Original article

Selecting the covariance structure for the seemingly unrelated regression models

Ali Hussein AL-Marshadi, Muhammad Aslam *, Abdullah Alharbey

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia

ABSTRACT

Objective: This paper is concerned with evaluating suggested approach of selecting the suitable covariance structure for fitting the seemingly unrelated regression equations (SURE) models efficiently.
Method: The paper assessed AL-Marshadi (2014) methodology in terms of its percentage of times that it identifies the right covariance structure for mixed model analysis of SURE models using simulated data.
Application: The simulated equations of SURE models have identical explanatory variables, the regressors in one block of equations are a subset of those in another, and different regressors in the equations with various settings of covariance structures of $\sum$. Moreover, the percentage of times that REML fail to converge under normal situation are reported. The application of the proposed methodology is given using a panel of data.
Conclusions: In short, AL-Marshadi (2014) methodology provided an excellent tool for selecting the right covariance structure for SURE models using restricted maximum likelihood (REML) estimation method in order to fit the SURE models more efficiently than the existing method that considering the stander unstructured covariance structure in fitting SURE models.

1. Introduction

A lot of studies come across econometric models having several equations to model a real-life situation. But it so happens that the disturbance terms for such equations are somewhat correlated meaning thereby that variables affecting the disturbance term in one equation may also simultaneously affect the disturbance term in some other equation in the system of equations under study. Numerous econometrically estimated theoretical models consist of more than one equation. Overlooking such a correlation of the disturbance terms results in producing inefficient estimates of the coefficients. But estimating all such equations simultaneously with generalized least squares (GLS (estimator, considering the suitable covariance structure of the residuals, leads to efficient estimates. The nomenclature for such is commonly known as "seemingly unrelated regression equations "(SURE), Zellner (1962). Foschi and Kontoghiorghes (2003) worked on estimating regression with autoregressive disturbance. Banterle et al. (2018), Feng and Polson (2020), and Bottolo et al. (2021) worked on regression theory using the Bayesian approach.

The paper assessed methodology was given by AL-Marshadi (2014) in terms of its percentage of times that it identifies the right covariance structure for mixed model analysis of SURE models using simulated data. The simulated equations of SURE models have identical explanatory variables, the regressors in one block of equations are a subset of those in another, and different regressors in the equations with various settings of covariance structures of $\sum$. It is expected that the proposed method will be more efficient than the existing method of considering the stander unstructured covariance structure in fitting SURE models' since the number of the parameters in the covariance structure will be reduced when the suitable covariance structure is selected other than the stander unstructured covariance structure.

2. The SURE model

The basic model we are concerned with comprises multiple regression equations, Greene (2003).

* Corresponding author.
 E-mail addresses: aalmarshadi@kau.edu.sa (A.H. AL-Marshadi), aslam_ravian@hotmail.com (M. Aslam), aharbey@kau.edu.sa (A. Alharbey).

$$y_{ti} = \sum_{j=1}^{K_i} x_{tij}\beta_{ij} + \epsilon_{ti} \tag{1}$$

$$(t = 1, \cdots, T; \quad i = 1, \cdots, M)$$

$$(j = 1, 2, \cdots, K_i)$$

where y_{it} is the t'th observation on the i'th dependent variable (the variable to be "explained" by the i'th regression equation); x_{tij} is the t'th observation on the j'th regressor or explanatory variable appearing in the i'th equation; β_{ij} is the coefficient associated with x_{tij} at each observation; and ϵ_{ti} is the t'th value of the random disturbance term associated with the i'th equation of the model.

In matrix notation, this M-equation model may be expressed more compactly as.

$$y_{i=X_i\beta_i+\epsilon_i} \quad (i = 1, \cdots, M) \tag{2}$$

where y_i is a $(T \times 1)$ vector with a typical element y_{it}; X_i is a $(T \times K_i)$ matrix, each column of which comprises the T observations on a regressor in the i'th equation of the model; β_i is a $(K_i \times 1)$ vector with the typical element β_{ij}; and ϵ_i is the corresponding $(T \times 1)$ disturbances vector.

By writing (2) as.

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_M \end{bmatrix}_{TM\times1} = \begin{bmatrix} X_1 & 0 & . & . & . & 0 \\ 0 & X_2 & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & . & X_M \end{bmatrix}_{TM\times K^*} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_M \end{bmatrix}_{K^*\times1} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ . \\ . \\ . \\ \epsilon_M \end{bmatrix}_{TM\times1}$$

the model may be expressed in the compact form.

$$y = X\beta + \epsilon \tag{3}$$

where y is $(TM \times 1)$, X is $(TM \times K^*)$, β is $(K^* \times 1)$, ϵ is $(TM \times 1)$, and $K^* = \sum_i K_i$

We assume that elements of the disturbance vector ϵ_i follow a multivariate probability distribution with

$$E(\epsilon_i) = 0 \quad \text{for all } i, \quad (i = 1, \cdots, M) \tag{4}$$

$$E(\epsilon_i\epsilon'_i) = \sigma_{ii}I_T \tag{5}$$

$$E(\epsilon_i\epsilon'_j) = \sigma_{ij}I_T \quad (i,j = 1, \cdots, M) \tag{6}$$

Consider the non-singular matrices, as Q_{ii} and Q_{ij}, which have the finite elements.

i.e. $Q_{ii} = \lim_{T\to\infty} \left(\frac{1}{T}X\prime_iX_i\right)$

and $Q_{ij} = \lim_{T\to\infty} \left(\frac{1}{T}X\prime_iX_j\right)$

Writing (4), (5) and (6) in compact form, we have

$$E(\in \in') = \begin{bmatrix} \sigma_{11}I_T & . & . & . & \sigma_{1M}I_T \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \sigma_{M1}I_T & . & . & . & \sigma_{MM}I_T \end{bmatrix}_{MT\times MT} = \left(\sum \otimes I_T\right) = \Psi$$

where Ψ is an $(MT \times MT)$ matrix.

2.1. Approaches for estimating the SURE model

For estimating the parameters of the SURE model, various estimators have been proposed, by considering those based on the principle of least squares. By observing these estimators, we recognize two basic approaches: the first approach is estimating the parameters of the SURE model equation by equation perhaps using ordinary least squares (OLS); the second approach is to estimate the parameters in all of the equations jointly, perhaps using the GLS estimator or one of its feasible variants in the case where $\sum$ is unknown Greene, (2003).

2.2. OLS estimation of the SURE model

The OLS estimation method ignores the essential jointness of the relationships that make up the SURE model. This method implicitly assumes that the SURE model (3) comprises a set of regression equations that are independent of one another.

The most frequently used method for the general linear model is OLS estimation.

Consider the system of the SURE model in matrix notation as

$$y_{TM\times1=X_{TM\times K^*}\beta_{K^*\times1}+\epsilon_{TM\times1}} \tag{7}$$

With the assumptions,

$$E(\epsilon_i) = 0$$

$$E(\varepsilon\varepsilon') = \left(\overset{\otimes}{\sum} I_T\right) = \Psi$$

The OLS estimation method is applied to the combined Eq. (7) is identical to OLS applied to each equation separately.

The Eq. (7) can be rewritten as a general linear model,

$$y = X\beta + \epsilon \tag{8}$$

where β is regression co-efficient. Let $\widehat{\beta}$ denote a column vector of the estimates of β.

It provided that the coefficients exit if the inverse exists.

The OLS estimator of β is given by

$$\hat{\beta} = (x'x)^{-1}x'y$$

It provided that the coefficients exit if the inverse exists. This form is obtained from the assumption of the SURE model is $\rho(X) = K^*$ where K^* is constant and this estimator treats the variance-covariance matrix as scalar; the equation of this procedure is OLS normal equations, i.e., $\widehat{\beta}$ is the OLS estimator of β. The OLS estimator is unbiased for β and, the variance- covariance matrix of $\widehat{\beta}$ is Greene (2003).

$$V\left(\hat{\beta}\right) = \sigma^2(X^\circ X)^{-1}$$

2.3. The GLS estimation of the SURE model

Consider the SURE model as

$$y_{TM\times1=X_{TM\times K^*}\beta_{K^*\times1}+\epsilon_{TM\times1}}$$

With the assumptions

$$E(\epsilon) = 0,$$

and $E(\varepsilon\varepsilon') = \left(\sum{}^{\otimes}I_T\right) = \Psi$.

The generalized least squares estimation may take the jointness of the relationships that make up the SURE model.

The GLS estimator of β will be

$$\hat{\beta}_{SURE} = \left(X'\Psi^{-1}X\right)^{-1}X'\Psi^{-1}Y = \left(X'\left(\sum \otimes I_T\right)X\right)^{-1}X'\left(\sum \otimes I_T\right)Y \tag{9}$$

and the variance-covariance matrix of $\widehat{\beta}_{SURE}$ is given by Greene (2003).

$$V\left(\hat{\beta}_{SURE}\right) = \left(X'\Psi^{-1}X\right)^{-1} = \left(X'\left(\sum \otimes I_T\right)X\right)^{-1}$$

2.4. GLS estimators of the SURE model reduces to OLS estimator

There are two special cases where the GLS estimator of the SURE model reduces to the OLS estimator.

We know that the theoretical GLS estimator of β is given by

$$\beta^* = E\left[\left(X^{*\prime} \sum{}^{*-1} X^* \right)^{-1} X^{*\prime} \sum{}^{*-1} Y^* \right] \tag{10}$$

And

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & . & . & \sigma_{1M} \\ \sigma_{21} & \sigma_{22} & . & . & \sigma_{2M} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \sigma_{M1} & \sigma_{M2} & . & . & \sigma_{MM} \end{bmatrix}_{M \times M} \tag{11}$$

The first case is when the unstructured covariance matrix (11) is diagonal so that there is no contemporaneous correlation i.e. the equations are actually unrelated. As a consequence, Eq. (10) will also be diagonal.

The second case is when the regressor matrices X_i are the same for all equations i = 1,2,...,m i.e. equations have identical explanatory variables. In that case $X = \begin{bmatrix} X & 0 & . & . & 0 \\ 0 & X & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & X \end{bmatrix} = I_M \otimes X.$

where $\otimes$ is the Kronecker product, Greene (2003).

3. Comparison between OLS and GLS estimator

The OLS and GLS estimators are unbiased, and GLS is at least as efficient as OLS when estimating β in the SURE model. $\widehat{\beta}_{SURE}$ is the BLUE of β in the SURE model, since $\sum$ is non-stochastic and observable; this follows Aitken's theorem, Aitken (1935). Looking at the expression for $\widehat{\beta}_{SURE}$ in (9), it is not an operational or feasible estimator of β because in general Σ, and hence Ψ, will be unobservable. Recognizing this, Zellner (1963) proposed an estimate of β in the SURE model, basing this on (9), but with Σ replaced by an observable $M \times M$ matrix S. In particular, the elements of S are chosen to be estimators of the corresponding elements of Σ. With this replacement for Σ, and hence Ψ, we now have a feasible generalized least squares FGLS estimator of β in (2):

$$b_F = \left(X' \left(S^{-1} \otimes I_T \right) X \right)^{-1} X' \left(S^{-1} \otimes I_T \right) Y. \tag{12}$$

We are assuming that the matrix S = $[s_{ij}]$ is non-singular, where s_{ij} is some estimator of σ_{ij}. Although there are many possible choices of S, two ways of obtaining the s_{ij} are popular. Each of these is based on residuals obtained by the application of OLS in one way or another. Oberhofer and Kmenta (1974) suggested a general procedure for obtaining maximum likelihood estimates by iterating the FGLS. Direct maximum likelihood estimation can be used by inserting the special form of Σ in the log-likelihood function for the generalized regression model instead of iterated FGLS, William H. Greene, (2003). Risto and Neudecker (1997) presented the essentials of parameters estimate for the coefficients β of the SURE model using least squares (LS), generalized least squares (GLS), and maximum likelihood (ML) (under normality). Also, parameters estimate for the variance-covariance matrix Ψ using an LS-related estimator and a maximum likelihood estimator (under normality) was presented with their asymptotic properties.

4. Restricted maximum likelihood estimation of the SURE model

The MIXED procedure of the SAS System has advantages over the standard multivariate procedures in fitting multiple design multivariate models (the seemingly unrelated regression models of econometrics) such as it uses observations that have incomplete responses in the calculation of the fitted models where most multivariate procedures discard an entire observation if it has any missing data. Also, it allows one to select one suitable covariance structure of the available collection covariance structures to fit the best models to data instead of restricting our analysis to unstructured covariance matrix, $\sum$ in (11), when fitting the SURE model, Wright (1998), Littell et al. (1999), and Khattree and Naik (2000). The MIXED procedure of the SAS System can be used to develop either maximum likelihood (ML) or restricted maximum likelihood (REML) estimates in order to complete the analysis of the data, where REML estimation are generally preferred to ML, Kenward and Rogers (1997). Selecting the suitable covariance structure to fit the best models to data is the concerned issue. In this paper, we suggest using AL-Marshadi (2014) method to select the suitable covariance structure to fit the best seemingly unrelated regression models to the data using the MIXED procedure with restricted maximum likelihood (REML) estimates method. Fitting the seemingly unrelated regression models with unsuitable covariance structure may impact the quality of the fitted model using the MIXED procedure, AL-Marshadi (2008). In order to evaluate the proposed suggestion, a simulation study was conducted for variety of settings of seemingly unrelated regression models such as equations that have identical explanatory variables, the regressors in one block of equations are a subset of those in another, and different regressors in the equations with various settings of covariance structures of $\sum$. The first covariance structure was considered compound symmetry (CS) covariance structures. The second covariance structure was considered heterogeneous compound symmetry (CSH) covariance structure. The third covariance structure was considered, first-order autoregressive (AR (1)) covariance structure. The fourth covariance structure was considered heterogeneous first-order autoregressive (ARH (1)) covariance structure. The fifth covariance structure was considered, Toeplitz (TOEP) covariance structure. The sixth covariance structure was considered, Unstructured (UN) covariance structure, Littell et al. (1999). Wright (1998) explained the way of fitting the seemingly unrelated regression equations SURE models format using a MIXED procedure with MIXED Model format with restricted maximum likelihood (REML) estimates method which was used in this study.

5. The simulation study

In this section, a simulation study is carried out to assess AL-Marshadi (2014) methodology in terms of the percentage of times that it identifies the right covariance structure for mixed model analysis of SURE models' data. Moreover, it is also reported that when the PROC MIXED procedure used REML without any intervention, the percentage of times that REML failed to converge under normal situation, Robert and Casella (2004).

From four SURE models given below, correlated multivariate normal data were generated by specifically developing SAS PROC IML code using the MIXED Model format, Wright (1998). Using two different sample sizes $T = 60$ and 100 with six covariance structures, resulting in 12 scenarios and for each scenario 4000 data sets were simulated. Thereafter, on each data set AL-Marshadi (2014) algorithm was employed and the percentage of times the right covariance structure is identified is reported in Table 1.

The first simulated SURE model:

Table 1
Six covariance matrix structure settings used in simulations.

Setting #.	Covariance Matrix
1 Compound Symmetry (CS)	$\begin{bmatrix} 16 & 12.8 & 12.8 & 12.8 & 12.8 & 12.8 & 12.8 \\ 12.8 & 16 & 12.8 & 12.8 & 12.8 & 12.8 & 12.8 \\ 12.8 & 12.8 & 16 & 12.8 & 12.8 & 12.8 & 12.8 \\ 12.8 & 12.8 & 12.8 & 16 & 12.8 & 12.8 & 12.8 \\ 12.8 & 12.8 & 12.8 & 12.8 & 16 & 12.8 & 12.8 \\ 12.8 & 12.8 & 12.8 & 12.8 & 12.8 & 16 & 12.8 \\ 12.8 & 12.8 & 12.8 & 12.8 & 12.8 & 12.8 & 16 \end{bmatrix}$
2 First-Order Autoregressive (AR(1))	$\begin{bmatrix} 16 & 14.4 & 12.96 & 11.664 & 10.4976 & 9.44784 & 8.503056 \\ 14.4 & 16 & 14.4 & 12.96 & 11.664 & 10.4976 & 9.44784 \\ 12.96 & 14.4 & 16 & 14.4 & 12.96 & 11.664 & 10.4976 \\ 11.664 & 12.96 & 14.4 & 16 & 14.4 & 12.96 & 11.664 \\ 10.4976 & 11.664 & 12.96 & 14.4 & 16 & 14.4 & 12.96 \\ 9.44784 & 10.4976 & 11.664 & 12.96 & 14.4 & 16 & 14.4 \\ 8.503056 & 9.44784 & 10.4976 & 11.664 & 12.96 & 14.4 & 16 \end{bmatrix}$
3 Toeplitz (TOEP)	$\begin{bmatrix} 16 & 1.6 & 8 & 6.4 & 4.8 & 3.2 & 11.2 \\ 1.6 & 16 & 1.6 & 8 & 6.4 & 4.8 & 3.2 \\ 8 & 1.6 & 16 & 1.6 & 8 & 6.4 & 4.8 \\ 6.4 & 8 & 1.6 & 16 & 1.6 & 8 & 6.4 \\ 4.8 & 6.4 & 8 & 1.6 & 16 & 1.6 & 8 \\ 3.2 & 4.8 & 6.4 & 8 & 1.6 & 16 & 1.6 \\ 11.2 & 3.2 & 4.8 & 6.4 & 8 & 1.6 & 16 \end{bmatrix}$
4 Heterogeneous Compound Symmetry (CSH)	$\begin{bmatrix} 4 & 4.8 & 6.4 & 8 & 9.6 & 11.2 & 12.8 \\ 4.8 & 9 & 9.6 & 12 & 14.4 & 16.8 & 19.2 \\ 6.4 & 9.6 & 16 & 16 & 19.2 & 22.4 & 25.6 \\ 8 & 12 & 16 & 25 & 24 & 28 & 32 \\ 9.6 & 14.4 & 19.2 & 24 & 36 & 33.6 & 38.4 \\ 11.2 & 16.8 & 22.4 & 28 & 33.6 & 49 & 44.8 \\ 12.8 & 19.2 & 25.6 & 32 & 38.4 & 44.8 & 64 \end{bmatrix}$
5 Heterogeneous First-Order Autoregressive (ARH(1))	$\begin{bmatrix} 4 & 4.8 & 5.12 & 5.12 & 4.9152 & 4.58752 & 4.194304 \\ 4.8 & 9 & 9.6 & 9.6 & 9.216 & 8.60160 & 7.86432 \\ 5.12 & 9.6 & 16 & 16 & 15.36 & 14.336 & 13.1072 \\ 5.12 & 9.6 & 16 & 25 & 24 & 22.4 & 20.48 \\ 4.9152 & 9.216 & 15.36 & 24 & 36 & 33.6 & 30.72 \\ 4.58752 & 8.6016 & 14.336 & 22.4 & 33.6 & 49 & 44.8 \\ 4.194304 & 7.86432 & 13.1072 & 20.48 & 30.72 & 44.8 & 64 \end{bmatrix}$
6 Unstructured (UN)	$\begin{bmatrix} 4 & 2.4 & 4.8 & 8 & 8.4 & 7 & 4.96 \\ 2.4 & 9 & 2.4 & 1.5 & 2.7 & 7.35 & 10.8 \\ 4.8 & 2.4 & 16 & 3.4 & 10.08 & 15.4 & 6.48 \\ 8 & 1.5 & 3.4 & 25 & 18.9 & 16.45 & 9.2 \\ 8.4 & 2.7 & 10.08 & 18.9 & 36 & 4.62 & 22.56 \\ 7 & 7.35 & 15.4 & 16.45 & 4.62 & 49 & 16.24 \\ 4.96 & 10.8 & 6.48 & 9.2 & 22.56 & 16.24 & 64 \end{bmatrix}$

$$y_{t1} = 3 + 3x_{t0} + 3x_{t1} + 2x_{t2} + \epsilon_{t1}$$

$$y_{t2} = 8 + 6x_{t0} + 6x_{t1} + 4x_{t2} + \epsilon_{t2}$$

$$y_{t3} = 5 + 9x_{t0} + 9x_{t1} + 6x_{t2} + \epsilon_{t3}$$

$$y_{t4} = 2 + 2x_{t0} + 2x_{t1} + 3x_{t2} + \epsilon_{t4}$$

$$y_{t5} = -9 - 9x_{t0} - 9x_{t1} - 6x_{2t} + \epsilon_{t5}$$

$$y_{t6} = -3 - 6x_{t0} - 6x_{t1} - 4x_{t2} + \epsilon_{t6}$$

$$y_{t7} = -9 - 3x_{t0} - 3x_{t1} - 2x_{t2} + \epsilon_{t7}$$

where x_0 is the dummy variable which takes the values 0 or 1, x_1 and x_2 explanatory variables following the Normal distribution with $\mu = 5$ and $\sigma = 2$. For simulation study, correlated disturbance terms (ϵ_{ti}) $(t = 1, \cdots, T; i = 1, \cdots, M = 7)$ of the given equations were simulated using a different setting of the covariance matrix Σ.

The second simulated SURE model:

$$y_{t1} = 3 + 3x_{t1} + 2x_{t2} + \epsilon_{t1}$$

$$y_{t2} = 8 + 6x_{t1} + 4x_{t2} + \epsilon_{t2}$$

$$y_{t3} = 5 + 9x_{t1} + 6x_{t2} + \epsilon_{t3}$$

$$y_{t4} = 2 + 2x_{t1} + 3x_{t2} + \epsilon_{t4}$$

$$y_{t5} = -9 - 9x_{t1} - 6x_{2t} + \epsilon_{t5}$$

$$y_{t6} = -3 - 6x_{t1} - 4x_{t2} + \epsilon_{t6}$$

$$y_{t7} = -9 - 3x_{t1} - 2x_{t2} + \epsilon_{t7}$$

where x_1 and x_2 are independent variables follow Normal distribution with $\mu = 5$ and $\sigma = 2$. For simulation study correlated disturbance terms (ϵ_{ti}) $(t = 1, \cdots, T; i = 1, \cdots, M = 7)$ of the given equations were simulated using the different settings of the covariance matrix Σ.

The third simulated SURE model:

$$y_{t1} = 3 + \epsilon_{t1}$$

$$y_{t2} = 9 + 6x_{t0} + 3x_{t1} + \epsilon_{t2}$$

$$y_{t3} = 7 + 5x_{t1} + \epsilon_{t3}$$

$$y_{t4} = 12 + 2x_{t2} + \epsilon_{t4}$$

$$y_{t5} = 5 + 22x_{t1} + 2x_{2t} + \epsilon_{t5}$$

$$y_{t6} = 1 + 18x_{t1} + 52x_{t2} + \epsilon_{t6}$$

$$y_{t7} = 5 + 12x_{t0} + 2x_{t1} + 2x_{t2} + \epsilon_{t7}$$

where x_0 is a dummy variable taking values 0 or 1, x_1 and x_2 explanatory variables following normal distribution with $\mu = 5$ and $\sigma = 2$. For simulation study, correlated disturbance terms (ϵ_{ti}) $(t = 1, \cdots, T; i = 1, \cdots, M = 7)$ of the given equations were simulated using a different setting of the covariance matrix Σ.

The fourth simulated SURE model:

$$y_{t1} = 3 + \epsilon_{t1}$$

$$y_{t2} = 8 + 3x_{t1} + \epsilon_{t2}$$

$$y_{t3} = 7 + 5x_{t1} + \epsilon_{t3}$$

$$y_{t4} = 12 + 2x_{t2} + \epsilon_{t4}$$

$$y_{t5} = 5 + 22x_{t1} + 2x_{2t} + \epsilon_{t5}$$

$$y_{t6} = 1 + 18x_{t1} + 52x_{t2} + \epsilon_{t6}$$

$$y_{t7} = 12 + 2x_{t1} + 2x_{t2} + \epsilon_{t7}$$

where x_1 and x_2 are explanatory variables following the normal distribution with μ = 5 and σ = 2. For simulation study correlated disturbance terms (ϵ_{ti}) $(t = 1, \cdots, T; i = 1, \cdots, M = 7)$ of the given equations were simulated using a different setting of the covariance matrix Σ.

6. Results

Table 2 exhibits summarized results of AL-Marshadi (2014) approach selecting the right covariance structure percentage of times from the six Covariance structures, when *W = 10*, and *T = 60*, where '*W*' represents the number of the bootstrap samples. Table 3 exhibits summarized results of AL-Marshadi (2014) approach selecting the right covariance structure percentage of times from the six Covariance structures, *W = 10*, and *T = 100*, where '*W*' represents the number of the bootstrap samples and this is in line with the suggestions provided in AL-Marshadi (2014). Results of Table 2 and Table 3 show similar reliable performance with a high percent of success in selecting the suitable covariance structure with AL-Marshadi (2014) approach across all the simulated SURE models with the different covariance structures which will be allowed the SURE models fitted more efficient than the existing method of considering the stander unstructured covari-

ance structure in fitting SURE models' since the number of the parameters in the suitable selected covariance structure will be reduced when the suitable covariance structure is selected other than the stander unstructured covariance structure. The comparison of the efficiency of fitting the SURE models using the stander unstructured covariance structure or using the suitable selected covariance structure with AL-Marshadi (2014) approach will be clear in the application using a panel of data. Results of Tables 2 and 3 also suggest that the performance of the approach has enhanced with the increase in the sample size with constant bootstrap samples.

Tables 4 and 5 exhibit the percentage of times that the PROC MIXED procedure failed to converge when the PROC MIXED procedure used REML without any interfering for all the investigated settings of the covariance matrix and *W = 10, and T = 60 and W = 10, and T = 100*. To put it in a nutshell, the results in Tables 4 and 5 suggest that increasing the sample size addresses the convergence problem (Table 6).

7. Application using grunfeld's data

To illustrate the AL-Marshadi (2014) approach to select the right covariance structure among the seven covariance structure considered in the study using a panel of data that serve as a useful tool for investigating multiple equation estimators for a long period of time in the literature. The data consist of time series of twenty yearly observations for five firms and three variables:

I_t = gross investment.
F_t = market value of the firm at the end of the previous year.
C_t = value of the stock of plant and equipment at the end of the previous year, William H. Greene, (2003), and Grunfeld (1958). The estimated model with these data is.

$$I_{it} = \beta_{1i} + \beta_{2i}F_{it} + \beta_{3i}C_{it} + \epsilon_{it}$$

where *i* indexes firms and *t* indexes years.

Therefore, the AL-Marshadi (2014) approach was applied to select the best covariance structure among the seven covariance structure considered in the study for the data as follow:

Table 2
True % of times covariance structures from possible covariance structures when *W = 10 and T = 60* selected through with *ACSMSCCS* approach.

The correct model	Best set of Covariance Structure Clusters	The percent of success			
		Model 1	Model 2	Model 3	Model 4
		%	%	%	%
CS	CS, CSH,TOEP,TOEPH,UN	98.625	98.925	98.7	99.025
CSH	CSH, ARH(1),TOEPH,UN	97.1	97.45	97.875	97.825
AR(1)	AR(1), ARH(1), TOEP, TOEPH, UN	100	100	100	100
ARH(1)	ARH(1),TOEPH,UN	85.45	86.35	87.75	87.375
TOEP	TOEP, TOEPH, UN	84.925	85.6	84.975	85.325
UN	UN	92.95	93.125	84.875	85.425
Over all the percent of success		93.175	93.575	92.3625	92.4958

Table 3
True % of times covariance structures from possible covariance structures when *W = 10 and T = 100* selected through with *ACSMSCCS* approach.

The correct model	Best set of Covariance Structure Clusters	The percent of success			
		Model 1 %	Model 2 %	Model 3 %	Model 4 %
CS	CS, CSH, TOEP, TOEPH,UN	100	100	100	100
CSH	CSH, ARH(1), TOEPH, UN	99.425	99.225	99.55	99.475
AR(1)	AR(1), ARH(1), TOEP, TOEPH, UN	100	100	100	100
ARH(1)	ARH(1), TOEPH, UN	94.6	95.025	95.475	95.5
TOEP	TOEP, TOEPH,UN	93.175	93.375	92.8	93
UN	UN	97.875	97.65	92.325	93.1
Over all the percent of success		97.5125	97.5458	96.6917	96.8458

1. We generated the bootstrap samples on a case-by-case using the original data (i.e., based on resampling from twenty-year observations). The bootstrap sample size is taken to be the same as the size of the original data (i.e., 20 years).

2. The model was fitted with the candidate covariance structures that we like to select the best covariance structure from them, to each of the bootstrap samples, thereby obtaining the bootstrap

Table 4
Percentage of times the PROC MIXED procedure failed to converge when the PROC MIXED procedure used REML without any interference for all investigated Covariance Matrix settings and W = 10 and T = 60.

The fitted structure	The right covariance structure						
	AR(1) %	ARH(1) %	CS %	CSH %	TOEP %	TOEPH %	UN %
CS	0	0	0	0	0.0000125	0.00000625	0
AR(1)	0	0	0	0	0	0.00000625	0
TOEP	0	0	0	0	0	0	0
CSH	0	0	0	0	0	0.0000375	0
ARH(1)	0	0	0	0	0	0.00001875	0
UN	0	0	0	0	0	0	0.04378

Table 5
Percentage of times the PROC MIXED procedure failed to converge when the PROC MIXED procedure used REML without any interference for all investigated Covariance Matrix settings and W = 10 and T = 100.

The fitted structure	The right covariance structure						
	AR(1) %	ARH(1) %	CS %	CSH %	TOEP %	TOEPH %	UN %
CS	0	0	0	0	0	0	0
AR(1)	0	0	0	0	0	0	0
TOEP	0	0	0	0	0.00000625	0	0
CSH	0	0	0	0	0	0.0000625	0
ARH(1)	0	0	0	0	0	0.000025	0
UN	0	0	0	0	0	0	0.02739

Table 6
The average of each information criteria for each model and two clusters according to the five correlated variables.

Structure	AIC	AICC	HQIC	BIC	CAIC	Cluster
CSH	600.854	603.031	599.847	603.857	609.857	1
ARH (1)	600.408	602.585	599.401	603.411	609.411	1
TOEPH	601.753	606.817	600.242	606.257	615.257	1
UN	596.818	613.123	594.299	604.325	619.325	1
CS	656.678	656.963	656.326	657.659	659.659	2
AR (1)	661.855	662.137	661.519	662.856	664.856	2
TOEP	665.208	666.723	664.368	667.710	672.710	2

Table 7
The parameter estimates of the model and the standard error of the parameter estimates using the two covariance structures.

Effect	Firm	Estimate	Standard Error	
			Using CSH	Using UN
Firm (β_{11})	1	−149.78	96.3157	105.84
Firm (β_{12})	2	−6.1900	13.5062	13.5065
Firm (β_{13})	3	−9.9563	31.3228	31.3742
Firm (β_{14})	4	−0.5094	7.9684	8.0153
Firm (β_{15})	5	−30.3685	129.34	157.05
F*Firm (β_{21})	1	0.1193	0.02351	0.02583
F*Firm (β_{22})	2	0.07795	0.01997	0.01997
F*Firm (β_{23})	3	0.02655	0.01554	0.01557
F*Firm (β_{24})	4	0.05289	0.01561	0.01571
F*Firm (β_{25})	5	0.1566	0.06497	0.07889
C*Firm (β_{31})	1	0.3714	0.03374	0.03707
C*Firm (β_{32})	2	0.3157	0.02881	0.02881
C*Firm (β_{33})	3	0.1517	0.02566	0.02570
C*Firm (β_{34})	4	0.09241	0.05577	0.05610
C*Firm (β_{35})	5	0.4239	0.1278	0.1552

$AIC^*BIC^*CAIC^*HQIC^*AICC^*$ for the model with the candidate covariance structures.

3. Repeat steps (1) and (2) (10) times
4. Bootstrapping of the original data is given us the opportunity to have (10) replication values for each model and each information criteria (from steps (1 to 3). we used the average of each information criteria for each model sepyarately in the algorithm as a random vector that follows 5-dimensional multivariate normal distribution: $\left[\overline{AIC}\ \overline{BIC}\ \overline{CAIC}\ \overline{HQIC}\ \overline{AICC}\right]_{Coveriancstructure-i}$.

In this stage, clustering method was used to cluster the candidate covariance structures s two clusters according to the five correlated variables (the averages of the five information criteria). One of the two clusters is called the cluster of the best set of covariance structures according to the cluster that contains the general covariance structure UN (Unstructured covariance structure). Then the best covariance structure is the simplest covariance structure in the cluster of the best set of covariance structures. The results for the data are given in Table 7 that suggest CSH is the best covariance structure for the data. The efficiency of using the best suitable covariance structure for the data instead of using the stander unstructured covariance structure for the data is shown in Table 7 which compares the stander error for estimated parameters for the two covariance structures. The comparison of the stander error for estimated parameters in Table 7 shows the stander error for estimated parameters with suitable selected covariance structure with AL-Marshadi (2014) approach are less than the stander error for estimated parameters with the stander unstructured covariance structure.

8. Conclusions

In our simulation, we considered SURE models, examining the performance of the AL-Marshadi (2014) approach to select the right covariance structure with different covariance structure settings. Overall, the AL-Marshadi (2014) approach provided an excellent tool for selecting the right covariance structure for SURE models in order to fit the SURE models more efficiently than the existing method that considering the stander unstructured covariance structure in fitting SURE models that can be seen with the application of the panel of data when the standard errors compared in Table 7. Future work can be considered to examine the performance of the AL-Marshadi (2014) approach under the existent of multicollinearity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aitken, A.C., 1935. On Least Squares and Linear Combinations of Observations. Proc. R. Stat. Soc. 55, 42–48.

AL-Marshadi, Ali Hussein, 2008. The Impact of Restricted Our Analysis of Repeated Measures Design to the Two Stander Covariance Structures with and Without Missing Data. Aust. J. Basic Appl. Sci. 2 (4), 1228–1238.

AL-Marshadi, Ali Hussein, 2014. Selecting the covariance structure in mixed model using statistical methods calibration. J. Math. Stat. 10 (3), 309–315.

Banterle, M., Bottolo, L., Richardson, S., Ala-Korpela, M., Järvelin, M.R., Lewin, A., 2018. Sparse variable and covariance selection for high-dimensional seemingly unrelated Bayesian regression. bioRxiv 467019.

Bottolo, L., Banterle, M., Richardson, S., Ala-Korpela, M., Järvelin, M.-R., Lewin, A., 2021. A computationally efficient Bayesian seemingly unrelated regressions model for high-dimensional quantitative trait loci discovery. J. Roy. Stat. Soc.: Ser. C (Appl. Stat.) 70 (4), 886–908.

Feng, G., Polson, N., 2020. Regularizing Bayesian predictive regressions. J. Asset Manage. 21 (7), 591–608.

Foschi, P., Kontoghiorghes, E.J., 2003. Estimating seemingly unrelated regression models with vector autoregressive disturbances. J. Econ. Dyn. Control 28 (1), 27–44.

Greene, W.H., 2003. Econometric Analysis. Prentice-Hall Inc, USA, New Jersey.

Grunfeld, Y., 1958. The Determinants of Corporate Investment. Department of Economics University of Chicago. Unpublished Ph.D. thesis.

Kenward, M.G., Rogers, J.H., 1997. Small sample inference for fixed effect from restricted maximum likelihood. Biometrics 53, 983–997.

Khattree, R., Naik, N.D., 2000. Multivariate Data Reduction and Discrimination with SAS Software. SAS Institute Inc., Cary NC, USA.

Littell, R.C., Milliken, G.A., Wolfinger, R.D., 1999. SAS System for Mixed Models. SAS Institute Inc., Cary, NC, USA.

Oberhofer, W., Kmenta, J., 1974. A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models. Econometrica 42, 579–590.

Risto, H., Heinz, N., 1998. Estimation of the SURE model. Stat. Pap. 39 (4), 423–430.

Robert, C., Casella, G., 2004. Monte Carlo Statistical Methods. Springer-Verlag, New York, NY, USA.

Wright, S. Paul, 1998. Multivariate analysis using the mixed procedure. In: Proceedings of the SUGI-23 Conference in statistics, data analysis, and modeling (pp. 1–5). Nashville, Tennessee.

Zellner, A., 1962. An efficient method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias. J. Am. Stat. Assoc. 57, 500–509.

Zellner, A., 1963. Estimators for Seemingly unrelated regression Equations: Some Exact Finite Sample Results. J. Am. Stat. Assoc. 58 (304), 977–992.