Data Science and Engineering
https://doi.org/10.1007/s41019-025-00318-6

RESEARCH PAPERS

Evaluation of Estimation Methods for Simultaneous Equations Models Across Varying Levels of Data Variability

Belén Pérez-Sánchez[1] · **Carmen Perea**[1] · **Martin Gonzalez**[1] · **Jose J. López-Espín**[1]

Received: 22 October 2024 / Revised: 25 August 2025 / Accepted: 1 September 2025

Abstract

Simultaneous Equations Model (SEM) is a set of regression equations where bidirectional relationships exist between variables. SEMs are widely used to model complex systems, capture the interdependencies between different variables, and make predictions about future outcomes in a wide range of fields such as economics, markets, or health sciences. In the literature, the performance of numerous methods, both classical and Bayesian, has been widely studied in various aspects such as endogeneity or correlation. To our knowledge, the study of estimator performance under varying levels of data variability in simultaneous equation models is not well-developed. This paper aims to evaluate the performance of methods for estimating SEMs of different sizes, considering the number of variables and the variability of endogenous variables. An experimental study has been conducted applying different estimation methods, including Two Stage Least Squares ($2SLS$) and the Optimized Bayesian Method of Moments ($Bmom_{OPT}$), to evaluate their performance across different SEMs. Based on our computational results, the main finding is that the performance of the methods depends on the variability of the data, with $Bmom_{OPT}$ being more accurate at lower levels of variability. These results could interest researchers aiming to apply SEMs in practical cases as they offer insights into selecting the estimation method while considering both the model size and data variability.

Keywords Simultaneous equation models · Optimized Bayesian method of moments · Entropy · Computational statistics

1 Introduction

A Simultaneous Equation Model (SEM) [1] is a statistical model that represents a system of regression equations with bidirectional relationships and interdependencies among the variables. Various methods are employed to estimate SEMs, including Full Information Maximum Likelihood (FIML) and Three-Stage Least Squares (3SLS), as well as Ordinary Least Squares (OLS), Indirect Least Squares (ILS), and Two-Stage Least Squares (2SLS) [2]. Among these classical methods, 2SLS is one of the most widely used due to its simplicity and the consistency of its estimators.

On the other hand, Bayesian methods eliminate the need for sampling assumptions but introduce significant complexity, particularly in the specification of prior distributions and the derivation of posterior distributions. The Bayesian Method of Moments (BMOM) [3] offers an alternative approach by avoiding these sampling assumptions, and the $Bmom_{OPT}$ method enhances BMOM by optimizing parameters to achieve minimum entropy values [4]. Chao and Phillips [5] investigated the behavior of posterior distributions under Jeffreys prior within SEMs, while Geweke [6] developed general methods for Bayesian inference using non-informative reference priors. Kleibergen and Van Dijk [7] advanced Bayesian SEMs by incorporating reduced rank structures. Additionally, the Markov Chain Monte Carlo (MCMC) method, including the Metropolis-Hastings

✉ Belén Pérez-Sánchez
m.perezs@umh.es

Carmen Perea
perea@umh.es

Martin Gonzalez
martin.gonzaleze@umh.es

Jose J. López-Espín
jlopez@umh.es

1 Miguel Hernández University, Avda. Universidad s/n, 03202 Elche, Spain

algorithm and Gibbs Sampling, enables efficient computation in large models with numerous unknown parameters [8].

SEM is a versatile statistical technique that finds applications in different fields. In tax research, it has been utilised to investigate the effects of fiscal decentralisation on regional income inequality in Indonesia [9]. Additionally, SEM has been employed to explore the relationship between employment and mental health [10]. In the context of China's textile industry, studies have examined the impact of foreign trade on energy efficiency [11]. SEM has also been utilised to model prescriptions in primary care settings [12]. Recently, a study focused on the impact of physical activity on electronic media use among Chinese adolescents, emphasising urban-rural differences [13].

The Bayesian approach has also been applied in many different fields. For example, in agricultural science, a Bayesian SEM has been developed to model energy intake and distribution in growing pigs [14], the study of the impact of product information on third-party websites on the feedback mechanism between internal Word-of-Mouth and retail sales on Download.com and Amazon.com [15]. They have also been applied to explore the effects of peers on the behavior of casino games [16], to model the interaction between the perception of risk to people's health and betel chewing habits in Taiwan [17] and to study the effects of repetitive iodine blockade of the thyroid on fetal brain and thyroid development in rats [18].

The literature mainly compares Bayesian and classical estimators in the context of method selection, with findings suggesting that Bayesian methods tend to outperform classical estimators, particularly in small sample sizes [3]. More recently, a comparative study has been conducted on large models, examining variations in the number of variables and sample size [4].

While the review covers both classical and Bayesian methods applied to simultaneous equations models, it is helpful to provide a more structured comparison between them. In this study, we analyze these approaches under varying levels of data variability, which allows us to highlight differences in predictive consistency, sensitivity to variability, and computational requirements.

On the other hand, the criteria of model selection serve as valuable tools for choosing among a set of estimated models that differ in goodness of fit and complexity, providing a relative estimate of the information loss when a particular model is employed to represent the underlying data-generating process. A more complex model is better equipped to capture the relationships under analysis, yet it risks losing its explanatory power. However, an excessively complex model may overfit the data and fail to generalise well beyond the training data. Thus, the objective must be to balance complexity and goodness of fit. Various parameter criteria are available for comparing models, such as the Akaike Information Criteria (AIC) [19, 20], its corrected version (AICc) [21], the Schwarz or Bayesian Information Criterion (SIC/BIC) [22], Hannan and Quinn (HQ) [23], and the Model Selection Criterion based on Kullback-Leibler's Symmetric Divergence [24].

In this context, the concept of entropy is introduced. Initially, the entropy has been employed in thermodynamics as a foundation for the second law of thermodynamics. Entropy later found application in statistical mechanics, connecting the macroscopic properties of entropy with system states [25], and also in various fields, such as finance [26], environmental and water engineering [27], urban systems [28], and customer satisfaction surveys [29]. A recent study introduces an entropy-based measure as an information criterion for the estimation method, serving as a selection criterion for the method with more homogeneous prediction errors [4].

Compared to previous literature, the main contribution of this paper is to provide a comparison of the accuracy of SEM estimation methods by considering data variability. Furthermore, the paper compares measures obtained through classical and Bayesian methods. From a practical viewpoint, this contribution may be interesting for researchers of health organizations, market research, government, etc. who need to understand and estimate complex relationships between multiple variables.

The paper is structured as follows. In Sect. 2, the model and the estimation methods used in this study are briefly reviewed. Section 3 describes some information criteria for selecting estimation methods. The experimental study and accuracy are presented in Sect. 4. Finally, Sect. 5 concludes.

2 The SEM and Estimation Methods

2.1 Model Definition

An SEM (1) is a model consisting of m interdependent (endogenous) variables that are influenced by k independent (exogenous) variables. Each endogenous variable is expressed as a linear combination of the other endogenous variables, the exogenous variables, and white noise representing stochastic interference [1].

$$
\begin{aligned}
y_1 &= B_{1,2}y_2 + B_{1,3}y_3 + \cdots + B_{1,m}y_m \\
&\quad + \Gamma_{1,1}x_1 + \cdots + \Gamma_{1,k}x_k + u_1 \\
y_2 &= B_{2,1}y_1 + B_{2,3}y_3 + \cdots + B_{2,m}y_m \\
&\quad + \Gamma_{2,1}x_1 + \cdots + \Gamma_{2,k}x_k + u_2 \\
&\vdots \\
y_m &= B_{m,1}y_1 + B_{m,2}y_2 + \cdots + B_{m,m-1}y_{m-1} \\
&\quad + \Gamma_{m,1}x_1 + \cdots + \Gamma_{m,k}x_k + u_m
\end{aligned}
\tag{1}
$$

The SEM can be represented in matrix form (2),

$$
YB^T + X\Gamma^T + U = 0 \tag{2}
$$

where $B \in \mathbb{R}^{m\times m}$ and $\Gamma \in \mathbb{R}^{m\times k}$ are coefficient matrices. Certain coefficients of $B_{i,j}$ and $\Gamma_{k,r}$ are pre-defined as zeros. The variables *x*, *y* and *u* represent exogenous, endogenous and white noise variables respectively and they are vectors of dimension *n*, where *n* is the sample size. Then, $Y = (y_1, \ldots, y_m)$, $X = (x_1, \ldots, x_k)$ and $U = (u_1, \ldots, u_m)$.

The solution of the model involves obtaining the matrices *B* and Γ in (2) by analyzing a representative sample of the model. This analysis aims to explain the relationship between the two sets of variables by solving a well-defined matrix equation (2).

2.2 SEM Estimation Methods

This section presents a brief description of the estimation methods used in this work.

(i) Two-stage Least Squares (2SLS) is a statistical method commonly used in econometrics for the estimation of SEM [1]. Ordinary Least Squares (OLS) is used in two steps. First, instrumental variables are used to predict the values of the endogenous variables. Then, these predicted values are inserted into the model in place of the original endogenous variables, and OLS is used again to estimate the parameters.

(ii) Bayesian Method of Moments (BMOM) is a double K-class estimator [3]. When there is insufficient information to obtain the likelihood function, this method allows a data analysis without specifying a probability function and without sampling assumption. The estimator is defined by two parameters, K_1 and K_2, whose values determine whether the estimation emphasizes goodness of fit, precision, or a balance of both.

(iii) The Optimized BMOM Method ($Bmom_{OPT}$) [4] is a variation of the original BMOM approach, designed to optimize the estimation process by adjusting the parameters K_1 and K_2 in order to minimize the AIC value. The expression for the AIC in the context of an SEM is given by:

$$
AIC = n \ln \left|\hat{\Sigma}_e\right| + 2\sum_{i=1}^{m}(m_i + k_i - 1) + m(m+1) \tag{3}
$$

where *n* is the sample size, *m* is the number of equations, m_i and k_i are the number of endogenous and exogenous variables in *i*-equation, and $\hat{\Sigma}_e$ the matrix of variance-covariance of the errors $e_j = Y_j - \hat{Y}_j$, j $= 1, \ldots, m.$

(iv) Bayesian Approach in Two-stages ($Bayes_{2S}$): This approach employs a Normal-Inverse Gamma prior to derive precise analytic expressions for the posterior distribution of the structural *B* and Γ coefficients in the SEM. It follows a two-stage procedure, similar to 2SLS, but replaces the use of OLS with Bayesian estimation in both stages.

(v) Markov Chain Monte Carlo (MCMC) is a computational technique used to generate samples from complex probability distributions. It is particularly valuable when direct sampling is challenging due to high-dimensional spaces or complex dependencies among variables.

3 Parameter Criteria

A parameter criterion as a measure for comparing and evaluating methods, provides an objective way of determining the quality, adequacy y effectiveness of a statistical approach to a specific objective. Choosing which parameter criteria is appropriate, is crucial to making informed decisions about which method to use in a given context. Akaike Information Criteria, expressed in (3), is the most extensed used parameter criteria. It is widely used in multivariate regression models and has been adapted to SEM [30] as a method selection criteria.

On the other hand, a method selection based on entropy involves evaluating this metric across various methods, prioritizing methods with the lowest entropy values. This approach aligns with the principle of Occam's razor, which favors simpler methods that balance complexity with fitting accuracy. In an experimental study by [4], the authors employed an entropy-based measure as a selection criteria for an SEM estimation method. This measure is represented by equation (4):

$$
H_2(e) = \frac{\prod_{j=1}^{m}\left(2 - (p_{ij})^{p_{ij}}\right)}{m} \tag{4}
$$

where *m* is the number of endogenous variables, and the p_{ij} values for each endogenous variable have been obtained as follows (5):

$$p_{ij} = \frac{e_{ij}}{\sum_{i=i}^{n} e_{ij}} \quad j = 1, 2, \dots m, \tag{5}$$

where $e_{ij} = Y_{ij} - \widehat{Y}_{ij}$, n is the sample size and, p_{ij} represents the error mass associated with each endogenous variable.

This entropy-based measure (5) provides an alternative to traditional criteria such as the AIC. It evaluates how estimation errors are distributed across the system. Lower entropy values indicate more concentrated and predictable residuals, while higher values reflect greater dispersion. This is particularly valuable in SEMs, where not only the magnitude but the structure of the error matters. Entropy thus serves as a selection criterion that captures both goodness-of-fit and internal consistency.

4 Experimental Study

In the experiments, SEMs (2) have been generated as follows: matrices B and Γ are randomly generated from an Uniform distribution, X is generated from a multivariate Normal distribution, and Y is calculated as $Y \sim \Pi X + U$, where $U \sim N(0, \sigma_u)$. Values of σ_u ranging from 0.10 to 10.00 were used, leading to different levels of variability in the SEM, denoted by σ_y. Tests were carried out in C code and imported some functions from statistical software R (GNU R version 3.5.2).

For the estimation of the SEM, 7 methods have been used: 2SLS, 3 versions of BMOM, Goodness of Fit (B_{gf}), Precision of Estimation (B_{pe}), and the Traditional Bayesian Approach (B_{ta}), the Optimized BMOM ($Bmom_{OPT}$), $Bayes_{2S}$, and MCMC. Specifically, for $Bmom_{OPT}$ to obtain optimal values of K_1 and K_2, the *optim* function from R has been imported. In the case of $Bayes_{2S}$ method, a Normal-Inverse Gamma prior with a mean of 5.0 and precision of 0.2 as the initial parameter has been employed. Finally, Gibbs Sampling has been chosen for MCMC to simulate the posterior distribution importing the *MCMCpack* package from R. For the comparison, the Euclidean distance ($D_{\delta,\hat{\delta}}$) between the coefficient matrix $\delta = [B; \Gamma]$ and its estimation $\hat{\delta}$, the AIC, the entropy $H_2(e)$, and the execution time in seconds have been calculated.

In order to facilitate the reader's subsequent comprehension, we provide a table which summarizes the descriptions of subsequent recurring symbols (Table 1).

Table 1 Statements of important symbols

Symbol	Statements
Y_{ij}	Observed value of endogenous variable *j* for observation *i*
$\hat{Y}_{ij}$	Estimated value of Y_{ij}
e_{ij}	Estimation error: $Y_{ij} - \hat{Y}_{ij}$
p_{ij}	Proportion of error mass for variable *j* and observation *i*
n	Sample size
m	Number of endogenous variables
k	Number of exogenous variables
σ_u	Standard deviation of the structural error term U
σ_y	Standard deviation of the response variable *Y*

Tables 2, 3 and 4 present the results of three SEMs with different number of endogenous and exogenous variables and sample size. The first two columns denote the model size and the name of measure, while the remaining columns display the mean and standard deviation of the measures for each method. The experiments were conducted for 30 trials in each SEM.

Table 2 presents the results of SEMs with σ_y = 0.1. The first SEM with m = 10, k = 20 and n = 100 shows that the minimum value of $D_{\delta,\hat{\delta}}$ (23.306) has been obtained by $Bmom_{OPT}$, followed by 2*SLS*, $Bayes_{2S}$, the BMOMs, and finally, MCMC. The AIC indicates that the BMOM traditional is the method that obtains the minimum value (2056.552) compared to the rest. Regarding the entropy, the minimum value is reached by $Bmom_{OPT}$ (4.54), followed by 2*SLS*, $Bayes_{2S}$, BMOMs and MCMC. And, in terms of the execution time, 2*SLS* is the method that requires the least time (0.043), while $Bmom_{OPT}$ needs the most time for the estimation (316.834).

By increasing the complexity of the SEM in terms of the number of variables and sample size (m = 10, k = 40, n = 400), the model has obtained the minimum values of $D_{\delta,\hat{\delta}}$ (11.146) with $Bmom_{OPT}$, AIC with BMOM B_{ta} (15255.394), and $H_2(e)$ with both $Bmom_{OPT}$ and 2*SLS* (5.567). $Bayes_{2S}$ is the method that requires the least time for the estimation (0.219). And for the largest SEM (m = 20, k = 100, n = 1000), the same results as the previous SEM have been obtained, being $Bmom_{OPT}$ the method with minimum $D_{\delta,\hat{\delta}}$ (19.849). The minimum $H_2(e)$ (6.558) has been obtained for both $Bmom_{OPT}$ and 2*SLS*.

These results indicate that $Bmom_{OPT}$ consistently achieves the lowest mean values for both $D_{\delta,\hat{\delta}}$ and $H_2(e)$ across model sizes. This comparative performance is also illustrated in Fig. 1, which summarizes the results graphically for $\sigma_u = 0.1$. Regarding computational efficiency, 2*SLS*, $Bayes_{2S}$, and BMOMs are the methods that require the least estimation time.

Furthermore, all methods are affected by the increase in complexity of the model associated with the number of variables and sample size.

Table 3 presents the results for SEMs with $\sigma_u = 1.0$. In all models, $Bmom_{OPT}$ consistently achieved the lowest

Table 2 Mean and standard deviation of $D_{\delta,\hat{\delta}}$, AIC, $H_2(e)$ and time (seconds) for 30 trials when varying the problem size and the estimation method. $\sigma_u = 0.1$

(m, k, n)	Measure	2SLS	BMOM			$Bmom_{OPT}$	$Bayes_{2S}$	MCMC
			B_{qf}	B_{pe}	B_{td}			
(10, 20, 100)	$D_{\delta,\hat{\delta}}$	$29.570_{7.292}$	$43.233_{6.553}$	$43.509_{6.567}$	$44.653_{6.511}$	$23.306_{8.462}$	$31.962_{10.139}$	$68.373_{4.435}$
	AIC	$2511.737_{567.795}$	$2085.189_{539.692}$	$2080.392_{538.473}$	$2056.552_{535.776}$	$2853.010_{627.993}$	$2684.088_{571.362}$	$4438.846_{235.293}$
	$H_2(e)$	$4.056_{0.014}$	$4.068_{0.013}$	$4.068_{0.013}$	$4.070_{0.013}$	$4.054_{0.014}$	$4.065_{0.015}$	$4.097_{0.012}$
	Time	$0.043_{0.013}$	$0.603_{0.168}$	$0.603_{0.168}$	$0.603_{0.168}$	$316.834_{148.244}$	$0.047_{0.010}$	$288.370_{24.462}$
(10, 40, 400)	$D_{\delta,\hat{\delta}}$	$19.486_{5.255}$	$37.365_{6.768}$	$37.391_{6.771}$	$37.784_{6.782}$	$11.146_{3.504}$	$24.802_{17.350}$	$89.056_{4.633}$
	AIC	$16538.739_{1844.884}$	$15279.305_{1598.832}$	$15277.803_{1598.426}$	$15255.394_{1593.465}$	$17443.120_{1991.016}$	$16814.836_{2095.821}$	$21809.350_{705.884}$
	$H_2(e)$	$5.567_{0.005}$	$5.569_{0.007}$	$5.569_{0.007}$	$5.570_{0.007}$	$5.567_{0.007}$	$5.571_{0.010}$	$5.601_{0.006}$
	Time	$0.232_{0.035}$	$4.437_{0.769}$	$4.437_{0.769}$	$4.437_{0.769}$	$2448.428_{919.992}$	$0.219_{0.035}$	$503.499_{41.883}$
(20, 100, 1000)	$D_{\delta,\hat{\delta}}$	$42.808_{6.468}$	$88.271_{8.806}$	$88.262_{8.800}$	$88.999_{8.814}$	$19.849_{3.787}$	$47.486_{7.043}$	$197.665_{3.093}$
	AIC	$99642.703_{4799.050}$	$92828.617_{4459.255}$	$92829.991_{4458.503}$	$92738.984_{4454.168}$	$104887.960_{5183.736}$	$99824.959_{4824.378}$	$126819.415_{1197.312}$
	$H_2(e)$	$6.558_{0.004}$	$6.559_{0.003}$	$6.559_{0.003}$	$6.559_{0.003}$	$6.558_{0.003}$	$6.559_{0.003}$	$6.564_{0.003}$
	Time	$1.555_{0.088}$	$50.089_{2.990}$	$50.089_{2.990}$	$50.089_{2.990}$	$22795.800_{8931.299}$	$1.397_{0.220}$	$2650.516_{85.397}$

values for both $D_{\delta,\hat{\delta}}$ and $H_2(e)$, while BMOMs were associated with the lowest AIC values. This performance is also reflected in Fig. 2, which graphically summarizes $D_{\delta,\hat{\delta}}$ and $H_2(e)$ across model sizes for $\sigma_u = 1.0$. On the other hand, 2*SLS* and $Bayes_{2S}$ required the least time for estimation.

Finally, Table 4 shows the results for SEMs with $\sigma_u = 10$. Under such conditions of high variability, none of the applied methods consistently outperforms the others across the three models, either in terms of $D_{\delta,\hat{\delta}}$ or $H_2(e)$. Specifically, in the largest SEM, the minimum value of $H_2(e)$ is obtained with both BMOMs and $Bmom_{OPT}$ (6.555), followed closely by 2*SLS* and *MCMC* (6.556). In this model, the $Bayes_{2S}$ method did not produce results due to convergence issues. Figure 3 offers a graphical overview of the methods' performance for $\sigma_u = 10$.

Having reviewed the tabular results for the three SEMs, Fig. 4 shows the results for all six models, including the three previously discussed and three additional ones. The figure illustrates the results for SEMs with varying sizes and σ_u values ranging from 0.1 to 10. Each SEM was generated across 30 trials, and the mean standard deviation of the endogenous variables, $\bar{\sigma}y$, was calculated. The color squares in the figure indicate the method that achieves the minimum mean for $D_{\delta,\hat{\delta}}$, AIC, $H_2(e)$, and time.

It is observed that $Bmom_{OPT}$ (green color) achieves the minimum value of $D_{\delta,\hat{\delta}}$ across a significant number of SEMs when $\bar{\sigma}y$ is below 5.4. However, when variability exceeds 5.4, BMOM methods yield the lowest $D_{\delta,\hat{\delta}}$ values.

In terms of AIC, BMOM methods consistently yield the lowest mean values, while the 2*SLS* method also performs well across all $\bar{\sigma}_y$ values.

Regarding $H_2(e)$, $Bmom_{OPT}$ reaches the minimum value for $\bar{\sigma}_y$ values below 4.6, but as variability increases, the method achieving the minimum value changes.

Finally, 2*SLS* and $Bayes_{2S}$ are the methods that require the least time across all the models studied.

Table 5 summarizes the performance of $Bmom_{OPT}$ for $D_{\delta,\hat{\delta}}$ and $H_2(e)$. It categorizes variability into three levels and shows the percentage of SEMs in which $Bmom_{OPT}$ reaches the minimum values for these measures. For example, in the first row, out of 30 SEMs with low variability ($\bar{\sigma}y < 3.00$), *BmomOPT* achieved the minimum $D_{\delta,\hat{\delta}}$ in 96.67% of cases and the minimum $H_2(e)$ in 76.67% of cases. When variability ranges between 3 and 6, the percentages are 88.57% and 58.62%, respectively. For SEMs with high variability, $Bmom_{OPT}$ does not achieve the minimum $D_{\delta,\hat{\delta}}$ compared to other methods but does achieve the minimum $H_2(e)$ in 41.18% of cases.

The results indicate that the performance of the estimation methods varies with the data's variability. Specifically, $Bmom_{OPT}$ is more effective in intervals of lower variability for $D_{\delta,\hat{\delta}}$, although it still achieves the minimum value

Table 3 Mean and standard deviation of $D_{\delta,\hat{\delta}}$, AIC, $H_2(e)$ and time (seconds) for 30 trials when varying the problem size and the estimation method. $\sigma_u = 1.0$

(m, k, n)	Measure	2SLS	BMOM			$Bmom_{OPT}$	$Bayes_{2S}$	MCMC
			B_{gf}	B_{pe}	B_{ta}			
(10, 20, 100)	$D_{\delta,\hat{\delta}}$	$64.243_{4.592}$	$65.812_{3.864}$	$65.711_{3.880}$	$65.826_{3.868}$	$62.303_{4.772}$	$63.981_{5.653}$	$84.628_{43.130}$
	AIC	$3389.415_{905.619}$	$3245.597_{914.478}$	$3230.032_{916.655}$	$3233.181_{916.539}$	$3894.438_{938.251}$	$3479.163_{940.086}$	$4904.369_{430.502}$
	$H_2(e)$	$4.058_{0.010}$	$4.069_{0.011}$	$4.071_{0.014}$	$4.071_{0.014}$	$4.054_{0.014}$	$4.063_{0.017}$	$4.083_{0.016}$
	Time	$0.047_{0.011}$	$0.712_{0.238}$	$0.712_{0.238}$	$0.712_{0.238}$	327.569	$0.054_{124.536}$	$292.828_{16.768}$
(10, 40, 400)	$D_{\delta,\hat{\delta}}$	$75.404_{7.037}$	$81.800_{5.701}$	$81.998_{5.707}$	$82.094_{5.678}$	$64.009_{10.438}$	$76.075_{7.571}$	$94.743_{13.298}$
	AIC	$18755.204_{3414.020}$	$17417.064_{3297.987}$	$17400.546_{3296.650}$	$17391.812_{3296.950}$	$22630.862_{4240.619}$	$18940.266_{3388.913}$	$22764.141_{1110.199}$
	$H_2(e)$	$5.577_{0.005}$	$5.581_{0.005}$	$5.581_{0.005}$	$5.582_{0.005}$	$5.575_{0.006}$	$5.579_{0.007}$	$5.596_{0.008}$
	Time	$0.242_{0.045}$	$4.538_{0.863}$	$4.538_{0.863}$	$4.538_{0.863}$	$3968.580_{1490.813}$	$0.227_{0.041}$	$514.619_{42.887}$
(20, 100, 1000)	$D_{\delta,\hat{\delta}}$	$180.287_{6.136}$	$190.691_{4.463}$	$190.646_{4.466}$	$190.735_{4.451}$	$159.196_{11.157}$	$180.287_{6.852}$	$202.415_{6.844}$
	AIC	$97999.498_{10710.142}$	$92340.942_{10683.055}$	$92345.481_{10683.908}$	$92318.846_{10682.283}$	$118914.540_{13403.792}$	$98592.443_{11085.038}$	$129176.009_{1934.991}$
	$H_2(e)$	$6.549_{0.003}$	$6.553_{0.003}$	$6.553_{0.003}$	$6.553_{0.003}$	$6.546_{0.003}$	$6.549_{0.004}$	$6.563_{0.003}$
	Time	$1.632_{0.113}$	$51.529_{3.407}$	$51.529_{3.407}$	$51.529_{3.407}$	$34335.860_{10226.573}$	$1.643_{0.913}$	$2652.106_{77.169}$

Table 4 Mean and standard deviation of $D_{\delta,\hat{\delta}}$, AIC, $H_2(e)$ and time (seconds) for 30 trials when varying the problem size and the estimation method. $\sigma_u = 10$

(m, k, n)	Measure	2SLS	BMOM			$Bmom_{OPT}$	$Bayes_{2S}$	MCMC
			B_{gf}	B_{pe}	B_{ta}			
(10, 20, 100)	$D_{\delta,\hat{\delta}}$	$83.084_{12.748}$	$76.477_{6.039}$	$76.643_{6.184}$	$76.699_{6.279}$	$92.806_{22.442}$	$82.445_{8.761}$	$91.919_{14.065}$
	AIC	$4489.501_{202.329}$	$4533.026_{177.548}$	$4480.348_{172.137}$	$4483.763_{170.444}$	$4859.169_{689.871}$	$4526.854_{224.714}$	$5003.599_{140.176}$
	$H_2(e)$	$4.085_{0.013}$	$4.083_{0.012}$	$4.080_{0.015}$	$4.080_{0.015}$	$4.087_{0.014}$	$4.085_{0.014}$	$4.086_{0.012}$
	Time	$0.046_{0.013}$	$0.736_{0.217}$	$0.736_{0.217}$	$0.736_{0.217}$	$354.003_{276.866}$	$0.051_{0.014}$	$290.513_{22.513}$
(10, 40, 400)	$D_{\delta,\hat{\delta}}$	$91.457_{5.470}$	$91.628_{5.126}$	$91.137_{5.203}$	$91.149_{5.206}$	$104.975_{17.365}$	$91.409_{5.437}$	$100.542_{11.608}$
	AIC	$22961.128_{805.790}$	$23234.376_{795.089}$	$23181.079_{795.44}$	$23185.416_{793.832}$	$26358.952_{3440.265}$	$22979.676_{831.064}$	$24295.830_{326.455}$
	$H_2(e)$	$5.573_{0.004}$	$5.576_{0.004}$	$5.575_{0.004}$	$5.575_{0.004}$	$5.571_{0.006}$	$5.574_{0.004}$	$5.576_{0.006}$
	Time	$0.226_{0.022}$	$4.364_{0.800}$	$4.364_{0.800}$	$4.364_{0.800}$	$2941.702_{946.765}$	$0.204_{0.023}$	$498.964_{37.740}$
(20, 100, 1000)	$D_{\delta,\hat{\delta}}$	$201.168_{2.378}$	$201.138_{2.521}$	$201.100_{2.560}$	$201.108_{2.565}$	$204.053_{6.275}$	*N/A*	$217.002_{16.618}$
	AIC	$130375.200_{5522.081}$	$131510.812_{5740.622}$	$131449.818_{5731.829}$	$131466.196_{5731.970}$	$134941.445_{10868.319}$	*N/A*	$139485.114_{740.127}$
	$H_2(e)$	$6.556_{0.003}$	$6.555_{0.002}$	$6.555_{0.002}$	$6.555_{0.002}$	$6.555_{0.003}$	*N/A*	$6.556_{0.003}$
	Time	$1.633_{0.225}$	$50.813_{3.104}$	$50.813_{3.104}$	$50.813_{3.104}$	$22192.915_{14299.055}$	*N/A*	$2634.509_{119.424}$

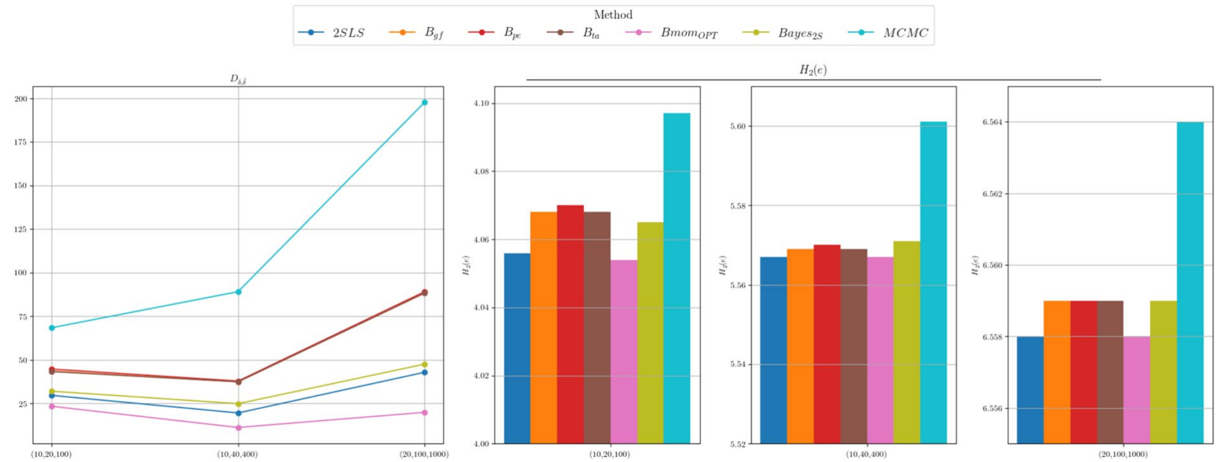

Fig. 1 Graphical summary of $D_{\delta,\hat{\delta}}$ and $H_2(e)$ for the SEMs in Table 2 ($\sigma_u = 0.1$)

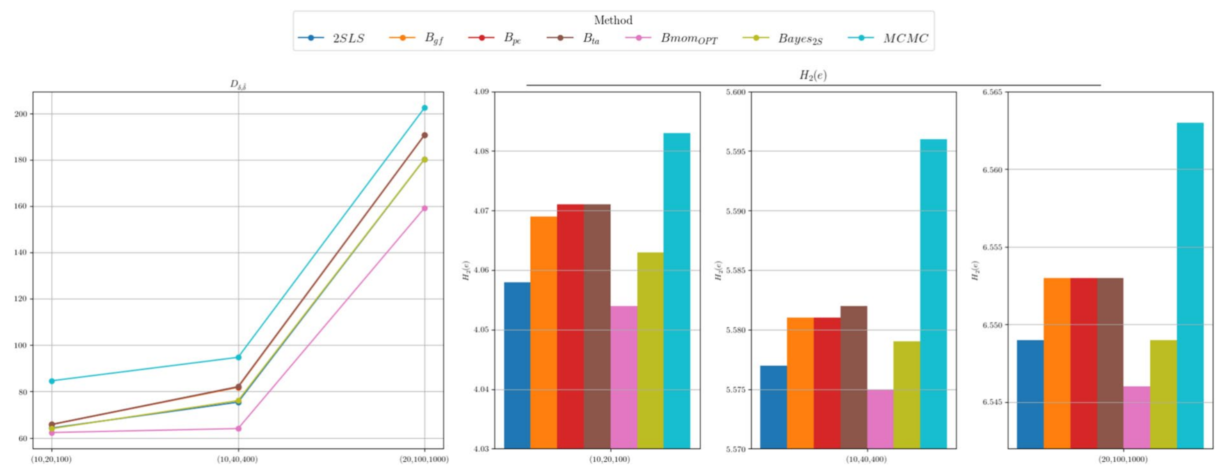

Fig. 2 Graphical summary of $D_{\delta,\hat{\delta}}$ and $H_2(e)$ for the SEMs in Table 3 ($\sigma_u = 1.0$)

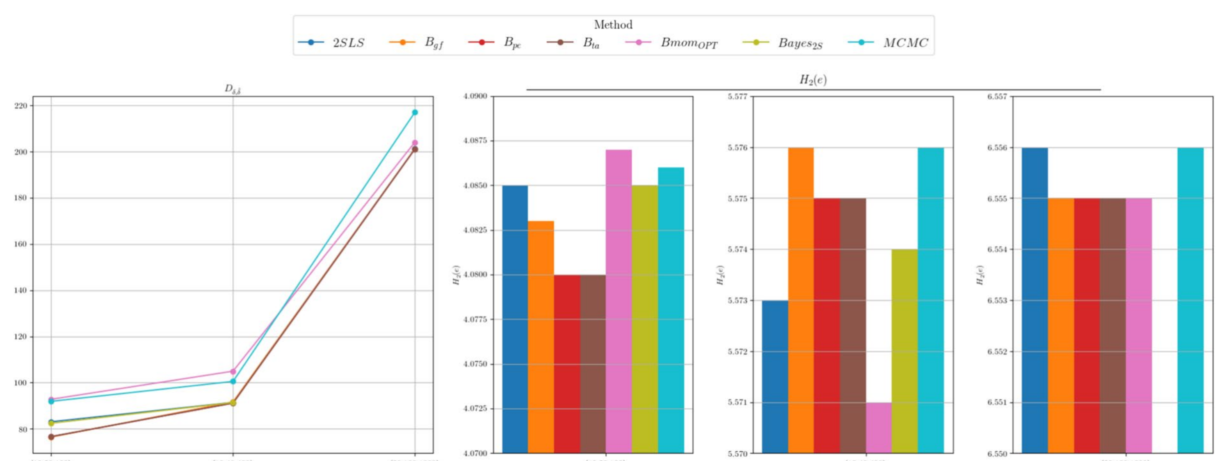

Fig. 3 Graphical summary of $D_{\delta,\hat{\delta}}$ and $H_2(e)$ for the SEMs in Table 4 ($\sigma_u = 10$)

for $H_2(e)$ in some cases of high variability. Regarding the scalability of $Bmom_{OPT}$ to very large datasets, it is worth noting that the method exhibits superior performance as the sample size increases, particularly in terms of estimation accuracy. However, the choice of method ultimately depends on the context and whether the priority lies in precision or

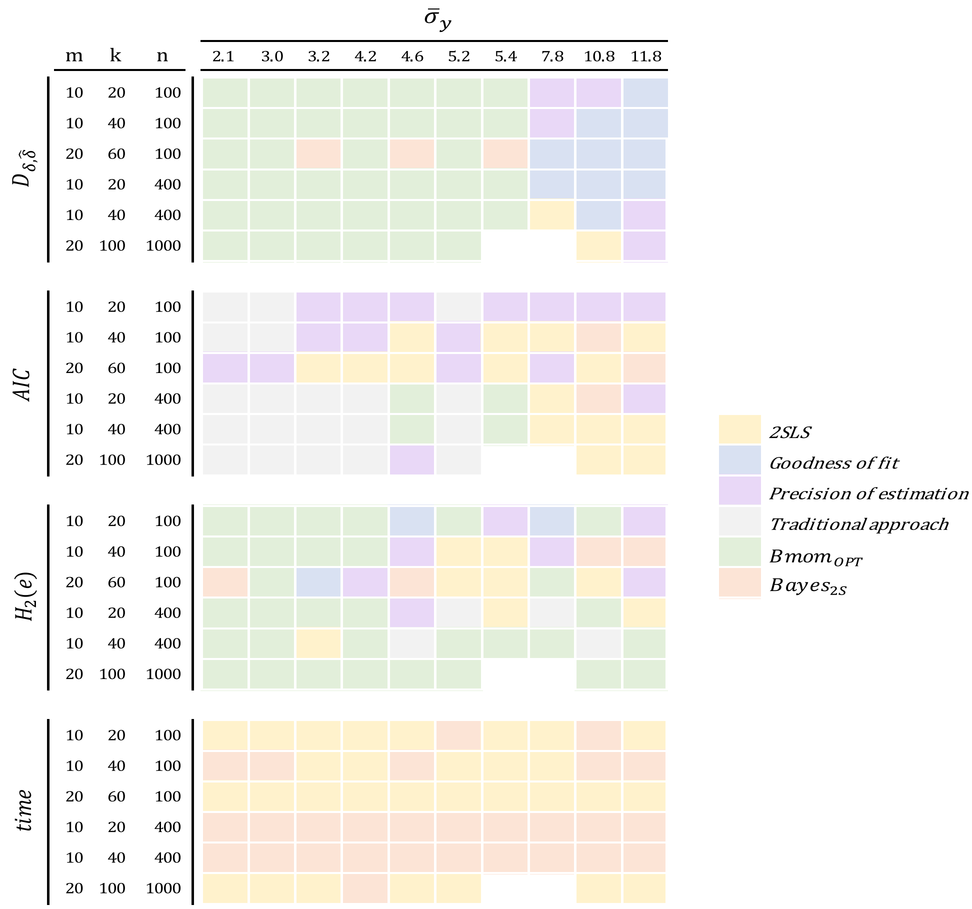

Fig. 4 Minimum mean values of $D_{\delta,\hat{\delta}}$, AIC, $H_2(e)$, and time represented by colors across 30 trials

Table 5 Percentage of minimum values for $D_{\delta,\hat{\delta}}$ and $H_2(e)$ achieved by $Bmom_{OPT}$ by variability levels

$\bar{\sigma}y$	SEMs	$D_{\delta,\hat{\delta}}$	$H_2(e)$
1.00 - 2.99	30	96.67%	76.67%
3.00 - 6.00	35	88.57%	58.62%
>6.00	17	0.00%	41.18%

computational efficiency. For instance, in medical applications, where precision is often more critical than estimation time, $Bmom_{OPT}$ might offer a favorable trade-off despite its higher computational cost.

Conversely, the $2SLS$ method and the BMOMs B_{ta} and B_{pe} methods attain the lowest values for AIC in many cases. However, this does not align with the methods that achieve a lower $H_2(e)$, which suggests that these methods may produce estimates closer to the true model parameters. Despite this, a correspondence is observed between $D_{\delta,\hat{\delta}}$ and $H_2(e)$ for $Bmom_{OPT}$, particularly at variability levels below 4.6.

These findings could be valuable in real-world SEM applications, helping to identify the most appropriate estimation method based on data variability. Additionally, they offer a potential criterion for selecting the method with the lowest prediction error.

5 Conclusions

In research, the choice of an appropriate estimation method depends on factors such as the number of variables, the data size, and the variability within the data. This study compares various estimation methods for SEMs, including $2SLS$, BMOM approaches, $Bmom_{OPT}$, $Bayes_{2S}$, and MCMC, focusing on how data variability influences their performance. The methods were evaluated using measures such as $D_{\delta,\hat{\delta}}$, $H_2(e)$, AIC, and execution time.

The findings reveal that data variability significantly impacts method performance. Specifically, $Bmom_{OPT}$ performs well at low variability intervals for $D_{\delta,\hat{\delta}}$ but requires the most execution time. It also shows promising results for $H_2(e)$, even at higher variability levels. A strong correlation between $D_{\delta,\hat{\delta}}$ and $H_2(e)$ suggests that $H_2(e)$ could be a valuable criterion for selecting an estimation method.

Aware of the limitations of our work, future research will focus on expanding the range of model configurations and complexity to improve the generalizability of the results. This includes exploring the influence of different parameter settings on the performance of the evaluated methods–particularly Bayesian techniques and BMOM-based approaches. A limitation of the present study is the absence of practical validation on real-world datasets. Such validation would require identifying and modeling problems with varying levels of variability, applying the compared methods, and analyzing the outcomes in terms of accuracy and efficiency. Given the scope and theoretical focus of this research, we deliberately restricted the analysis to controlled scenarios to establish a clear framework. We regard this as a necessary first step before addressing the complexity of real-world data, which will be the subject of future work. Real-world applications, such as those in health sciences or economics, will also be considered as illustrative contexts. The findings presented here offer a valuable comparative foundation that can help future researchers select appropriate estimation methods depending on the nature and size of the problem.

Acknowledgements Not applicable.

Author Contributions B.P-S. conceived and designed the experiments, performed the experiments, performed the computation work, analyzed the data, prepared figures and/or tables, and approved the final draft. C.P. performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft. M.G. performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft. JJ.L-E. conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Funding This research was also partially supported by the CIPROM/2024/34 grant, funded by the Conselleria de Educación, Cultura, Universidades y Empleo, Generalitat Valenciana.

Data Availability https://github.com/mbps-umh/variability

Declarations

Competing interests The authors declare that they have no competing interests.

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Gujarati DN, Porter DC (2004) Econometría. McGraw-Hill Interamericana Editores SA, México DF, 5th edn
2. Greene WH (1998) Econometric Analysis. Prentice Hall, 3rd edn
3. Zellner A (1998) The finite sample properties of simultaneous equations' estimates and estimators bayesian and non-bayesian approaches. J Econom 83(1):185–212
4. Pérez-Sánchez B, González M, Perea C, López-Espín JJ (2021) A new computational method for estimating simultaneous equations models using entropy as a parameter criteria. Mathematics 9(7):700
5. Chao JC, Phillips PCB (2002) Jeffreys prior analysis of the simultaneous equations model in the case with n+1 endogenous variables. J Econom 111(2):251–283 (**Finite Sample and Asymptotic Methods in Econometrics**)
6. Geweke J (1996) Bayesian reduced rank regression in econometrics. J Econom 75(1):121–146
7. Kleibergen F, van Dijk HK (1998) Bayesian simultaneous equations analysis using reduced rank structures. Econom Theory 14(6):701–743
8. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2015) Bayesian Data Analysis. Chapman and Hall/CRC, 3rd edn
9. Siburian ME (2020) Fiscal decentralization and regional income inequality: evidence from Indonesia. Appl Econ Lett 27(17):1383–1386
10. Steele F, French R, Bartley M (2013) Adjusting for selection bias in longitudinal analysis using simultaneous equations modeling. The relationship between employment transitions and mental health. Epidemiology 24:703–711 (**Cambridge, Mass.**)
11. Zhao H, Lin B (2020) Impact of foreign trade on energy efficiency in China's textile industry. J Clean Prod 245:118878
12. Guadalajara Olmeda MN, Barrachina Martínez I (2014) Application of simultaneous equation models to temporary disability prescriptions in primary healthcare centres. Int J Comput Math 91(2):252–260
13. Bai S, Yin Y, Chen S (2023) The impact of physical activity on electronic media use among Chinese adolescents and urban-rural differences. BMC Public Health 23(1):1264
14. Strathe AB, Jørgensen H, Kebreab E, Danfær A (2012) Bayesian simultaneous equation models for the analysis of energy intake and partitioning in growing pigs. J Agric Sci 150:764–774
15. Zhou W, Duan W (2015) An empirical study of how third-party websites influence the feedback mechanism between online word-of-mouth and retail sales. Decis Support Syst 76:14–23
16. Park HM, Manchanda P (2015) When harry bet with sally: an empirical analysis of multiple peer effects in casino gambling behavior. Mark Sci 34(2):179–194
17. Chen CM, Chang KL, Lin L, Lee JL (2013) Health risk perception and betel chewing behavior. The evidence from Taiwan. Addict Behav 38:2714–2717
18. Cohen DPA, Benadjaoud MA, Lestaevel P, Lebsir D, Benderitter M, Souidi M (2019) Effects of repetitive iodine thyroid blocking on the development of the foetal brain and thyroid in rats: a systems biology approach. bioRxiv. https://doi.org/10.1101/710764
19. Akaike H (1998) Information theory and an extension of the maximum likelihood principle. Selected Papers of Hirotugu Akaike. Springer Series in Statistics. Springer, New York, NY, pp 199–213
20. Keerativibool W (2015) New criteria for selection in simultaneous equations model. Thailand Stat 10(2):163–181
21. Hurvich CM, Tsai CL (1989) Regression and time series model selection in small samples. Biometrika 76:297–397

22. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464
23. Hannan EJ, Quinn BG (1979) The determination of the order of an autoregression. J Roy Stat Soc: Ser B (Methodol) 41(2):190–195
24. Keerativibool W, Jitthavech J (2015) Model selection criterion based on Kullback-Leibler's symmetric divergence for simultaneous equations model. Chiang Mai J Sci 42(3):761–773
25. Cover TM, Thomas JA (1991) Elements of Information Theory. John Wiley and Sons, Inc
26. Zhou R, Cai R, Tong G (2013) Applications of entropy in finance: a review. Entropy 15(11):4909–4931
27. Cui H, Sivakumar B, Singh VP (2019) Entropy Applications in Environmental and Water Engineering. Academic Press, New York
28. Purvis B, Mao Y, Robinson D (2019) Entropy and its application to urban systems. Entropy 21(1):56
29. Oruç ÖE, Kuruoglu E, Gündüz A (2011) Entropy applications for customer satisfaction survey in information theory. Front Sci 1(1):1–4
30. Gorobets A (2005) The optimal prediction simultaneous equations selection. Econ Bull 36(3):1–8