Efficiency Comparisons of Robust and Non-Robust Estimators for Seemingly Unrelated Regressions Model

AHMED H. YOUSSEF[1], MOHAMED R. ABONAZEL[1], AMR R. KAMEL[1,2]
[1, 2]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research (FGSSR), Cairo University, Giza 12613, EGYPT
[2]Data Processing and Tabulation at Central Agency for Public Mobilization and Statistics (CAPMAS), Nasser City 2086, EGYPT

Abstract: - This paper studies and reviews several procedures for developing robust regression estimators of the seemingly unrelated regressions (SUR) model, when the variables are affected by outliers. To compare the robust estimators (M-estimation, S-estimation, and MM-estimation) with non-robust (traditional maximum likelihood and feasible generalized least squares) estimators of this model with outliers, the Monte Carlo simulation study has been performed. The simulation factors of our study are the number of equations in the system, the number of observations, the contemporaneous correlation among equations, the number of regression parameters, and the percentages of outliers in the dataset. The simulation results showed that, based on total mean squared error (TMSE), total mean absolute error (TMAE) and relative absolute bias (RAB) criteria, robust estimators give better performance than non-robust estimators; specifically, the MM-estimator is more efficient than other estimators. While when the dataset does not contain outliers, the results showed that the unbiased SUR estimator (feasible generalized least squares estimator) is more efficient than other estimators.

Key-Words: - Asymptotic efficiency, Breakdown point, Contemporaneous correlation, Feasible generalized least squares estimator, Maximum likelihood estimator, Monte Carlo simulation, Non-robust estimators, Outliers, Robust SUR estimators.

Received: June 12, 2021. Revised: March 15, 2022. Accepted: April 13, 2022. Published: May 6, 2022.

1 Introduction

The seemingly unrelated regressions (SUR) model proposed by Zellner [1] is considered to be one of the most successful and efficient methods for estimating SUR and tests of aggregation bias. Many studies in econometrics are based on regression models containing more than one equation. Unconsidered factors that influence the error term in one equation often also influence the error terms in other equations. Ignoring this dependence structure of the error terms and estimating these equations separately using ordinary least squares (OLS) leads to inefficient estimates. Therefore, the SUR model has been developed. This model is composed of several regression equations that are linked by the fact that their error terms are contemporaneously correlated. This system of structurally related equations is simultaneously estimated with a feasible generalized least squares (FGLS) estimator that takes the covariance structure of the error terms into account. Each equation satisfies the assumptions of the classical linear regression model.

The SUR model is a special case of the simultaneous equations model where no endogenous variables appear as regresses in any of the equations. Also, the SUR which considers joint modeling is a special case of the multivariate regression models (MLMs), see [1,2]. It is used to capture the effect of different covariates allowed in the regression equations. In all the estimation procedures developed for different SUR situations as reported above, FGLS basic recommendation for high contemporaneous correlation between the error vectors with uncorrelated explanatory variables within each response equations was also maintained. However, SUR Model depends on the FGLS estimator and assumes data without outliers but in some cases, this cannot be achieved. If the dataset contains outliers and influential observations, the FGLS estimator is not efficient. The SUR model assumption is used in a variety of econometric applications (or models), including panel data models and related fields, see [3,4], and many more.

The robust estimation methods are considered the one important approach to dealing with outliers. In the SUR model, it is necessary to use robust estimates to detect outliers and to provide resistant stable results in the presence of outliers, see [5,6]. The main purpose of this paper is to propose robust SUR estimators that can resist the potentially damaging effect of outliers in the dataset, and that do not require a separate estimation of the residual scale. To achieve these goals we investigate the efficiency of three robust estimators of the SUR model with outliers and compare them with (non-robust) FGLS and maximum likelihood estimators.

The remainder of this paper is organized as follows: Section 2 provides the SUR model and some methods of estimations. While in Section 3 robust estimation methods for the SUR model have been discussed. Section 4 contains the Monte Carlo simulation study. Finally, Section 5 presents the concluding remarks.

2 Classical SUR Model and Estimation

The SUR model explains the variation of not just one dependent variable, as in the univariate multiple regression model, but the variation of a set of m dependent variables, The m equations have no link or relationship with one another except that their disturbances are said to be correlated, this is the simplest version of a linear. Moreover, by joint analysis of the set of regression equations rather than equation by equation analysis, more precise estimates and predictions are obtained that lead to better solutions to many applied problems. For textbook and other analyses of the SUR model and its applications of it, see [7]. The SUR is used to reflect the fact that the individual equations are related to one another even though, superficially, they may not seem to be but are related through their error terms.

Zellner [1] developed the SUR estimator for estimating models with dependent variables that allow for different regressor matrices in each equation $\left(\text{e.g.} X_i \neq X_j\right)$. and account for contemporaneous correlation; $i.e. E\left(u_i u_j\right) \neq 0$. Now we can assume that if there is a m number of equations that are related to each other because the error terms are correlated. The regression equations in a SUR model can be combined into two equivalent single matrix form equations. Let diag $(\cdot)$ denote the operator that constructs a block diagonal matrix from its arguments. Moreover, let $\otimes$ denote the kronecker product and let Σ be a symmetric matrix with elements σ_{ij}. First, we can express it as a multiple linear regression model:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{pmatrix}_{mn\times 1} = \begin{pmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_m \end{pmatrix}_{mn\times K} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}_{K\times 1} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}_{mn\times 1} \quad (1)$$

This multiple equation can be simply re-written compactly as:

$$Y = X\beta + U, \quad (2)$$

where the $Y = (y_1', \dots, y_m')'$ is the column vector of observation on the i^{th} endogenous variable, $X = \text{diag}[X_i]$; with X_i(for $\text{i} = 1,2, \dots, m$)is a block diagonal design matrix of the exogenous non-stochastic variables of equation number i with dimension $n \times k_i$, and $\beta = (\beta_1', \dots, \beta_m')'$ is the column vector of the stacked coefficient vectors of all equations, the total number of parameters estimated for all k sub models is $K = \sum_{i=1}^{m} k_i$, while $U = (u_1', \dots, u_m')'$ is the column vector of contemporaneous correlated random error. Second, the SUR model can be rewritten as another equivalent formulation uses the MLMs:

$$\underset{n\times m}{\tilde{Y}} = \underset{n\times K}{\tilde{X}} \; \underset{K\times m}{\mathcal{B}} + \underset{n\times m}{\mathcal{U}}, \quad (3)$$

where $\tilde{Y} = (y_1, \dots, y_m)$ is the response matrix, $\tilde{X} = (X_1, \dots, X_m)$ is the design matrix, the coefficient matrix here has a constrained structure:

$$\mathcal{B} = \begin{pmatrix} \mathcal{B}_1 & 0 & \dots & 0 \\ 0 & \mathcal{B}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathcal{B}_m \end{pmatrix} = \text{diag}(\mathcal{B}_1, \dots, \mathcal{B}_m).$$

The structured is a $K \times m$ parameter matrix, and $\mathcal{U} = (u_1, \dots, u_m)$ is the error matrix. Equivalently we can write the error matrix as $\mathcal{U} = \tilde{Y} - \tilde{X}\hat{\mathcal{B}} = (\mathcal{u}_1, \dots, \mathcal{u}_n)'$ with $\mathcal{u}_\text{i}$ the m-dimensional vector containing the errors of the i^{th} observation in each block. For an estimate $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_m)'$, $\hat{\Sigma}$ uses the inner product matrix of residuals;

$$\hat{\Sigma} = \frac{1}{n} \begin{pmatrix} \hat{u}_1' \\ \vdots \\ \hat{u}_m' \end{pmatrix} (\hat{u}_1, \dots, \hat{u}_m), \quad (4)$$

2.1 SUR Model Assumptions

A1: $E(U) = 0$, error term has a normal distribution, and to be independent across individuals.
A2: X_i is fixed in repeated samples (non-stochastic matrix) and $cov(X, U) = 0$.
A3: X is full column rank matrix, i.e., $rank(X) = K$.
A4: The random errors of SUR model are assumed to have the following variance-covariance matrix of errors;

$$E(UU') = cov(U) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_{mm} \end{bmatrix} \otimes I_n = \Sigma \otimes I_n,$$

where I_n is an $n \times n$ identity matrix and $\Sigma = (\sigma_{ij})$ with positive definite and symmetric matrices (PDS) of dimension $m \times m$.Thus it must satisfy the assumptions that;

- The error variance for every individual equation which is a part of SUR is constant (no heteroscedasticity).
- The error variance may be different for every individual equation.
- The errors for every individual equation which is a part of SUR are uncorrelated (no autocorrelation).
- The errors for different individual equations are contemporaneously correlated.

2.2 Methods of Estimation

Each equation in Eq. (1) could be estimated separately using the OLS estimator but this would ignore the covariance structure of the errors. Consequently, it is generally less efficient and may yield inefficient estimates. The generalized least squares (GLS) estimator is a modification of the OLS estimator that can deal with any type of correlation, including contemporaneous correlation, GLS estimator is efficient and also fulfill the maximum likelihood requirement. Because it gives the best linear unbiased estimators (BLUEs). For the SUR model, the GLS estimator takes the form;

$$\hat{\beta}_{GLS} = (\mathrm{X}'\Omega^{-1}\mathrm{X})^{-1}(\mathrm{X}'\,\Omega^{-1}\,\mathrm{Y}), \qquad (5)$$

since $\Omega^{-1} = \Sigma^{-1} \otimes I_n$, GLS estimator is more efficient than the OLS estimator, but in most situations the covariance Σ needed in GLS estimator is unknown. FGLS estimator the elements of Σ by $\hat{\Sigma} = \sum_{i=1}^{n} \hat{u}_i\, \hat{u}_i'$,where $\hat{u}_i$ is the residual vector of the i^{th} block obtained from OLS and then replace Σ in FGLS estimator by the resulting estimator $\hat{\Sigma}$. The FGLS estimator takes the form;

$$\hat{\beta}_{\mathrm{FGLS}} = \left(\mathrm{X}'\hat{\Omega}^{-1}\mathrm{X}\right)^{-1}\left(\mathrm{X}'\,\hat{\Omega}^{-1}\,\mathrm{Y}\right). \qquad (6)$$

The variance-covariance matrix of the FGLS estimator can be evaluated by the following form;

$$cov\,(\hat{\beta}_{\mathrm{FGLS}}) = \left(\mathrm{X}'\hat{\Omega}^{-1}\mathrm{X}\right)^{-1}. \qquad (7)$$

Although the asymptotic efficiency of both GLS and FGLS methods is identical. The variance-covariance matrix $\hat{\Sigma}$ can then be re-estimated using the SUR residuals, and continue iterating the procedure until convergence is achieved. This is the iterated FGLS (IFGLS), see [8].

Alternatively, a maximum likelihood (ML) estimator can be considered; see [9]. Assuming that the disturbances are normally distributed, and retaining all the basic assumptions specified in the introductory section. The log-likelihood of the SUR model is given by;

$$\ell(\beta, \Sigma | X, Y) = -\frac{mn}{2}\ln(2\pi) - \frac{n}{2}ln(|\Sigma|) - \frac{1}{2}(Y - X\beta)'(\Sigma^{-1} \otimes I_n)(Y - X\beta). \qquad (8)$$

Maximizing this log-likelihood with respect to (β, Σ) yields the estimates $\left(\hat{\beta}_{ML}, \hat{\Sigma}_{ML}\right)$ which are the solutions of the equations:

$$\hat{\beta}_{ML} = \left\{X'\left(\hat{\Sigma}_{ML}^{-1} \otimes I_n\right)X\right\}^{-1} X'\left(\hat{\Sigma}_{ML}^{-1} \otimes I_n\right)Y, \qquad (9)$$

$$\hat{\Sigma}_{ML} = (\tilde{Y} - \tilde{X}\hat{\mathcal{B}}_{ML})'\,\left(\tilde{Y} - \tilde{X}\hat{\mathcal{B}}_{ML}\right)/n, \qquad (10)$$

with $\hat{\mathcal{B}}_{ML}$ the block diagonal form of $\hat{\beta}_{ML}$. Hence, the ML estimator correspond to the fully IFGLS estimator. The resulting ML estimator is, under general conditions, consistent, asymptotically efficient, and asymptotically normally distributed. Thus the asymptotic properties of the ML estimator are the same as those of the previous estimates, see [10].

3 Robust Estimators for SUR Model

It is well-known that traditional procedures like OLS, ML, and FGLS methods are all very sensitive to outliers in the data (observations that deviate from the main pattern in the data). Small anomalies in the data such as the presence of a few contaminated observations suffice to have a large impact on the resulting estimates. Outliers can appear in the data for several reasons. For example, some observations can be governed by a different data generating process other than the majority of the data while yet interest is in modeling the bulk of the data. Also, outliers can originate from an incorrect recording of the true dataset. Hence, these estimates are expected to yield non-robust estimates. Therefore, we introduce robust estimates for the SUR model which can combine high robustness with high efficiency, and obtained efficient and powerful robust tests. The main purpose of robust estimation is to provide resistant results in the presence of outliers. To achieve this stability, robust regression limits the influence of outliers, see [11,12].

Many robust methods have been proposed to achieve high robustness or high efficiency or both in several regression models, see e.g. [13-17]. In this section, we will review and compare some robust methods to determine the best robust method, and provides a detailed description of algorithms for these methods.

3.1 M-Estimation Method

Koenker and Portnoy [18] proposed the M-estimation method of the MLMs; these weighted M-estimates achieve an asymptotic covariance matrix analogous to that of the SUR estimator. The M-estimation method is a generalization of the ML estimator in the context of location models. That is nearly as efficient as traditional methods such as ML and FGLS. As the objective, the M-estimation method principle is minimizing the residual function; M-estimation is based on the residual scale of the FGLS estimator. It can be introduced the M-estimation method for the context of SUR models.

Definition 3.1: Let $(X_j, Y_j) \in \mathbb{R}^{n\times(k_j+1)}$ for $j = 1,2,\dots,m$ with $n \geq k+m$, and let ρ_0 be a ρ-function with parameter c_0. Then, the M-estimator of the SUR model $(\widehat{\mathcal{B}}, \widehat{\Sigma})$ are the solutions that minimize $|\Sigma_0|$ of the optimization problem;

$$\min_{(\beta,\Sigma)} |\Sigma|, \text{ subject to } \frac{1}{n}\sum_{i=1}^{n} \rho_0 \left\{[(\tilde{Y}_i - \tilde{X}_i B)'\Sigma_0^{-1}(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}\right\} = \Upsilon. \quad (11)$$

Where the minimization is over all $B = \text{diag}(B_1, \dots, B_m) \in \mathbb{R}^{k\times m}$, and $\Sigma_0 \in$ PDS (m) of dimension $m \times m$, since B and Σ_0 are initial estimates. The determinant of Σ is denoted by $|\Sigma|$, and Υ is a positive constant. In order to obtain estimates which can resist outliers ρ should satisfy the following conditions:

Condition 3.1: ρ is symmetric, twice continuously differentiable and satisfies $\rho(0) = 0$;

Condition 3.2: ρ is strictly increasing on $[0, c]$ and constant on $[c, \infty[$ for some $c > 0$.

Here the constant Υ is given by $\Upsilon = E_F\{\rho_0(|e|)\}$, to obtain a consistent estimator at an assumed error distribution F. A popular choice is Tukey's biweight ρ-function:

$$\rho(u) = \begin{cases} \frac{u^2}{2} - \frac{u^4}{2c^2} + \frac{u^6}{6c^4}, & |u| \leq c; \\ \frac{c^2}{6}, & |u| > c. \end{cases}$$

Where c is an appropriate tuning constant, the smaller value of c produce more resistance to outliers but comes at the price of loss in efficiency under the normal distribution. Usually, the tuning constant is picked to give reasonably high efficiency in the normal case for the Tukey's bisquare function, which generally, $c = 4.685$ is used to produces 95% efficiency, see [19].The derivative of this function is known as Tukey's bisquare function:

$$\psi(u) = \rho'(u) = \begin{cases} u\left[1 - \left(\frac{u}{c}\right)^2\right]^2, & |u| \leq c; \\ 0, & |u| > c. \end{cases}$$

Additionally, the minimization condition mentioned above the robust SUR estimators of β and Σ also satisfy the following equations:

$$\hat{\beta}_{\text{M}} = \left\{X'(\widehat{\Sigma}_{\text{FGLS}}^{-1} \otimes W_{\text{M}})X\right\}^{-1} X'(\widehat{\Sigma}_{\text{FGLS}}^{-1} \otimes W_{\text{M}})Y, \quad (12)$$

$$\widehat{\Sigma}_{\text{FGLS}} = m(\tilde{Y} - \tilde{X}\widehat{\mathcal{B}})' W_{\text{M}} (\tilde{Y} - \tilde{X}\widehat{\mathcal{B}})\{\textstyle\sum_{i=1}^{n} \upsilon_0(w_{\text{M}i})\}^{-1} \quad (13)$$

where $W_{\text{M}} = \text{diag}\{u(w_{M1}),, \dots, u(w_{Mn})\}$ is a diagonal matrix of weights, $w_{\text{M}i}^2 = e_i(\widehat{\mathcal{B}})' \sum_{\text{FGLS}}^{-1} e_i(\widehat{\mathcal{B}})$, where $e_i(\widehat{\mathcal{B}})'$represents the i^{th} row of the residual matrix $\tilde{Y} - \tilde{X}\widehat{\mathcal{B}}$, $u(w_M) = \psi_0(w_M)/w_M$; $\psi_0(w_M) = \rho_0'(w_M)$, and $\upsilon_0(w_{\text{M}}) = \psi_0(w_M)w_M - \rho_0(w_M) + \Upsilon$.

The efficiency and breakdown point (BDP) [20] are two traditionally used important criteria to compare different robust methods. The efficiency is used to measure the relative efficiency of the robust estimates compared to the non-robust (ML and FGLS) estimates when the error distribution is exactly normal and there are no outliers. BDP is to measure the proportion of outliers an estimate can tolerate before it goes to infinity. Thus the higher the BDP of an estimator, the more robust is. Intuitively, a BDP cannot exceed 0.5. In fact, the

BDP of the M-estimator is $\text{BDP} = 1/n \to 0$, see [21].

Algorithm 3.1: M-Estimation

Since the weights depend on the unknown parameter β and $\sum$, we cannot calculate the weighted mean explicitly. But this weighted-means representation of the M-estimator leads to a simple iterative algorithm for calculating the M-estimator. By developing the algorithms for [22, 23], our algorithm 3.1 can be described in the following steps:

Step (1): Let $\tilde{\beta}_1^{(0)}, \dots, \tilde{\beta}_j^{(0)}; j = 1,2, \dots, J$ is initial candidates estimate for β, and set an initial variance-covariance matrix $\sum_0 \in \text{PDS}\ (m)$.

Step (2): Design the variable $(X_i, U_i); i = 1,2, \dots, m$. For each $\tilde{\beta}_j^{(0)}$:

a. Estimate SUR model coefficients with all factors using a non-robust FGLS estimator, and test all assumptions.
b. Detect the presence of outliers in the dataset.
c. Calculate residuals matrix $\hat{\Sigma} = \frac{1}{n}\hat{U}'\hat{U}$.

Step (3): Calculate the variance-covariance matrix $\hat{\Sigma}(\hat{\beta}_{\text{FGLS}})$, and weighted matrix $W_M(\hat{\beta}_{\text{FGLS}})$.

Step (4): Calculate M-estimator as in Eq. (11) for some ρ-function ρ_0 by set $q = 0$, where q is a number of iteration and get the iterate following steps:

(i) Let $\hat{\beta}_{j\,(\text{M})}^{(q+1)} = \left\{X'\left(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\text{FGLS})}) \otimes W_{\text{M}}(\hat{\beta}_{j\,(\text{FGLS})})\right)X\right\}^{-1} X'\left(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\text{FGLS})}) \otimes W_{\text{M}}(\hat{\beta}_{j\,(\text{FGLS})})\right)Y.$

(ii) If either $q = maxit$ (maximum number of iterations) or
$$\left\|\hat{\beta}_{j\,(\text{M})}^{q} - \hat{\beta}_{j\,(\text{M})}^{(q+1)}\right\| < \text{T}\left\|\hat{\beta}_{j\,(\text{M})}^{q}\right\|,$$
where $\text{T} > 0$ is a fixed small constant (the tolerance level), then set $\hat{\beta}_{j\,(\text{M})}^{\text{F}} = \hat{\beta}_{j\,(\text{M})}^{q}$ and break.

(iii) Else, Calculate $W_{\text{M}}\left(\hat{\beta}_{j\,(\text{M})}^{(q+1)}\right), \hat{\Sigma}\left(\hat{\beta}_{j\,(\text{M})}^{(q+1)}\right)$ and set $q \leftarrow q + 1$.

Step (5): Calculate the objective function for each $\hat{\beta}_{j\,(\text{M})}^{\text{F}}; j = 1,2, \dots J$, and select the one with the lowest value, that is, Select $\hat{\beta}_{\text{M}}$ which active;
$$\min_{1 \le j \le J} \left[\frac{1}{n}\sum_{i=1}^{n} \rho_0\{[(\tilde{Y}_i - \tilde{X}_i B)'\hat{\Sigma}^{-1}(\hat{\beta}_{\text{j}\,(\text{M})}^{\text{F}})\,(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}\}\right]$$

Step (6): Repeat steps 3 and 4 until the algorithm converges to obtain a convergent value of $\hat{\beta}_{\text{M}}$ using Eq. (12).

The J initial candidates $\tilde{\beta}_j^{(0)}$ in Step 1 can be chosen in several ways. Intuitively we want them to correspond to different regions of the optimization domain. In linear regression problems, these initial points are generally chosen based on the sample, see [24].

3.2 S-Estimation Method

Bilodeau and Duchesne [25] introduced a new class of robust SUR estimates; in response to the low BDP of the M-estimator, the regression estimates associated with M-estimator is the S-estimator is a member of the class of high BDP estimates. S-estimator is based on the residual scale of M-estimator. This method uses the residual standard deviation to overcome the weaknesses of the median; the idea behind the method is simple. For OLS, the objective is to minimize the variance of the residuals. [26] Gives an improved resampling algorithm for S-estimator for multivariate regression. They studied the robustness of the estimates in terms of their BDP and influence function in the context of univariate regression and multivariate location and scatter, and developed a fast and robust bootstrap method for the multivariate S-estimator to obtain inference for the regression parameters. With this algorithm, S-estimator is easier to calculate. In the next section, we will discuss how to adapt that algorithm to the context of SUR model.

Definition 3.2: Let $\hat{\sum}_{\text{M}}$ denote the M-estimator of covariance in Definition 3.1, and Let $(X_j, Y_j) \in \mathbb{R}^{n\times(k_j+1)}$ for $j = 1,2, \dots, m$ and let ρ_1 be a ρ-function with parameter c_1 in Condition 3.2. Then, the S-estimator of the SUR model $(\hat{\mathcal{B}}, \hat{\sum})$ are the solutions that minimize $|\sum|$ subject to the condition;
$$\min_{(\beta,\Sigma)} |\Sigma|, \text{ subject to } \frac{1}{n}\sum_{i=1}^{n} \rho_1 \left\{[(\tilde{Y}_i - \tilde{X}_i B)'\textstyle\sum_0^{-1}(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}\right\} = \mu. \qquad (14)$$

Where the minimization is over all $B = \text{diag}(B_1, \dots, B_m) \in \mathbb{R}^{k\times m}$, $\sum_0 \in \text{PDS}\ (m)$, since B and $\sum_0$ are initial estimates and μ is a positive constant.

This formulation is between the S-estimator of regression and the multivariate S-estimator since we have to minimize a multivariate measure of scale in the presence of m regression models. As before, the regression coefficient estimates in the matrix $\hat{\mathcal{B}}$ can also be collected in the vector $\hat{\beta} = (\hat{\beta}_1', \dots, \hat{\beta}_m')'$. The first-order conditions corresponding to the above minimization problem yield the following fixed-point equations for S-estimator;

$$\hat{\beta}_{\mathrm{S}} = \left\{X'(\hat{\Sigma}_{\mathrm{M}}^{-1} \otimes W_{\mathrm{S}})X\right\}^{-1} X'(\hat{\Sigma}_{\mathrm{M}}^{-1} \otimes W_{\mathrm{S}})Y, \quad (15)$$

$$\hat{\Sigma}_{\mathrm{M}} = m(\tilde{Y} - \tilde{X}\hat{\mathcal{B}})' W_{\mathrm{S}}(\tilde{Y} - \tilde{X}\hat{\mathcal{B}}) \{\textstyle\sum_{i=1}^{n} v_1(w_{\mathrm{S}i})\}^{-1}. \quad (16)$$

With diagonal matrix;
$W_{\mathrm{S}} = \mathrm{diag}\{u(w_{\mathrm{S}1}), \ldots, u(w_{\mathrm{S}n})\}$, where
$w_{\mathrm{S}i}^2 = e_i(\hat{\mathcal{B}})' \sum_{\mathrm{M}}^{-1} e_i(\hat{\mathcal{B}})$;
$u(w_{\mathrm{S}}) = \psi_1(w_{\mathrm{S}})/w_{\mathrm{S}}$; $\psi_1(w_{\mathrm{S}}) = \rho_1'(w_{\mathrm{S}})$, and $v_1(w_{\mathrm{S}}) = \psi_1(w_{\mathrm{S}})w_{\mathrm{S}} - \rho_1(w_{\mathrm{S}}) + \mu$.

Starting from the initial M-estimator, the S-estimator is calculated easily by iterating these estimating equations until convergence. The S-estimating equations (15) and (16) reduce to the normal ML estimating equations, and similarities to the FGLS, see [10]. Unlike the ML and M-estimator. S-estimator satisfies the first-order conditions of M-estimator see [27], so they are asymptotically normal. However, the choice of the tuning parameter c_0 involves a trade-off between BDP (robustness) and efficiency in the central model. For this reason, S-estimators are less adequate for robust inference. The choice of BDP affects the efficiency of the estimator under a Gaussian model. The higher the BDP, the lower the efficiency and vice versa. Hence, S-estimator can attain the maximal BDP of 50%.

Algorithm 3.2: S-Estimation

In this algorithm, we compute S-estimator for the SUR model, our algorithm 3.2 can be described in the following steps:

Step (1): Let $\tilde{\beta}_1^{(0)}, \ldots, \tilde{\beta}_j^{(0)}$; $\mathrm{j} = 1,2,\ldots,J$ is initial candidates estimate for β, and set an initial variance-covariance matrix $\sum_0 \in \mathrm{PDS}\ (m)$.

Step (2): Generate, and design the variable(X_i, U_i); $i = 1,2,\ldots,m$. For each $\tilde{\beta}_j^{(0)}$:

a. Estimate SUR model coefficients with all factors using a non-robust FGLS estimator, and test all assumptions.

b. Detect the presence of outliers in the dataset.

c. Calculate residuals matrix $\hat{\Sigma} = \frac{1}{n}\hat{U}'\hat{U}$.

Step (3): Calculate $\hat{\beta}_{\mathrm{M}}$ using Algorithm 3.1, and Calculate $W_{\mathrm{S}}(\hat{\beta}_M), \hat{\Sigma}(\hat{\beta}_M)$.

Step (4): Calculate S-estimator as in Eq. (14) for some ρ-function ρ_1 by set $\hbar = 0$, where $\hbar$ is a number of iteration and get the iterate following steps:

(i) Let $\hat{\beta}_{j\,(\mathrm{S})}^{(\hbar+1)} =$
$\left\{X'(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\mathrm{M})}) \otimes W_S(\hat{\beta}_{j\,(\mathrm{M})}))X\right\}^{-1}$
$X'(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\mathrm{M})}) \otimes W_S(\hat{\beta}_{j\,(\mathrm{M})}))Y$.

(ii) If either $\hbar = maxit$ (maximum number of iterations) or;

$$\left\|\hat{\beta}_{j\,(\mathrm{S})}^{\hbar} - \hat{\beta}_{j\,(\mathrm{S})}^{(\hbar+1)}\right\| < \mathrm{T}\left\|\hat{\beta}_{j\,(\mathrm{S})}^{\hbar}\right\|,$$

where $\mathrm{T} > 0$ is a fixed small constant (the tolerance level), then set $\hat{\beta}_{j\,(\mathrm{S})}^{\mathrm{F}} = \hat{\beta}_{j\,(\mathrm{S})}^{\hbar}$ and break.

(iii) Else, Calculate $W_{\mathrm{S}}\left(\hat{\beta}_{j\,(\mathrm{S})}^{(\hbar+1)}\right), \hat{\Sigma}\left(\hat{\beta}_{j\,(\mathrm{S})}^{(\hbar+1)}\right)$ and set $\hbar \leftarrow \hbar + 1$

Step (5): Calculate the objective function for each $\hat{\beta}_{j\,(\mathrm{S})}^{\mathrm{F}}$; $j = 1,2,\ldots J$, and select the one with the lowest value, that is, Select $\hat{\beta}_{\mathrm{S}}$ which active;

$$\min_{1\le j\le J} \left[\frac{1}{n}\sum_{i=1}^{n} \rho_1\{[(\tilde{Y}_i - \tilde{X}_i B)' \hat{\Sigma}^{-1}(\hat{\beta}_{\mathrm{j}\,(\mathrm{S})}^{\mathrm{F}})\,(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}\}\right].$$

Step (6): Repeat steps 3 and 4 until the algorithm converges to obtain a convergent value of $\hat{\beta}_{\mathrm{S}}$ using Eq. (15).

3.3 MM-Estimation Method

Peremans and Van Aelst [28] proposed the MM-estimator in the context of the SUR model to obtain estimates that have both high BDP and a high normal efficiency. A fast and robust bootstrap procedure is developed to obtain robust inference for these estimates, by combining S-estimation with M-estimation. The initial estimate is a high BDP estimate using S-estimator, and the second stage computes an M-estimator of the scale of the errors from the initial high BDP estimate residuals matrix. Recently, [29, 30] studied the efficiency of some robust estimates by different applications (Economy and insurance), and showed that MM-estimator is highly efficient, and not sensitive to leverage points compared to other robust estimates.

Let $\hat{\sum}_{\mathrm{S}}$ denote the S-estimator of variance covariance matrix. Decompose $\hat{\sum}_{\mathrm{S}}$ into a scale component $\hat{\sigma}$ and a shape matrix $\hat{\Gamma}$ such that $\hat{\sum}_{\mathrm{S}} = \hat{\sigma}^2\,\hat{\Gamma}$ with $|\hat{\Gamma}| = 1$.

<u>Definition 3.3:</u> Let $(X_j, Y_j) \in \mathbb{R}^{n\times(k_j+1)}$ for $j = 1,2,\ldots,m$ and let ρ_2 be a ρ-function with parameter c_2 in Condition 3.2. Given the S-scale $\hat{\sigma}$.

Then the MM-estimator of the SUR model $(\hat{\mathcal{B}}, \hat{\Gamma})$ are the solutions that minimize $|\Sigma|$ subject to the condition;

$$\min_{(\beta,\Sigma)} |\Sigma|, \text{ subject to } \frac{1}{n}\sum_{i=1}^{n} \rho_2 \left\{[(\tilde{Y}_i - \tilde{X}_i B)'\Sigma_0^{-1}(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}/\hat{\sigma}\right\} \quad (17)$$

where the minimization is over all $B = \text{diag}(B_1,,\ldots,B_m) \in \mathbb{R}^{k\times m}$, $\sum_0 \in$ PDS (m) with $|\Sigma| = 1$, since B and $\sum_0$ are initial estimates, The MM-estimator for covariance is defined as $\hat{\Sigma}_{\text{MM}} = \hat{\sigma}^2\,\hat{\Gamma}$.

The MM-estimator of the regression coefficients $\hat{\mathcal{B}}$ can also be written as $\hat{\beta} = (\hat{\beta}_1', \ldots, \hat{\beta}_m')'$ in vector form. Similarly, as for S-estimator, the first-order conditions corresponding to the above minimization problem yield a set of fixed-point equations:

$$\hat{\beta}_{\text{MM}} = \left\{X'(\hat{\Sigma}_S^{-1} \otimes W_{\text{MM}})X\right\}^{-1} X'(\hat{\Sigma}_S^{-1} \otimes W_{\text{MM}})Y, \quad (18)$$

$$\hat{\Sigma}_S = m(\tilde{Y} - \tilde{X}\hat{\mathcal{B}})' W_{\text{MM}}(\tilde{Y} - \tilde{X}\hat{\mathcal{B}}) \{\textstyle\sum_{i=1}^{n} \psi_2(w_{\text{MM}i})\, w_{\text{MM}i}\}^{-1} \quad (19)$$

With diagonal matrix;
$W_{\text{MM}} = \text{diag}\{u(w_{\text{MM}1}), \ldots, u(w_{\text{MM}n})\}$;
Where; $w_{\text{MM}i}^2 = e_i(\hat{\mathcal{B}})' \sum_S^{-1} e_i(\hat{\mathcal{B}})$,
$u(w_{\text{MM}}) = \psi_2(w_{\text{MM}})/w_{\text{MM}}$; $\psi_2(w_{\text{MM}}) = \rho_2'(w_{\text{MM}})$.

Starting from the initial S-estimator, the MM-estimator is calculated easily by iterating these estimating equations until convergence. MM-estimator inherits the BDP of the initial S-estimator. Hence, they can attain the maximal BDP if an initial high-BDP S-estimator is used, see [31].

Algorithm 3.3: MM-Estimation

In this algorithm, we compute MM-estimator for the SUR model. Our algorithm 3.3 can be described in the following steps:

Step (1): Let $\tilde{\beta}_1^{(0)}, \ldots, \tilde{\beta}_j^{(0)}$; $\text{j} = 1,2,\ldots,J$ is initial candidates estimate for β, and set an initial variance-covariance matrix $\sum_0 \in$ PDS (m).

Step (2): Generate, and design the variable (X_i, U_i); $i = 1,2,\ldots,m$. For each $\tilde{\beta}_j^{(0)}$:

a. Estimate SUR model coefficients with all factors using a non-robust FGLS estimator, and test all assumptions.
b. Detect the presence of outliers in the dataset.
c. Calculate residuals matrix $\hat{\Sigma} = \frac{1}{n}\hat{U}'\hat{U}$.

Step (3): Calculate $\hat{\beta}_S$ using Algorithm 3.2, and Calculate $W_{\text{MM}}(\hat{\beta}_S), \hat{\Sigma}(\hat{\beta}_S)$.

Step (4): Calculate MM-estimator as in Eq. (16) for some ρ-function ρ_2 by set $g = 0$, where g is a number of iteration and get the iterate following steps:

(i) Let $\hat{\beta}_{\text{j (MM)}}^{(g+1)} = \left\{X'(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\text{S})}) \otimes W_{\text{MM}}(\hat{\beta}_{j\,(\text{S})}))X\right\}^{-1} X'(\hat{\Sigma}^{-1}(\hat{\beta}_{j\,(\text{S})}) \otimes W_{\text{MM}}(\hat{\beta}_{j\,(\text{S})}))Y.$

(ii) If either $g = maxit$ (maximum number of iterations) or;

$$\left\|\hat{\beta}_{j\,(\text{MM})}^{g} - \hat{\beta}_{j\,(\text{MM})}^{(g+1)}\right\| < \text{T}\left\|\hat{\beta}_{j\,(\text{MM})}^{g}\right\|,$$

where $\text{T} > 0$ is a fixed small constant (the tolerance level), then set; $\hat{\beta}_{j\,(\text{MM})}^{\text{F}} = \hat{\beta}_{j\,(\text{MM})}^{g}$ and break.

(iii) Else, Calculate $W_{\text{MM}}\left(\hat{\beta}_{j\,(\text{MM})}^{(m+1)}\right), \hat{\Sigma}\left(\hat{\beta}_{j\,(\text{MM})}^{(m+1)}\right)$ and set $g \leftarrow g + 1$.

Step (5): Calculate the objective function for each $\hat{\beta}_{j\,(\text{MM})}^{\text{F}}$; $j = 1,2,\ldots J$, and select the one with the lowest value, that is, select $\hat{\beta}_{\text{MM}}$ which active;

$$\min_{1\le j\le J}\left[\frac{1}{n}\sum_{i=1}^{n} \rho_2\{[(\tilde{Y}_i - \tilde{X}_i B)'\,\hat{\Sigma}^{-1}(\hat{\beta}_{\text{j (M)}}^{\text{F}})\,(\tilde{Y}_i - \tilde{X}_i B)]^{\frac{1}{2}}/\hat{\sigma}\}\right]$$

Step (6): Repeat steps 3 and 4 until the algorithm converges to obtain a convergent value of $\hat{\beta}_{\text{MM}}$ using Eq. (18).

Practically, while MM-estimator has maximal BDP, there is some loss of robustness because the bias due to contamination is generally higher as compared to S-estimator. However, it turns out that more accurate and powerful tests are obtained if a more efficient MM-estimator is used. Now using these algorithms it became easy to calculate the three robust estimators (M-estimation, S-estimation, and MM-estimation), which we will use in the simulation study.

4 Monte Carlo Simulation Study

In this section, we conduct a comparative study between the classical non-robust (ML and FGLS) estimators and the three robust (M-estimation, S-estimation, and MM-estimation) methods for the SUR model, through the Monte Carlo simulation study. In our simulation study, Monte Carlo experiments were performed based on the model in equations (2) and (3).To investigate the performance of these estimates in different situations, we will use different simulation factors as shown in Table 1. R software "version 4.1.2" is used to perform this study. For further information on how to make Monte Carlo simulation studies using R, see e.g. [32, 33].

Table 1. The simulation factors of our study

No.	Simulation factor	Symbol	Levels
1	The number of parameters ($\boldsymbol{\beta}$) in each equation (without intercept)	k_i	4 or 6
2	The number of equations	m	3,6 or 8
3	The true values of the parameters ($\boldsymbol{\beta}$) (as [34])	β	$\beta_1 = \cdots = \beta_m = \mathbf{1}$
4	The values of sample size in each equation	n	20,30,50,80 or100
5	The exogenous variables: $X \sim MVN(1, \Sigma_x)$, where diag $(\Sigma_x) = 1$ and The percentages of outliers ($\tau\%$) in the endogenous variables	$\tau\%$	0, 10, 20, 30 or 40
6	The error term: $U \sim MVN(0, \Omega)$, the variance-covariance matrix of $U(\Omega = \Sigma \otimes I_n)$ is defined as diag $(\Sigma_u) = 1$, and off-diag $(\Sigma_u) = \rho_\Sigma$	ρ_Σ	0.70 or 0 .90
7	The outliers generated from normal distribution with $(\delta, 1)$; where $\delta = 4 \times \text{IQR}(Y)$, and IQR is interquartile range (as [13-17]).		

All Monte Carlo experiments involved 1000 replications and all the results of all separate experiments are obtained by precisely the same series of random numbers. To compare the performance of the estimates with different n, m, k_i, ρ_Σ, and $\tau\%$, we evaluated their total mean squared error (TMSE) and total mean absolute error (TMAE) for $\hat{\beta}$.

$$\text{TMSE} = \frac{1}{L}\sum_{l=1}^{L} \left(\hat{\beta}_l - \beta\right)'\left(\hat{\beta}_l - \beta\right);$$

$$\text{TMAE} = \frac{1}{L}\sum_{l=1}^{L} \left|\left(\hat{\beta}_l - \beta\right)\right|,$$

where $\hat{\beta}_l$ is the vector of estimated parameters at l^{th} experiment of $\text{L} = 1000$ Monte Carlo experiments, while β is the vector of true parameters.

4.1 The Simulation Algorithm

The simulation study is based on the following algorithm:

Step (1): Generate the exogenous non-stochastic variables, $X = \text{diag}[X_i]$; with X_i (for $i = 1, \dots, m$) is a block diagonal design matrix, from $MVN(1, \Sigma_x)$.

Step (2): Set the true values of β_i.

Step (3): Simulate the vector of random errors (U) from $MVN(0, \Omega)$.

Step (4): The outliers are generated from contaminated normal distribution under different scenarios.

Step (5): The endogenous variables are then generated from the values already obtained for the X_i's (step 1), the values assigned to β_i (step 2), and the error term U's (step 3), according the following formula;

$$Y_i = X_i\beta_i + U_i; \qquad i = 1, \dots, m.$$

Step (6): Estimate SUR Parameters using non-robust (ML and FGLS) estimators.

Step (7): Estimate the SUR model using robust estimators (M-estimation, S-estimation, and MM-estimation), through the proposed algorithm for each method in Section 3.

Step (8): Repeat steps from step (3) to step (7) 1000 times and then calculate the parameter estimates ($\hat{\beta}_i$), TMSE and TMAE criteria for different estimators.

4.2 Simulation Results

The simulation results are presented in Tables 2 to 9. Specifically, Tables 2, 3,6 and 7 present the TMSE and TMAE values of the estimates when $k_i = 4$, while the case of $k_i = 6$ is presented in Tables 4, 5, 8 and 9 with different percentages of outliers ($\tau\%$) for the SUR model. Each table has five sections that represent the percentages of outliers in which each row represents a different sample size. Moreover, our simulation study has revealed four factors that have a bearing on the performance of the multivariate robust parameters in terms of TMSE and TMEA criteria. These factors are the number of equations (m), the number of observations (n), the percentages of outliers ($\tau\%$) and the percentages of contemporaneous correlation among equations (ρ_Σ). In all cases the performance of the multivariate robust parameters, in terms of the above factors, From Tables 2 to 9, we can summarize the effects of the main simulation factors on TMSE and TMAE values for all estimates (robust and non-robust) as follows:

- As m increases, the values of TMSE and TMAE are increases for all simulation situations.
- As n increases, the values of TMSE and TMAE are decreases in all situations.
- As $\tau\%$ increases, the values of TMSE and TMAE are increases in all situations.
- As ρ_Σ increases, the values of TMSE and TMAE are decreases (almost).

However, if the values of k_i is increased, the TMSE and TMAE values of ML and FGLS estimates are increased more than robust estimates. In all simulation cases, it is noticeable that the values of TMSE and TMAE for robust estimates are smaller than those of TMSE and TMAE for non-robust estimates. In another word, we can conclude that robust estimates are more efficient than ML and FGLS estimates. Specifically, among the robust estimates MM-estimator is the best estimator because it has minimum TMSE and TMAE values in all simulation situations.

Moreover, it is noticeable that for the percentages of outliers $\tau = 0\%$, the TMSE and TMAE for FGLS estimator are smaller than those of TMSE and TMAE for ML and robust estimates. We can conclude that in the absence of outliers FGLS estimator is more efficient than robust and non-robust (ML) estimates.

Graphically, we illustrate the average TMSE values for different estimates in all cases with different main factors by 3D graphs are shown in Figures 1 to 3. It is clear that, the FGLS estimator has the largest average TMSE values, followed by M-estimator, S-estimator, and finally MM-estimator. Moreover, the Figures confirm that MM-estimator is the best estimator for this model, especially when $\tau\%$ increases.

On the other hand, we also depend on another comparative performance level called relative efficiency (RE). The RE values are given by dividing the TMSE of ML by the TMSE of the estimator. The RE values of the estimates for $\rho_\Sigma = 0.70$ and $\rho_\Sigma = 0.90$ by 2D graphs are shown in Figures 4 and 5, respectively.

Figure 4 indicates that RE values of the MM-estimator are greater than RE values of different robust estimates for all n values, since MM-estimator has the largest RE values. This suggests that the MM-estimator is more efficient than the robust estimates in different n and $\tau\%$ values. However, when n and $\tau\%$ increase, the efficiency of the MM-estimator increases. In Figure 5, the efficiency of the robust estimates is close, but the MM-estimator is still more efficient than different robust estimates.

4.3 Relative Absolute Bias

To compare the performance of the selected estimators under different scenarios, we also depend on another comparative level is called relative absolute bias (RAB); it indicates a comparative performance level of an estimate based on its inputs and outputs with those of others in the collection. It can be considered as the absolute bias divided by its true value, which is calculated as:

$$\text{RAB} = \frac{\left|(\hat{\beta}_l - \beta)\right|}{\beta}.$$

The RAB results are presented in Tables 10 to 13 revealing the estimation results for each parameter $\hat{\beta}_\text{i}$ and RAB values to show the efficiency of the different estimators. The previous simulation algorithm presented in Section 4.1 has been used, when $m = 3, k_i = 5, n = 20, \tau = 10\%$ or 40%, $\rho_\Sigma = 0.70$ or 0.90, and the true values of the parameters β is $(1,2,3,4,5)'$.

According to results, specifically Tables 10-11 present the estimation for each parameter $\hat{\beta}_i$ and RAB values of the estimates when $\rho_\Sigma = 0.70$. While case of $\rho_\Sigma = 0.90$ is presented in Tables 12-13. It can be noted that, the robust estimates improve the efficiency of the estimates for the SUR model when the dataset contains outliers. It is clear that when $\tau\%$ increases, the values of RAB are increases for all estimates. This increase is somehow large for non-robust (ML and FGLS) estimates. Robust estimates still have minimum RAB values; we can conclude that robust estimates are more efficient than non-robust estimates. Specifically, the MM-estimator is the best robust estimator because it has minimum RAB values.

Finally, the Final conclusion from the simulation study along with the results of RE values is that MM-estimator outperforms the other estimates in the sense of RAB, TMSE and TMAE criteria. Moreover, MM-estimator has the best performance in the simulation in most or all cases.

Table 2. TMSE and TMAE values for different estimates when $m = 6$, $k_i = 4$ and $\rho_\Sigma = 0.70$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	0.6449	**0.5485**	1.7336	0.6418	0.6372	4.1114	**3.8222**	6.6976	4.1629	4.1480
50	0.2939	**0.2839**	0.9697	0.3535	0.3512	2.9606	**2.7606**	5.0758	3.0971	3.0860
80	0.1736	**0.1624**	0.6207	0.2108	0.2090	2.4628	**1.9802**	4.0578	2.3941	2.3843
100	0.1374	**0.1253**	0.5215	0.1690	0.1673	1.9750	**1.8125**	3.7239	2.1373	2.1270
					$\tau = 10\%$					
30	145.5312	117.4314	69.9230	17.2961	**14.1768**	44.0170	39.7458	23.9825	15.4728	**13.7377**
50	84.4440	80.4268	16.4373	9.0381	**8.4713**	34.8697	31.0649	12.3955	11.3988	**10.8913**
80	62.5679	60.6152	8.1222	5.7643	**4.6562**	30.3540	28.1370	9.7702	9.5720	**9.0952**
100	54.9215	52.4693	7.0713	4.5548	**4.0643**	28.7037	26.5993	8.9193	8.4496	**8.1644**
					$\tau = 20\%$					
30	281.9233	231.5541	108.5037	56.6098	**39.2795**	63.1778	57.7044	35.6429	24.5447	**22.4001**
50	181.5560	175.1858	63.3612	33.9467	**24.4077**	52.2595	51.4116	28.5852	21.4592	**18.4585**
80	151.6483	150.0597	49.1791	25.6107	**17.2553**	48.7713	45.5762	25.6778	18.6915	**15.5372**
100	138.0080	134.4415	40.6304	21.7928	**14.3843**	47.1391	43.0848	24.1521	17.3645	**14.2700**
					$\tau = 30\%$					
30	404.6679	340.7385	139.6607	89.6089	**75.8728**	76.9916	71.0476	53.1248	29.7552	**26.8209**
50	293.6038	285.8037	106.6324	56.5487	**42.0847**	67.4903	66.7943	43.9418	28.1067	**24.4463**
80	260.1241	258.9050	87.7246	48.4947	**32.5188**	65.6506	63.6540	40.1426	26.1948	**21.4918**
100	242.2294	240.6620	79.2175	43.4430	**28.1141**	63.5898	60.7654	38.1352	25.0312	**20.1323**
					$\tau = 40\%$					
30	507.2812	438.1510	194.3324	96.8271	**84.0558**	86.2553	81.1094	63.5899	48.8008	**43.5644**
50	405.0117	398.6224	126.1240	85.3366	**63.5409**	80.7761	74.5166	54.4618	45.0508	**41.9157**
80	383.6447	380.2413	107.0231	59.6454	**48.2619**	73.7036	68.1605	50.4742	36.9624	**33.8848**
100	360.5484	357.8497	98.1965	46.9293	**35.1803**	70.9126	64.6742	47.9862	34.8114	**32.7669**

Note: The best performance for each percentage of outliers is given in bold.

Table 3. TMSE and TMAE values for different estimates when $m = 6$, $k_i = 4$ and $\rho_\Sigma = 0.90$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	0.5276	**0.5171**	2.2213	0.6535	0.6461	5.7666	**5.7308**	11.6167	6.4528	6.4197
50	0.2382	**0.2255**	1.2693	0.3499	0.3463	3.9028	**4.0350**	8.7805	4.7361	4.7143
80	0.1358	**0.1241**	0.7297	0.2032	0.2010	2.9544	**3.0209**	6.6719	3.6190	3.6009
100	0.1141	**0.1055**	0.6757	0.1738	0.1721	2.5759	**2.6183**	6.4033	3.3386	3.3237
					$\tau = 10\%$					
30	129.1042	104.4606	47.0596	15.0850	**13.9512**	41.1930	37.2495	22.9168	15.6322	**15.0993**
50	76.1349	72.5108	15.9527	11.8782	**8.5732**	32.7935	30.0472	11.9435	12.8531	**12.8980**
80	56.8294	51.9891	9.8027	9.1482	**7.5975**	28.8939	26.6952	10.3911	10.3687	**9.2051**
100	48.8852	43.4880	7.4582	7.0432	**5.3791**	27.0419	25.9422	9.2942	9.2872	**8.0323**
					$\tau = 20\%$					
30	244.9528	204.0295	91.3063	52.7848	**35.2396**	58.5062	53.5759	33.4293	26.6371	**23.6838**
50	164.5525	158.7848	57.2550	39.7133	**19.7736**	49.4495	45.6445	26.9896	22.5045	**20.5202**
80	136.1537	134.7512	43.3156	22.0107	**12.1175**	46.1006	40.9166	24.0619	19.7443	**16.7637**
100	123.2675	120.7956	37.4221	19.5295	**10.6728**	43.4377	38.3891	20.5885	18.2691	**14.3103**
					$\tau = 30\%$					
30	351.8149	296.3762	135.8456	86.0209	**73.9701**	71.0489	65.5477	49.1075	38.6158	**33.5693**
50	265.4046	258.3169	113.2356	62.8905	**37.9236**	63.7553	60.1135	41.4111	32.5360	**30.5304**
80	234.1565	230.0934	99.5767	49.9201	**28.7060**	62.1860	58.1886	37.8176	24.6202	**20.2042**
100	214.4017	211.8465	86.6459	37.3062	**21.1810**	60.6683	56.8491	35.5299	20.1300	**18.6629**
					$\tau = 40\%$					
30	437.9722	379.0437	167.8719	95.8737	**82.4057**	79.2538	74.4384	58.7349	53.9845	**44.7871**
50	370.5123	364.3977	127.1102	77.8285	**61.2421**	76.4940	73.2409	51.5080	35.3245	**28.5749**
80	342.3235	339.7887	105.8272	56.4329	**43.8695**	72.0450	70.4726	47.2605	26.4356	**23.7126**
100	321.7174	316.7481	101.5156	46.1344	**35.3904**	67.3612	64.0871	44.9760	22.6495	**20.6532**

Note: The best performance for each percentage of outliers is given in bold.

Table 4. TMSE and TMAE values for different estimates when $m = 6$, $k_i = 6$ and $\rho_\Sigma = 0.70$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	0.9468	**0.9206**	2.5687	1.0504	1.0448	7.7293	**7.3832**	14.3156	9.2018	9.1796
50	0.4970	**0.4397**	1.1760	0.5040	0.5031	6.3933	**5.9525**	9.7531	6.4371	6.4307
80	0.2689	**0.2469**	0.5939	0.2936	0.2936	4.7372	**4.4945**	6.9970	4.9430	4.9424
100	0.2174	**0.2074**	0.5153	0.2329	0.2330	4.1395	**3.8523**	6.5090	4.3953	4.3972
					$\tau = 10\%$					
30	236.7644	208.1461	81.1941	25.9744	**19.0315**	67.6543	64.3024	36.7668	26.9935	**22.4755**
50	131.9086	126.3575	39.6762	18.4750	**15.2936**	55.3219	52.0475	24.4006	22.1796	**19.7311**
80	88.0985	84.7778	17.7650	11.6467	**10.8537**	44.0977	43.1315	15.5958	17.7444	**13.9878**
100	75.5432	73.3401	16.4025	10.6078	**9.6791**	42.4872	40.3896	14.0191	12.1528	**11.6530**
					$\tau = 20\%$					
30	426.4927	387.8449	123.4940	62.8194	**55.5711**	96.0983	91.2297	66.4781	35.3564	**33.9758**
50	352.4068	241.4068	110.3439	49.8590	**40.5694**	78.1093	73.0477	46.2102	31.9344	**29.2316**
80	184.2984	179.3090	69.0886	38.2695	**31.5299**	66.3529	63.9161	37.3146	28.2812	**24.5428**
100	165.0145	160.5570	58.3372	32.2898	**23.6757**	62.5092	60.6745	34.5523	26.0526	**22.6162**
					$\tau = 30\%$					
30	607.5306	547.3879	146.1978	96.3106	**84.1927**	115.2914	109.2588	85.1167	49.2880	**42.3600**
50	363.2953	356.4610	121.8265	73.4655	**66.5410**	92.5863	89.3676	69.5004	38.9070	**33.5978**
80	284.0429	279.6431	93.9534	65.3239	**47.4196**	84.5215	80.8659	57.0963	35.4167	**31.0785**
100	261.0352	258.3064	86.2610	59.2066	**40.4774**	80.3973	78.3635	50.6596	30.6766	**28.0041**
					$\tau = 40\%$					
30	749.6527	680.5225	217.1058	124.2992	**106.3162**	130.4198	122.3600	97.3313	58.5610	**52.2773**
50	468.3829	460.1729	149.9310	108.3619	**93.6996**	108.5413	102.7551	89.4199	54.0609	**50.2988**
80	391.1478	386.0914	113.1978	82.7282	**65.4907**	98.9015	96.6531	81.7926	44.3549	**40.6617**
100	367.9219	363.8788	107.1333	65.0909	**51.5284**	95.6713	93.0857	77.6091	41.7737	**39.3203**

Note: The best performance for each percentage of outliers is given in bold.

Table 5. TMSE and TMAE values for different estimates when $m = 6$, $k_i = 6$ and $\rho_\Sigma = 0.90$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	0.7262	**0.7117**	3.0577	0.8996	0.8893	6.7844	**6.2742**	13.6670	7.5917	7.5527
50	0.3279	**0.3104**	1.7472	0.4816	0.4767	4.5917	**4.3747**	10.3303	5.5720	5.5464
80	0.1870	**0.1708**	1.0044	0.2797	0.2767	3.4759	**3.2554**	7.8495	4.2578	4.2364
100	0.1571	**0.1452**	0.9300	0.2392	0.2369	3.1305	**3.0805**	7.5335	3.9279	3.9104
					$\tau = 10\%$					
30	221.1813	197.1476	61.9059	26.1573	**20.2306**	75.5191	63.9641	33.3135	22.0346	**21.8194**
50	120.4342	113.7169	34.3213	19.9130	**15.2921**	53.1203	49.4990	19.9642	17.2164	**16.3173**
80	77.3602	74.9280	20.9429	11.0718	**11.3283**	44.9712	40.4753	13.7421	12.5258	**11.5515**
100	65.7502	62.2731	15.2447	10.1456	**9.1258**	39.5759	37.0838	12.3404	11.3834	**10.8263**
					$\tau = 20\%$					
30	369.6532	333.2891	109.7337	68.8135	**49.0722**	91.5352	84.0985	57.8944	33.4833	**32.4411**
50	201.9114	195.0401	95.9744	42.4803	**35.5455**	72.5917	64.9762	43.1790	29.6925	**27.6404**
80	153.2585	149.2244	64.3770	34.4412	**25.7216**	56.0660	53.3351	36.1790	26.7104	**23.4563**
100	141.1818	137.4720	54.8934	28.4822	**20.4071**	51.6723	48.9139	33.7412	24.5314	**20.6179**
					$\tau = 30\%$					
30	568.7293	484.9511	146.5433	94.7937	**81.9606**	104.2307	93.9518	80.6831	56.7814	**48.5773**
50	354.6912	349.8161	122.8340	73.8079	**58.1147**	85.1963	73.8546	63.6807	42.1507	**35.0507**
80	268.1570	262.7963	113.1883	57.8095	**42.7152**	73.6069	69.8979	53.2112	34.8496	**30.2510**
100	243.3130	237.2691	106.7808	51.8249	**36.8945**	70.1026	65.3408	50.4770	31.3564	**28.3533**
					$\tau = 40\%$					
30	622.3340	552.2631	180.7879	103.2528	**93.5939**	103.7687	97.7644	93.3615	72.9710	**67.9497**
50	484.7617	471.6830	150.3779	86.6461	**64.7116**	89.2152	81.0080	73.3443	64.1589	**56.8442**
80	386.4686	381.0374	120.8739	62.7029	**52.0914**	85.8742	79.2827	63.3082	45.9977	**43.4831**
100	351.1662	349.6917	118.5967	58.5677	**45.5407**	81.1460	77.7343	59.6490	38.2049	**32.4237**

Note: The best performance for each percentage of outliers is given in bold.

Table 6. TMSE and TMAE values for different estimates when $m = 8$, $k_i = 4$ and $\rho_\Sigma = 0.70$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	1.8813	**1.8293**	5.1040	2.0872	2.0760	11.4657	**10.9523**	21.2357	13.6499	13.6170
50	0.9875	**0.8737**	2.3368	1.0014	0.9998	9.4839	**8.8300**	14.4678	9.5488	9.5392
80	0.5342	**0.4906**	1.1801	0.5834	0.5834	7.0272	**6.6672**	10.3794	7.3324	7.3315
100	0.4320	**0.4121**	1.0240	0.4627	0.4630	6.1406	**5.7145**	9.6555	6.5200	6.5228
					$\tau = 10\%$					
30	291.5543	253.6436	97.4179	40.7617	**31.8662**	96.1699	89.9097	57.6981	42.3609	**35.2708**
50	207.0042	198.2928	62.2639	38.9929	**24.0003**	80.8166	74.6782	38.2919	34.8065	**30.9640**
80	138.2530	133.0418	30.6014	23.2772	**17.0326**	69.2024	67.6862	24.4745	21.9510	**27.8463**
100	118.5500	115.0926	21.3246	18.6469	**15.1895**	66.6752	61.3835	22.1422	18.2871	**15.9329**
					$\tau = 20\%$					
30	537.7993	476.3786	142.2134	78.9893	**69.6445**	123.8937	117.6170	85.7062	45.5829	**43.8029**
50	425.0590	383.6534	121.2598	61.2802	**52.3037**	100.7016	94.1760	59.5761	41.1711	**37.6865**
80	292.8944	284.9651	87.0718	49.3386	**36.1698**	85.5448	82.4032	48.1074	36.4612	**31.6416**
100	263.2477	255.1636	65.2106	41.6293	**30.5236**	80.5894	78.2240	44.5462	33.5880	**29.1576**
					$\tau = 30\%$					
30	643.5927	581.0710	168.5980	121.1478	**95.5208**	158.1913	149.9140	106.7886	67.6280	**58.1221**
50	498.4775	487.1001	147.1581	101.8020	**81.3009**	127.0377	122.6213	95.3615	53.3843	**46.0996**
80	389.7352	383.6983	118.9135	87.6309	**64.0644**	115.9719	110.9561	78.3419	48.5952	**42.6429**
100	358.1664	354.4703	98.3587	76.2374	**53.5390**	109.3131	102.5226	69.5100	42.0913	**38.4245**
					$\tau = 40\%$					
30	910.4704	826.5102	185.3905	143.3941	**113.7550**	171.4399	160.8450	127.9442	76.9798	**68.7198**
50	662.8618	658.8906	163.0877	122.4442	**99.1703**	142.6800	135.0739	117.5445	69.0643	**62.1190**
80	475.0580	468.9169	129.8012	108.7481	**86.0890**	130.0083	127.0527	93.5183	58.3055	**51.4508**
100	446.8496	439.9391	121.8292	92.5635	**67.7352**	125.7621	120.3633	88.0189	50.9125	**46.6875**

Note: The best performance for each percentage of outliers is given in bold.

Table 7. TMSE and TMAE values for different estimates when $m = 8$, $k_i = 4$ and $\rho_\Sigma = 0.90$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	1.4727	**1.4433**	4.2007	1.8243	1.8035	8.7946	**8.1333**	17.7166	9.8411	9.7906
50	0.6649	**0.6294**	2.0543	0.9767	0.9667	5.9522	**5.6709**	13.3912	7.2230	7.1897
80	0.3791	**0.3464**	1.0369	0.5672	0.5612	4.5058	**4.2200**	10.1753	5.5194	5.4917
100	0.3186	**0.2945**	0.9886	0.4851	0.4805	4.0581	**3.9932**	9.7657	5.0917	5.0690
					$\tau = 10\%$					
30	273.2766	232.7162	74.4755	38.7085	**31.6434**	97.4499	81.9490	56.2215	37.1866	**36.8234**
50	203.2508	191.9143	57.9224	27.4827	**23.2460**	79.6485	73.5370	33.6926	29.0552	**27.5379**
80	130.5569	126.4523	30.3443	18.6853	**18.8058**	75.8957	68.3081	23.1919	21.1392	**19.4949**
100	110.9633	105.0953	22.7276	17.1222	**15.0401**	66.7902	62.5844	20.8262	19.2111	**18.2710**
					$\tau = 20\%$					
30	486.7583	359.0368	117.2892	81.7037	**58.2644**	117.8352	108.2617	74.5287	43.1037	**41.7621**
50	320.4982	309.5913	98.9524	54.6858	**45.7585**	93.4488	83.6452	55.5852	38.2237	**35.5820**
80	211.5240	205.1207	76.4362	40.8927	**30.5397**	66.5684	63.3259	42.9561	31.7138	**27.8501**
100	208.4812	202.5926	65.1761	33.8175	**24.2298**	61.3516	58.0765	40.0616	29.1266	**24.4801**
					$\tau = 30\%$					
30	600.8170	512.3121	161.4656	109.6213	**94.7809**	124.9101	112.5918	96.6906	68.0468	**53.2150**
50	474.7028	469.5527	140.0477	85.3529	**67.2050**	102.0992	88.5074	76.3149	50.5134	**42.0048**
80	293.2864	287.6233	130.8932	66.8521	**49.3967**	88.2105	83.7656	63.7682	41.7637	**36.2528**
100	277.0407	270.6558	121.4834	59.9313	**42.6655**	84.0110	78.3044	60.4916	37.5775	**33.9785**
					$\tau = 40\%$					
30	808.2251	717.2241	193.6317	123.4439	**102.8316**	133.0300	125.3326	113.8879	81.0144	**72.8891**
50	629.5601	612.5747	165.2202	95.1981	**71.0987**	114.3726	103.8512	89.4698	70.2648	**63.3419**
80	501.9068	494.8532	132.8041	68.8917	**57.2328**	110.0895	101.6393	77.2271	56.1107	**48.0433**
100	456.0596	454.1446	130.3022	64.3484	**50.0356**	104.0281	99.6543	72.7634	46.6047	**39.5524**

Note: The best performance for each percentage of outliers is given in bold.

Table 8. TMSE and TMAE values for different estimates when $m = 8$, $k_i = 6$ and $\rho_\Sigma = 0.70$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	2.8025	**2.7251**	7.6034	3.1093	3.0927	15.1863	**14.5063**	28.1267	18.0793	18.0357
50	1.4710	**1.3015**	3.4811	1.4918	1.4893	12.5614	**11.6953**	19.1626	12.6474	12.6347
80	0.9077	**0.8335**	2.0049	0.9913	0.9911	9.3075	**8.8307**	13.7475	9.7118	9.7106
100	0.7339	**0.7001**	1.7397	0.7861	0.7866	8.1332	**7.5689**	12.7887	8.6357	8.6394
					$\tau = 10\%$					
30	462.9620	402.7633	124.6908	64.7259	**50.6006**	132.7091	121.7686	91.6193	67.2652	**56.0069**
50	328.7041	314.8711	98.8694	51.9172	**38.1103**	118.3296	113.5822	60.8041	55.2695	**49.1681**
80	219.5334	211.2585	48.5923	36.9620	**27.0463**	109.8872	107.4796	38.8633	34.8562	**44.2174**
100	215.2467	208.7568	33.8616	29.6096	**22.1195**	105.8742	97.4714	35.1599	29.0383	**25.2999**
					$\tau = 20\%$					
30	698.4399	618.6729	161.4712	92.6845	**76.4828**	146.0088	138.6116	118.1460	62.8360	**60.3823**
50	552.0242	498.2507	135.3541	73.4566	**62.6964**	118.6769	110.9865	82.1256	56.7543	**51.9508**
80	380.3820	370.0842	94.3729	59.1422	**43.3567**	100.8146	97.1122	66.3161	50.2618	**43.6180**
100	341.8798	331.3810	78.1679	49.9011	**36.5887**	94.9746	92.1869	61.4070	46.3011	**40.1938**
					$\tau = 30\%$					
30	835.6408	754.4626	175.1880	133.0687	**104.9200**	173.7099	164.6206	127.9434	81.0251	**69.6361**
50	647.2232	632.4508	158.6384	111.8193	**89.3009**	139.5001	134.6504	114.2526	63.9597	**55.2319**
80	506.0322	498.1939	120.6145	96.2538	**70.3684**	127.3488	121.8409	93.8614	58.2219	**51.0904**
100	465.0433	460.2442	111.0372	83.7392	**58.8073**	120.0367	115.5800	84.2800	50.4296	**46.0363**
					$\tau = 40\%$					
30	1203.0045	1092.0679	196.7326	157.3710	**129.1224**	207.9565	195.1049	143.2771	92.2218	**82.3263**
50	875.8393	870.5921	177.4563	139.7492	**110.2852**	173.0709	163.8447	120.8183	82.7391	**74.4185**
80	627.6942	619.5799	158.7468	113.9990	**98.2869**	157.7001	154.1149	102.0349	69.8500	**61.6380**
100	590.4223	581.2916	151.9971	107.2052	**82.8402**	152.5494	146.6592	97.4466	60.9932	**55.9316**

Note: The best performance for each percentage of outliers is given in bold.

Table 9. TMSE and TMAE values for different estimates when $m = 8$, $k_i = 6$ and $\rho_\Sigma = 0.90$

n	TMSE					TMAE				
	ML	**FGLS**	**M**	**S**	**MM**	**ML**	**FGLS**	**M**	**S**	**MM**
					$\tau = 0\%$					
30	2.2599	**2.2148**	6.4459	2.7994	2.7675	13.8379	**12.7974**	27.8763	15.4846	15.4051
50	1.0203	**0.9659**	3.1523	1.4988	1.4835	9.3656	**8.9230**	21.0705	11.3651	11.3128
80	0.6576	**0.6008**	1.5912	0.8703	0.8611	7.0896	**6.6400**	16.0104	8.6845	8.6410
100	0.5525	**0.5109**	1.5170	0.7445	0.7373	6.3853	**6.2832**	15.3659	8.0116	7.9759
					$\tau = 10\%$					
30	366.8192	312.3749	92.3347	47.9908	**39.2315**	118.1949	99.3943	68.1899	45.1029	**42.6624**
50	272.8236	257.6066	71.8122	34.0730	**28.8203**	96.6039	89.1915	40.8651	35.2405	**33.4002**
80	175.2465	169.7369	37.6208	23.1661	**23.3155**	92.0523	82.8494	28.1290	25.6393	**23.6450**
100	148.9461	141.0694	28.1777	21.2281	**18.6467**	81.0085	75.9073	25.2597	23.3008	**22.1606**
					$\tau = 20\%$					
30	598.6494	441.5685	128.5215	92.3146	**65.8312**	143.6128	131.9451	98.9014	52.6008	**50.5468**
50	394.1711	380.7570	111.8033	61.7879	**51.7011**	113.8916	101.9434	68.7483	47.5097	**43.4887**
80	260.1470	252.2718	86.3630	46.2034	**34.5059**	98.1309	96.1791	55.5140	42.0748	**36.5131**
100	256.4048	249.1626	73.6405	38.2094	**26.3766**	87.7728	84.7813	51.4045	38.7592	**33.6467**
					$\tau = 30\%$					
30	780.2810	665.3397	175.6144	120.7319	**98.3873**	159.8265	144.0648	108.4751	76.3402	**59.7008**
50	616.4966	609.8081	154.0544	94.0038	**74.0165**	130.6392	113.2481	85.6161	56.6699	**47.1243**
80	380.8910	373.5363	143.9843	73.6278	**58.4033**	112.8682	107.1808	71.5403	46.8538	**40.6713**
100	359.7928	351.5007	133.6334	66.0056	**47.9898**	106.4948	99.1930	67.8643	42.1574	**38.1198**
					$\tau = 40\%$					
30	985.2264	874.2961	195.7165	134.8985	**112.8708**	172.3883	162.4135	116.6462	85.5495	**77.6303**
50	767.4337	746.7286	172.5972	110.9724	**82.8797**	148.2109	134.5766	96.0638	78.4069	**66.3277**
80	611.8244	603.2261	154.8098	80.3070	**66.7163**	142.6606	131.7104	87.2347	63.3819	**53.9170**
100	555.9366	553.6022	150.8933	75.0109	**58.3265**	134.8058	129.1381	72.1925	52.6440	**44.6778**

Note: The best performance for each percentage of outliers is given in bold.

Table 10. Estimation results and RAB values for different estimates when $n = 20$, $\rho_\Sigma = 0.70$ and $\tau = 10\%$
The true value of β is $(1,2,3,4,5)'$

Equations	Parameters	ML		FGLS		M		S		MM	
		Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB
Equation 1											
	β_{11}	3.5781	2.5781	3.5013	2.5013	1.6956	0.6956	1.3956	0.3956	0.9299	**0.0701**
	β_{12}	6.2300	2.1150	6.3030	2.1515	3.1152	0.5576	2.3948	0.1974	1.9248	**0.0376**
	β_{13}	0.5800	0.8067	0.6159	0.7947	1.9670	0.3443	2.4404	0.1865	2.7210	**0.0930**
	β_{14}	0.5154	0.8711	2.7603	0.3099	4.1986	0.0496	4.0662	0.0165	3.9465	**0.0134**
	β_{15}	8.4105	0.6821	8.2618	0.6524	3.9317	0.2137	4.0732	0.1654	4.4955	**0.1009**
Equation 2											
	β_{21}	-0.1413	1.1413	0.1844	0.8156	0.6111	0.3890	1.2427	0.2427	0.9283	**0.0717**
	β_{22}	4.4380	1.2190	3.3552	0.6776	2.8854	0.4427	2.5407	0.2703	1.9035	**0.0482**
	β_{23}	1.3254	0.5582	4.3075	0.4358	4.2300	0.4100	3.6733	0.2244	3.1628	**0.0543**
	β_{24}	-0.8510	1.2128	-1.0132	1.2533	2.3393	0.4152	3.6095	0.0976	4.0912	**0.0228**
	β_{25}	9.3800	0.8760	9.0175	0.8035	7.6545	0.5309	6.2571	0.2514	5.1298	**0.0260**
Equation 3											
	β_{31}	3.1950	2.1950	2.6386	1.6386	0.6520	0.3480	0.8324	0.1676	0.9486	**0.0514**
	β_{32}	0.3080	0.8460	0.3599	0.8201	2.3579	0.1789	1.4862	0.2569	1.7563	**0.1219**
	β_{33}	4.6952	0.5651	3.8782	0.0588	3.1765	0.0588	3.1362	0.0454	3.0456	**0.0152**
	β_{34}	5.3467	0.3367	5.6180	0.4045	4.8462	0.2116	4.4371	0.1093	4.1038	**0.0259**
	β_{35}	9.2022	0.8404	6.1314	0.2263	6.1457	0.2291	5.3607	0.0721	4.6938	**0.0612**

Note: The best performance is given in bold.

Table 11. Estimation results and RAB values for different estimates when $n = 20$, $\rho_{\Sigma} = 0.70$ and $\tau = 40\%$
The true value of β is $(1,2,3,4,5)'$

Equations	Parameters	ML		FGLS		M		S		MM	
		Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB
Equation 1											
	β_{11}	4.1142	3.1142	4.0205	3.0205	2.4257	1.4257	1.3892	0.3892	1.1825	**0.1825**
	β_{12}	8.0690	3.0345	7.8837	2.9419	5.0352	1.5176	1.8532	0.0734	2.2489	**0.1245**
	β_{13}	3.6605	0.2202	3.2681	0.0894	3.7235	0.2412	3.8727	0.2909	2.9568	**0.0144**
	β_{14}	3.5332	0.1167	3.1267	0.2183	5.3817	0.3454	4.6905	0.1726	4.0872	**0.0218**
	β_{15}	-0.1007	1.0201	-1.1943	1.2389	4.6905	0.0619	4.5852	0.0830	4.6193	**0.0761**
Equation 2											
	β_{21}	5.5691	4.5691	5.1302	4.1302	3.4755	2.4755	0.8918	0.1082	1.1384	**0.1384**
	β_{22}	6.8121	2.4061	5.9068	1.9534	2.7553	0.3777	2.4283	0.2142	1.8935	**0.0533**
	β_{23}	2.1074	0.2975	2.0212	0.3263	2.4452	0.1849	2.7985	0.0672	2.5834	**0.1389**
	β_{24}	8.2854	1.0713	5.0367	0.2592	5.4912	0.3728	4.7360	0.1840	3.5946	**0.1014**
	β_{25}	14.8642	1.9728	13.4798	1.6960	6.6065	0.3213	5.9740	0.1948	4.6787	**0.0643**
Equation 3											
	β_{31}	9.5633	8.5633	8.5704	7.5704	3.6054	2.6054	1.0905	0.0905	1.2531	**0.2531**
	β_{32}	3.9257	0.9628	4.4539	1.2270	1.7285	0.1358	1.5037	0.2481	1.7627	**0.1187**
	β_{33}	5.8704	0.9568	5.1775	0.7258	3.3082	0.1027	3.5785	0.1928	2.9363	**0.0212**
	β_{34}	5.7541	0.4385	6.8425	0.7106	5.4272	0.3568	3.8959	0.0260	4.3064	**0.0766**
	β_{35}	6.3037	0.2607	7.4357	0.4871	6.5601	0.3120	4.5713	0.0857	5.0669	**0.0134**

Note: The best performance is given in bold.

Table 12. Estimation results and RAB values for different estimates when $n = 20$, $\rho_{\Sigma} = 0.90$ and $\tau = 10\%$
The true value of β is $(1,2,3,4,5)'$

Equations	Parameters	ML		FGLS		M		S		MM	
		Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB
Equation 1											
	β_{11}	3.4240	2.4240	3.4742	2.4742	1.5787	0.5787	2.9822	1.9822	0.9365	**0.0635**
	β_{12}	2.6304	0.3152	2.7377	0.3689	2.3995	0.1998	2.0690	0.0345	2.2978	**0.1489**
	β_{13}	-0.8983	1.2994	-1.0462	1.3487	2.5970	0.1343	3.2850	0.0950	2.7241	**0.0920**
	β_{14}	4.4219	0.1055	4.9867	0.2467	3.7380	0.0655	4.1477	0.0369	4.1206	**0.0302**
	β_{15}	8.0989	0.6198	8.3205	0.6641	4.8353	0.0329	5.2344	0.0469	4.7901	**0.0420**
Equation 2											
	β_{21}	-0.1622	1.1622	1.2389	0.2389	1.1553	0.1553	0.8578	0.1422	0.8964	**0.1036**
	β_{22}	6.0623	2.0311	3.1781	0.5890	2.5371	0.2686	2.6117	0.3059	2.0371	**0.0186**
	β_{23}	3.9057	0.3019	4.0876	0.3625	2.7637	0.0788	2.8045	0.0652	2.9148	**0.0284**
	β_{24}	-0.6517	1.1629	-0.7834	1.1959	2.4205	0.3949	3.0699	0.2325	3.5748	**0.1063**
	β_{25}	9.6272	0.9254	8.8977	0.7795	5.5196	0.1039	4.4435	0.1113	4.8486	**0.0303**
Equation 3											
	β_{31}	3.1525	2.1525	2.5950	1.5950	0.7965	0.2035	0.8590	0.1410	0.8746	**0.1254**
	β_{32}	-2.0528	2.0264	-0.3481	1.1741	1.3070	0.3465	1.4489	0.2755	1.7650	**0.1175**
	β_{33}	3.8804	0.2935	3.6991	0.2330	3.1417	0.0472	3.1593	0.0531	3.0420	**0.0140**
	β_{34}	6.1190	0.5298	5.6147	0.4037	4.6416	0.1604	4.2450	0.0613	4.4538	**0.1135**
	β_{35}	10.1303	1.0261	8.6816	0.7363	6.0919	0.2184	5.4760	0.0952	5.1368	**0.0274**

Note: The best performance is given in bold.

Table 13. Estimation results and RAB values for different estimates when $n = 20$, $\rho_{\Sigma} = 0.90$ and $\tau = 40\%$
The true value of β is $(1,2,3,4,5)'$

Equations	Parameters	ML		FGLS		M		S		MM	
		Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB	Estimate	RAB
Equation 1											
	β_{11}	3.7416	2.7416	3.6695	2.6695	2.2403	1.2403	0.5768	0.4232	1.1510	**0.1510**
	β_{12}	-0.4599	1.2300	-0.7879	1.3939	3.5526	0.7763	1.8392	0.0804	2.0912	**0.0456**
	β_{13}	3.5703	0.1901	3.2170	0.0723	3.4361	0.1454	2.8112	0.0629	2.6963	**0.1012**
	β_{14}	5.6196	0.4049	5.0466	0.2616	4.2526	0.0631	4.4901	0.1225	4.1527	**0.0382**
	β_{15}	7.4023	0.4805	7.2352	0.4470	4.3472	0.1306	3.2583	0.3483	4.8516	**0.0297**
Equation 2											
	β_{21}	5.0903	4.0903	4.6945	3.6945	2.9804	1.9804	1.5639	0.5639	1.1874	**0.1874**
	β_{22}	6.5210	2.2605	5.4194	1.7097	3.3424	0.6712	2.4715	0.2357	1.7815	**0.1093**
	β_{23}	2.2483	0.2506	2.1725	0.2758	2.5627	0.1458	3.0728	0.0243	2.8153	**0.0616**
	β_{24}	8.2770	1.0692	5.3216	0.3304	4.9076	0.2269	4.6894	0.1724	4.0464	**0.0116**
	β_{25}	13.7882	1.7576	12.5479	1.5096	6.0136	0.2027	5.2807	0.0561	5.1787	**0.0357**
Equation 3											
	β_{31}	8.5755	7.5755	7.7164	6.7164	1.5376	0.5376	1.4845	0.4845	0.8742	**0.1258**
	β_{32}	3.8490	0.9245	4.2817	1.1408	3.5793	0.7896	2.7164	0.3582	2.1418	**0.0709**
	β_{33}	5.5357	0.8452	4.9268	0.6423	3.1674	0.0558	0.5432	0.8189	3.0574	**0.0191**
	β_{34}	5.5064	0.3766	6.4702	0.6176	4.9688	0.2422	1.2153	0.6962	3.9026	**0.0243**
	β_{35}	7.0574	0.4115	6.0721	0.2144	4.7276	0.0545	4.5584	0.0883	4.8913	**0.0217**

Note: The best performance is given in bold.

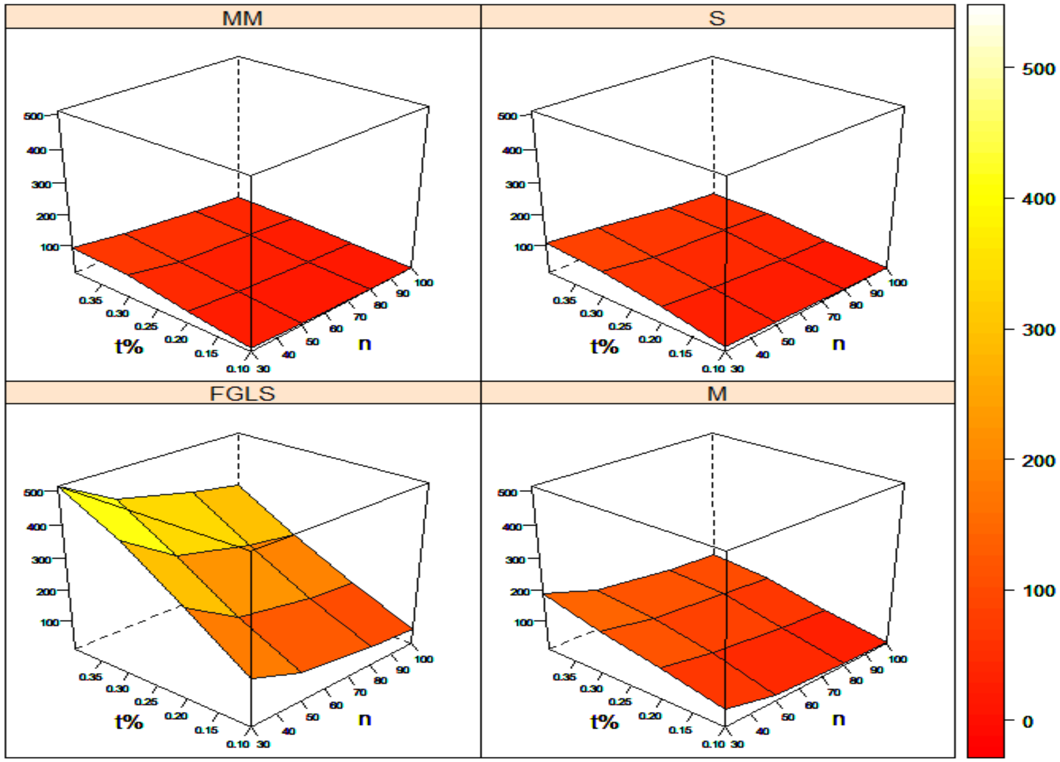

Fig. 1: Average TMSE values for different estimates in all cases when $m = 6$

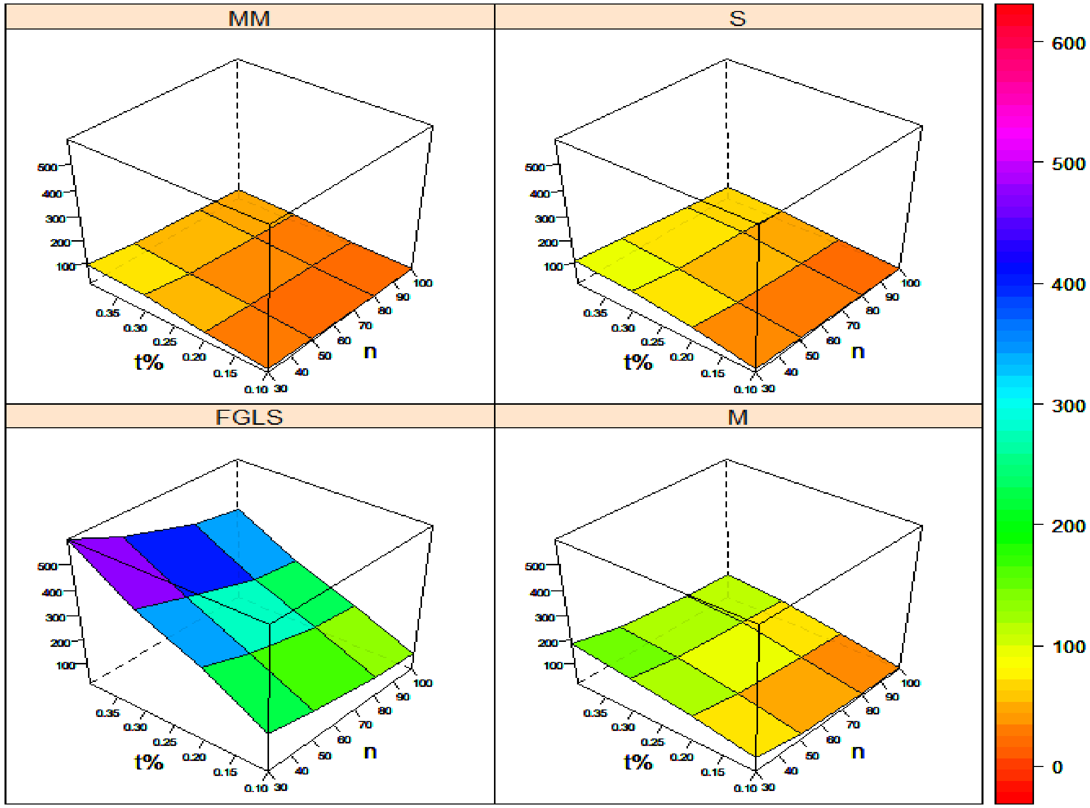

Fig. 2: Average TMSE values for different estimates in all cases when $k_i = 4$

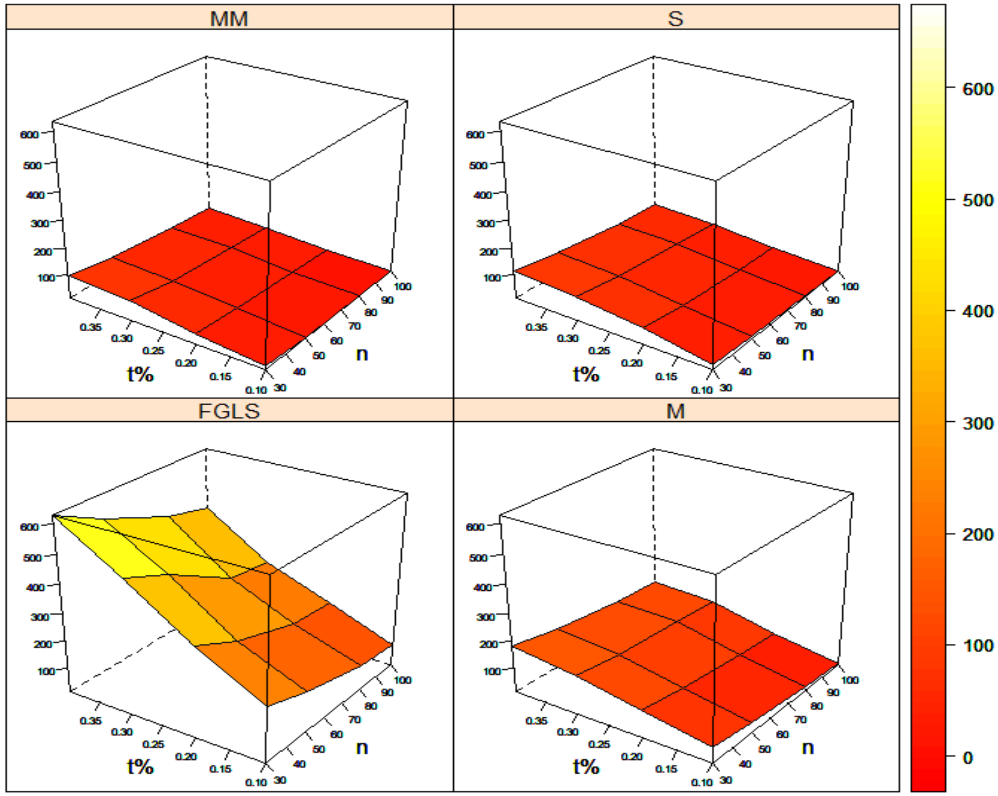

Fig. 3: Average TMSE values for different estimates in all cases when $\rho_{\Sigma} = 0.90$

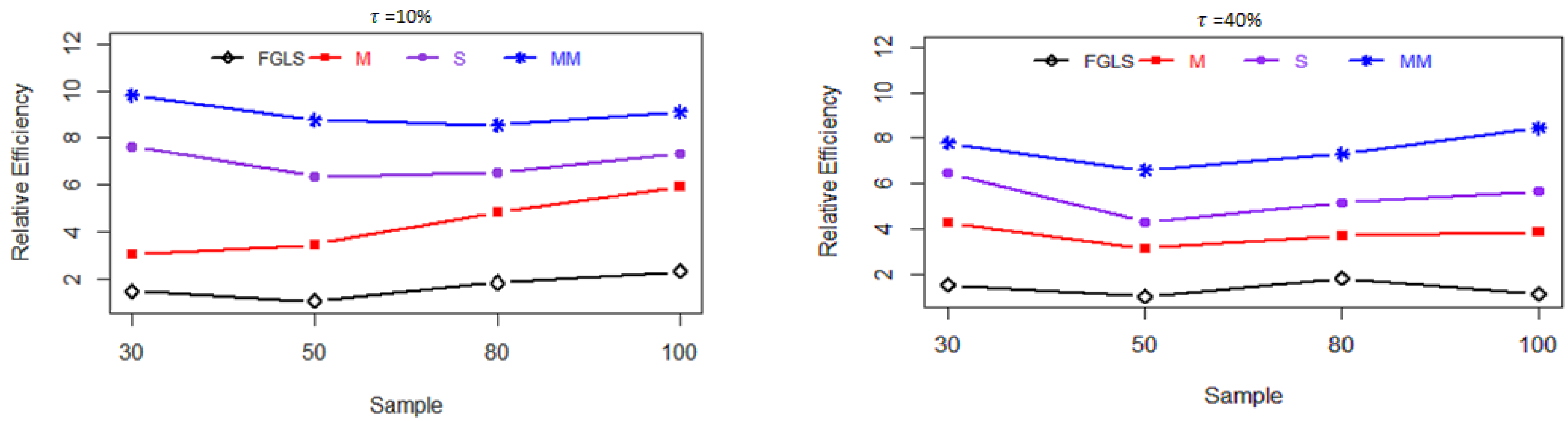

Fig. 4: Relative efficiency for the different estimates when $\rho_{\Sigma} = 0.70$

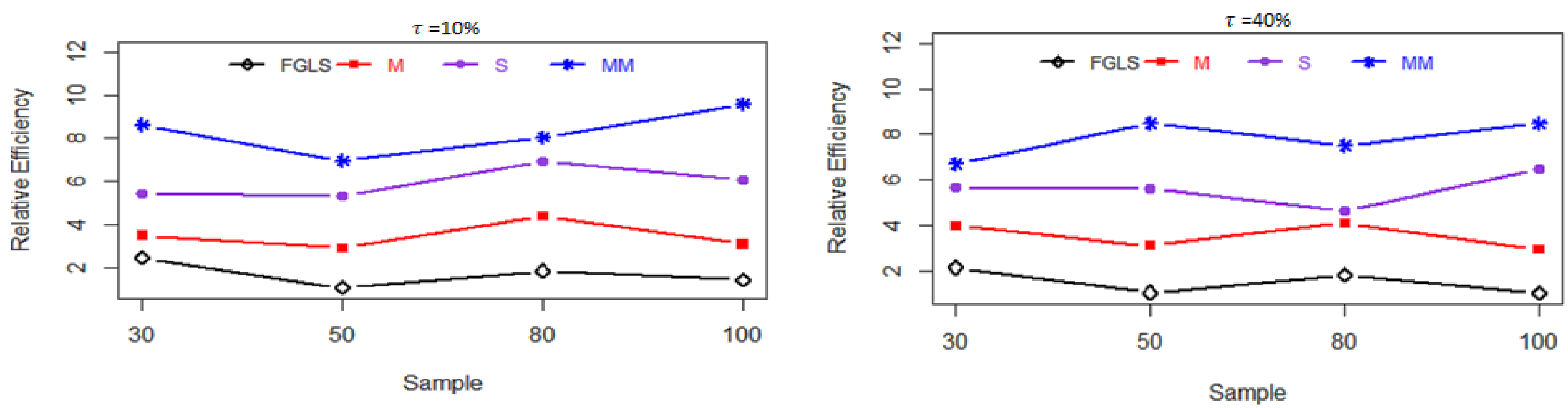

Fig. 5: Relative efficiency for the different estimates when $\rho_{\Sigma} = 0.90$

5 Conclusions

In this paper, we have reviewed three robust (M, S, and MM) estimators of the SUR model and compared these estimators with non-robust (ML and FGLS) estimators when the outliers are present. Moreover, our new algorithm for robust SUR provides robust parameter estimates and useful outlier diagnostics, as illustrated in the simulation study. Simulation study results indicated that, in general, non-robust estimators are very sensitive to outliers, while robust estimators are more effective. In addition, the MM-estimator is more efficient than other robust estimators because it has minimum RAB, TMSE, and TMAE values in all simulation situations. Also, the results showed that in the absence of outliers the FGLS estimator is more efficient than ML, M, S, and MM estimators.

In future work, we plan to study the efficiency of the robust estimators in other models, such as semi-parametric regression models [35,36] and the autoregressive integrated moving average (ARIMA) model [37,38]. Moreover, we can study how to combine robust estimators with neural networks (NN) or artificial intelligence (AI) methods [39].

References:

[1] A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American statistical Association*, 57, 348-368 (1962).

[2] A. Zellner, Estimators for seemingly unrelated regression equations: Some exact finite sample results, *Journal of the American Statistical Association*, 58, 977-992 (1963).

[3] M.R. Abonazel, Different estimators for stochastic parameter panel data models with serially correlated errors. *Journal of Statistics Applications & Probability* 7.3, 423-434 (2018).

[4] M.R. Abonazel, Generalized estimators of stationary random-coefficients panel data models: Asymptotic and small sample properties." *Revstat Statistical Journal* 17.4, 493-521(2019).

[5] A.R. Kamel, *Handling outliers in seemingly unrelated regression equations model*, MSc thesis, Faculty of graduate studies for statistical research (FGSSR), Cairo University, Egypt, (2021).

[6] G. Saraceno, F. Alqallaf and C. Agostinelli, A Robust Seemingly Unrelated Regressions for Row-Wise and Cell-Wise Contamination. *arXiv preprintarXiv:2107.00975*, (2021). Available at: https://arxiv.org/pdf/2107.00975.pdf

[7] V. K. Srivastava and T. D. Dwivedi, Estimation of seemingly unrelated regression equations: A brief survey, *Journal of Econometrics*, 10, 15-32 (1979).

[8] P. Schmidt, A note on the estimation of seemingly unrelated regression systems, *Journal of Econometrics*, 7, 259-261(1978).

[9] V. K. Srivastava and D. E. Giles, *Seemingly unrelated regression equations models: Estimation and inference*. 2nd Edition, CRC Press, (2020).

[10] J. Kmenta and R. Gilbert, Small sample properties of alternative estimators of seemingly unrelated regressions, *Journal of the American Statistical Association*, 63, 1180-1200 (1968).

[11] A. A. Alharbi, A. R. Kamel, and S. A. Atia, A new robust molding of heat and mass transfer process in MHD based on adaptive-network-based fuzzy inference system, *WSEAS Transactions on Heat and Mass Transfer*, vol. 17, pp. 80-96, (2022).

[12] P. Rousseeuw, S. Van Aelst, K. Van Driessen and J. Gulló, Robust multivariate regression, *Technometrics*, 46, 293-305 (2004).

[13] I. Dawoud and M. R. Abonazel, Robust Dawoud–Kibria estimator for handling multicollinearity and outliers in the linear regression model. *Journal of Statistical Computation and Simulation*, 91, 3678-3692 (2021).

[14] F. A. Awwad, I. Dawoud and M. R. Abonazel, Development of robust Özkale–Kaçiranlar and Yang–Chang estimators for regression models in the presence of multicollinearity and outliers. *Concurrency and Computation: Practice and Experience*, *34*, e6779 (2022).

[15] M. R. Abonazel, S. M. El-Sayed, and O. M. Saber, Performance of robust count regression estimators in the case of overdispersion, zero inflated, and outliers: simulation study and application to German health data. *Commun. Math. Biol. Neurosci.*, 2021, Article-ID 55 (2021).

[16] M.R. Abonazel, and O. M. Saber, Comparative study of robust estimators for Poisson regression model with outliers. *Journal of Statistics Applications and Probability*, 9, 279-286 (2020).

[17] M.R. Abonazel, and I. Dawoud, Developing robust ridge estimators for Poisson regression model. *Concurrency and Computation Practice and Experience*. (2022) DOI**:** 10.1002/cpe.6979

[18] R. Koenker and S. Portnoy, M-estimation of multivariate regressions, *Journal of the American Statistical Association*, 85, 1060-1068 (1990).
[19] J. Jurečková, J. Picek and M. Schindler, *Robust statistical methods with R*, CRC Press, (2019). Available at: https://b-ok.africa/book/5405741/23690c.
[20] D. L. Donoho and P. J Huber, The notion of breakdown point, *A festschrift for Erich L. Lehmann*, 157-184 (1983).
[21] P. J. Rousseeuw and A. M. Leroy, *Robust regression and outlier detection*, Vol. 589, John wiley & Sons, (2005).
[22] M. Salibian-Barrera and V. Yohai, A fast algorithm for S-regression estimates, *Journal of computational and Graphical Statistics*, 15, 414-427 (2006).
[23] M. Hubert, T. Verdonck and O. Yorulmaz, Fast robust SUR with economical and actuarial applications, *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 10, 77-88 (2017).
[24] K. Tharmaratnam, G. Claeskens, C. Croux and M. Salibián-Barrera, S-estimation for penalized regression splines, *Journal of Computational and Graphical Statistics*, 19, 609-625(2010).
[25] M. Bilodeau and P. Duchesne, Robust estimation of the SUR model, *Canadian Journal of Statistics*, 28, 277-288 (2000).
[26] S. Van Aelst and Willems, G, Multivariate regression S-estimators for robust estimation and inference, *Statistica Sinica*, 981-1001(2005).
[27] R. A. Maronna, R. D. Martin and V. J. Yohai, *Robust statistics: theory and methods (with R)*. John Wiley & Sons, (2019).
[28] K. Peremans and S. Van Aelst, Robust inference for seemingly unrelated regression models, *Journal of Multivariate Analysis*, 167, 212-224 (2018).
[29] A. H.Youssef, M. R. Abonazel and A. R. Kamel, Robust SURE estimates of profitability in the Egyptian insurance market, *Statistical journal of the IAOS*, 37, 1275-1287 (2021). DOI:10.3233/SJI-200734.
[30] M.R. Abonazel and A. R. Rabie, The impact for using robust estimations in regression models: an application on the Egyptian economy, *Journal of Advanced Research in Applied Mathematics and Statistics*, 4, 8-16 (2019).
[31] N. L. Kudraszow and R. A. Maronna, Estimates of MM type for the multivariate linear model, *Journal of multivariate analysis*, 102, 1280-1292 (2011).
[32] M. R Abonazel, Handling Outliers and Missing Data in Regression Models Using R: Simulation Examples, *Academic Journal of Applied Mathematical Sciences*, 6, 187-203 (2020). Available at: https://doi.org/10.32861/ajams.68.187.203.
[33] M. R. Abonazel, New ridge estimators of SUR model when the errors are serially correlated, *International Journal of Mathematical Archive*, 10, 53-62 (2019).
[34] M. R. Abonazel, A practical guide for creating Monte Carlo simulation studies using R, *International Journal of Mathematics and Computational Science*, 4, 18-33 (2018).
[35] M. R. Abonazel, A.A. Gad, Robust partial residuals estimation in semiparametric partially linear model. Communications in Statistics-Simulation and Computation, 49, 1223-1236(2020).
[36] M.R. Abonazel, F. Shoaeb, M.N. Abdel Fattah, and S. Abdel-Rahman, Semi-parametric estimation of Engel curve and measuring food poverty in rural Egypt. *Journal of Applied Probability & Statistics*, 16, 31-53(2021).
[37] M.R. Abonazel and N.M. Darwish, Forecasting confirmed and recovered COVID-19 cases and deaths in Egypt after the genetic mutation of the virus: ARIMA Box-Jenkins approach. *Communications in Mathematical Biology and Neuroscience*,17(2022).
[38] F.A. Awwad, M.A. Mohamoud, and M.R. Abonazel. Estimating COVID-19 cases in Makkah region of Saudi Arabia: Space-time ARIMA modeling. *PLOS ONE*, 16(4), e0250149(2021).
[39] A.A El-Sheikh, M.R. Abonazel, and M.C. Ali, Proposed two variable selection methods for big data: simulation and application to air quality data in Italy. *Communications in Mathematical Biology and Neuroscience*, 16(2022).