



# Model Selection of Simultaneous Equations Model Based on Kullback Information Criterion for Small Sample Case

Warangkhan Riansut\*

*Mathematics and Data Management Program, Faculty of Science and Digital Innovation,  
Thaksin University, Phatthalung 93210, Thailand*

Received 2 July 2025; Received in revised form 30 October 2025  
Accepted 28 February 2026; Available online 30 June 2026

## ABSTRACT

This paper introduces a new criterion called SKICc, designed specifically for selecting simultaneous equations models (SEM) in small sample cases. The efficiency of SKICc was compared to that of SAIC, SBIC, and SKIC, considering the percentage of correct model selection and the  $L_2$  efficiency. The study's scope was defined under the assumption that the independent variables were normally distributed, had a mean of zero, a variance of one, and no issues of multicollinearity. Additionally, the errors in the model had a mean of zero but exhibited contemporaneous correlations among the equations. The small sample sizes examined in the study included 15, 20, and 25 units, with 1,000 simulated data sets generated for each size. The results of the study indicated that the SKICc created had the following formula:

$$SKICc = \log |\tilde{\Sigma}| + M \log \left( \frac{2T}{2T - 2K - M + 1} \right) + \frac{2M}{2T - 2K - M + 1} + \frac{M(2K + M + 1)}{T - M - 1} - \frac{KM(M + 1)}{T(T - M - 1)}.$$

The efficiency comparison revealed that SKICc demonstrated the highest efficiency for small sample sizes. It achieved the best performance in terms of correct model selection and had the highest mean and lowest standard deviation of the  $L_2$  efficiency. However, as the sample size increased, SKIC exhibited the best efficiency. Furthermore, the research findings indicated that SAIC and SBIC tended to select models with too many variables due to their low penalty function values. In contrast, SKIC often selects models with too few variables because of its high penalty function value. Overall, SKICc effectively addressed the issue of selecting models with either too many or too few variables, outperforming the other criteria studied.

**Keywords:**  $L_2$  efficiency; Model selection criterion; Percentage of correct model selection; Simultaneous Equations Model (SEM); SKICc

## 1. Introduction

Selecting the appropriate simultaneous equations model (SEM) from among all possible models is a crucial problem. In practice, we can't know what the true model is, but we expect to find a model that accurately and appropriately represents it. Simultaneous equations are widely used in statistical and economic models. The primary characteristic is that the dependent variable of one equation may serve as the independent variable in the remaining equation. Therefore, the dependent variable will become random and related to the error of the equation that appears as an independent variable [1]. An example of SEM is the analysis of costs and returns from rubber plantations; a case study of Pa Phayom district, Phatthalung province and Thung Song district, Nakhon Si Thammarat province, found that the cost variable, which is the dependent variable in the first equation, will be the independent variable in the second equation related to the returns variable [2]. Creating the SEM to fit an observed data set involves two main steps: determining the number of model variables and estimating parameters. The key part of the equation creation problem is to determine the number of model variables [3]. Such decisions are often made by using statistical methods that help identify models with an appropriate number of variables, neither too many nor too few. This selection process is often evaluated using the concept of 'Goodness of Fit and Parsimony', which refers to the balance between the model's complexity and its ability to represent the data accurately [4]. This paper will focus solely on the model selection criteria. The first known criterion for the regression model is AIC [5]. This criterion is an asymptotically unbiased estimator of the variance of Kullback's directed divergence, also known as Kullback-Leibler information, used to assess the difference between the true model and the fitted model. Other model selection criteria that have been introduced include BIC [6], KIC [7], and KICc [8-10].

Despite AIC's continued dominance in model selection, BIC, KIC, and KICc are

emerging as popular alternatives. BIC typically selects models with fewer variables compared to AIC [11]. KIC is a symmetric criterion that combines information from the two types of direct divergence, known as Kullback's symmetric divergence. Therefore, KIC is a more sensitive measure of model divergence than AIC, which only measures individual components [7-8]. However, when dealing with small sample sizes or a high ratio of parameters to sample size, AIC and KIC often underestimate Kullback's direct and symmetric divergence, respectively. This issue can result in model overfitting. To address this problem, several researchers have proposed an adjusted version known as AICc [12] and KICc [8-10].

After conducting an extensive literature review on model selection criteria for SEM, it was found that there are currently a few relevant criteria available: SAIC [13], SBIC [14], and SKIC [4]. However, there is a notable lack of established model selection criteria specifically for SEM in small sample-size scenarios. For this reason, this paper proposes a new criterion called SKICc (System of Simultaneous Equations KICc), designed to select the SEM when working with small sample sizes.

SKICc is derived based on Kullback's symmetric divergence, as many researchers have concluded that model selection criteria based on directed divergence, such as AIC [5] and AICc [12], tend to produce unreasonably low estimates, thus tending to select models with too many variables [7-10]. To prevent the problems of overfitting or underfitting variables, the use of symmetric divergence is suggested.

Additionally, the effectiveness of SKICc will be evaluated in comparison to SAIC [13], SBIC [14], and SKIC [4], focusing on both the percentage of correct model selections and  $L_2$  efficiency. The criterion that proves to be the best will be the one with the highest percentage of correct model selections, the lowest occurrence of models with either too many or too few variables, and the highest  $L_2$  efficiency. This performance encompasses both the highest

mean and the lowest standard deviation of the  $L_2$  efficiency.

## 2. Materials and Methods

### 2.1 Selection criteria for SEM based on KIC for small samples

The SEM's structural form is

$$\mathbf{Y}\mathbf{\Gamma} + \mathbf{X}\mathbf{B} = \mathbf{U}, \quad (2.1)$$

where  $\mathbf{Y}$  is a matrix of dependent variables of size  $T \times M$ ,  $\mathbf{\Gamma}$  is a matrix of coefficients of dependent variables of size  $M \times M$  and assumed to be nonsingular.  $\mathbf{X}$  is a matrix of independent variables of size  $T \times K$  with full-column rank,  $\mathbf{B}$  is a matrix of coefficients of independent variables of size  $K \times M$ , and  $\mathbf{U}$  is a matrix of contemporaneously correlated errors of size  $T \times M$ . The reduced form of Eq. (2.1) can be written as

$$\mathbf{Y} = \mathbf{X}\mathbf{\Pi} + \mathbf{E}, \quad (2.2)$$

where  $\mathbf{\Pi} = -\mathbf{B}\mathbf{\Gamma}^{-1}$  is a matrix of unknown parameters of size  $K \times M$  and  $\mathbf{E} = \mathbf{U}\mathbf{\Gamma}^{-1}$  is a matrix of contemporaneously correlated errors of size  $T \times M$ . All  $M$  equations of Eq. (2.2) can be represented as a nested form:

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\pi} + \boldsymbol{\varepsilon}, \quad (2.3)$$

where  $\mathbf{y}$  is a vector of the dependent variables of size  $TM \times 1$  consisting of  $M$  vectors  $\mathbf{y}_j$  of size  $T \times 1$ ,  $\mathbf{X}^*$  is a matrix of the independent variable of size  $TM \times KM$  consisting of  $M$  diagonal matrices  $\mathbf{X}$  of size  $T \times K$ ,  $\boldsymbol{\pi}$  is a vector of the parameter of size  $KM \times 1$  consisting of  $M$  vectors  $\boldsymbol{\pi}_j$  of size  $K \times 1$ , and  $\boldsymbol{\varepsilon}$  is a vector of the error term of size  $TM \times 1$  consisting of  $M$  vectors  $\boldsymbol{\varepsilon}_j$  of size  $T \times 1$ . Given that the errors  $\boldsymbol{\varepsilon}_{ij}$  is normally distributed, have a mean of zero, and a variance of  $\sigma_{jj}$ , i.e.,  $\boldsymbol{\varepsilon}_{ij} \sim N(0, \sigma_{jj})$ , then we have

$$\boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \\ \vdots \\ \varepsilon_{tM} \end{bmatrix} \sim N_M(\mathbf{0}, \boldsymbol{\Sigma}), \quad (2.4)$$

where  $\boldsymbol{\Sigma}$  is a matrix of contemporaneous covariance of the error terms of size  $M \times M$ . The selection criterion for the model starts with the consideration of the multivariate normal distribution [15] to obtain the minus twice of the logarithmic function:

$$\begin{aligned} -2 \log L(\boldsymbol{\theta} | \mathbf{y}) &= TM \log(2\pi) + T \log |\boldsymbol{\Sigma}| \\ &+ (\mathbf{y} - \mathbf{X}^* \boldsymbol{\pi})' (\boldsymbol{\Sigma}^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \boldsymbol{\pi}), \end{aligned}$$

where  $\boldsymbol{\theta} = (\boldsymbol{\pi}, \boldsymbol{\Sigma})$  represents the unknown parameter and  $\otimes$  represents the Kronecker product.

We refer to the model in Eq. (2.3) as the fitted model, while Eq. (2.5) is referred to as the true model. We specify that the fitted model covers the true model. Any parameter in the true model that has a non-zero value is denoted as  $K_0M$ , where  $0 < K_0M < KM$ . The remaining parameters,  $(K - K_0)M$ , are set to zero.

$$\mathbf{y} = \mathbf{X}^* \boldsymbol{\pi}_0 + \boldsymbol{\varepsilon}_0. \quad (2.5)$$

Kullback's symmetric divergence can be written as

$$2J(\boldsymbol{\theta}_0, \boldsymbol{\theta}) = d(\boldsymbol{\theta}_0, \boldsymbol{\theta}) - d(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0) + d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) - d(\boldsymbol{\theta}, \boldsymbol{\theta}), \quad (2.6)$$

where  $d(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j) = E_{\boldsymbol{\theta}_i} \{-2 \log L(\boldsymbol{\theta}_j | \mathbf{y})\}$ .

Since SKIC [4] is based on KIC and KICc, which utilize an unbiased estimator (UE) for estimating the contemporaneous covariance matrix of the errors, many researchers have pointed out a significant issue: the application of UE to generate KICc often leads to overestimation of model selection criteria [16-18]. This problem can result in the selection of models that include too few variables or are underfitted. Therefore,

this paper emphasizes the crucial role of accurate model selection in our field and argues that using the maximum likelihood estimator (MLE) instead of UE can help us achieve this goal. MLE, with its lower value than UE, is calculated from the sum of squared errors divided by the total number of observations (T), whereas UE is calculated from the sum of squared errors divided by the difference between the total number of observations and the number of parameters to be estimated (T - K). UE of  $\Sigma$  is

$$\hat{\Sigma} = \frac{1}{T-K} (\mathbf{Y} - \mathbf{X}\hat{\Pi})' (\mathbf{Y} - \mathbf{X}\hat{\Pi}).$$

MLE of  $\Sigma$  is

$$\tilde{\Sigma} = \frac{1}{T} (\mathbf{Y} - \mathbf{X}\tilde{\Pi})' (\mathbf{Y} - \mathbf{X}\tilde{\Pi}).$$

Kullback's symmetric divergence in Eq. (2.6), when replaced by MLE, becomes

$$\begin{aligned} & 2J(\theta_0, \tilde{\theta}) \\ &= d(\theta_0, \tilde{\theta}) - d(\theta_0, \theta_0) + d(\tilde{\theta}, \theta_0) - d(\tilde{\theta}, \tilde{\theta}) \\ &= TM \log(2\pi) + T \log |\tilde{\Sigma}| \\ &+ E_{\theta_0} \left\{ (\mathbf{y} - \mathbf{X}^* \tilde{\pi})' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \tilde{\pi}) \right\} \\ &- TM \log(2\pi) - T \log |\Sigma_0| \\ &- E_{\theta_0} \left\{ (\mathbf{y} - \mathbf{X}^* \pi_0)' (\Sigma_0^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \pi_0) \right\} \\ &+ TM \log(2\pi) + T \log |\Sigma_0| \\ &+ E_{\tilde{\theta}} \left\{ (\mathbf{y} - \mathbf{X}^* \pi_0)' (\Sigma_0^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \pi_0) \right\} \\ &- TM \log(2\pi) - T \log |\tilde{\Sigma}| \\ &- E_{\tilde{\theta}} \left\{ (\mathbf{y} - \mathbf{X}^* \tilde{\pi})' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \tilde{\pi}) \right\}. \quad (2.7) \end{aligned}$$

Ignore  $d(\theta_0, \theta_0)$  in Eq. (2.7) as it remains constant and does not depend on  $\tilde{\theta}$ . We utilize these fundamental principles:

$$\begin{aligned} & E_{\theta_0} \left\{ (\mathbf{y} - \mathbf{X}^* \tilde{\pi})' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \tilde{\pi}) \right\} \\ &= (\pi_0 - \tilde{\pi})' \mathbf{X}'' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\pi_0 - \tilde{\pi}) + T \text{tr}(\tilde{\Sigma}^{-1} \Sigma_0), \\ & E_{\tilde{\theta}} \left\{ (\mathbf{y} - \mathbf{X}^* \pi_0)' (\Sigma_0^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \pi_0) \right\} \\ &= (\tilde{\pi} - \pi_0)' \mathbf{X}'' (\Sigma_0^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\tilde{\pi} - \pi_0) + T \text{tr}(\Sigma_0^{-1} \tilde{\Sigma}), \\ & E_{\tilde{\theta}} \left\{ (\mathbf{y} - \mathbf{X}^* \tilde{\pi})' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) (\mathbf{y} - \mathbf{X}^* \tilde{\pi}) \right\} = TM. \end{aligned}$$

Therefore, the estimate of Kullback's symmetric divergence in Eq. (2.7) can be reduced to

$$\begin{aligned} & 2J(\theta_0, \tilde{\theta}) = TM \log(2\pi) + T \log |\tilde{\Sigma}| \\ &+ (\pi_0 - \tilde{\pi})' \mathbf{X}'' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\pi_0 - \tilde{\pi}) \\ &+ T \text{tr}(\tilde{\Sigma}^{-1} \Sigma_0) + T \log |\Sigma_0| \\ &+ (\tilde{\pi} - \pi_0)' \mathbf{X}'' (\Sigma_0^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\tilde{\pi} - \pi_0) \\ &+ T \text{tr}(\Sigma_0^{-1} \tilde{\Sigma}) - T \log |\tilde{\Sigma}| - TM. \quad (2.8) \end{aligned}$$

The expected value of Eq. (2.8) is

$$\begin{aligned} & E_{\theta_0} \left\{ 2J(\theta_0, \tilde{\theta}) \right\} \\ &= TM [\log(2\pi) + 1] + T E_{\theta_0} \left\{ \log |\tilde{\Sigma}| \right\} \\ &+ E_{\theta_0} \left\{ (\pi_0 - \tilde{\pi})' \mathbf{X}'' (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\pi_0 - \tilde{\pi}) \right\} \\ &+ T E_{\theta_0} \left\{ \text{tr}(\tilde{\Sigma}^{-1} \Sigma_0) \right\} \\ &+ E_{\theta_0} \left\{ (\tilde{\pi} - \pi_0)' \mathbf{X}'' (\Sigma_0^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\tilde{\pi} - \pi_0) \right\} \\ &+ T E_{\theta_0} \left\{ \text{tr}(\Sigma_0^{-1} \tilde{\Sigma}) \right\} \\ &- T E_{\theta_0} \left\{ \log (|\tilde{\Sigma}| / |\Sigma_0|) \right\} - 2TM. \quad (2.9) \end{aligned}$$

Ignore  $TM [\log(2\pi) + 1]$  in Eq. (2.9) as it is irrelevant to our analysis. We know that  $\tilde{\pi}$  and  $T\tilde{\Sigma}$  are independent where  $\tilde{\pi}$  has Gaussian distribution with mean  $\pi$  and the covariance matrix  $[\mathbf{X}'' (\Sigma_0^{-1} \otimes \mathbf{I}_T) \mathbf{X}^*]^{-1}$ . Then

$T\tilde{\Sigma}$  has a Wishart distribution with degrees of freedom  $T$ , i.e.,  $T\tilde{\Sigma} \sim W_{KM}(\Sigma_0, T)$  [19]. The Wishart distribution is a probability distribution that generalizes the chi-squared distribution to multiple dimensions. It is commonly used in multivariate statistics, especially in the estimation of covariance matrices. From the above requirements, we will get,

$$\begin{aligned} E_{\theta_0} \{ \tilde{\Sigma}^{-1} \} &= \frac{T}{T-M-1} \Sigma_0^{-1}, \\ E_{\theta_0} \{ (\tilde{\pi} - \pi_0)' \mathbf{X}^{*'} (\Sigma_0^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\tilde{\pi} - \pi_0) \} &= KM, \\ E_{\theta_0} \{ (\pi_0 - \tilde{\pi})' \mathbf{X}^{*'} (\tilde{\Sigma}^{-1} \otimes \mathbf{I}_T) \mathbf{X}^* (\pi_0 - \tilde{\pi}) \} \\ &= \text{tr} \left\{ \left[ E_{\theta_0} (\tilde{\Sigma}^{-1}) \otimes \mathbf{I}_T \right] E_{\theta_0} \left[ \mathbf{X}^* (\pi_0 - \tilde{\pi}) (\pi_0 - \tilde{\pi})' \mathbf{X}^{*'} \right] \right\} \\ &= \frac{TKM}{T-M-1}, \\ E_{\theta_0} \{ T \text{tr} (\tilde{\Sigma}^{-1} \Sigma_0) \} &= \frac{T^2}{T-M-1} \text{tr} \{ \Sigma_0^{-1} \Sigma_0 \} = \frac{T^2 M}{T-M-1} \\ E_{\theta_0} \{ T \text{tr} (\Sigma_0^{-1} \tilde{\Sigma}) \} &= \text{tr} \{ \Sigma_0^{-1} E_{\theta_0} (T \tilde{\Sigma}) \} = TM, \\ T E_{\theta_0} \{ \log(|\tilde{\Sigma}|/|\Sigma_0|) \} \\ &= T E_{\theta_0} \{ \log(|T \tilde{\Sigma}|/|\Sigma_0|) \} - TM \log T. \end{aligned}$$

Using the knowledge outlined above, we substitute values into Eq. (2.9) to obtain the result

$$\begin{aligned} E_{\theta_0} \{ 2J(\theta_0, \tilde{\theta}) \} \\ &= T E_{\theta_0} \{ \log|\tilde{\Sigma}| \} - T E_{\theta_0} \{ \log(|T \tilde{\Sigma}|/|\Sigma_0|) \} \\ &\quad + TM \log T + \frac{TM(2K+M+1)}{T-M-1} - \frac{KM(M+1)}{T-M-1}. \end{aligned} \quad (2.10)$$

Since  $|T \tilde{\Sigma}|/|\Sigma_0| \sim \prod_{i=1}^M \chi_{T-K-M+i}^2$ , then  $\log(|T \tilde{\Sigma}|/|\Sigma_0|) \sim \sum_{i=1}^M \log \chi_{T-K-M+i}^2$ . Using the second-order Taylor series expansion to expand the term  $\log(\chi_p^2)$  around the mean  $p$ , we have

$$\begin{aligned} \log(\chi_p^2) &\doteq \log(p) + \frac{1}{p}(\chi_p^2 - p) - \frac{1}{2p^2}(\chi_p^2 - p)^2, \\ E[\log(\chi_p^2)] &\doteq \log(p) - \frac{1}{p}, \\ E_{\theta_0} \{ \log(|T \tilde{\Sigma}|/|\Sigma_0|) \} \\ &\doteq \sum_{i=1}^M \left[ \log(T-K-M+i) - \frac{1}{T-K-M+i} \right]. \end{aligned} \quad (2.11)$$

McQuarrie and Tsai [20] determined that for  $T-K-M > M$ , Eq. (2.11) has the formulas as Eq. (2.12) and Eq. (2.13):

$$\sum_{i=1}^M \log(T-K-M+i) \doteq M \log \left( T-K-\frac{M-1}{2} \right), \quad (2.12)$$

and

$$\sum_{i=1}^M \frac{1}{T-K-M+i} \doteq \frac{M}{T-K-\frac{M-1}{2}}. \quad (2.13)$$

Substituting Eq. (2.11), (2.12), and (2.13) into Eq. (2.10), we obtain the expected value of Kullback's symmetric divergence estimate as

$$\begin{aligned} E_{\theta_0} \{ 2J(\theta_0, \tilde{\theta}) \} \\ &= T E_{\theta_0} \{ \log|\tilde{\Sigma}| \} + TM \log \left( \frac{2T}{2T-2K-M+1} \right) \\ &\quad + \frac{2TM}{2T-2K-M+1} + \frac{TM(2K+M+1)}{T-M-1} \\ &\quad - \frac{KM(M+1)}{T-M-1}. \end{aligned} \quad (2.14)$$

Based on Eq. (2.14), we propose a new criterion for selecting the SEM after dividing by the sample size ( $T$ ). This criterion is developed using Kullback's symmetric divergence and employs MLE to estimate the model parameters. We have named this new criterion SKICc, and its formula is as follows:

$$SKICc = \log|\tilde{\Sigma}| + M \log \left( \frac{2T}{2T-2K-M+1} \right)$$

$$+ \frac{2M}{2T-2K-M+1} + \frac{M(2K+M+1)}{T-M-1} - \frac{KM(M+1)}{T(T-M-1)}. \quad (2.15)$$

Performance comparison of SKICc, as described in Eq. (2.15), is conducted against SAIC [13] in Eq. (2.16), SBIC [14] in Eq. (2.17), and SKIC [4] in Eq. (2.18). This comparison is carried out using data simulation, following the steps outlined in Section 2.2. The formulas for the previously mentioned model selection criteria are as follows:

$$SAIC = \log |\hat{\Sigma}| + \frac{M}{T}(K+M+1), \quad (2.16)$$

$$SBIC = \log |\hat{\Sigma}| + \left[ \frac{2K+M+1}{2} \right] \frac{M \log(T)}{T} - \frac{KM}{T}, \quad (2.17)$$

$$SKIC = \log |\hat{\Sigma}| + M \log \left( \frac{2T}{2T-2K-M+1} \right) + \frac{2M}{2T-2K-M+1} + \frac{M(2K+M+1)}{T-K-M-1}. \quad (2.18)$$

Consider the efficiency at percentage of correct model selection and  $L_2$  efficiency as follows:

$$L_2 \text{ efficiency} = \frac{\min_{1 \leq K \leq K^*} L_2(K)}{L_2(K_s)}, \quad (2.19)$$

where  $\min_{1 \leq K \leq K^*} L_2(K)$  is the minimum of  $L_2(K)$  in class  $K^*$  of all possible fitted models.

$$L_2(K) = \frac{1}{T} (\pi - \tilde{\pi})' X^{**} (\tilde{\Sigma}^{-1} \otimes I_T) X^* (\pi - \tilde{\pi}),$$

is the distance between the true and fitted models, and  $L_2(K_s)$  is the distance between the true and the fitted models using the considering model selection criterion.

## 2.2 Simulation

This study employs an SEM in the reduced form of Eq. (2.5) that contains three equations ( $M=3$ ). The choice of three equations was made to strike a balance between a model that is complex enough to illustrate the relationships among endogenous variables, yet simple enough to be interpreted clearly. Two sets of coefficients are considered: the first uses low parameter values, representing weak relationships between the independent and dependent variables—when coefficients approach zero, the independent variable contributes little to explaining the variation in the dependent variable, as shown in Eq. (2.20). The second set, shown in Eq. (2.21), uses larger parameter values to reflect stronger relationships between the variables. Each equation contains four independent variables, which gives the model enough detail to explain the relationships clearly while still keeping it simple enough to understand.

$$\begin{aligned} y_{t1} &= 1 + 0.9x_{t2} + 0.8x_{t3} + 0.7x_{t4} + \varepsilon_{t1} \\ y_{t2} &= 1 - 0.25x_{t2} - 0.13x_{t3} - 0.05x_{t4} + \varepsilon_{t2} \\ y_{t3} &= 1 + 0.5x_{t2} + 0.5x_{t3} + 0.5x_{t4} + \varepsilon_{t3}, \end{aligned} \quad (2.20)$$

and

$$\begin{aligned} y_{t1} &= 1 + 3x_{t2} + 2x_{t3} + x_{t4} + \varepsilon_{t1} \\ y_{t2} &= 1 - 0.5x_{t2} - 0.25x_{t3} - 0.1x_{t4} + \varepsilon_{t2} \\ y_{t3} &= 1 + x_{t2} + x_{t3} + x_{t4} + \varepsilon_{t3}. \end{aligned} \quad (2.21)$$

The steps for simulating data are as follows:

**Step 1.** Generate a vector of 1,000,000 multivariate normal errors in Eq. (2.4) of size  $3 \times 1$  as follows:

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \\ \varepsilon_{t3} \end{bmatrix} \sim N_3 \left( \mathbf{0}, \Sigma = \begin{bmatrix} 0.81 & 0.648 & 0.441 \\ 0.648 & 0.64 & 0.448 \\ 0.441 & 0.448 & 0.49 \end{bmatrix} \right).$$

**Step 2.** Split the error vector obtained from Step 1 into 1,000 simulated data sets, each with

three different sample sizes:  $T = 15, 20$ , and  $25$  units.

**Step 3.** Generate independent variables:  $x_{t2}, x_{t3}, \dots, x_{t8}$  based on 1,000,000 observations that are normally distributed random variables with a zero mean and variance equal to one. The independent variables used in creating the SEM in Eq. (2.20) and Eq. (2.21) are referred to as  $x_{t2}, x_{t3}, x_{t4}$ , while those not used in the SEM are labeled as  $x_{t5}, x_{t6}, x_{t7}, x_{t8}$ .

**Step 4.** Divide the independent variables obtained from Step 3 into 1,000 simulated data sets, each with three different sample sizes:  $T = 15, 20$ , and  $25$  units.

**Step 5.** Test for multicollinearity in Step 4 to avoid issues that could distort the interpretation of the model's performance and lead to inaccurate conclusions regarding the relationships between variables. Ensure that the tolerance value is greater than or equal to 0.1 and the variance inflation factor is less than 10. Discard the sample sets that fail the test. Collect 1,000 sets of simulation data samples that do not exhibit multicollinearity for creating the SEM in the next step.

**Step 6.** Create dependent variables in the SEM as in Eq. (2.20) and Eq. (2.21) using the relevant independent variables from Step 5 and the errors from Step 2.

**Step 7.** The independent variables were entered into the model one at a time, following a specific order, until all seven independent variables in Step 3 were included. This study did not incorporate all possible independent variables because the literature review indicated that the method of sequentially entering the independent variables or importing all variables yielded no significant difference in experimental results [18]. For each time the independent variables were entered into the model, we defined  $K = 2$  for

the model contained constants and  $x_{t2}$ ;  $K = 3$  for the model contained constants,  $x_{t2}$ , and  $x_{t3}$ , until  $K = 8$  for the model contained constants,  $x_{t2}, x_{t3}, x_{t4}, x_{t5}, x_{t6}, x_{t7}$  and  $x_{t8}$ .

**Step 8.** Each time the independent variables are entered into the model, calculate the values of SKICc in Eq. (2.15), SAIC in Eq. (2.16), SBIC in Eq. (2.17), and SKIC in Eq. (2.18). Determine which model order,  $K = 2, 3, \dots, 8$ , yields the lowest value for the model selection criterion. This model will be identified as the fitted model. Repeat this process for 1,000 simulated data sets.

**Step 9.** Calculate the percentage of correct model selection from 1,000 simulated data sets. Additionally, calculate the  $L_2$  efficiency as in Eq. (2.19). Determine which criterion has the highest probability of correct model selection—specifically, the highest percentage of correct model selection, the lowest percentage of models that contain too many or too few variables in the SEM, and the highest  $L_2$  efficiency, represented by the highest mean and the lowest standard deviation of  $L_2$  efficiency.

### 3. Results and Discussion

The effectiveness of the model selection criteria was measured by the percentage of correct model selection, as reported in Tables 1 and 2. For models with weak correlations between independent and dependent variables (Table 1), SKICc performed best for sample sizes of 15 and 20, outperforming SBIC, SAIC, and SKIC. However, when the sample size increased to 25, SKIC proved to be the most effective criterion. In models with strong correlations (Table 2), SKICc again had the highest accuracy at a sample size of 15; however, as the sample size increased to 20 and 25, SKIC emerged as the most effective criterion, clearly surpassing the others.

**Table 1.** Percentage of correctly selected models with a weak correlation between independent and dependent variables.

| T  | Criteria | Model order (K) |     |             |     |     |     |     |
|----|----------|-----------------|-----|-------------|-----|-----|-----|-----|
|    |          | 2               | 3   | 4*          | 5   | 6   | 7   | 8   |
| 15 | SAIC     | 0.3             | 0.3 | 84          | 7.2 | 1.8 | 1.4 | 5   |
|    | SBIC     | 0.6             | 0.6 | 90.7        | 5.2 | 0.6 | 0.8 | 1.5 |
|    | SKIC     | 81.9            | 5.1 | 13          | 0   | 0   | 0   | 0   |
|    | SKICc    | 2.3             | 1.6 | <b>93.1</b> | 2.6 | 0.1 | 0.1 | 0.2 |
| 20 | SAIC     | 0               | 0   | 90.3        | 7.1 | 2   | 0.3 | 0.3 |
|    | SBIC     | 0               | 0   | 96.1        | 3.5 | 0.4 | 0   | 0   |
|    | SKIC     | 7.3             | 2.6 | 90.1        | 0   | 0   | 0   | 0   |
|    | SKICc    | 0               | 0.1 | <b>96.8</b> | 2.7 | 0.4 | 0   | 0   |
| 25 | SAIC     | 0               | 0   | 93.5        | 5   | 0.8 | 0.4 | 0.3 |
|    | SBIC     | 0               | 0   | 98          | 1.9 | 0.1 | 0   | 0   |
|    | SKIC     | 0.1             | 0.4 | <b>99.4</b> | 0.1 | 0   | 0   | 0   |
|    | SKICc    | 0               | 0   | 98.2        | 1.7 | 0.1 | 0   | 0   |

**Note:** \* denotes the true model. Bold text indicates the highest percentage of correct model selection.

**Table 2.** Percentage of correctly selected models with a strong correlation between independent and dependent variables.

| T  | Criteria | Model order (K) |   |             |     |     |     |     |
|----|----------|-----------------|---|-------------|-----|-----|-----|-----|
|    |          | 2               | 3 | 4*          | 5   | 6   | 7   | 8   |
| 15 | SAIC     | 0               | 0 | 84.5        | 7.4 | 1.9 | 1.4 | 4.8 |
|    | SBIC     | 0               | 0 | 91.9        | 5.2 | 0.6 | 0.8 | 1.5 |
|    | SKIC     | 4.3             | 6 | 89.7        | 0   | 0   | 0   | 0   |
|    | SKICc    | 0               | 0 | <b>97</b>   | 2.6 | 0.1 | 0.1 | 0.2 |
| 20 | SAIC     | 0               | 0 | 90.3        | 7.1 | 2   | 0.3 | 0.3 |
|    | SBIC     | 0               | 0 | 96.1        | 3.5 | 0.4 | 0   | 0   |
|    | SKIC     | 0               | 0 | <b>100</b>  | 0   | 0   | 0   | 0   |
|    | SKICc    | 0               | 0 | 96.9        | 2.7 | 0.4 | 0   | 0   |
| 25 | SAIC     | 0               | 0 | 93.5        | 5   | 0.8 | 0.4 | 0.3 |
|    | SBIC     | 0               | 0 | 98          | 1.9 | 0.1 | 0   | 0   |
|    | SKIC     | 0               | 0 | <b>99.9</b> | 0.1 | 0   | 0   | 0   |
|    | SKICc    | 0               | 0 | 98.2        | 1.7 | 0.1 | 0   | 0   |

**Note:** \* denotes the true model. Bold text indicates the highest percentage of correct model selection.

The efficiency of the model selection criteria was also evaluated using  $L_2$  efficiency, as presented in Tables 3 and 4. Table 3 shows that, for models with weak correlations between independent and dependent variables and sample sizes of 15 and 20, SKICc achieved the highest mean  $L_2$  efficiency and the lowest standard deviation, indicating the most stable performance. When the sample size increased to 25, SKIC

showed superior performance compared to the other criteria, followed by SKICc, SBIC, and SAIC. Likewise, Table 4 indicates that for models with strong correlations and a sample size of 15, SKICc achieved the best results. However, as the sample size grew to 20 and 25, SKIC emerged as the most efficient criterion, with SKICc, SBIC, and SAIC ranking next in that order.

**Table 3.** Mean and standard deviation of  $L_2$  efficiency from models with a weak correlation between independent and dependent variables.

| T  | Criteria | $L_2$ efficiency |                    |
|----|----------|------------------|--------------------|
|    |          | Mean             | Standard deviation |
| 15 | SAIC     | 0.7988           | 0.3018             |
|    | SBIC     | 0.8458           | 0.2525             |
|    | SKIC     | 0.7090           | 0.2157             |
|    | SKICc    | <b>0.8698</b>    | <b>0.2137</b>      |
| 20 | SAIC     | 0.9168           | 0.1979             |
|    | SBIC     | 0.9502           | 0.1471             |
|    | SKIC     | 0.9288           | 0.1667             |
|    | SKICc    | <b>0.9542</b>    | <b>0.1403</b>      |
| 25 | SAIC     | 0.9210           | 0.1572             |
|    | SBIC     | 0.9326           | 0.1431             |
|    | SKIC     | <b>0.9881</b>    | <b>0.0630</b>      |
|    | SKICc    | 0.9807           | 0.0928             |

**Note:** Bold text indicates the  $L_2$  efficiency with the highest mean and lowest standard deviation.

**Table 4.** Mean and standard deviation of  $L_2$  efficiency from models with a strong correlation between independent and dependent variables.

| T  | Criteria | $L_2$ efficiency |                    |
|----|----------|------------------|--------------------|
|    |          | Mean             | Standard deviation |
| 15 | SAIC     | 0.8274           | 0.2945             |
|    | SBIC     | 0.8800           | 0.2369             |
|    | SKIC     | 0.8848           | 0.1959             |
|    | SKICc    | <b>0.9089</b>    | <b>0.1895</b>      |
| 20 | SAIC     | 0.9158           | 0.1655             |
|    | SBIC     | 0.9305           | 0.1463             |
|    | SKIC     | <b>0.9849</b>    | <b>0.0638</b>      |
|    | SKICc    | 0.9645           | 0.1304             |
| 25 | SAIC     | 0.9211           | 0.1567             |
|    | SBIC     | 0.9326           | 0.1431             |
|    | SKIC     | <b>0.9935</b>    | <b>0.0400</b>      |
|    | SKICc    | 0.9833           | 0.0884             |

**Note:** Bold text indicates the  $L_2$  efficiency with the highest mean and lowest standard deviation.

Comparison of the penalty function of each model selection criterion presented in Eqs. (2.15)-(2.18) is as follows:

penalty function of SAIC is  $\frac{M}{T}(K+M+1)$ ,

penalty function of SBIC is

$$\left[ \frac{2K+M+1}{2} \right] \frac{M \log(T)}{T} - \frac{KM}{T},$$

penalty function of SKIC is

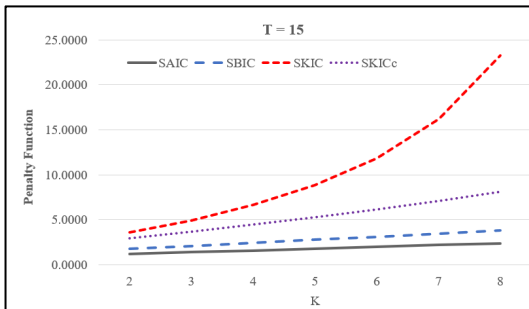
$$M \log \left( \frac{2T}{2T-2K-M+1} \right) + \frac{2M}{2T-2K-M+1} + \frac{M(2K+M+1)}{T-K-M-1},$$

and penalty function of SKICc is

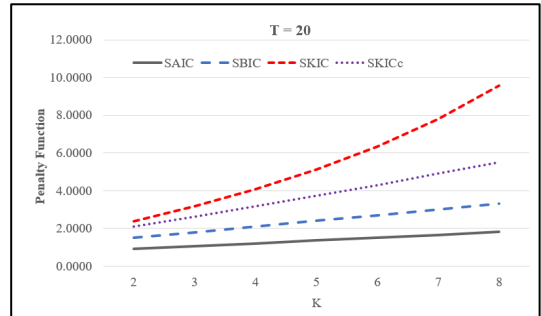
$$M \log \left( \frac{2T}{2T-2K-M+1} \right) + \frac{2M}{2T-2K-M+1} + \frac{M(2K+M+1)}{T-M-1} - \frac{KM(M+1)}{T(T-M-1)}.$$

The results of comparing the penalty function when the number of equations in the SEM is three equations ( $M = 3$ ) and across three small sample sizes ( $T = 15, 20$ , and  $25$ ) are presented in Figs. 1-3. It was observed that the penalty function for SKIC increases rapidly as the number of independent variables ( $K$ ) in the model increases, especially when the sample size is very small, as illustrated in Fig.

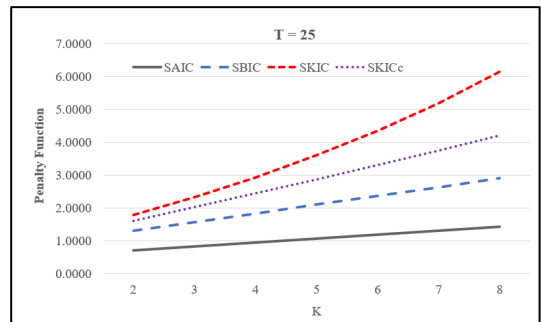
1. The rapid increase in the penalty function also resulted in a significant increase in the value of the model selection criterion, as shown in Tables 5-6. Consequently, this criterion tends to select models with too few variables [16-18]. In contrast, the penalty functions of SAIC and SBIC exhibit unrealistically low values, leading to a tendency to select models with too many variables, as indicated by the percentage of correctly selected models presented in Tables 1 and 2. However, when comparing SAIC and SBIC, SBIC typically selects models with fewer variables than SAIC [11, 13, 21].



**Fig. 1.** Comparison of the penalty functions of SAIC, SBIC, SKIC, and SKICc for a sample size of 15.



**Fig. 2.** Comparison of the penalty functions of SAIC, SBIC, SKIC, and SKICc for a sample size of 20.



**Fig. 3.** Comparison of the penalty functions of SAIC, SBIC, SKIC, and SKICc for a sample size of 25.

**Table 5.** Mean and standard deviation of SAIC, SBIC, SKIC, and SKICc from models with a weak correlation between independent and dependent variables.

| T  | Criterion | Stat.     | Model order (K) |         |                |         |         |         |         |
|----|-----------|-----------|-----------------|---------|----------------|---------|---------|---------|---------|
|    |           |           | 2               | 3       | 4*             | 5       | 6       | 7       | 8       |
| 15 | SAIC      | $\bar{X}$ | -0.6247         | -1.0454 | <b>-2.5762</b> | -2.1622 | -1.7258 | -1.2873 | -0.8426 |
|    |           | s         | 0.7268          | 0.7850  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SBIC      | $\bar{X}$ | -0.0583         | -0.3373 | <b>-1.7266</b> | -1.1709 | -0.5930 | -0.0128 | 0.5735  |
|    |           | s         | 0.7268          | 0.7850  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SKIC      | $\bar{X}$ | <b>1.7614</b>   | 2.5078  | 2.4830         | 4.9036  | 8.1350  | 12.7277 | 20.0063 |
|    |           | s         | 0.7268          | 0.7850  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SKICc     | $\bar{X}$ | 0.7018          | 0.5975  | <b>-0.6085</b> | 0.1418  | 0.9298  | 1.7418  | 2.5926  |
|    |           | s         | 0.7268          | 0.7850  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
| 20 | SAIC      | $\bar{X}$ | -0.7843         | -1.3110 | <b>-2.9070</b> | -2.6025 | -2.2782 | -1.9442 | -1.6062 |
|    |           | s         | 0.6033          | 0.6353  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SBIC      | $\bar{X}$ | -0.1868         | -0.5642 | <b>-2.0109</b> | -1.5570 | -1.0833 | -0.6000 | -0.1126 |
|    |           | s         | 0.6033          | 0.6353  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SKIC      | $\bar{X}$ | 0.6940          | 0.8036  | <b>-0.0440</b> | 1.1500  | 2.5449  | 4.1883  | 6.1600  |
|    |           | s         | 0.6033          | 0.6353  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SKICc     | $\bar{X}$ | 0.0887          | -0.2291 | <b>-1.6134</b> | -1.0937 | -0.5501 | 0.0084  | 0.5775  |
|    |           | s         | 0.6033          | 0.6353  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |

| T  | Criterion | Stat.     | Model order (K) |         |                |         |         |         |         |
|----|-----------|-----------|-----------------|---------|----------------|---------|---------|---------|---------|
|    |           |           | 2               | 3       | 4*             | 5       | 6       | 7       | 8       |
| 25 | SAIC      | $\bar{X}$ | -0.8773         | -1.4548 | <b>-3.1128</b> | -2.8659 | -2.6074 | -2.3488 | -2.0813 |
|    |           | s         | 0.5317          | 0.5449  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SBIC      | $\bar{X}$ | -0.2922         | -0.7235 | <b>-2.2352</b> | -1.8420 | -1.4373 | -1.0324 | -0.6187 |
|    |           | s         | 0.5317          | 0.5449  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SKIC      | $\bar{X}$ | 0.1857          | 0.0378  | <b>-1.1357</b> | -0.3397 | 0.5447  | 1.5218  | 2.6204  |
|    |           | s         | 0.5317          | 0.5449  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SKICc     | $\bar{X}$ | -0.2304         | -0.6524 | <b>-2.1535</b> | -1.7484 | -1.3300 | -0.9094 | -0.4777 |
|    |           | s         | 0.5317          | 0.5449  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |

Note: \* denotes the true model. Bold text indicates the lowest mean of the selection criteria.

**Table 6.** Mean and standard deviation of SAIC, SBIC, SKIC, and SKICc from models with a strong correlation between independent and dependent variables.

| T  | Criterion | Stat.     | Model order (K) |         |                |         |         |         |         |
|----|-----------|-----------|-----------------|---------|----------------|---------|---------|---------|---------|
|    |           |           | 2               | 3       | 4*             | 5       | 6       | 7       | 8       |
| 15 | SAIC      | $\bar{X}$ | 1.3367          | -0.1369 | <b>-2.5762</b> | -2.1622 | -1.7258 | -1.2873 | -0.8426 |
|    |           | s         | 0.7210          | 0.7894  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SBIC      | $\bar{X}$ | 1.9032          | 0.5711  | <b>-1.7266</b> | -1.1709 | -0.5930 | -0.0128 | 0.5735  |
|    |           | s         | 0.7210          | 0.7894  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SKIC      | $\bar{X}$ | 3.7228          | 3.4162  | <b>2.4830</b>  | 4.9036  | 8.1350  | 12.7277 | 20.0063 |
|    |           | s         | 0.7210          | 0.7894  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
|    | SKICc     | $\bar{X}$ | 2.6632          | 1.5059  | <b>-0.6085</b> | 0.1418  | 0.9298  | 1.7418  | 2.5926  |
|    |           | s         | 0.7210          | 0.7894  | 0.8544         | 0.8996  | 0.9569  | 1.0311  | 1.1511  |
| 20 | SAIC      | $\bar{X}$ | 1.1689          | -0.4014 | <b>-2.9070</b> | -2.6025 | -2.2782 | -1.9442 | -1.6062 |
|    |           | s         | 0.5916          | 0.6335  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SBIC      | $\bar{X}$ | 1.7663          | 0.3454  | <b>-2.0109</b> | -1.5570 | -1.0833 | -0.6000 | -0.1126 |
|    |           | s         | 0.5916          | 0.6335  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SKIC      | $\bar{X}$ | 2.6472          | 1.7133  | <b>-0.0440</b> | 1.1500  | 2.5449  | 4.1883  | 6.1600  |
|    |           | s         | 0.5916          | 0.6335  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
|    | SKICc     | $\bar{X}$ | 2.0418          | 0.6805  | <b>-1.6134</b> | -1.0937 | -0.5501 | 0.0084  | 0.5775  |
|    |           | s         | 0.5916          | 0.6335  | 0.6592         | 0.6940  | 0.7216  | 0.7527  | 0.7943  |
| 25 | SAIC      | $\bar{X}$ | 1.0771          | -0.5471 | <b>-3.1128</b> | -2.8659 | -2.6074 | -2.3488 | -2.0813 |
|    |           | s         | 0.5204          | 0.5468  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SBIC      | $\bar{X}$ | 1.6622          | 0.1842  | <b>-2.2352</b> | -1.8420 | -1.4373 | -1.0324 | -0.6187 |
|    |           | s         | 0.5204          | 0.5468  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SKIC      | $\bar{X}$ | 2.1402          | 0.9455  | <b>-1.1357</b> | -0.3397 | 0.5447  | 1.5218  | 2.6204  |
|    |           | s         | 0.5204          | 0.5468  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |
|    | SKICc     | $\bar{X}$ | 1.7240          | 0.2553  | <b>-2.1535</b> | -1.7484 | -1.3300 | -0.9094 | -0.4777 |
|    |           | s         | 0.5204          | 0.5468  | 0.5658         | 0.5932  | 0.6071  | 0.6231  | 0.6449  |

Note: \* denotes the true model. Bold text indicates the lowest mean of the selection criteria.

The mean and standard deviation of the SAIC, SBIC, SKIC, and SKICc from 1,000 simulated data sets are presented in Tables 5 and 6. Although all four model selection criteria shared the same standard deviation,

their means varied in descending order: SKIC, SKICc, SBIC, and SAIC. When comparing these criteria with the percentage of correct model selections shown in Tables 1 and 2, it was observed that criteria yielding

unrealistically high values, like SKIC, tended to select models with too few variables, especially in cases with small sample sizes. Conversely, criteria yielding unrealistically low values, like SAIC and SBIC, tended to select models with too many variables. In the case of SKICc, there is rarely a problem with selecting a model that has either too few or too many variables.

#### 4. Conclusion

Our paper introduces a new criterion for selecting the SEM with small sample sizes, based on Kullback's symmetric divergence, and employs the MLE to estimate the model parameters. The new criterion is called SKICc (System of Simultaneous Equations KICc). The formula is as follows:

$$\text{SKICc} = \log |\tilde{\Sigma}| + M \log \left( \frac{2T}{2T - 2K - M + 1} \right) + \frac{2M}{2T - 2K - M + 1} + \frac{M(2K + M + 1)}{T - M - 1} - \frac{KM(M + 1)}{T(T - M - 1)},$$

where  $\tilde{\Sigma} = \frac{1}{T} (\mathbf{Y} - \mathbf{X}\tilde{\Pi})' (\mathbf{Y} - \mathbf{X}\tilde{\Pi})$  represents MLE of  $\Sigma$ .

The comparison of SKICc's efficiency with SAIC, SBIC, and SKIC revealed notable insights. When adding one independent variable to the model at a time and selecting based on the lowest model selection criterion, results from 1,000 simulated data sets show that under relatively low parameter values—indicating weak relationships between the independent and dependent variables in the SEM—SKICc performs better than the other methods for sample sizes of 15 and 20. It achieves the highest percentage of correct model selections and demonstrates the highest mean, along with the lowest standard deviation of  $L_2$  efficiency. However, as the sample size increases to 25, SKIC shows the best performance.

In cases where SEM uses larger parameter values to reflect stronger relationships between variables, SKICc performs very well

with a sample size of 15, accurately selecting the correct model. However, as the sample size increases to 20 or 25, SKIC becomes a more effective option. This study also found that the small sample case study revealed the disadvantages of SKIC, which tends to select models with too few variables because the penalty function and the model selection criterion value increase rapidly as the number of independent variables increases, especially when the sample size is very small, such as 15 units. Consequently, SKIC will converge slowly to the true value. On the other hand, the disadvantages of SAIC and SBIC include a tendency to select models that contain too many variables, and they can be inconsistent due to a low penalty function value. However, SBIC usually chooses models that have fewer variables compared to SAIC. The performance evaluation of  $L_2$  efficiency revealed that SKICc is more efficient than other criteria, especially with very small sample sizes. As the sample size increases, SKIC demonstrates the best performance among the criteria. Notably, SKICc rarely encounters issues with selecting models that contain either too many or too few variables. Therefore, SKICc is an appealing alternative for selecting the SEM in cases involving small samples. However, as the sample size increases, SKIC becomes a more appropriate criterion than the others. In future studies, we plan to explore situations in which the standard assumptions do not hold, such as when errors and independent variables deviate from normality or when multicollinearity is present among the independent variables. Additionally, we will test the proposed new criteria on real data.

#### Acknowledgements

This research was supported by the National Science, Research, and Innovation Fund (NSRF) and Thaksin University (Research project grant), Fiscal Year 2024.

## References

- [1] Perez-Sanchez B, Gonzalez M, Perea C, Lopez-Espin JJ. A new computational method for estimating simultaneous equations models using entropy as a parameter criteria. *Mathematics*. 2021; 9(700): 1-9.
- [2] Keerativibool W. Costs and benefits gained from planting para rubber in the form of system of simultaneous equations: a case study of Pa Phayom district, Phatthalung province and Thung Song district, Nakhon Si Thammarat province. *NDJ*. 2014; 54(3): 95-124.
- [3] Zhang J, Yang Y, Ding J. Information criteria for model selection. *WIREs Comp Stats*. 2023; 15(5): 1-27.
- [4] Keerativibool W, Jitthavech J. Model selection criterion based on Kullback-Leibler's symmetric divergence for simultaneous equations model. *Chiang Mai J Sci*. 2015; 42(3): 761-73.
- [5] Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974; 19(6): 716-23.
- [6] Schwarz G. Estimating the dimension of a model. *Ann. Stat*. 1978; 6(2): 461-64.
- [7] Cavanaugh JE. A large-sample model selection criterion based on Kullback's symmetric divergence. *Stat Probabil Lett*. 1999; 42(4): 333-43.
- [8] Cavanaugh JE. Criteria for linear model selection based on Kullback's symmetric divergence. *Aust N Z J Stat*. 2004; 46(2): 257-74.
- [9] Seghouane AK, Bekara M. A small sample model selection criterion based on Kullback's symmetric divergence. *IEEE Trans Signal Process*. 2004; 52(12): 3314-23.
- [10] Hafidi B, Mkhadri A. A corrected Akaike criterion based on Kullback's symmetric divergence: applications in time series, multiple and multivariate regression. *Comput Stat Data Anal*. 2006; 50(6): 1524-50.
- [11] Neath AA, Cavanaugh JE. Regression and time series model selection using variants of the Schwarz information criterion. *Commun Stat - Theory Methods*. 1997; 26(3): 559-80.
- [12] Hurvich CM, Tsai CL. Regression and time series model selection in small samples. *Biometrika*. 1989; 76(2): 297-307.
- [13] Keerativibool W, Jitthavech J, Lorchira-choonkul V. Model selection in a system of simultaneous equations model. *Commun Stat-Theory Methods*. 2011; 40(3): 373-93.
- [14] Keerativibool W. New criteria for selection in simultaneous equations model. *Thail Stat Thail*. 2012; 10(2): 163-81.
- [15] Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. 4<sup>th</sup> ed. New Jersey: Prentice Hall; 1998.
- [16] Keerativibool W. Unifying the derivations of Kullback information criterion and corrected versions. *Thail Stat Thail*. 2014; 12(1): 37-53.
- [17] Keerativibool W. Study on the penalty functions of model selection criteria. *Thail Stat Thail*. 2014; 12(2): 161-78.
- [18] Keerativibool W, Siripanich P. Comparison of the model selection criteria for multiple regression based on Kullback-Leibler's information. *Chiang Mai J Sci*. 2017; 44(2): 699-714.
- [19] Anderson TW. *An introduction to multivariate statistical analysis*. 3<sup>rd</sup> ed. New Jersey: Wiley; 2003.
- [20] McQuarrie AD, Tsai CL. *Regression and time series model selection*. Singapore: World Scientific Pub Co Inc; 1998.

- [21] Acquah HG. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship. *J Dev Agric Econ.* 2010; 2(1): 1-6.