Journal of Statistical Theory and Applications ISSN 1538-7887
Volume 11, Number 1, 2012, pp.63-86

GENERAL STOCHASTIC RESTRICTED SUR RIDGE AND SUR ROBUST RIDGE ESTIMATORS WITH ROBUST CROSS VALIDATION

[1] El- Houssainy A.Rady,[2] Sayed M. El Saye,[3] Alaa A. Abdel-Aziz,[4]Naglaa A. Morad, Tarek M. Omara

Tarek_em@yahoo.com

Department of Statistics, Mathematics and Insurance,

Faculty of Commerce Kaferisheikh University

Abstract

If we consider that, the SUR model suffered from outliers, multicollinarity and it involved a degree of uncertainty associated with restrictions on the parameters. Three estimators introduced to avoid this problem. The first estimator , General Stochastic Restricted SUR Ridge Estimator with ridge parameter depends on resent robust cross validation instead of classic robust cross validation .The second estimator, General Stochastic Restricted SUR Robust Ridge Estimator which uses S-estimators with Ridge Regression under stochastic restriction and ridge parameter depends on resent robust cross validation . The third estimator, General Stochastic Restricted SUR Robust Estimator which depends on S-estimators. We introduced algorithm to compute General Stochastic Restricted SUR Ridge estimator, General Stochastic Restricted SUR Robust estimator and General Stochastic Restricted SUR Robust Ridge estimator. And we conducted a set of simulation study which used the ASE(average squared error) criterion to measure the goodness of fit at the several factors.

Key words: General Stochastic Restricted SUR Estimator, General Stochastic Restricted SUR Robust estimators, General Stochastic Restricted SUR Ridge Estimators, General Stochastic Restricted SUR Robust Ridge estimators.

[1,2,3,4] Department of Applied statistics, Institute of Statistical of Studies and Research ,Cairo University

1. Introduction

The seemingly unrelated regression (SUR) Model was introduced by zellner in 1962 which described as a set of linear regressions equations where the disturbance in the different equations was correlated. The SUR model contains many assumptions, one of them, the model without outliers which involves large residuals and often effect on the fettle least squares regression function.

In addition, the (SUR) model assumes the non-existence of multicollinearity which lead to that the standard errors of the estimates become extremely large .So, parallel, many studies discussed the problem of multicollinarity and outliers in SUR model separately . (Koenker and Portnoy,1990) used M-estimators to remove effect of outliers in SUR model, and (Bilodeau and Duchese ,2000) suggested S-estimator to deal with outliers for SUR model .

On the other hand , (Srivastava and Giles,1987) suggested ridge and general ridge estimator as a solution for multicollinarity . For features the canonical form, Alkhamisi (2007) suggested ridge estimator for SUR model in canonical form and chosen ridge parameter at several way to reach to efficiency estimator.

To reach the best ridge estimator, Firinguetti, (1997), Kibria (2003) suggested the different choice for ridge parameter to get the ridge estimator more efficiently. In contrast , there are many results to choice the penalty parameter in nonparametric regression .Stone (1974)chose ridge parameter which minimize classic cross validation criteria .Wang(1994) suggested minimizing the absolute cross validation to chose penalty parameter .Park(2005) used several form of robust cross validation in nonparametric regression , Tharmaratnam(2008) used classic cross validation to choose penalty parameter in penalized S-estimation. Kang-Mo Jung(2009) used robust cross validation to choose the ridge parameters in linear regression .

There is one study used robust cross validation to choose ridge parameter in SUR model. El-Houssainy et al (2010) suggested median cross validation to choose ridge parameter in SUR Ridge and SUR Robust Ridge regression .

On the other side, the parameter of SUR model may be restricted with degree of uncertain . The important study which deals with stochastic restricted in SUR model, suggested by Srivastava and Giles,(1987) which solved the problem of stochastic restricted without discussion the problem of multicollinearity and outliers at the same time . In fact ,we can't found any study used resent robust cross validation in SUR Ridge or SUR Robust Ridge Regression under stochastic restriction .

In this study , we propose robust cross validation for SUR Ridge , SUR Robust Ridge regression under stochastic restriction and SUR Robust regression under stochastic restriction

2. General Stochastic Restricted Estimator for SUR model

Consider the SUR model:

$$Y = X\,\beta + \varepsilon \qquad (1)$$

where $Y_i \in R^n$, i=1,2,…..,m ,n is number of observations , m is number of equation , $X_i \in R^{n\times k_i}$, i=1,2,.......,m, k_i is number of variables in i equation ,$k_i = k$, $\beta_i \in R^{k_i}$,i=1,2,……m and $\varepsilon_i \in R^n$.

This model has many assumption that X_i a full- column rank with $n\times k_i$,i=1,2,…..,m and the data without outliers .

We can be replaced the form (1) in the form of multiple regression

$$\ddot{Y} = \ddot{X}\ddot{\beta} + \ddot{\varepsilon} \qquad (2)$$

Where $\ddot{Y}$ is a matrix of $n \times m$ observations , $\ddot{\varepsilon}$ is a matrix of n residual , $\varepsilon_i \in R^n$

, $\ddot{\varepsilon}_i$ a $m\times 1$ vector, $\ddot{X}$ are $n \times mk$ matrix and coefficient and $\ddot{\beta}$ matrix are $mk\times m$ matrix

If we assume there are stochastic restriction of the system of equation in the following:

$$\begin{pmatrix} r_1 \\ r_2 \\ . \\ . \\ . \\ r_m \end{pmatrix} = \begin{pmatrix} \Lambda_1 & 0 & \ldots\ldots & 0 \\ 0 & \Lambda_2 & \ldots\ldots & 0 \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \ldots\ldots & \Lambda_m \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_m \end{pmatrix} + \begin{pmatrix} \upsilon_1 \\ \upsilon_2 \\ . \\ . \\ . \\ \upsilon_m \end{pmatrix}$$

where Λ_i is a $l_i\times k_i$ prior information of m equations ,l_i is number of restriction for ith equation , ,i=1,2,….,m ,r_i is a $l_i\times 1$ vector of with non-stochastic element , υ_i is a $l_i\times 1$ random disturbance vector.

We can write this form in stacked form as

$$r = \Lambda\,\beta + \upsilon \qquad (3)$$

where v distributed as

$$v \sim MN(\H{U}, \ddot{\Omega})$$

Where

$$\H{U}=E(v)=0$$

And
$$\ddot{\Omega} = E(vv') = \begin{pmatrix} \Omega_{11}I & \Omega_{12}I\ldots\ldots & \Omega_{1m}I \\ \Omega_{21}I & \Omega_{22}I\ldots\ldots & \Omega_{2m}I \\ . & . & . \\ . & . & . \\ \Omega_{m1}I & \Omega_{m2}I\ldots\ldots & \Omega_{mm}I \end{pmatrix} = \Omega\otimes I$$

Where

$$\Omega_{ij}=E(v_{is}\ ,v_{jh})$$
$$,i,j=1,2,\ldots\ldots m$$
$$,s,h=1,2,\ldots\ldots n$$

For (1) and (3)we design the model:

$$\begin{bmatrix} Y \\ r \end{bmatrix} = \begin{bmatrix} X \\ \Lambda \end{bmatrix} \beta + \begin{bmatrix} \varepsilon \\ v \end{bmatrix} \tag{4}$$

And ,we can write this form in stacked form as

$$Y_r = X_r\,\beta + \varepsilon_r \tag{5}$$

Where

$$E(\varepsilon_r) = E\begin{bmatrix} \varepsilon \\ v \end{bmatrix} = 0$$

,
$$E(\varepsilon_r\,\varepsilon_r) = E\begin{bmatrix} \varepsilon \\ v \end{bmatrix}\begin{bmatrix} \varepsilon' & v' \end{bmatrix} = E\begin{bmatrix} \varepsilon\varepsilon' & \varepsilon v' \\ v\varepsilon' & vv' \end{bmatrix} = \begin{bmatrix} \Sigma & 0 \\ 0 & \Omega \end{bmatrix} = \Sigma_r \tag{6}$$

Rewrite the formula (5) as

$$\ddot{Y}_r = \ddot{X}_r\ddot{\beta} + \ddot{\varepsilon}_r \tag{7}$$

Where $\ddot{Y}_r$ is a matrix of (n+l) × m observations and $\ddot{\varepsilon}_r$ is a matrix of (n+l) × m residual, $\varepsilon i \in R^n$

with a m×1 vector ,$\ddot{X}_r$ are (n+l) × mk matrix and coefficient $\ddot{\beta}$ are mk × m matrix
We can design the model (7) as :

$$\begin{pmatrix} \ddot{Y} \\ \ddot{r} \end{pmatrix} = \begin{pmatrix} \ddot{X} \\ \ddot{\Lambda} \end{pmatrix} \ddot{\beta}_r + \begin{pmatrix} \ddot{\varepsilon} \\ \ddot{v} \end{pmatrix} \quad (8)$$

Srivastava and Giles,(1987) suggested Stochastic Restricted estimator for SUR model defined as:

$$\beta_{\text{SR Ridge SUR}} = (X'X + \Lambda'\Lambda)^{-1}[X'Y + \Lambda'r] \quad (9)$$

Also Srivastava and Giles,(1987) suggested General stochastic restricted estimator for SUR model defined as:

$$\beta_{\text{GSR Ridge SUR}} = (X'(\sum^{-1} \otimes I_n)X + \Lambda'(\Omega^{-1} \otimes I_l)\Lambda)^{-1}[X'(\sum^{-1} \otimes I_n)Y + \Lambda'(\Omega^{-1} \otimes I_l)r] \quad (10)$$

3. General Stochastic Restricted SUR Robust estimators

Since the least squares estimators extremely sensitive for outliers. (Bilodeau and Duchese ,2000) suggested S-estimators for SUR model. .Extending of these methods ,if we applied S-estimators for (8), we can get General Stochastic Restricted SUR Robust estimators by

$$\min_{(\beta,\Sigma_r)} |\Sigma_r| \text{ subject } \frac{1}{n+l}\sum_{i=1}^{n+l} \rho\left[(\ddot{\varepsilon}_{ri}{}' \, \Sigma^{-1}\ddot{\varepsilon}_{ri})^{\frac{1}{2}}\right] = \kappa \quad (11)$$

Where

$$\ddot{\varepsilon}_{ri} = A_i\varepsilon_{ri} = A_i(Y_r - X_r\beta)$$

Then we minimize

$$L = \log|\Sigma_r| - \lambda_{rs}\left[\frac{1}{n+l}\sum_{i=1}^{n+l}\rho\left[\{\varepsilon_{ri}'A_i'\Sigma^{-1}A_i\varepsilon_{ri}\}^{\frac{1}{2}}\right] - \varsigma\right]$$

(12)

Where ϛ ,is constant, can be chosen such that ϛ =EΦ[ρ(ε)] ,Φ is standard normal distribution Φ=N(0,1) and Ai=diag[a_i,……,a_i] be a (n+l) m× m matrix where a_i is the vector with one in position i and zero elsewhere .

we can choice ρ as Tukey's biweight function (Beaton and Tukey,1974) :

$$\rho(x) = \begin{cases} x^2/2 - x^4/2c^2 + x^6/6c^4 & |x| \le c \\ x^2/6 & |x| \ge c \end{cases}$$

Where c is positive number yields percentage of breakdown point and Ḱ depend on c .when choice c=2.56 then $\varsigma = 0.3278$ and breakdown point is 30% and when c=1.547 , ϛ became 0.1995 and breakdown point is 50%. On the other hand , there are a Inverse relationship between efficiency of the S-estimator under a Gaussian model and breakdown .If breakdown reach 50% and 30%,the efficiency of the S-estimator under a Gaussian model 28.7% and 66.1% respectively.

Proposition (1)

The General Stochastic Restricted SUR Robust estimators(new) is

$$\beta_{\text{Robust SUR}} = \left(X_r'\left(\Sigma_r^{-1} \otimes W_{r\varepsilon}\right)X_r\right)^{-1} X_r'\left(\Sigma_r^{-1} \otimes W_{r\varepsilon}\right)Y_r$$

(13)

Where

$$\sum_r = m_r(\ddot{Y}_r - \ddot{X}_r\ddot{\beta}_r)'W_{r_\varepsilon}(\ddot{Y}_r - \ddot{X}_r\ddot{\beta}_r) / \sum_{i=1}^{n+l}\upsilon(w_{r_i})$$

(14)

$$w_{r_i} = \varepsilon_r'A_i'\Sigma_r^{-1}A_i\varepsilon_r$$

$$\sum_{i=1}^{n+l} \frac{\rho'(w_{r_i})}{w_{r_i}} A_i{}' \Sigma_r{}^{-1} A_i' = \left(\Sigma_r{}^{-1} \otimes W_{r_\varepsilon}\right)$$

$$, W_{r_\varepsilon} = \text{diag}(w_{r1}, w_{r2}, \ldots\ldots, w_{r(n+l)})$$

For (6),(13) we get :

$$\beta_{\text{GSRRobust SUR}} = \left([X \quad \Lambda]' \left[\begin{pmatrix} \Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \otimes W_{r_\varepsilon} \right] \begin{bmatrix} X \\ \Lambda \end{bmatrix} \right)^{-1} [X \quad \Lambda]' \left[\begin{pmatrix} \Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \otimes W_{r_\varepsilon} \right] \begin{bmatrix} Y \\ r \end{bmatrix}$$

$$\beta_{\text{GSRRobust SUR}} = \left[X' \left(\Sigma^{-1} \otimes W_{r_\varepsilon} \right) X + \Lambda' \left(\Omega^{-1} \otimes W_{r_\varepsilon} \right) \Lambda \right]^{-1}$$

$$\left[X' \left(\Sigma^{-1} \otimes W_{r_\varepsilon} \right) Y + \Lambda' \left(\Omega^{-1} \otimes W_{r_\varepsilon} \right) r \right]$$

(15)

$$, \lambda_{rsij} = -m_r \left(\frac{1}{2(n+l)} \sum_{i=1}^{n+l} \rho'[W_{r_\varepsilon}] W_{r_\varepsilon} \right)^{-1}$$

(16)

Where m_r is number of equation in (7),l is a number of restricted

4. General Stochastic Restricted SUR Ridge Estimators

To get Stochastic Restricted SUR Ridge Estimators we minimize

$$L = \min_{\beta \in R^{K \times m}} (Y_r - X_r \beta)'(Y_r - X_r \beta) + \lambda_{rd} \sum_{i=1}^{m} \sum_{j=1}^{K} \beta_{ij}^2$$

(17)

$$\beta_{\text{Ridge SUR}} = (X_r{}' X_r + \lambda_{rd} I_{km})^{-1} X_r{}' Y_r$$

(18)

The General Stochastic Restricted SUR Ridge estimator for (5) gets by minimize

$$L = \min_{\beta \in R^{K\times m}} (Y_r{}^* - X_r{}^*\beta)'(Y_r{}^* - X_r{}^*\beta) - \lambda_{rd} \sum_{i=1}^{m} \sum_{j=1}^{k} \beta_{ij}{}^2$$

(19)

Where $Y_r{}^* = Y_r\left(\sum_r{}^{-1/2} \otimes I_{(n+l)}\right)$, $X_r{}^* = X_r\left(\sum_r{}^{-1} \otimes I_{(n+l)}\right)$

Proposition (2)

The General Stochastic Restricted SUR Ridge estimator (new)

$$\beta_{\text{GSRRidge SUR}} = \left(X_r{}'\left(\sum_r{}^{-1} \otimes I_{(n+l)}\right)X_r + \lambda_{rd} I_{km}\right)^{-1} X_r{}'\left(\sum_r{}^{-1} \otimes I_{(n+l)}\right)Y_r$$

(20)

For (6) ,(20) we get

$$\beta_{\text{GSRRidge SUR}} = \left([X \quad \Lambda]' \left[\begin{pmatrix} \Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \otimes \begin{pmatrix} I_n & 0 \\ 0 & I_l \end{pmatrix}\right] \begin{bmatrix} X \\ \Lambda \end{bmatrix} + \lambda_{rd} I_{km}\right)^{-1} [X \quad \Lambda]' \left[\begin{pmatrix} \Sigma & 0 \\ 0 & \Omega \end{pmatrix}^{-1} \otimes \begin{pmatrix} I_n & 0 \\ 0 & I_l \end{pmatrix}\right] \begin{bmatrix} Y \\ r \end{bmatrix}$$

Then

$$\beta_{\text{GSR Ridge SUR}} = (X'(\Sigma^{-1} \otimes I_n)X + \Lambda'(\Omega^{-1} \otimes I_l)\Lambda + \lambda_{rd} I_{km})^{-1}$$

$$[X'(\Sigma^{-1} \otimes I_n)Y + \Lambda'(\Omega^{-1} \otimes I_l)r]$$

(21)

We chose λ_{rd} for Stochastic Restricted General SUR Ridge estimator by develop a version of the Ruppert(2003) criterion to apply on SUR model . And then, we suggest (new) chosen λ(new) via minimize:

$$1 - \text{SGCV}(\lambda_{rd}) = (n+l)m \sum_{j=1}^{m} \sum_{i=1}^{n+l} \frac{\left(Y_{rij}{}^* - X_{rij}{}^*\hat{\beta}\right)^2}{\left((n+l)m - \text{tr}\left(H_r{}^*(\lambda_{rd})_{i,i\,j}\right)\right)^2}$$

(22)

Where

$$H_r^{*}(\lambda_d)_{i,i\ j} = X_{rj}^{*}\left(X_{rj}^{*\prime}X_{rj}^{*} + \lambda_{rd}I_{mk}\right)^{-1}X_{rj}^{*\prime}$$

For access to the criterion to choose λ_{rd} which take outliers in mind, we minimize the robust Cross Validation (new)

$$2 - \quad \mathrm{SGMCV}(\lambda_{rd}) = \mathrm{median}\left[\frac{\left(Y^{*}_{rij} - X^{*}_{rij}\hat{\beta}\right)^2}{\left(1 - \frac{\mathrm{tr}\left(H_r^{*}(\lambda_{rd})_{i,i\ j}\right)}{(n+l)m}\right)^2}\right]$$

i=1,2,……(n+l), j=1,2,……m

(23)

5. General Stochastic Restricted Robust Ridge SUR model

We can get Stochastic Restricted General SUR Robust Ridge estimator for (5)by minimize

$$L = \log|\Sigma_r| - \lambda_{rs}\left[\frac{1}{n+l}\sum_{i=1}^{n+l}\rho\left[\left\{\varepsilon_r{}'A_i{}'\Sigma_r{}^{-1}A_i\varepsilon_r\right\}^{\frac{1}{2}}\right] - \varsigma\right] + \lambda_{rd}\sum_{i=1}^{m}\sum_{j=1}^{k}\beta_{ij}{}^2$$

(24)

Where

$$\ddot{\varepsilon}_{ri} = A_i\varepsilon_r = A_i(Y_r - X_r\beta)$$

,$A_i = \mathrm{diag}[a_i, \dots \dots, a_i]$ be a (n+l)m×m matrix where a_i is the vector with one in position i and zero elsewhere.

Proposition (3)

If l+n=k, then the Stochastic Restricted General SUR Robust Ridge estimator (new)

$$\beta_{\text{GSRRobust Ridge SUR}} = \left(X_r'\left(\Sigma_r^{-1} \otimes W_{r_\varepsilon}\right)X_r + R_r I_{km}\right)^{-1} X_r'\left(\Sigma_r^{-1} \otimes W_{r_\varepsilon}\right)Y_r$$

$$= (X'(\Sigma^{-1} \otimes W_\varepsilon)X + \Lambda'(\Omega^{-1} \otimes W_\upsilon)\Lambda + R_r I_{km})^{-1} \qquad (25)$$

Where

$$R_r = 2(n+l)\lambda_{rs}^{-1}\lambda_{rd} \qquad (26)$$

$$,\Sigma_r = m_r(\ddot{Y}_r - \ddot{X}_r\ddot{\beta}_r)'W_{r_\varepsilon}(\ddot{Y}_r - \ddot{X}_r\ddot{\beta}_r)/\sum_{i=1}^{n+l}\upsilon(w_{r_i}),\lambda_{rsij} = -m_r\left(\frac{1}{2(n+l)}\sum_{i=1}^{n+l}\rho'[W_{r_\varepsilon}]\,W_{r_\varepsilon}\right)^{-1}$$

For (6),(26) we get :

$\beta_{\text{GSRRobust Ridge SUR}}$

$$= \left([X \quad \Lambda]'\left[\begin{pmatrix}\Sigma & 0\\ 0 & \Omega\end{pmatrix}^{-1} \otimes W_{r_\varepsilon}\right]\begin{bmatrix}X\\ \Lambda\end{bmatrix} + R_r I_{km}\right)^{-1}[X \quad \Lambda]'\left[\begin{pmatrix}\Sigma & 0\\ 0 & \Omega\end{pmatrix}^{-1} \otimes W_{r_\varepsilon}\right]\begin{bmatrix}Y\\ r\end{bmatrix} \qquad (27)$$

Then

$$\beta_{\text{GSRRobust Ridge SUR}} = \left(X'(\Sigma^{-1} \otimes W_{r_\varepsilon})X + \Lambda'\left(\Omega^{-1} \otimes W_{r_{\varepsilon_\upsilon}}\right)\Lambda + R_r I_{km}\right)^{-1}$$

$$\left[X'(\Sigma^{-1} \otimes W_{r_\varepsilon})Y + \Lambda'(\Omega^{-1} \otimes W_{r_\varepsilon})r\right]$$

We chose λ_{rd} for Stochastic Restricted General SUR Robust Ridge estimator by minimize

$$SGMCV(\lambda_{rd}) = \text{median}\left[\frac{\left(Y^{**}{}_{rij} - X^{**}{}_{rij}\hat{\beta}\right)^2}{\left(1 - \frac{tr\left(H_r{}^{*}(\lambda_{rd})_{i,i\,j}\right)}{(n+l)m}\right)^2}\right]$$

(n+l,) , j=1,

(28)

Where

$$H_r{}^{**}(\lambda_{rd})_{i,i\,j} = X_{rj}^{**}\left(X_{rj}^{**\prime}X_{rj}^{**} + R_r I_{mk}\right)^{-1}X_{rj}^{**\prime}$$

and

$$Y_r{}^{**} = \left(\Sigma_r{}^{-1/2} \otimes W_{r\varepsilon}\right)Y_r \quad , X_r{}^{**} = \left(\Sigma_r{}^{-1/2} \otimes W_{r\varepsilon}\right)X_r$$

6. Simulation study

6.1 Simulation settings

We setting the simulations by assumed that :

1- The explanatory variables X and coefficient of restriction parameters Λ, was distributed normality according to $MVN_k(0,\Sigma_X)$, $MVN_k(0,\Sigma_\Lambda)$, where Σ_X and ,Σ_Λ the variance covariance matrix of X and Λ respectively which defined as diag(Σ_X)= diag(Σ_Λ)=1 , off-diag(Σ_X)= ρ_X and off-diag(Σ_Λ)= ρ_Λ . Σ_X and ,Σ_Λ were chosen as $\Sigma_{X(high)}$ then ρ_{Xij}= 0.90, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\Lambda\ (high)}$ then $\rho_{\Lambda ij}$= 0.90, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji or $\Sigma_{X(low)}$ then ρXij= 0.20, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\Lambda\ (low)}$ then $\rho_{\Lambda ij}$= 0.20, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji

2- We generate the percentages of error between equations ε and percentages of error of restriction equation υ when data without outliers according to $MVN_m(0,\Sigma_\varepsilon)$, MVN_k (0, Σ_υ) where Σ_ε and ,Σ_υ the variance covariance matrix of ε and υ respectively which defined as diag(Σ_υ)= diag(Σ_ε)=1 , off-diag(Σ_ε)= ρ_ε and off-diag(Σ_υ)= ρ_υ . Σ_ε and ,Σ_υ were chosen as

$\Sigma_{\varepsilon\,(high)}$ then $\rho_{\varepsilon ij}$= 0.90, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\upsilon\,(high)}$ then $\rho_{\upsilon ij}$= 0.90, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji or $\Sigma_{\upsilon\,(low)}$ then $\rho_{\varepsilon ij}$= 0.20, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\upsilon\,(low)}$ then $\rho_{\upsilon ij}$= 0.20, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji.

3- We generate the percentages of error between equations ε and percentages of error of restriction equation υ when data within outliers according to $MVN_{m=2}$([10 5],,Σ_ε), $MVN_{m=2}$ ([10 15],, Σ_υ) where Σ_ε and ,Σ_υ the variance covariance matrix of ε and υ respectively which defined as diag(Σ_υ)= diag(Σ_ε)=1 , off-diag(Σ_ε)= ρ_ε and off-diag(Σ_υ)= ρ_υ . Σ_ε and ,Σ_υ were chosen as $\Sigma_{\varepsilon\,(high)}$ then $\rho_{\varepsilon ij}$= 0.95, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\upsilon\,(high)}$ then $\rho_{\upsilon ij}$= 0.95, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji or $\Sigma_{\upsilon\,(low)}$ then $\rho_{\varepsilon ij}$= 0.10, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j ,ij=ji and $\Sigma_{\upsilon\,(low)}$ then $\rho_{\upsilon ij}$= 0.10, i=1,2 ,…,k, ,j=1,2…..,k,, i≠j.

4- The number of variable , the number of observations and the number of restricteds as a same form all equations .

5- For the condition k=n+l which mean if the matrix X and Λ are singular matrix , the matrix X_r may be nonsingular. This fact leads to the use of Least Squares method as a initial point is impossible. So ,we choose initial point arbitrary then we choose it by two way :

i- If k=20, β_0=[1,2,……,20], and we assumed n=10and l=10.

ii- If k=40, β_0=[1,2,……,40] , and we assumed n=20and l=20 .

6- The presence of outliers , it can be chosen as10%,20%,40%.

7- The criterion to Compare and measure the goodness of estimators of β in the second simulation are quantified by average squared error(ASE) :

$$\mathrm{ASE} = \frac{1}{(n+l)m}\left(X_r\beta_0 - \left(X_r\hat{\hat{\beta}}_r\right)\right)'\left(X_r\beta_0 - \left(X_r\hat{\hat{\beta}}_r\right)\right) \tag{29}$$

Where β_0 is initial point of β and $\hat{\hat{\beta}}_r$ any Stochastic Restricted for β.

6.2 Algorithm:

The steps of algorithm to compute general Stochastic Restricted SUR Ridge estimator , general Stochastic Restricted SUR Robust estimators and general Stochastic Restricted SUR Robust Ridge estimator for (SUR) model are:

a- Generate the variable$\left(X^{(i)}, Y^{(i)}\right)$,i=1,2,……m from ($MVN_k(0,\Sigma_X)$, $MVN_m(0,\Sigma_Y)$),

b- Generate the variable$\left(\Lambda^{(i)}, r^{(i)}\right)$,i=1,2,……m from ($MVN_k(0,\Sigma_\Lambda)$, $MVN_m(0,\Sigma_r)$), then design matrix (X_r, Y_r) .

c- let $\hat{\beta}_{j,0}, \hat{\beta}_{rj,0}$ be initial candidates.

d- Generate the vector random error $\varepsilon^{*(i)}$,i=1,2,…… m from MVNm(0,Σε), and Generate the different percentages $\varepsilon^{**(i)}$of outliers from MVNm=2([10 5],Σε) then design matrix (ε) .

e- Generate the error of restriction equation $\upsilon^{*(i)}$,i=1,2,…… m from MVNm(0,Συ), and Generate the different percentages $\upsilon^{**(i)}$of outliers from MVNm=2([10 15],Συ) then design matrix (ε_r) .

f- Set S=0,and get the following steps:

i- Compute $W_{r\varepsilon}^{(S)}$, ${\Sigma_r}^{(S)}$, ${R_r}^{(S)}$

ii- Let

$$\beta_{\text{GSRRobust Ridge SUR}}{}^{(S+1)} = \left(X_r'\left({\sum_r}^{-1^{(S)}} \otimes {W_{r\varepsilon}}^{(S)}\right)X_r + {R_r}^{(S)} I_{km}\right)^{-1} \left[X_r'\left({\sum_r}^{-1^{(S)}} \otimes {W_{r\varepsilon}}^{(S)}\right)Y_r\right]$$

iii- If either S equal the maximum number of iterations

or $$\left\|\beta^{(S)}_{\text{GSRRobustRidge SUR}} - \beta^{(S+1)}_{\text{GSRRobust RidgeSUR}}\right\| = {}^\circ C\beta^{(S)}_{\text{RGSRobust RidgeSUR}}$$

where °C > 0 is affixed small constant

Then $\beta^{T}_{\text{GSRRobustRidge SUR}} = \beta^{(S)}_{\text{GSRRobustRidge SUR}}$

and break

Else S←S+1

j- Set h=0,and get the following steps:

i- compute ${W_{r\varepsilon}}^{(h)}$, ${\Sigma_r}^{(h)}$

ii- let ${\beta_{GSRRobust\ SUR}}^{(S+1)} =$

$$\left[X_r'\left(\Sigma^{-1^{(S)}} \otimes W_\varepsilon{}^{(S)}\right)X_r\right]^{-1}\left[X_r'\left(\Sigma^{-1^{(S)}} \otimes W_\varepsilon{}^{(S)}\right)Y_r\right]$$

iii- If either h equal the maximum number of iterations

or

$$\left\|\beta^{(h)}_{GSRRobust\ SUR} - \beta^{(h+1)}_{GSRRobust\ SUR}\right\| = {}^{\circ}C\beta^{(h)}_{GSRRobust\ SUR}$$

where $^{\circ}C > 0$ is affixed small constant

Then $\beta^{T}_{GSRRobust\ SUR} = \beta^{(h)}_{GSRRobust\ SUR}$ and

break

Else h←h+1

k- let $\beta_{GSR\ Ridge\ SUR}{}^{(S+1)} =$

$$\left[X_r'\left({\Sigma_r}^{-1^{(S)}} \otimes I_{(n+l)}\right)X_r + \lambda_{dr}{}^{(S)}{}_j I_{km}\right]^{-1}\left[X_j'\left({\Sigma_r}^{-1^{(S)}} \otimes I_{(n+l)}\right)Y_r\right]$$

The value °C , h , S were chosen as 10-5,100,100

The matlab code for algorithm is available from the authors.

6.3 The simulation results

We can summarize the values of the MASE(median average squared error) for general Stochastic Restricted Ridge(SGCV) ,general Stochastic Restricted Ridge(SGMCV) ,Robust and general Stochastic Restricted Robust Ridge estimator for SUR model in the table (1),(2).

In table (1) , we show the result of ASE for General Stochastic Restricted Robust, General Exact Restricted Ridge, and General Stochastic Restricted Robust Ridge estimators at (k=20)(n=10)(l=10).

In table (2),we take the same factor form table (1) except if we assume the number of variables equal 40 , the number of observations equal 20 and it equal the number of restriction .

The result of table (1) show that ,In case (1),(2) Stochastic Restricted General SUR Robust Ridge estimator is more efficient than other estimator and, in case (3),(4) ,Stochastic Restricted General SUR Robust estimator is more efficient than other estimator.

The result of table (2) show that, Stochastic Restricted General SUR Robust estimator is more efficient when the presence of outliers reach to 40% for all cases and Stochastic Restricted General SUR Robust Ridge estimator is more efficient when the presence of outliers reach to 20% for all cases. In table (1),(2),The Stochastic Restricted General SUR Robust Ridge estimator is the worst efficient among the estimators at all cases.

7. Conclusions:

To solve the problem of multicollinarity and outliers in SUR model when there is degree of uncertainty associated with restrictions on the parameters , we present three new estimators ,the first is general Stochastic Restricted ridge estimators with ridge parameter depend on robust cross validation and the second general Stochastic Restricted Robust Ridge estimators which depend on the S-estimation in ridge regression with the new robust cross validation . The third estimator , general Stochastic Restricted Robust estimators which used S-estimator as a robust method .We set the simulation to compare the estimators depend on ASE(average squared error) criterion to determined which estimator is more efficient for other estimators. Many factors taken into account in the simulation ,like degree of correlation between variables and between restrictions ,correlation between errors for equation and equation of restriction ,number of variables ,number of observations ,number of restricted and presence of outliers , it can be chosen as10%,20%,40%.

The results of the simulation indicate that, in the case of decrease the number of variables, General Stochastic Restricted Robust Ridge estimators is more efficient among the estimators , when correlated between variables , correlation between coefficient of restriction, correlated between errors between equations and correlated between restricted errors between equations are high .If the Previous relationships were low ,the General Stochastic Restricted Robust estimator is more efficient among the estimators.

When the number of variables increase , if the presence of outliers reach to 20% and 40% for all cases , sequentially ,General Stochastic Restricted Robust Ridge estimator and General Stochastic Restricted Robust estimator are more efficient among the estimators,.

The Stochastic Restricted General SUR Robust Ridge estimator is the worst efficient among the estimators at all cases.

References

[1] Alkhamisi M.A. & Shukur, G., (2007). "Developing Ridge Parameters for SUR Models,", Working Paper Series in Economics and Institutions of Innovation 80, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.

[2] Beaton A. and Tukey, J. (1974). "The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data". Technometrics, Vol 16.

[3] Bilodeau M. and Duchesne P.,(1991), "Robust estimation of the SUR model ",The Canadian Journal of Statistics, Vol 28.

[4] Firinguetti L. (1997), “Ridge Regression in the Context of a System of Seemingly Unrelated Regression Equations”, J. Statist. Comput. Simul., Vol. 56, pp. 145-62

[5] Hoerl A. E. and Kennard, R. W. (1970a). “Ridge Regression: Biased Estimation for Non-Orthogonal Problems”, Tech., Vol 12, 55-67.

[6] Jung K.(2009),"Robust cross validation in ridge regression ",J. Appl. Math &Informatics,Vol.27,No.3-4,pp903-908.

[7] Koenker R., & Portnoy, S. (1990)." M-estimation of multivariate regressions". J. Amer .Statist. Assoc, Vol. 63.

[8] Kibria B.& Golam M. (2003). “Performance of Some New Ridge Regression Estimators ”,Commun. Statist. B(2), Vol. 32, 419-435.

[9] Lopuhaa H. P. (1989), "On the Relation Between S-estimators and M-estimators of Multivariate Location and Covariance," The Annals of Statistics, Vol.17,1662-1683.

[10] Nagakura D. (2008) "A note on the relationship between the information matrix test and a score test for parameter constancy" Economics Bulletin, Vol. 3, No. 5 , 1-7.

[11] Park D. (2005), "Robust cross validation score", The Korean Communications In Statistics, Vol 12 , No. 413-423.

[12] Rousseeuw P. J., and Yohai, V. J. (1984), "Robust Regression by Means of S-estimators, in Robust and Nonlinear Time Series Analysis", eds. J. Franke , W. Statist. Assoc., Vol. 85.

[13] Rousseeuw P. J., & Leroy, A. M. (1987). Robust Regression and Outlier Detection. John Wiley, New York.

[14] Ruppert D. (1992)." Computing S estimators for regression and multivariate location/dispersion". J. of computational and graphical statist., Vol. 32,.

[15] Ruppert D., Wand, M., and Carroll, R. (2003)." Semiparametric Regression" Cambridge University Press, Cambridge, UK.

[16] Srivastava V. & Giles, D. (1987). “Seemingly Unrelated Regression Equations Models: Estimation and Inference” ,New York: Marcel Dekker.

[17] Stone M., (1974), "Cross-validatory choice and assessment of statistical predictions ", J. of the Royal Statist. Soc., B, 36 ,111-147.

[18] El- Houssainy A., Sayed M. , Alaa A., Naglaa A. , Tarek m. , (2011)" Robust Cross Validation in SUR Ridge Estimators and SUR Robust Ridge Estimators", Journal of statistic theory and application ,Vol.10,No.1.

[19] Tharmaratnam K. & Claeskens, G (2008)"S-Estimation for Penalized Regression Splines", Compute. Statist., Vol.30.

[20] Wang F. & Scott D. , (1994),The L1 method for robust nonparametric regression, J. of the Amer. Statist. Assoc., 89, 249-260.

[21] Zellner A. (1962): “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias”, J. Am. Statist. Assoc., Vol. 57..

Appendix: Proofs of the Proposition

Proofs of the proposition (1)

If we differentiate (24) with respect to β and equalize the result to zero then:

$$-\frac{\lambda_{rs}}{2(\,n+l)}\sum_{i=1}^{n}\rho'\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{-\frac{1}{2}}\right]$$

$$2[X_r{}'A_i{}'\Sigma_r^{-1}A_iX_r\beta_i - X_r{}'A_i{}'\Sigma_r^{-1}A_iY_r] + 2\lambda_{rd}\sum_{i=1}^{m}\sum_{j=1}^{K}\beta_{ij} = 0$$

$$\frac{\lambda_{rs}}{2(\,n+l)}\sum_{i=1}^{n}\rho'\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{-\frac{1}{2}}\right]$$

$$2[-X_r{}'A_i{}'\Sigma_r^{-1}A_iY_r + X_r{}'A_i{}'\Sigma_r^{-1}A_iX_r\beta_i] + 2\lambda_{rd}\sum_{i=1}^{m}\sum_{j=1}^{K}\beta_{ij} = 0$$

$$-\frac{\lambda_{rs}}{(\,n+l)}\sum_{i=1}^{(\,n+l)}\frac{\rho'(w_{r_i})}{w_{r_i}}X_r{}'A_i{}'\Sigma_r^{-1}A_iY_r + \frac{\lambda_{rs}}{(\,n+l)}\sum_{i=1}^{(\,n+l)}\frac{\rho'(w_{r_i})}{w_{r_i}}X_r{}'A_i{}'\Sigma_r^{-1}A_iX_r\beta$$
$$+2\lambda_{rd}\sum_{i=1}^{m}\sum_{j=1}^{K}\beta_{ij} = 0$$

(30)

Where

$$w_{r_i}{}^2 = \varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r = \ddot{\varepsilon}_r{}'\Sigma_r^{-1}\ddot{\varepsilon}_r$$

Then

$$\sum_{i=1}^{(\,n+l)}\frac{\rho'(w_{r_i})}{w_{r_i}}A_i{}'\Sigma_r^{-1}A_i = (\Sigma_r^{-1}\otimes W_{r\varepsilon})$$

(31)

and

$$W_{r\varepsilon}=\mathrm{diag}\left(\frac{\rho'(w_{r_1})}{w_{r_1}},\frac{\rho'(w_{r_2})}{w_{r_2}},\ldots\ldots,\frac{\rho'\left(w_{r(n+l)}\right)}{w_{r(n+l)}}\right),$$

For (30),(31) we get

$$\frac{\lambda_{rs}}{(\,n+l)}X_r{}'(\Sigma_r^{-1}\otimes W_{r\varepsilon})Y_r = \frac{\lambda_{rs}}{(\,n+l)}X_r{}'(\Sigma_r^{-1}\otimes W_{r\varepsilon})X_r\beta + 2\lambda_{rd}I_{km}\beta$$

$$\frac{\lambda_{rs}}{(n+l)} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})Y_r = \left(\frac{\lambda_{rs}}{(n+l)} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})X_r + 2\lambda_{rd} I_{km}\right)\beta$$

Then

$$\beta_{GSRRobust\ Ridge\ SUR} = \left(\frac{\lambda_{rs}}{(n+l)} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})X_r + 2\lambda_{rd} I_{km}\right)^{-1} \frac{\lambda_{rs}}{(n+l)} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})Y_r$$

Let K= $(n+l)$ then

$$\beta_{GSRRobust\ Ridge\ SUR} = (X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})X_r + RI_{km})^{-1} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})Y_r$$

$$= (X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})X_r + RI_{km})^{-1} X_r'(\Sigma_r^{-1} \otimes W_{r\varepsilon})Y_r$$

Where

$$R = 2(n+l)\lambda_{rs}{}^{-1}\lambda_{rd}$$

We can get λ_{rs}, λ_{rd} and Σ_r^{-1} by differentiate (24) with respect to Σ_r^{-1} and equalize the result to zero then:

$$\Sigma_r^{-1} - \frac{\lambda_{rs}}{2(n+l)} \sum_{i=1}^{(n+l)} \frac{\rho'\left[\{\varepsilon_r' A_i' \Sigma_r^{-1} A_i \varepsilon_r\}^{\frac{1}{2}}\right]}{\{\varepsilon_r' A_i' \Sigma_r^{-1} A_i \varepsilon_r\}^{-\frac{1}{2}}} \frac{\partial\, \varepsilon_r' A_i' \Sigma_r^{-1} A_i \varepsilon_r}{\partial \Sigma_r} = 0 \tag{32}$$

Lemma 1 : (see ,Nagakura (2008))

$$\frac{\partial\, \varepsilon_r' A_i' \Sigma_r^{-1} A_i \varepsilon_r}{\partial \Sigma_r} = \begin{bmatrix} -h_{11} & \cdots & -h_{1m} \\ \vdots & \ddots & \vdots \\ -h_{m1} & \cdots & -h_{mm} \end{bmatrix} = -\Sigma_r^{-1} A_i \varepsilon_r \varepsilon_r' A_i' \Sigma_r^{-1} \tag{33}$$

Where h_{ij} is (i , j) element of the matrix $\Sigma_r^{-1} A_i \varepsilon_r \varepsilon_r' A_i' \Sigma_r^{-1}$

Proof lemma 1

$$\frac{\partial\, \varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r}{\partial\Sigma_{r_{ij}}} = \mathrm{tr}\left[\varepsilon_r{}'A_i\frac{\partial\Sigma_r^{-1}}{\partial\Sigma_{r_{ij}}}A_i\varepsilon_r\right]$$

$$= -\mathrm{tr}\left[\varepsilon_r{}'A_i{}'\Sigma^{-1}\frac{\partial\Sigma_r}{\partial\Sigma_{rij}}\Sigma^{-1}A_i\varepsilon_r\right]$$

$$= -\mathrm{tr}\left[\frac{\partial\Sigma_r}{\partial\Sigma_{rij}}\Sigma_r^{-1}A_i\varepsilon_r\varepsilon_r{}'A_i{}'\Sigma_r^{-1}\right]$$

$$= -h_{ij}$$

Where Σ_{ij} is a (i , j) element of Σ , $\frac{\partial\Sigma_r}{\partial\Sigma_{rij}}$ is m×m matrix whose (i , j) element is one and all other elements are zero and $\mathrm{tr}[\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r] = \varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r$, (since $\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r$ is a scalar).

For (33), we have (32).

Then for emma1 ,we can write the equation (32) as

$$,\Sigma_r^{-1} + \frac{\lambda_{rs}}{2(\,n+l)}\sum_{i=1}^{(\,n+l)}\frac{\rho'\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]}{\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}}\Sigma_r^{-1}\varepsilon_r{}'A_i{}'\varepsilon_r A_i\Sigma_r^{-1} = 0$$

If Σ is symmetric positive definite matrix, a nonsingular matrix M can be found such that Σ =MM´, then Σ-1= M´-1M-1 and

$$M'^{-1}M^{-1} + \frac{\lambda_{rs}}{2(\,n+l)}\sum_{i=1}^{(\,n+l)}\frac{\rho'\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]}{\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}}M'^{-1}M^{-1}A_i{}'\varepsilon_r{}'\varepsilon_r A_i M'^{-1}M^{-1} = 0 \tag{34}$$

Pre –multiply (34) by M´ and post –multiply (34) by M then this equation can be written as:

$$I_m + \frac{\lambda_{rs}}{2(\,n+l)}\sum_{i=1}^{(\,n+l)}\rho'\left[\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]\{\varepsilon_r{}'A_i{}'\Sigma_r^{-1}A_i\varepsilon_r\}^{-\frac{1}{2}}M'^{-1}A_i{}'\varepsilon_r{}'\varepsilon_r A_i M^{-1} = 0 \tag{35}$$

Take the trace of (35)

$$m+\frac{\lambda_s}{2(n+l)}\sum_{i=1}^{n}\frac{\rho'\left[\{\varepsilon_r'A_i'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]}{[\varepsilon_r'A_i'M'^{-1}M^{-1}A_i\varepsilon_r]^{\frac{1}{2}}}\operatorname{tr}[M^{-1}A_i'\varepsilon_r'\varepsilon_rA_iM'^{-1}]=0 \tag{36}$$

Then

$$m+\frac{\lambda_s}{2(n+l)}\sum_{i=1}^{n}\frac{\rho'\left[\{\varepsilon_r'A_i'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right]}{[\varepsilon_r'A_i'M'^{-1}M^{-1}A_i\varepsilon_r]^{\frac{1}{2}}}[\varepsilon_r'A_i'M'^{-1}M^{-1}A_i\varepsilon_r]=0$$

Then

$$m+\frac{\lambda_s}{2n}\sum_{i=1}^{n}\rho'\left[\{\varepsilon_r'A_i'\Sigma_r^{-1}A_i\varepsilon_r\}^{\frac{1}{2}}\right][\varepsilon_r'A_i'\Sigma_r^{-1}A_i\varepsilon_r]^{\frac{1}{2}}=0$$

$$m+\frac{\lambda_s}{2(n+l)}\sum_{i=1}^{(n+l)}\rho'[w_{ri}]\,w_{ri}=0$$

$$\lambda_{rsij}=-m\left(\frac{1}{2(n+l)}\sum_{i=1}^{(n+l)}\rho'[w_{ri}]\,w_{ri}\right)^{-1} \tag{37}$$

i=1,2,……n , j=1,2,……,m

Using (37) in (35) then:

$$I_m-\frac{m\left(\frac{1}{2(n+l)}\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)^{-1}}{2(n+l)}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}M^{-1}A_i'\varepsilon_r'\varepsilon_rA_iM'^{-1}=0$$

$$I_m-m\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)^{-1}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}M^{-1}A_i'\varepsilon_r'\varepsilon_rA_iM'^{-1}=0 \tag{38}$$

Multiply (38)from $\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)$ then

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_r\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}M^{-1}A_i'\varepsilon_r'\varepsilon_rA_iM'^{-1}=0$$

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_rM^{-1}M'^{-1}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}A_i'\varepsilon_r'\varepsilon_rA_i=0$$

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_r\Sigma_r^{-1}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}A_i'\varepsilon_r'\varepsilon_rA_i=0$$

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_r\Sigma_r^{-1}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}\ddot{\varepsilon}_r'\ddot{\varepsilon}_r=0$$

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_r\Sigma_r^{-1}\sum_{i=1}^{(n+l)}\frac{\rho(w_{ri})}{w_{ri}}(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})'(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})=0$$

$$\Sigma_r\left(\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}\right)-m\Sigma_r\Sigma_r^{-1}(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})'W_{r\varepsilon}(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})=0$$

Then

$$\Sigma_r=m(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})'W_{r\varepsilon}(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})/\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}$$

And then

$$\Sigma_r=m(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})'W_{r\varepsilon}(\ddot{Y}_r-\ddot{X}_r\ddot{\beta})/\sum_{i=1}^{(n+l)}\upsilon(w_{ri})$$

Where:

$$,\ \sum_{i=1}^{(n+l)}\upsilon(w_{ri})=\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}-\sum_{i=1}^{(n+l)}\rho[w_{ri}]-\varsigma$$

The term $-\sum_{i=1}^{(n+l)}\rho[w_{ri}]-\varsigma$ is added to $\sum_{i=1}^{(n+l)}\rho(w_{ri})\,w_{ri}$ because merely substituting λ_{rsij} ,i=1,2,……n , j=1,2,……,m into(37) would give us a system of linear dependent equations .

Table (1)

ASE(average squared error) for Stochastic Restricted General SUR Ridge(SGCV) , Stochastic Restricted General SUR Ridge(SGMCV) , Stochastic Restricted General SUR Robust and Stochastic Restricted General Robust Ridge (SGMCV) estimator
(h=S=100) (m=2)(k=20)(n=10)(l=10)

No.	Distribution	v	SR Ridge (SGCV)	SR Ridge (SGMCV)	SR Robust	SR Robust Ridge (SGMCV)
(1)	$X_i \sim MVN_{k=10}(0,\sum_{x(high)})$ $\Lambda_i \sim MVN_{k=10}(0,\sum_{\Lambda\,(high)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(high)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(high)})$ $\varepsilon_\Lambda{}^* \sim MVN_{m=2}(0,\sum_{\upsilon(high)})$ $\varepsilon_\Lambda{}^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon\,(high)})$	10% 20% 40%	$3.521e^{10}$ $1.245e^{11}$ $6.125e^{06}$	$2.292e^{04}$ $3.657e^{05}$ 24.117	$6.637e^{-17}$ $4.312e^{-17}$ 258.389	$1.877e^{-22}$ $1.43\ e^{-22}$ $4.363e^{-23}$
(2)	$X_i \sim MVN_{k=10}(0,\sum_{x(high)})$, $\Lambda_i \sim MVN_{k=10}(0,\sum_{\Lambda\,(high)})$ $\varepsilon^* \sim MVN_{q=2}(0,\sum_{\varepsilon(low)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(low)})$ $\varepsilon_\Lambda{}^* \sim MVN_{m=2}(0,\sum_{\upsilon\,(low)})$ $\varepsilon_\Lambda{}^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon\,(low)})$	10% 20% 40%	$1.920e^{04}$ $1.023e^{08}$ $9.021e^{05}$	17.663 $6.4017e^{04}$ 20.786	$1.024e^{-15}$ $1.859e^{-17}$ 66.781	$1.462e^{-23}$ 824.345 0.814
(3)	$Xi \sim MVN_{k=10}(0,\sum_{x(low)})$ $\Lambda_i \sim MVN_{k=10}(0,\sum_{\Lambda\,(low)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(high)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(high)})$ $\varepsilon_\Lambda{}^* \sim MVN_{m=2}(0,\sum_{\upsilon\,(high)})$ $\varepsilon_\Lambda{}^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon\,(high)})$	10% 20% 40%	$2.014e^{04}$ $7.982e^{05}$ $2.031e^{10}$	5.479 2.481 $9.355e^{06}$	$3.391e^{-20}$ $1.202e^{-08}$ 87.311	0.5299 $2.641e^{06}$ $9.354e^{06}$
(4)	$Xi \sim MVN_{k=10}(0,\sum_{x(low)})$ $\Lambda_i \sim MVN_{k=10}(0,\sum_{\Lambda\,(low)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(low)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(low)})$ $\varepsilon_\Lambda{}^* \sim MVN_{m=2}(0,\sum_{\upsilon\,(low)})$ $\varepsilon_\Lambda{}^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon\,(low)})$	10% 20% 40%	$8.032e^{04}$ $6.102e^{09}$ $2.014e^{05}$	611.976 $1.238e^{03}$ 1.454	$1.0525e^{-18}$ $1.243e^{-19}$ $1.029e^{03}$	681.859 $1.516e^{03}$ $7.093e^{05}$

Table (2)

ASE(average squared error) for Stochastic Restricted General SUR Ridge(SGCV) , Stochastic Restricted General SUR Ridge(SGMCV) , Stochastic Restricted General SUR Robust and Stochastic Restricted General Robust Ridge (SGMCV) estimator
(h=S=100) (m=2)(k=40)(n=20)(l=20)

No.	Distribution	υ	SR Ridge (SGCV)	SR Ridge (SGMCV)	SR Robust	SR Robust Ridge (SGMCV)
(5)	$X_i \sim MVN_{k=20}(0,\sum_{x(high)})$ $\Lambda_i \sim MVN_{k=20}(0,\sum_{\Lambda(high)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(high)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(high)})$ $\varepsilon_\Lambda^* \sim MVN_{m=2}(0,\sum_{\upsilon(high)})$ $\varepsilon_\Lambda^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon(high)})$	10%	$1.210e^{-03}$	$4.756e^{-22}$	$2.682e^{-17}$	$3.280e^{-22}$
		20%	$5.108e^{12}$	$2.166e^{03}$	$1.103e^{-17}$	$2.149e^{-18}$
		40%	$2.90e^{10}$	$2.031e^{06}$	28.382	$2.03e^{06}$
(6)	$X_i \sim MVN_{k=20}(0,\sum_{x(high)})$, $\Lambda_i \sim MVN_{k=20}(0,\sum_{\Lambda(high)})$ $\varepsilon^* \sim MVN_{q=2}(0,\sum_{\varepsilon(low)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(low)})$ $\varepsilon_\Lambda^* \sim MVN_{m=2}(0,\sum_{\upsilon(low)})$ $\varepsilon_\Lambda^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon(low)})$	10%	$1.891e^{11}$	$4.680e^{06}$	$8.815e^{-17}$	17.508
		20%	$1.0258e^{09}$	$3.601e^{03}$	$9.998e^{-17}$	$2.292e^{-21}$
		40%	$9.315e^{10}$	$1.227e^{07}$	210.805	$1.227e^{07}$
(7)	$Xi \sim MVN_{k=20}(0,\sum_{x(low)})$ $\Lambda_i \sim MVN_{k=20}(0,\sum_{\Lambda(low)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(high)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(high)})$ $\varepsilon_\Lambda^* \sim MVN_{m=2}(0,\sum_{\upsilon(high)})$ $\varepsilon_\Lambda^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon(high)})$	10%	$5.024e^{04}$	288.393	$7.126e^{-20}$	401.9082
		20%	$2.145e^{04}$	0.457	$8.281e^{-20}$	$2.148e^{-25}$
		40%	$6.124e^{03}$	$4.742e^{05}$	242.111	$9.199e^{-25}$
(8)	$Xi \sim MVN_{k=20}(0,\sum_{x(low)})$ $\Lambda_i \sim MVN_{k=20}(0,\sum_{\Lambda(low)})$ $\varepsilon^* \sim MVN_{m=2}(0,\sum_{\varepsilon(low)})$ $\varepsilon^{**} \sim MVN_{m=2}([10\ 5],\sum_{\varepsilon(low)})$ $\varepsilon_\Lambda^* \sim MVN_{m=2}(0,\sum_{\upsilon(low)})$ $\varepsilon_\Lambda^{**} \sim MVN_{m=2}([10\ 15],\sum_{\upsilon(low)})$	10%	$8.024e^{04}$	3.1019	$5.149e^{-19}$	549.821
		20%	$3.120e^{08}$	$1.913e^{03}$	$1.386e^{-19}$	$1.913e^{-20}$
		40%	$4.512e^{09}$	$6.899e^{06}$	929.532	$6.9005e^{06}$