International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

On Sampling Distribution of Improved Estimators for Coefficients in Seemingly Unrelated Regression SUR Models

A. M. A. Fattah[a]*, E. A. Rady[b]

[a]*Assistant Lecturer, Department of Applied Statistics and Econometrics, Cairo University, Giza-5, Egypt*

[b]*Professor, Department of Applied Statistics and Econometrics, Cairo University, Giza-5, Egypt*

[a]*Email: asoubh84@gmail.com*

[b]*Email: elhoussainy@yahoo.com*

Abstract

The main objective of this study is to estimate the parameters of the SUR models. A new family of biased estimators, called k-class that has the same asymptotic normal distribution as the Aitken generalized least squares (GLS) with the assumption that the covariance matrix is known. The exact bias have been studied and derived.

Keywords: Seemingly unrelated regression models; K-Class estimators; Non Central Chi-Square distribution; Exact Bias.

1. Introduction

It is well known that the generalized least squares method provides best linear unbiased estimates of the parameters of seemingly unrelated regression model under the assumptions of Gauss-Markov theorem.

---

* Corresponding author.

However, if got out of the class of linear functions and relax the condition of unbiasedness, it is possible to obtain an estimator (which is a nonlinear function of observations on the dependent variable and is in fact biased) which has a smaller mean squared error. In fact, such an estimator has been given by [1], and more explicitly by [2] while considering the problem of estimating the mean vector of a multivariate normal distribution. It has been shown by James and Stein that their estimator is better than the usual one in the sense that the sum of its component wise mean squared errors is smaller than that of the other, for all parameter values provided at least three parameters are to be estimated.

However, it should be noted that the sampling distribution of the improved James and Stein estimator is not known.

This paper considered the estimation of the parameters of the general seemingly unrelated regression model. Also, the development a new family of biased estimators, namely the k-class, by using an operational variant of the minimum mean square error estimator which depends on unknown parameters (see [3]). The procedure of developing this family is simple and straightforward. The model and the unbiased estimator are presented in Section 2. In section 3, important functions and notations are listed. Section 4 presented the new proposed estimator using the k-class method. In section 5, the analysis of the exact bias is derived.

2. The Seemingly Unrelated Regression (SUR) Model

The basic philosophy of the SURE model is as follows. The jointness of the equations is explained by the structure of the SURE model and the covariance matrix of the associated disturbances. Such jointness introduces additional information which is over and above the information available when the individual equations are considered separately. So it is desired to consider all the separate relationships collectively to draw the statistical inferences about the model parameters. In this case, the model is,

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{U} \tag{2.1}$$

where $\boldsymbol{Y}$ is an $\boldsymbol{n} \times \mathbf{1}$ vector of observations on the dependent variable, $\boldsymbol{X}$ is an $\boldsymbol{n} \times \boldsymbol{p}$ matrix of observations on $\boldsymbol{p}$ explanatory variables, $\boldsymbol{\beta}$ is a $\boldsymbol{p} \times \mathbf{1}$ parameter vector, and $\boldsymbol{U}$ is an $\boldsymbol{n} \times \mathbf{1}$ disturbance vector. Let $\boldsymbol{n} = \boldsymbol{TL}$ represents $\boldsymbol{T}$ observations for each of the $\boldsymbol{L}$ equations.

The following are conventional assumptions:

Assumption 1: The $\boldsymbol{X}$ matrix is nonstochastic and of rank $\boldsymbol{p}$.

Assumption 2: The disturbance vector $\boldsymbol{U}$ of order $\boldsymbol{n} \times \mathbf{1}$ is distributed as multivariate normal with mean vector zero and variance covariance matrix

$$\mathbf{U} \sim \mathbf{N}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}) \tag{2.2}$$

where the variance covariance matrix $\boldsymbol{Var}\ (\boldsymbol{U}) = (\boldsymbol{\Sigma} \otimes \boldsymbol{I_T})$ with order $\boldsymbol{n} \times \boldsymbol{n}$ and $\boldsymbol{\Sigma}$ is a known positive definite matrix with order $\boldsymbol{L} \times \boldsymbol{L}$.

Assumption 3: The sample size $\boldsymbol{n}$ is greater than the total number of explanatory variables $\boldsymbol{p}$; i.e. $\boldsymbol{n} > p$.

[4] introduced an unbiased estimator called the generalized least squares (GLS) estimator $\widehat{\boldsymbol{\beta}}_{SURE}$ of β in (2.1) as follows:

$$\widehat{\boldsymbol{\beta}}_{\mathbf{SURE}} = [\mathbf{X}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\mathbf{X}]^{-1}\mathbf{X}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I}) \qquad \textbf{(2.3)}$$

[5] derived the variance covariance matrix for the $\widehat{\boldsymbol{\beta}}_{SURE}$, which takes the following form:

$$\boldsymbol{Var}(\widehat{\boldsymbol{\beta}}_{SURE}) = [\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\mathbf{X}]^{-1} \qquad \textbf{(2.4)}$$

and proved that $\widehat{\boldsymbol{\beta}}_{SURE}$ is best linear unbiased estimator.

Also, this estimator is unbiased in small samples assuming that the error terms $\boldsymbol{U}$ have symmetric distribution. As for the large samples, it is consistent and asymptotically normal with limiting distribution:

$$\sqrt{\boldsymbol{T}}(\widehat{\boldsymbol{\beta}}_{SURE} - \boldsymbol{\beta}) \xrightarrow{d} \mathbf{N}\left(\mathbf{0}, \left[\frac{\mathbf{1}}{\boldsymbol{T}}\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\mathbf{X}\right]^{-\mathbf{1}}\right)$$

In other words, the $E(\widehat{\boldsymbol{\beta}}_{SURE}) = \boldsymbol{\beta}$, and $E(\widehat{\boldsymbol{\beta}}_{SURE} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}}_{SURE} - \boldsymbol{\beta})' = \left[\frac{\mathbf{1}}{\boldsymbol{T}}\boldsymbol{X}'(\boldsymbol{\Sigma}^{-1}\otimes\mathbf{I})\mathbf{X}\right]^{-\mathbf{1}}$.

3. Preliminaries

This section is dedicated to list some definitions and notations needed for the research completion.

3.1. Definitions

Definition 1: Confluent Hypergeometric Function in the series form from [6]:

$${}_1\boldsymbol{F}_1(\boldsymbol{a};\boldsymbol{c};\boldsymbol{\lambda}) = \sum_{i=0}^{\infty}\frac{\boldsymbol{\lambda}^i}{i!}\frac{(\boldsymbol{a})_i}{(c)_i}$$

where $(a)_i$ and $(c)_i$ are Pochhammer symbols.

Also, ${}_1\boldsymbol{F}_1(\boldsymbol{a};\boldsymbol{a};\boldsymbol{\lambda}) = \boldsymbol{e}^{\boldsymbol{\lambda}}$

The Integral form is represented as follows,

$${}_1\boldsymbol{F}_1(\boldsymbol{a};\boldsymbol{c};\boldsymbol{\lambda}) = \frac{\Gamma(c)}{\Gamma(c-a)\Gamma(a)}\int_0^1 \boldsymbol{e}^{\boldsymbol{\lambda t}}\boldsymbol{t}^{\boldsymbol{a}-\mathbf{1}}(\mathbf{1}-\boldsymbol{t})^{\boldsymbol{c}-\boldsymbol{a}-\mathbf{1}}\,\boldsymbol{dt} \qquad \textbf{(3.1)}$$

Definition 2:

$$\boldsymbol{f_{\mu,\nu}} = \frac{\Gamma(\frac{r}{2}+\mu)}{\Gamma(\frac{r}{2}+\nu)} \boldsymbol{e^{-\lambda}}\,{}_1\boldsymbol{F}_1(\frac{r}{2}+\boldsymbol{\mu}; \frac{r}{2}+\boldsymbol{v}; \boldsymbol{\lambda}), \text{ and } \boldsymbol{v}-\boldsymbol{\mu} > 0 \qquad \textbf{(3.2)}$$

Definition 3:

$$\frac{\partial}{\partial\lambda}\boldsymbol{f}_{\mu,\nu} = \boldsymbol{f}_{\mu+1,\nu} - \boldsymbol{f}_{\mu,\nu} \qquad \textbf{(3.3)}$$

and

$$\frac{\partial^2}{\partial\lambda^2}\boldsymbol{f}_{\mu,\nu} = \boldsymbol{f}_{\mu+2,\nu} - 2\boldsymbol{f}_{\mu+1,\nu} + \boldsymbol{f}_{\mu,\nu} \qquad \textbf{(3.4)}$$

3.2. Notations

The following are some facts proven in the literature concerning the non-central chi square χ^2 distribution.

I. According to [7], [8] and [9], If $\boldsymbol{X} \sim \boldsymbol{N_n}(\boldsymbol{\mu},\boldsymbol{\Sigma}), \boldsymbol{\Sigma} > 0$, then $\boldsymbol{Q} = \boldsymbol{X'AX} + \boldsymbol{d} \sim \boldsymbol{\chi^2_{r,\lambda}}$ where $\boldsymbol{r}$ is the degree of freedom, $\boldsymbol{d}$ is a scalar and $\boldsymbol{\lambda} = \frac{1}{2}\boldsymbol{\mu' A\mu} + \boldsymbol{d}$ is the noncentrality parameter under the following conditions:

(i) $\boldsymbol{A} = \boldsymbol{A'}$

(ii) $\boldsymbol{A\Sigma A} = \boldsymbol{A}$

(iii) $\boldsymbol{trace}(\boldsymbol{A\Sigma}) = \boldsymbol{r}$ **(3.5)**

II. According to [8], and [10], the probability density function of the non-central chi square distribution (r,λ) is as follows:

$$\boldsymbol{f}(\boldsymbol{Q} = \boldsymbol{\chi^2_{r,\lambda}}) = \boldsymbol{e^{-\lambda}} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \frac{Q^{\left(\frac{r+2i}{2}\right)-1} e^{\frac{-Q}{2}}}{2^{\left(\frac{r+2i}{2}\right)} \Gamma\left(\frac{r+2i}{2}\right)}, \text{ and } Q > 0 \qquad \textbf{(3.6)}$$

where the degrees of freedom is r, and λ is the noncentrality parameter.

III. According to [11], provided that $\frac{r}{2} > s$, the s-th inverse moments of the non-central chi square $\boldsymbol{Q}$ is as follows,

$$\boldsymbol{E}(\boldsymbol{Q^{-s}}) = \boldsymbol{2^{-s}} \frac{\Gamma(\frac{r}{2}-s)}{\Gamma(\frac{r}{2})} \boldsymbol{e^{-\lambda}} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} \frac{(\frac{r}{2}-s)_i}{(\frac{r}{2})_i} = \boldsymbol{2^{-s}} \frac{\Gamma(\frac{r}{2}-s)}{\Gamma(\frac{r}{2})} \boldsymbol{e^{-\lambda}}\,{}_1\boldsymbol{F}_1(\frac{r}{2}-\boldsymbol{s}; \frac{r}{2}; \boldsymbol{\lambda}). \qquad \textbf{(3.7)}$$

for $\boldsymbol{s} = 0,1,2,\ldots$, and $\boldsymbol{r} = 1,2,\ldots$.

where ${}_1\boldsymbol{F}_1(\;\;)$ is the confluent hypergeometric function defined in (3.1)

IV. According to [9], the derivatives of $\boldsymbol{E(Q^{-s})}$ with respect to $\boldsymbol{\lambda}$ is as follows,

$$\frac{\partial^m}{\partial\lambda^m}\boldsymbol{E(Q^{-s}) = 2^{-s}(-1)^m \frac{\Gamma(s+m)}{\Gamma(s)}\frac{\Gamma(\frac{r}{2}-s)}{\Gamma(\frac{r}{2}+m)} e^{-\lambda}{}_1F_1(\frac{r}{2}-s;\frac{r}{2}+m;\lambda)} \quad \textbf{(3.8)}$$

for $\boldsymbol{s}$ = 1,2,…, and $\boldsymbol{m}$ = 1,2,...

4. The K-Class Estimator

The k-class estimators have been studied in the literature extensively by so many researchers ([11], [12], [13], etc). Latter, it was mentioned in the literature that the k-class estimator is biased and has a smaller mean squared error. In this section, a proposed k-class estimator is derived.

Proposition (1):

The family of k-class estimators of $\boldsymbol{\beta}$, $\widetilde{\boldsymbol{b}}$ is

$$\widetilde{\boldsymbol{b}} = \left[\boldsymbol{1 - k\frac{Y'HY}{1+Y'HY}}\right]\widehat{\boldsymbol{\beta}}_{Sure} \quad \textbf{(4.1)}$$

where $\mathbf{k}$ is an arbitrary constant. For $\mathbf{k = 0}$, the k-class estimator will be the Zellner's seemingly unrelated regression estimator $\widetilde{\boldsymbol{b}} = \widehat{\boldsymbol{\beta}}_{Sure}$ and note that

$$\boldsymbol{0 \le \frac{Y'HY}{1+Y'HY} \le 1}$$

Proof: Let the model,

$$\boldsymbol{PY = PX\beta + PU} \quad \textbf{(4.2)}$$

where P is $n \times n$ matrix such that $\boldsymbol{P'P = \Sigma^{-1}\otimes I}$. The existence of such P is from the fact that $\boldsymbol{\Sigma}$ is positive definite matrix.

Write the new model:

$$\boldsymbol{Y^* = X^*\beta + U^*} \quad \textbf{(4.3)}$$

where $\boldsymbol{Y^* = PY}$, $\boldsymbol{X^* = PX}$, and $\boldsymbol{U^* = PU}$.

$\boldsymbol{E(U^*) = PE(U) = 0_{n\times 1}}$ and $Cov(\boldsymbol{U^*}) = \boldsymbol{E(PUU'P') = PE(UU')P' = P(\Sigma\otimes I)P' = I_n}$.

Define a class of linear estimators $\boldsymbol{\beta^* = AY^*}$ where $\boldsymbol{A}$ is an arbitrary $\boldsymbol{p \times n}$ matrix. The mean squared error matrix $\boldsymbol{\Psi_{p\times p}}$ of $\boldsymbol{\beta^*}$ is given by

$$\boldsymbol{\Psi}_{\mathbf{p}\times\mathbf{p}} = \mathrm{E}(\boldsymbol{\beta}^* - \boldsymbol{\beta})(\boldsymbol{\beta}^* - \boldsymbol{\beta})' = \boldsymbol{E}\left[\left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)\boldsymbol{\beta} + \boldsymbol{AU}_*\right]\left[\left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)\boldsymbol{\beta} + \boldsymbol{AU}^*\right]'$$
$$= \left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)\boldsymbol{\beta\beta}'\left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)' + \left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)\boldsymbol{\beta E}(\boldsymbol{U}^*)\boldsymbol{A} + \boldsymbol{AE}(\boldsymbol{U}^*)\boldsymbol{\beta}'\left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)'$$
$$+ \boldsymbol{AE}(\boldsymbol{U}^*\boldsymbol{U}^{*\prime})\boldsymbol{A}'$$

and since $\boldsymbol{E}(\boldsymbol{U}^*) = \mathbf{0}_{\boldsymbol{n}\times\mathbf{1}}$ and $\boldsymbol{Cov}(\boldsymbol{U}^*) = \boldsymbol{I}_{\boldsymbol{n}}$ then, $\boldsymbol{\Psi}_{\mathbf{p}\times\mathbf{p}} = \left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)\boldsymbol{\beta\beta}'\left(\boldsymbol{AX}^* - \boldsymbol{I}_p\right)' + \boldsymbol{AA}'$.

Choose $\boldsymbol{A}$ such that $\boldsymbol{\Psi}_{\mathbf{p}\times\mathbf{p}}$ is minimum by differentiating $\boldsymbol{\Psi}_{\mathbf{p}\times\mathbf{p}}$with respect to $\boldsymbol{A}$,

$\frac{\partial\boldsymbol{\Psi}_{\mathbf{p}\times\mathbf{p}}}{\partial\boldsymbol{A}_{\mathbf{p}\times\mathbf{n}}} = \mathbf{2}\boldsymbol{AX}^*\boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime} - \boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime} + \mathbf{2}\boldsymbol{A} = \mathbf{0}$, and solve for **A**:

$$\boldsymbol{A} = \boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime}[\boldsymbol{X}^*\boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime} + \boldsymbol{I}]^{-\mathbf{1}} \quad \mathbf{(4.4)}$$

Hence, $\boldsymbol{\beta}^* = \boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime}(\boldsymbol{X}^*\boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime} + \boldsymbol{I}_{\boldsymbol{n}})^{-\mathbf{1}}\boldsymbol{Y}^*$

Further, $\boldsymbol{\beta}^*$ can be rewritten by expanding $(\boldsymbol{X}^*\boldsymbol{\beta\beta}'\boldsymbol{X}^{*\prime} + \boldsymbol{I}_{\boldsymbol{n}})^{-\mathbf{1}}$ as in [14] as follows:

$$\boldsymbol{\beta}^* = [\frac{\boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{Y}^*}{\mathbf{1} + \boldsymbol{\beta}'\boldsymbol{X}^{*\prime}\boldsymbol{X}^*\boldsymbol{\beta}}]\boldsymbol{\beta} = \left(\frac{\boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\boldsymbol{I})\boldsymbol{Y}}{\mathbf{1} + \boldsymbol{\beta}'\boldsymbol{X}'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\boldsymbol{I})\boldsymbol{X\beta}}\right)\boldsymbol{\beta}$$

Using the $\boldsymbol{X\beta} = \boldsymbol{Y} - \boldsymbol{U}$, then

$$\boldsymbol{\beta}^* = \left[\frac{(\boldsymbol{Y} - \boldsymbol{U})'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})\mathbf{Y}}{\mathbf{1} + (\boldsymbol{Y} - \boldsymbol{U})'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})(\boldsymbol{Y} - \boldsymbol{U})}\right]\boldsymbol{\beta}$$

Replace $\boldsymbol{U}$, $\boldsymbol{\beta}$ by $\boldsymbol{e}, \widehat{\boldsymbol{\beta}}_{\boldsymbol{sure}}$, this will result in $\widetilde{\boldsymbol{b}}$ as follows

$$\widetilde{\boldsymbol{b}} = \left[\frac{(\boldsymbol{Y}-\boldsymbol{e})'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})\mathbf{Y}}{\mathbf{1}+(\boldsymbol{Y}-\boldsymbol{e})'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})(\boldsymbol{Y}-\boldsymbol{e})}\right]\widehat{\boldsymbol{\beta}}_{\boldsymbol{sure}} \quad \mathbf{(4.5)}$$

Expanding (4.4):

$$\widetilde{\boldsymbol{b}} = \left(\frac{\boldsymbol{Y}'\boldsymbol{HY}}{\mathbf{1} + \boldsymbol{Y}'\boldsymbol{HY}}\right)\widehat{\boldsymbol{\beta}}_{\boldsymbol{sure}}$$

Note that, $\boldsymbol{H}$ is a positive definite matrix and takes the following form

$$\boldsymbol{H} = (\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})\mathbf{X}[\boldsymbol{X}'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})\mathbf{X}]^{-\mathbf{1}}\boldsymbol{X}'(\boldsymbol{\Sigma}^{-\mathbf{1}}\otimes\mathbf{I})$$

5. The Exact Bias of $\widetilde{\boldsymbol{b}}$

In this section, the exact formula for the bias of the k-class estimator for the seemingly unrelated regression (SUR) model is derived. This section along with the proposed estimator is considered as the major contribution of this study.

The sampling error of the estimator in (2.20) can be written in the following equation,

$$\widetilde{\boldsymbol{b}} - \boldsymbol{\beta} = \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} - \boldsymbol{k}\left(\frac{\boldsymbol{Y'HY}}{\mathbf{1}+\boldsymbol{Y'HY}}\right)\widehat{\boldsymbol{\beta}} \tag{5.1}$$

Let $\boldsymbol{c} = \frac{\boldsymbol{Y'HY}}{\mathbf{1}+\boldsymbol{Y'HY}}$

Then, $\widetilde{\boldsymbol{b}} - \boldsymbol{\beta} = \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta} - \boldsymbol{kc}\widehat{\boldsymbol{\beta}}$ **(5.2)**

Also $\boldsymbol{c}$ can be written as follows:

$\boldsymbol{c} = \frac{\boldsymbol{Y}^{*\prime}\mathbf{NY}^*}{\mathbf{1}+\boldsymbol{Y}^{*\prime}\mathbf{NY}^*} = \mathbf{1} - \frac{\mathbf{1}}{\mathbf{1}+\boldsymbol{Y}^{*\prime}\mathbf{NY}^*}$ where $\boldsymbol{N} = \mathbf{X}^*[\boldsymbol{X}^{*\prime}\mathbf{X}^*]^{-\mathbf{1}}\boldsymbol{X}^{*\prime}$ is an $\boldsymbol{n} \times \boldsymbol{n}$ idempotent matrix with rank $\boldsymbol{p}$.

Lemma (1): $E(\boldsymbol{c}) = \mathbf{1} - \mathbf{2}^{-\mathbf{1}}\boldsymbol{f}_{-\mathbf{1},\mathbf{0}}$

Proof: $\boldsymbol{E}(\boldsymbol{c}) = \boldsymbol{E}\left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{1}+\boldsymbol{Y}^{*\prime}\boldsymbol{NY}^*}\right) = \mathbf{1} - \boldsymbol{E}\left(\frac{\mathbf{1}}{\mathbf{1}+\boldsymbol{Y}^{*\prime}\boldsymbol{NY}^*}\right)$

From (3.5) assuming $d = 1$, $A = N$, $\Sigma = I$, and N is an idempotent matrix with rank $\boldsymbol{p}$, then

$$\boldsymbol{Q} = \mathbf{1} + \boldsymbol{Y}^{*\prime}\mathbf{NY}^* \sim \boldsymbol{\chi}^{\mathbf{2}}_{\mathbf{p},\boldsymbol{\lambda}} \tag{5.3}$$

where $\boldsymbol{\chi}^{\mathbf{2}}_{\mathbf{p},\boldsymbol{\lambda}}$ represents the noncentral chi-square distribution with the d.f. $\boldsymbol{p}$ and the noncentrality parameter $\boldsymbol{\lambda} = (\mathbf{1} + \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\mu}^{*\prime}\boldsymbol{\mu}^*)$.

According to (3.2) and (3.7) then, (5.3) can be rewritten as follows

$$\boldsymbol{E}(\boldsymbol{c}) = \mathbf{1} - \boldsymbol{E}(\boldsymbol{Q}^{-\mathbf{1}}) = \mathbf{1} - \mathbf{2}^{-\mathbf{1}}\frac{\Gamma(\frac{\mathbf{p}}{\mathbf{2}}-\mathbf{1})}{\Gamma(\frac{\mathbf{p}}{\mathbf{2}})}\boldsymbol{e}^{-\boldsymbol{\lambda}}{}_{\mathbf{1}}\boldsymbol{F}_{\mathbf{1}}\left(\frac{\boldsymbol{p}}{\mathbf{2}} - \mathbf{1}; \frac{\boldsymbol{p}}{\mathbf{2}}; \boldsymbol{\lambda}\right) = \mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{f}_{-\mathbf{1},\mathbf{0}} \tag{5.4}$$

given that $\frac{\boldsymbol{p}}{\mathbf{2}} > 1$.

Lemma (2): $\frac{\partial}{\partial\boldsymbol{\mu}^{*\prime}}\boldsymbol{E}(\boldsymbol{c}) = \boldsymbol{\mu}^*\boldsymbol{f}_{-\mathbf{1},\mathbf{1}}$

Proof: Using (3.3)

$$\frac{\partial}{\partial\boldsymbol{\mu}^{*\prime}}\boldsymbol{E}(\boldsymbol{c}) = -\frac{\partial}{\partial\boldsymbol{\mu}^{*\prime}}\boldsymbol{E}(\boldsymbol{Q}^{-\mathbf{1}}) = -\frac{\partial}{\partial\boldsymbol{\lambda}}\boldsymbol{E}(\boldsymbol{Q}^{-\mathbf{1}})\frac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{\mu}^{*\prime}} = \boldsymbol{\mu}^*\boldsymbol{f}_{-\mathbf{1},\mathbf{1}}. \tag{5.5}$$

Lemma (3): $\boldsymbol{E}(\boldsymbol{cY}^*) = \boldsymbol{\mu}^*\boldsymbol{f}_{-\mathbf{1},\mathbf{1}} + \boldsymbol{\mu}^*(\mathbf{1} - \mathbf{2}^{-\mathbf{1}}\boldsymbol{f}_{-\mathbf{1},\mathbf{0}})$

Proof:

Write $\mathbf{cY}^* = \mathbf{c}(\mathbf{Y}^* - \boldsymbol{\mu}^*) + \boldsymbol{\mu}^*\mathbf{c}$, then

$$\boldsymbol{E(c\mathbf{Y}^*) = E(Y^* - \mu^*)c + \mu^* E(c) = (2\pi)^{\frac{-n}{2}} \int_{Y^*} (Y^* - \mu^*)c\ exp\left\{-\frac{1}{2}(Y^* - \mu^*)'(Y^* - \mu^*)\right\} dY^* + \mu^* E(c)}$$

$$= \frac{\partial}{\partial \boldsymbol{\mu}^*} \boldsymbol{E(c) + \mu^* E(c)} \quad \textbf{(5.6)}$$

Using Lemma (1) and (2),

$\boldsymbol{E(c\mathbf{Y}^*) = \boldsymbol{\mu}^* f_{-1,1} + \mu^*(1 - 2^{-1} f_{-1,0})}$.

Theorem: The exact bias of the k-class estimator of β for $p > 2$ is given by

$$\boldsymbol{E(\tilde{b} - \beta) = -k\beta(1 - 2^{-1} f_{-1,0} + f_{-1,1})}$$

where $\mathbf{f_{-1,0}}$ is as given in (3.3) for $\boldsymbol{\mu} = \mathbf{-1}$, and $\boldsymbol{\nu} = \mathbf{0}$

Proof: From (5.2), write

$$\boldsymbol{E(\tilde{b} - \beta) = -kE(c\hat{\beta})} \quad \textbf{(5.7)}$$

By Expanding $\hat{\boldsymbol{\beta}}_{\mathbf{sure}} = [\mathbf{X}^{*\prime}\mathbf{X}^*]^{-1}\mathbf{X}^{*\prime}\mathbf{Y}^*$

Then,

$$\boldsymbol{E(\tilde{b} - \beta) = -k}[\mathbf{X}^{*\prime}\mathbf{X}^*]^{-1}\mathbf{X}^{*\prime}\boldsymbol{E(c\mathbf{Y}^*)} \quad \textbf{(5.8)}$$

where $\mathbf{Y}^* \sim \boldsymbol{N_{TL}}(\boldsymbol{\mu}^*, \mathbf{I_{TL}})$ and $\boldsymbol{\mu^* = X^*\beta = PX\beta}$ from (4.3). From lemma (3),

$$\boldsymbol{E(\tilde{b} - \beta) = -k}[\mathbf{X}^{*\prime}\mathbf{X}^*]^{-1}\mathbf{X}^{*\prime}\{\boldsymbol{\mu}^* \boldsymbol{f_{-1,1} + \mu^*(1 - 2^{-1} f_{-1,0})}\} = \boldsymbol{-k\beta(1 - 2^{-1} f_{-1,0} + f_{-1,1})}.$$

Lemma (4): The simplified form for the exact bias of the k-class estimator of β for $p > 2$ is given by

$$\boldsymbol{E(\tilde{b} - \beta) = -\frac{1}{2} k\beta}$$

Proof: From (3.3), write

$$\boldsymbol{E(c\mathbf{Y}^*) = \boldsymbol{\mu}^*\left(1 - \tfrac{1}{2} f_{-1,0}\right) - \tfrac{1}{2}\mu^*(f_{0,0} - f_{-1,0}) = \boldsymbol{\mu}^*(1 - \tfrac{1}{2} f_{0,0})} \quad \textbf{(5.9)}$$

According to (3.2), taking $\boldsymbol{\mu} = \mathbf{0}$, and $\boldsymbol{\nu} = \mathbf{0}$ then,

$$\boldsymbol{f}_{\mathbf{0,0}} = \frac{\Gamma(\frac{r}{2})}{\Gamma(\frac{r}{2})} \boldsymbol{e}^{-\boldsymbol{\lambda}} {}_1\boldsymbol{F}_1\left(\frac{r}{2}; \frac{r}{2}; \boldsymbol{\lambda}\right) = \boldsymbol{e}^{-\boldsymbol{\lambda}} \boldsymbol{e}^{\boldsymbol{\lambda}} = \mathbf{1}.$$

This means that the exact bias of the k-class estimator is written as follows:

$$\boldsymbol{E}\left(\widetilde{\boldsymbol{b}} - \boldsymbol{\beta}\right) = -\boldsymbol{k}[\mathbf{X}^{*\prime}\mathbf{X}^{*}]^{-\mathbf{1}}\mathbf{X}^{*\prime}\boldsymbol{\mu}^{*}\left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{f}_{\mathbf{0,0}}\right) = -\frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{k}\boldsymbol{\beta}.$$

Corollary:

(a)The k-class estimator $\widetilde{\boldsymbol{b}}$ is unbiased only for $\boldsymbol{k} = \mathbf{0}$ and in this case it is the GLS estimator $\widehat{\boldsymbol{\beta}}_{\boldsymbol{sure}}$.

(b) The exact bias of the $\widetilde{\boldsymbol{b}}$ is a decreasing function of $\boldsymbol{k}$.

6. Conclusion

This study has derived a new estimator to the seemingly unrelated regression models which is biased and have the same asymptotic normal distribution as the Aitken generalized least squares (GLS) with the assumption that the covariance matrix is known. It has been shown that the new estimator is a special case of the unbiased estimator called the generalized least squares (GLS) estimator $\widehat{\boldsymbol{\beta}}_{\boldsymbol{SURE}}$. Also the new estimator has a decreasing bias function in the $\boldsymbol{k}$.

References

[1] C. Stein, 'Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution', in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1956, pp. 197-206.

[2] W. James and C. Stein, 'Estimation with Quadratic Loss', in Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961, pp. 361-379.

[3] H. Theil, Principles of econometrics. New York: Wiley, 1971.

[4] A. Zellner, 'An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias', Journal of the American Statistical Association, vol. 57, no. 298, pp. 348-368, 1962.

[5] A. Zellner, 'Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results', Journal of the American Statistical Association, vol. 58, no. 304, pp. 977-992, 1963.

[6] L. Slater, Confluent hypergeometric functions. Cambridge [England]: University Press, 1960.

[7] A. Mathai and S. Provost, Quadratic forms in random variables. New York: M. Dekker, 1992.

[8] M. Srivastava and C. Khatri, An introduction to multivariate statistics. New York: North-Holland/New

York, 1979.

[9] A. Ullah, Finite sample econometrics. Oxford: Oxford University Press, 2004.

[10] C. Rao, Linear statistical inference and its applications. New York: Wiley, 1965.

[11] A. Ullah, 'On the sampling distribution of improved estimators for coefficients in linear regression', Journal of Econometrics, vol. 2, no. 2, pp. 143-150, 1974.

[12] A. Ullah and S. Ullah, 'Double k-Class Estimators of Coefficients in Linear Regression', Econometrica, vol. 46, no. 3, p. 705, 1978.

[13] R. Carter, M. Srivastava, V. Srivastava and A. Ullah, 'Unbiased Estimation of the MSE Matrix of Stein-Rule Estimators, Confidence Ellipsoids, and Hypothesis Testing', Econometric Theory, vol. 6, no. 01, p. 63, 1990.

[14] M. Srivastava and E. Carter, An introduction to applied multivariate statistics. New York: North Holland, 1983.