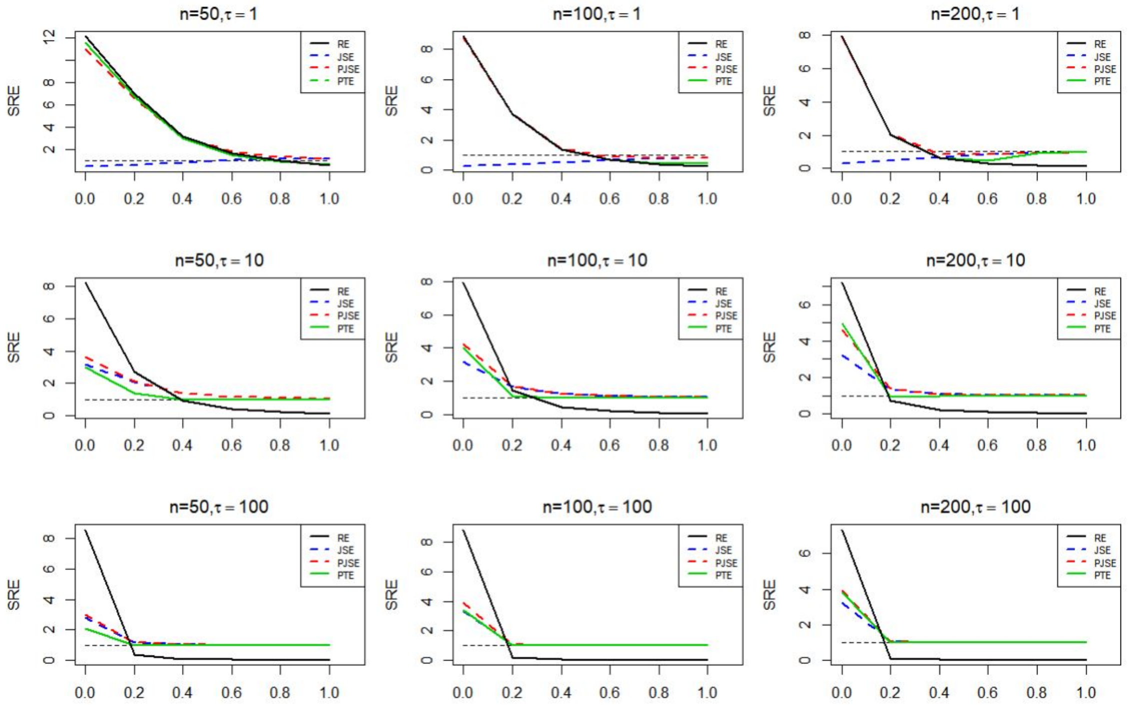

Figure 4. The SRE of proposed estimators for $p = 12$, $\tau = 1, 10, 100$ and $n = 50, 100, 200$.

The most important characteristics of the tables and the figures are summarized as follows:

When the restrictions are true:

(1) The simulated relative efficiency of the proposed estimators increases as the number of independent variables increases for all estimators.
(2) When the value of the precision parameter is small, the James-Stein estimator does not perform as well as the unrestricted estimator but the restricted estimator, positive James-Stein estimator and preliminary test estimators have better performance than the unrestricted estimator.
(3) In all the cases, the performance of the restricted estimator is superior to other estimators such that it has the utmost simulated relative efficiency value.

When the restrictions are not true:

(1) In all combinations of p, n, and τ, for the small value of κ, the restricted estimator dominates all estimators. However, when the value of κ increases, the performance of the restricted estimator decreases so that it gets worth more than other estimators.
(2) For all values of τ, as κ increases, the behavior of the restricted, positive James-Stein and preliminary test estimators are the same.
(3) When κ increases, for $\tau = 1$, the James-Stein estimator does not perform as well as the unrestricted estimator but for $\tau = 10$ and 100, the performance of the James-Stein estimator decreases and finally converges to the unrestricted estimator.
(4) For the small value of κ, the positive James-Stein and preliminary test estimators still dominate the unrestricted estimator and then converge to the unrestricted estimator as κ increases.

5. Real data application

We apply the proposed estimators to a dataset from the Turkish Statistics Association in this section. This dataset reported in 2015, includes the indicator values of the well-being index for provinces and was analyzed by [23]. The data is accessible from http://www.turkstat.gov.tr/PreHaberBultenleri.do?id=24561.

The independent variables that we chose to analyze are: the percentage of households having problems with the quality of dwellings (x_1), the percentage of households having problems with the quality of dwellings (x_2), average daily earnings (x_3), job satisfaction rate (x_4) savings deposit per capita (x_5), the percentage of households in middle or higher income groups (x_6), life expectancy at birth (Years) (x_7) satisfaction rate with public health services (x_8) murder rate (per million people) (x_9), the number of traffic accidents involving death or injury (per thousand people) (x_{10}) satisfaction rate with public safety services (%) (x_{11}) the number of cinema and theatre audience (per hundred persons) (x_{12}), satisfaction rate with social relations (%) (x_{13}), satisfaction rate with social life (%) (x_{14}) and response variable is the level of happiness (y) measured as a life satisfaction index that lies between zero and one and the values close to one refers to a better level of life.

Table 6. Estimates and standard error of the proposed estimators for the real data set.

Parameters		UNE	RE	JSE	PJSE	PTE
β_0	Estimates	2.4121	1.3538	1.6142	2.1833	2.0688
	SE	0.1571	0.0951	0.1102	0.1402	0.1465
β_1	Estimates	1.3318	0.7957	1.2283	1.2292	1.2505
	SE	0.0445	0.0262	0.0397	0.0397	0.0421
β_2	Estimates	-0.9247	-0.6594	0.5836	-0.8739	-0.8845
	SE	0.0335	0.0253	0.0341	0.0314	0.0327
β_4	Estimates	1.1926	1.5997	0.4848	1.2814	1.2927
	SE	0.0531	0.0410	0.0407	0.0502	0.0521
β_7	Estimates	-0.0779	-0.0660	0.7864	-0.0754	-0.0742
	SE	0.0017	0.0011	0.0262	0.0016	0.0016
β_8	Estimates	2.3260	2.2043	0.8774	2.2858	2.2583
	SE	0.0564	0.0382	0.0438	0.0516	0.0534
β_9	Estimates	-0.7940	-0.7941	0.7990	-0.7914	-0.7889
	SE	0.0147	0.0116	0.0276	0.0137	0.0142
β_{11}	Estimates	-2.0958	-2.2179	0.8953	-2.1139	-2.0890
	SE	0.0613	0.0426	0.0389	0.0567	0.0579
β_{13}	Estimates	3.1963	3.5572	0.5235	3.2787	3.2942
	SE	0.0561	0.0435	0.0401	0.0520	0.0538
β_{14}	Estimates	0.8120	0.6062	0.9618	0.7746	0.7786
	SE	0.0222	0.0173	0.0288	0.0208	0.0216

To analyze this dataset, we used the *betareg* package in the R programming language. The results show that the AIC of the full model is -235.86 and x_3, x_5, x_6, x_{10} and x_{12} are not significant. Thus in the next step, we consider the model without these independent variables. The AIC of the new model is -242.23, which shows that the new model is better than the full model. Therefore, we consider a sub-space that sets x_3, x_5, x_6, x_{10} and x_{12} to zero. Our aim here is to evaluate the performance of proposed estimators when this sub-space is correct. We use the bootstrap case resampling method to calculate the estimates, standard errors, and also the simulated relative efficiency of the proposed estimators to evaluate the performance of the proposed estimators. We choose a bootstrap sample size of 30 from the 81 observations of the dataset with 2000 replacements, and we compute the estimators, standard errors (SE), and relative efficiencies according to each bootstrap sample. The estimators are then determined based on the estimation methods described in Sections 2 and 3, and the SEs are estimated using the sample SE. Summary statistics for all parameters are presented in Table 6.

The SRE of the estimators is presented in Table 7. The results show that the SRE of the James-Stein estimator is less than one (0.9813). Therefore, this estimator is not an efficient estimator when the sub-space information is correct. On the other hand, the SRE of the restricted estimator, positive James-Stein estimator and preliminary test estimator are more significant than one, which indicates that these estimators outperform the unrestricted estimator.

Table 7. The SRE of the proposed estimators for the real data set.

	RE	JSE	PJSE	PTE
SRE	2.3233	0.9813	1.2259	1.1278

6. Conclusion

In this paper, the MLE (as unrestricted estimator) in beta regression model is improved using restricted estimators and shrinkage strategies when there is some information about the coefficients' values or their relationship. The properties of the proposed estimators, such as symptomatic distribution bias and quadratic risk, were theoretically and numerically studied. The numerical results confirm that if the sub-space is true ($\kappa = 0$), the restricted estimator outperforms the proposed estimators but when κ moves away from zero, the SRE of it gently decreases and gets even worse than the unrestricted estimator. The performance of the positive James-Stein estimator and preliminary test estimator, when the sub-space is true, is better than the unrestricted estimator. While κ increases, the SRE of these estimators decreases and eventually reaches one. About the James-Stein estimator, for the small value of the precision parameter when the sub-space is true, its performance is worse than the unrestricted estimator but as κ increases, the SRE of the James-Stein estimator increases. On the other hand, when the precision parameter is large, the James-Stein estimator behaves like the positive James-Stein and preliminary test estimator. Eventually, we analyze the application of the proposed estimators on an example of real-life data to evaluate the SRE of the proposed estimators. The results for real data also show better performance for the proposed estimators.

References

[1] M.R. Abonazel, Z.Y. Algamal, F.A. Awwad and I.M. Taha, *A new two-parameter estimator for beta regression model: method, simulation, and application*, Front. Appl. Math. Stat. **7**, 780322, 1-10, 2022.

[2] M.R. Abonazel, I. Dawoud, F.A. Awwad and A.F. Lukman, *Dawoud-Kibria estimator for beta regression model: simulation and application*, Front. Appl. Math. Stat. **8**, 775068, 1-12, 2022.

[3] M.R. Abonazel and I.M. Taha, *Beta ridge regression estimators: simulation and application*, Comm. Statist. Simulation Comput., Doi: 10.1080/03610918.2021.1960373, 2021.

[4] S.E. Ahmed, *Penalty, Shrinkage and Pretest Strategies: Variable Selection and Estimation*, Springer, 2014.

[5] M.N. Akram, M. Amin, A. Elhassanein and M. Aman Ullah, *A new modified ridge-type estimator for the beta regression model: simulation and application*, AIMS Math. **7** (1), 1035-1057, 2021.

[6] Z.Y. Algamal, *A particle swarm optimization method for variable selection in beta regression model*, Electron. J. Appl. Stat. Anal. **12** (2), 508-519, 2019.

[7] Z.Y. Algamal and M.R. Abonazel, *Developing a Liu-type estimator in beta regression model*, Concurr. Comput. Pract. Exp. **34** (5), 1-11, 2021.

[8] M. Arashi, *Preliminary test and Stein estimations in simultaneous linear equations*, Linear Algebra Appl. **436** (5), 1195-1211, 2012.

[9] T.A. Bancroft, *On biases in estimation due to the use of preliminary tests of significance*, Ann. Math. Stat. **15** (2), 190204, 1944.

[10] P.L. Espinheira, S.L.P. Ferrari and F. Cribari-Neto, *Influence diagnostics in beta regression*, Comput. Statist. Data Anal. **52** (9), 4417-4431, 2008.

[11] P.L. Espinheira, S.L.P. Ferrari and F. Cribari-Neto, *On beta regression residuals*, J. Appl. Stat. **35** (4), 407-419, 2008.

[12] L. Fahrmeir and H. Kaufmann, *Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models*, Ann. Statist. **13** (1), 342-368, 1985.

[13] S. Ferrari and F. Cribari-Neto, *Beta regression for modeling rates and proportions*, J. Appl. Stat. **31** (7), 799-815, 2004.

[14] S.L.P. Ferrari and E.C. Pinheiro, *Improved likelihood inference in beta regression*, J. Stat. Comput. Simul. **81** (4), 431-443, 2011.

[15] D. Fourdrinier, W.E. Strawderman and M.T. Wells, *Shrinkage Estimation*, Springer International Publishing, 2018.

[16] G.G. Judge and M.E. Bock, *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*, North Holland Publishing Company, 1978.

[17] P. Karlsson, K. Månsson and B.M.G. Kibria, *A Liu estimator for the beta regression model and its application to chemical data*, J. Chemom. **34** (10), 1-16, 2020.

[18] M. Qasim, K. Månsson and B.M.G. Kibria, *On some beta Ridge regression estimators: method,simulation and application*, J. Stat. Comput. Simul. **91** (9), 1699-1712, 2021.

[19] S.W. Mahmood, N.N. Seyala and Z.Y. Algamal, *Adjusted R^2-type measures for beta regression model*, Electron. J. Appl. Stat. Anal. **13** (2), 350-357, 2020.

[20] E. Saleh, *Theory of Preliminary Test and Stein-Type Estimation with Applications*, John Wiley and Sons, 2006.

[21] A.B. Simas, W. Barreto-Souza and A.V. Rocha, *Improved estimators for a general class of beta regression models*, Comput. Statist. Data Anal. **54** (2), 348-366, 2010.

[22] M. Smithson and J. Verkuilen, *A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables*, Psychol. Methods **11** (1), 54-71, 2006.

[23] H. Unlua and S. Aktasb, *Beta regression for the indicator values of well-being index for provinces in Turkey*, Res. J. Appl. Sci. Eng. Technol. **2** (2), 101-111, 2017.

Appendix A.

The proof of Theorem 3.1 and 3.2 are provided here. First, we obtain some statistical properties of $\hat{\boldsymbol{\beta}}^{UE}$ and $\hat{\boldsymbol{\beta}}^{RE}$. If we define $U_1 = \sqrt{n}(\hat{\boldsymbol{\beta}}^{UE} - \boldsymbol{\beta})$, $U_2 = \sqrt{n}(\hat{\boldsymbol{\beta}}^{RE} - \boldsymbol{\beta})$ and $U_3 = \sqrt{n}(\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE})$, by using (2.6), we can derive:

$$\begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ -J\xi \\ J\xi \end{bmatrix}, \begin{bmatrix} \phi\mathcal{C}_x & \phi(\mathcal{C}_x - J_0) & \phi J_0 \\ & \phi(\mathcal{C}_x - J_0) & \mathbf{0} \\ & & \phi J_0 \end{bmatrix} \right). \tag{A.1}$$

We also present the following lemma, which is proved by [16], that helps us to derive the results of the mentioned theorems.

Lemma A.1. *Let Z be a m-dimensional vector follows $N_m(\mu_Z, \boldsymbol{\Sigma}_Z)$. For any measurable function of ψ, we have*

$$\mathbb{E}\Big(Z\psi(Z'Z) \Big) = \mu_Z \mathbb{E}\Big(\psi(\chi^2_{m+2}(\Delta^2)) \Big)$$
$$\mathbb{E}\Big(Z'Z\psi(Z'Z) \Big) = \boldsymbol{\Sigma}_Z \mathbb{E}\Big(\psi\left(\chi^2_{m+2}(\Delta^2)\right) \Big) + \mu_Z'\mu_Z \mathbb{E}\Big(\psi\left(\chi^2_{m+4}(\Delta^2)\right) \Big),$$

where $\Delta^2 = \mu_Z'\boldsymbol{\Sigma}_Z^{-1}\mu_Z$ is the non-centrality parameter.

Appendix B. Proof of Theorem 3.1

By using (A.1), we have

$$ADB(\hat{\boldsymbol{\beta}}^{RE}) = \mathbb{E}(\sqrt{n}(\hat{\boldsymbol{\beta}}^{RE} - \boldsymbol{\beta})) = \mathbb{E}(U_2) = -J\xi. \tag{B.1}$$

We can write

$$\begin{aligned} ADB(\hat{\boldsymbol{\beta}}^{JSE}) &= \mathbb{E}(\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})) \\ &= \mathbb{E}\Big[\sqrt{n}\Big\{ (\hat{\boldsymbol{\beta}}^{RE} - \boldsymbol{\beta}) + \{1 - c\mathfrak{F}_n^{-1}\}(\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE}) \Big\} \Big] \\ &= \mathbb{E}\Big[U_2 + \{1 - c\mathfrak{F}_n^{-1}\}U_3 \Big] \\ &= \mathbb{E}(U_2) + \mathbb{E}\Big[\{1 - c\mathfrak{F}_n^{-1}\}U_3 \Big]. \end{aligned}$$

By using (A.1) and above Lemma, we have

$$\begin{aligned} \mathbb{E}(\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})) &= -J\xi + J\xi\mathbb{E}\Big[1 - c\chi^{-2}_{q+2}(\lambda) \Big] \\ &= -cJ\xi\mathbb{E}\Big[\chi^{-2}_{q+2}(\lambda) \Big]. \end{aligned} \tag{B.2}$$

To get the asymptotic distributional bias of $\hat{\boldsymbol{\beta}}^{PJSE}$, we can write

$$\begin{aligned}
ADB(\hat{\boldsymbol{\beta}}^{PJSE}) &= \mathbb{E}(\sqrt{n}(\hat{\boldsymbol{\beta}}^{PJSE} - \boldsymbol{\beta})) \\
&= \mathbb{E}\Big[\sqrt{n}\Big\{\hat{\boldsymbol{\beta}}^{RE} + \{1 - c\mathfrak{F}_n^{-1}\}^{+}(\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE}) - \boldsymbol{\beta}\Big\}\Big] \\
&= \mathbb{E}\Big[\sqrt{n}\Big\{(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta}) + (\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE})I_{(\mathfrak{F}_n < c)}\Big] \\
&= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta}) + \sqrt{n}(\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE})I_{(\mathfrak{F}_n < c)}\Big] \\
&= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta}) + U_3 I_{(\mathfrak{F}_n < c)}\Big] \\
&= ADB(\hat{\boldsymbol{\beta}}^{JSE}) + \mathbb{E}\Big[I_{(\mathfrak{F}_n < c)}U_3\Big] \\
&= ADB(\hat{\boldsymbol{\beta}}^{JSE}) + \mathbb{E}(U_3)\mathbb{E}\Big[I_{(\mathfrak{F}_n < c)}\Big] \\
&= ADB(\hat{\boldsymbol{\beta}}^{JSE}) + J\xi H_{q+2}(c;\lambda) \\
&= -cJ\xi\mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] + J\xi H_{q+2}(c;\lambda).
\end{aligned}$$

Finally, for the asymptotic distributional bias of $\hat{\boldsymbol{\beta}}^{PTE}$, we have

$$\begin{aligned}
ADB(\hat{\boldsymbol{\beta}}^{PTE}) &= \mathbb{E}(\sqrt{n}(\hat{\boldsymbol{\beta}}^{PTE} - \boldsymbol{\beta})) \\
&= \mathbb{E}\Big[\sqrt{n}\Big\{(\hat{\boldsymbol{\beta}}^{UE} - \boldsymbol{\beta}) - (\hat{\boldsymbol{\beta}}^{UE} - \hat{\boldsymbol{\beta}}^{RE})I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big\}\Big] \\
&= \mathbb{E}\Big[U_1 - U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\
&= \mathbb{E}(U_1) - \mathbb{E}\Big[U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\
&= -\mathbb{E}(U_3)\mathbb{E}\Big[I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\
&= -J\xi H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda).
\end{aligned}$$

Appendix C. Proof of Theorem 3.2

By definition of the asymptotic risk function (ADR), at first, we need to calculate the MSE of estimators:

$$\begin{aligned}
MSE(\hat{\boldsymbol{\beta}}^{UE}) &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{UE} - \boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{UE} - \boldsymbol{\beta})^T\Big] \\
&= \mathbb{E}\Big[U_1 U_1^T\Big] = \phi\mathfrak{C}_x.
\end{aligned}$$

Therefore,

$$ADR(\hat{\boldsymbol{\beta}}^{UE}) = tr(\phi\mathfrak{C}_x) = \phi tr(\mathfrak{C}_x). \tag{C.1}$$

Now, we calculate ADR of $\hat{\boldsymbol{\beta}}^{RE}$ as

$$\begin{aligned} MSE(\hat{\boldsymbol{\beta}}^{RE}) &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{RE}-\boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{RE}-\boldsymbol{\beta})^T\Big] \\ &= \mathbb{E}\Big[U_2U_2^T\Big] \\ &= cov(U_2) + \Big[\mathbb{E}(U_2)\Big]\Big[\mathbb{E}(U_2)\Big]^T \\ &= \phi[\mathcal{C}_x - J_0] + J\xi\xi^T J^T. \end{aligned} \tag{C.2}$$

Thus,

$$\begin{aligned} ADR(\hat{\boldsymbol{\beta}}^{RE}) &= tr\big(\phi[\mathcal{C}_x - J_0] + J\xi\xi^T J^T\big) \\ &= \phi tr(\mathcal{C}_x - J_0) + J^T\xi^T\xi J. \end{aligned}$$

For ADR of $\hat{\boldsymbol{\beta}}^{JSE}$, we can write

$$\begin{aligned} MSE(\hat{\boldsymbol{\beta}}^{JSE}) &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE}-\boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE}-\boldsymbol{\beta})^T\Big] \\ &= \mathbb{E}\Big[\Big(U_2 + \{1-c\mathfrak{F}_n^{-1}\}U_3\Big)\Big(U_2 + \{1-c\mathfrak{F}_n^{-1}\}U_3\Big)^T\Big] \\ &= \mathbb{E}\Big[U_2U_2^T\Big] + \mathbb{E}\Big[U_2\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] + \mathbb{E}\Big[U_2^T\{1-c\mathfrak{F}_n^{-1}\}U_3\Big] \\ &\quad + \mathbb{E}\Big[\{1-c\mathfrak{F}_n^{-1}\}^2 U_3U_3^T\Big] \\ &= cov(U_2) + \Big[\mathbb{E}(U_2)\Big]\Big[\mathbb{E}(U_2)\Big]^T + 2\underbrace{\mathbb{E}\Big[U_2\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big]}_{M_1} \\ &\quad + \underbrace{\mathbb{E}\Big[\{1-c\mathfrak{F}_n^{-1}\}^2 U_3U_3^T\Big]}_{M_2}. \end{aligned} \tag{C.3}$$

By using the lemma, we have

$$\begin{aligned} M_1 &= \mathbb{E}\Big[U_2\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] \\ &= \mathbb{E}\Big[\mathbb{E}\Big[U_2\{1-c\mathfrak{F}_n^{-1}\}U_3^T|U_3\Big]\Big] \\ &= \mathbb{E}\Big[\mathbb{E}\Big[U_2|U_3\Big]\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] \\ &= \mathbb{E}\Big[\Big\{\mathbb{E}(U_2) + cov(U_2,U_3)[cov(U_3)]^{-1}(U_3 - \mathbb{E}(U_2))\Big\}\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] \\ &= \mathbb{E}\Big[-J\xi\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] \\ &= -J\xi\mathbb{E}\Big[\{1-c\mathfrak{F}_n^{-1}\}U_3^T\Big] \\ &= -J\xi\mathbb{E}(U_3^T)\mathbb{E}\Big[1-c\chi_{q+2}^{-2}(\lambda)\Big] \end{aligned}$$

$$\begin{aligned} M_1 &= -J\xi\xi^T J^T \mathbb{E}\Big[1 - c\chi_{q+2}^{-2}(\lambda)\Big] \\ &= -J\xi\xi^T J^T + cJ\xi\xi^T J^T \mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] \end{aligned} \tag{C.4}$$

and by following the lemma for M_2, we have

$$\begin{aligned} M_2 &= \mathbb{E}\Big[\{1 - c\mathfrak{F}_n^{-1}\}^2 U_3 U_3^T\Big] \\ &= cov(U_3)\mathbb{E}\Big[\{1 - c\chi_{q+2}^{-2}(\lambda)\}^2\Big] + \Big[\mathbb{E}(U_3)\Big]\Big[\mathbb{E}(U_3)\Big]^T \mathbb{E}\Big[\{1 - c\chi_{q+4}^{-2}(\lambda)\}^2\Big] \\ &= \phi J_0 \mathbb{E}\Big[\{1 - c\chi_{q+2}^{-2}(\lambda)\}^2\Big] + J\xi\xi^T J^T \mathbb{E}\Big[\{1 - c\chi_{q+4}^{-2}(\lambda)\}^2\Big]. \end{aligned} \tag{C.5}$$

Consequently,

$$\begin{aligned} MSE(\hat{\boldsymbol{\beta}}^{JSE}) &= \phi[\mathfrak{C}_x - J_0] + J\xi\xi^T J^T - 2J\xi\xi^T J^T + 2cJ\xi\xi^T J^T \mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] \\ &\quad + \phi J_0 \mathbb{E}\Big[\{1 - c\chi_{q+2}^{-2}(\lambda)\}^2\Big] + J\xi\xi^T J^T \mathbb{E}\Big[\{1 - c\chi_{q+4}^{-2}(\lambda)\}^2\Big] \\ &= \phi\mathfrak{C}_x + 2cJ\xi\xi^T J^T \mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] + \phi c^2 J_0 \mathbb{E}\Big[\chi_{q+2}^{-4}(\lambda)\Big] - 2c\phi J_0 \mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] \\ &\quad + c^2 J\xi\xi^T J^T \mathbb{E}\Big[\chi_{q+4}^{-4}(\lambda)\Big] - 2cJ\xi\xi^T J^T \mathbb{E}\Big[\chi_{q+4}^{-2}(\lambda)\Big]. \end{aligned} \tag{C.6}$$

Therefore,

$$\begin{aligned} ADR(\hat{\boldsymbol{\beta}}^{JSE}) &= \phi tr(\mathfrak{C}_x) + 2c\xi^T J^T J\xi \mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] + \phi c tr(J_0)\mathbb{E}\Big[\chi_{q+2}^{-4}(\lambda)\Big] \\ &\quad - 2c\phi tr(J_0)\mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)\Big] + c^2\xi^T J^T J\xi \mathbb{E}\Big[\chi_{q+4}^{-4}(\lambda)\Big] \\ &\quad - 2c\xi^T J^T J\xi \mathbb{E}\Big[\chi_{q+4}^{-2}(\lambda)\Big]. \end{aligned}$$

For the ADR of $\hat{\boldsymbol{\beta}}^{PJSE}$, we have

$$\begin{aligned} MSE(\hat{\boldsymbol{\beta}}^{PJSE}) &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{PJSE} - \boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{PJSE} - \boldsymbol{\beta})^T\Big] \\ &= \mathbb{E}\Big[\Big\{\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta}) + U_3 I_{(\mathfrak{F}_n < c)}\Big\}\Big\{\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta}) + U_3 I_{(\mathfrak{F}_n < c)}\Big\}^T\Big] \\ &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})^T\Big] + \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})U_3^T I_{(\mathfrak{F}_n < c)}\Big] \\ &\quad + \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})^T U_3 I_{(\mathfrak{F}_n < c)}\Big] + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n < c)}\Big] \\ &= MSE(\hat{\boldsymbol{\beta}}^{JSE}) + 2\underbrace{\mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})U_3^T I_{(\mathfrak{F}_n < c)}\Big]}_{M_3} + \underbrace{\mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n < c)}\Big]}_{M_4}, \end{aligned} \tag{C.7}$$

where

$$\begin{aligned}
M_3 &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{JSE} - \boldsymbol{\beta})U_3^T I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[\Big(U_2 + \{1 - c\mathfrak{F}_n^{-1}\}U_3\Big)U_3^T I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[U_2U_3^T I_{(\mathfrak{F}_n<c)} + U_3U_3^T\{1 - c\mathfrak{F}_n^{-1}\}I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[U_2U_3^T I_{(\mathfrak{F}_n<c)}|U_3\Big]\Big] + \mathbb{E}\Big[U_3U_3^T\{1 - c\mathfrak{F}_n^{-1}\}I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[U_2|U_3\Big]U_3^T I_{(\mathfrak{F}_n<c)}\Big] + \mathbb{E}\Big[U_3U_3^T\{1 - c\mathfrak{F}_n^{-1}\}I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[\Big\{\mathbb{E}(U_2) + cov(U_2,U_3)[cov(U_3)]^{-1}(U_3 - \mathbb{E}(U_2))\Big\}U_3^T I_{(\mathfrak{F}_n<c)}\Big] \\
&\quad + \mathbb{E}\Big[U_3U_3^T\{1 - c\mathfrak{F}_n^{-1}\}I_{(\mathfrak{F}_n<c)}\Big] \\
&= \mathbb{E}\Big[\mathbb{E}(U_2)U_3^T I_{(\mathfrak{F}_n<c)}\Big] + \mathbb{E}\Big[U_3U_3^T\{1 - c\mathfrak{F}_n^{-1}\}I_{(\mathfrak{F}_n<c)}\Big] \\
&= -J\xi\mathbb{E}\Big[U_3^T I_{(\mathfrak{F}_n<c)}\Big] + cov(U_3)\mathbb{E}\Big[\{1 - c\chi_{q+2}^{-2}(\lambda)\}I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] \\
&\quad + \mathbb{E}(U_3)\mathbb{E}(U_3)^T\mathbb{E}\Big[\{1 - c\chi_{q+4}^{-2}(\lambda)\}I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] \\
&= -J\xi\xi^TJ^TH_{q+2}(c;\lambda) + \phi J_0\mathbb{E}\Big[\{1 - c\chi_{q+2}^{-2}(\lambda)\}I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] \\
&\quad + J\xi\xi^TJ^T\mathbb{E}\Big[\{1 - c\chi_{q+4}^{-2}(\lambda)\}I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] \\
&= -J\xi\xi^TJ^TH_{q+2}(c;\lambda) + \phi J_0H_{q+2}(c;\lambda) - c\phi J_0\mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] \\
&\quad + J\xi\xi^TJ^TH_{q+4}(c;\lambda) - cJ\xi\xi^TJ^T\mathbb{E}\Big[\chi_{q+4}^{-2}(\lambda)I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] \qquad \text{(C.8)}
\end{aligned}$$

and

$$\begin{aligned}
M_4 &= \mathbb{E}\Big[U_3U_3^T I_{(\mathfrak{F}_n<c)}\Big] \\
&= cov(U_3)\mathbb{E}\Big[I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] + \mathbb{E}(U_3)\mathbb{E}(U_3)^T\mathbb{E}\Big[I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] \\
&= \phi J_0\mathbb{E}\Big[I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] + J\xi\xi^TJ^T\mathbb{E}\Big[I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] \\
&= \phi J_0H_{q+2}(c;\lambda) + J\xi\xi^TJ^TH_{q+4}(c;\lambda). \qquad \text{(C.9)}
\end{aligned}$$

By replacing (C.8) and (C.9) in (C.7), we have

$$\begin{aligned}
MSE(\hat{\boldsymbol{\beta}}^{PJSE}) &= MSE(\hat{\boldsymbol{\beta}}^{JSE}) - 2J\xi\xi^TJ^T\Big[H_{q+2}(c;\lambda) - H_{q+4}(c;\lambda)\Big] \\
&\quad + 3\phi J_0H_{q+2}(c;\lambda) - 2c\phi J_0\mathbb{E}\Big[\chi_{q+2}^{-2}(\lambda)I_{(\chi_{q+2}^{-2}(\lambda)<c)}\Big] \\
&\quad - 2cJ\xi\xi^TJ^T\mathbb{E}\Big[\chi_{q+4}^{-2}(\lambda)I_{(\chi_{q+4}^{-2}(\lambda)<c)}\Big] + 2J\xi\xi^TJ^TH_{q+4}(c;\lambda). \qquad \text{(C.10)}
\end{aligned}$$

Therefore,

$$\begin{aligned} ADR(\hat{\boldsymbol{\beta}}^{PJSE}) &= ADR(\hat{\boldsymbol{\beta}}^{JSE}) - 2\xi^T J^T J\xi\Big[H_{q+2}(c;\lambda) - H_{q+4}(c;\lambda)\Big] \\ &\quad + 3\phi tr(J_0)H_{q+2}(c;\lambda) - 2c\phi tr(J_0)\mathbb{E}\Big[\chi^{-2}_{q+2}(\lambda)I_{(\chi^{-2}_{q+2}(\lambda)<c)}\Big] \\ &\quad - 2c\xi^T J^T J\xi\mathbb{E}\Big[\chi^{-2}_{q+4}(\lambda)I_{(\chi^{-2}_{q+4}(\lambda)<c)}\Big] + 2\xi^T J^T J\xi H_{q+4}(c;\lambda). \qquad (C.11) \end{aligned}$$

For $\hat{\boldsymbol{\beta}}^{PTE}$, we have

$$\begin{aligned} MSE(\hat{\boldsymbol{\beta}}^{PTE}) &= \mathbb{E}\Big[\sqrt{n}(\hat{\boldsymbol{\beta}}^{PTE} - \boldsymbol{\beta})\sqrt{n}(\hat{\boldsymbol{\beta}}^{PTE} - \boldsymbol{\beta})^T\Big] \\ &= \mathbb{E}\Big[\big(U_1 - U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\big)\big(U_1 - U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\big)^T\Big] \\ &= \mathbb{E}\Big[U_1 U_1^T\Big] - \mathbb{E}\Big[U_1 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] - \mathbb{E}\Big[U_1^T U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &\quad + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= cov(U_1) + \Big[\mathbb{E}(U_1)\Big]\Big[\mathbb{E}(U_1)\Big]^T - 2\mathbb{E}\Big[U_1^T U_3 I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &\quad + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - 2\mathbb{E}\Big[\mathbb{E}\Big[U_1 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}|U_3\Big]\Big] + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - 2\mathbb{E}\Big[\mathbb{E}\Big[U_1|U_3\Big]U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - 2\mathbb{E}\Big[\Big\{\mathbb{E}(U_1) + cov(U_1, U_3)\Big[cov(U_3)\Big]^{-1}(U_3 - \mathbb{E}(U_3))\Big\}U_3^T I_{(\mathfrak{F}_n \le k)}\Big] \\ &\quad + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - 2\mathbb{E}\Big[(U_3 - J\xi)U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] + \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - \mathbb{E}\Big[U_3 U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] + 2J\xi\mathbb{E}\Big[U_3^T I_{(\mathfrak{F}_n \le \mathfrak{F}_{n,\alpha})}\Big] \\ &= \phi\mathcal{C}_x - cov(U_3)H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) - \Big[\mathbb{E}(U_1)\Big]\Big[\mathbb{E}(U_1)\Big]^T H_{q+4}(\mathfrak{F}_{n,\alpha};\lambda) \\ &\quad + 2J\xi\xi^T J^T H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) \\ &= \phi\mathcal{C}_x - \phi J_0 H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) + J\xi\xi^T J^T\big(2H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) - H_{q+4}(\mathfrak{F}_{n,\alpha};\lambda)\big). \end{aligned}$$

Thus,

$$\begin{aligned} ADR(\hat{\boldsymbol{\beta}}^{PTE}) &= \phi tr(\mathcal{C}_x) - \phi tr(J_0)H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) \\ &\quad + \xi^T J^T J\xi\big(2H_{q+2}(\mathfrak{F}_{n,\alpha};\lambda) - H_{q+4}(\mathfrak{F}_{n,\alpha};\lambda)\big). \qquad (C.12) \end{aligned}$$