Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v11n2p577

A hybrid method combining the PLS and the Bayesian approaches to estimate the Structural Equation Models

By Ouazza, Rhomari, Zarrouk

Published: 14 October 2018

This work is copyrighted by Università del Salento, and is licensed under a `Creative Commons Attribuzione - Non commerciale - Non opere derivate 3.0 Italia License`.
For more information see:
`http://creativecommons.org/licenses/by-nc-nd/3.0/it/`

Electronic Journal of Applied Statistical Analysis
Vol. 11, Issue 02, October 2018, 577-607
DOI: 10.1285/i20705948v11n2p577

A hybrid method combining the PLS and the Bayesian approaches to estimate the Structural Equation Models

Ouazza Ahmed *[a], Rhomari Noureddine[a], and Zarrouk Zoubir[b]

[a] *Faculty of Sciences, Mohamed First University, Oujda, Morocco*
[b] *Faculty of Legal, Economic and Social Sciences, Mohamed First University, Oujda, Morocco*

Published: 14 October 2018

The purpose of this paper is to provide a semi parametric approach namely hybrid method combining the Partial Least Squares and the Bayesian approaches to estimate the Structural Equation Models. The aim advantage of this new method is to overcome the assumption of normality that is required in Bayesian approach. The results obtained from an application on simulated and on real data show that our proposed method outperforms both PLS and Bayesian approaches in terms of standard errors.

keywords: Structural Equation Models, Partial Least Squares approach, Bayesian approach, Hybrid method.

1 Introduction

Structural Equation Models (SEMs) are growing family of statistical methods for modeling the relations between endogenous and exogenous variables . These variables can be observed or unobserved (latent) (Hoyle, 2012). Thus, SEM is a collection of statistical techniques that provide a set of tools for researchers in social sciences, behavioral and other disciplines (Yanuar, 2014). The traditional methods in estimating SEM parameters can be divided into three families: parametric, nonparametric and semiparametric methods.

Concerning the parametric family methods, we find the LISREL (LInear Structural RELationships) (Jöreskog, 1970), which is considered the first method used in the context

*Corresponding author: ouazzaahmed@gmail.com

©Università del Salento
ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

of estimating SEM parameters. It is based essentially on the estimation of the covariance matrix using often the likelihood function that imposes the assumption of normality and the independence of observations. This method requires a large sample size (Reinartz et al., 2009) and a non complex model. Further, the model identification is always problematic.

Another estimation method in this family is the Bayesian approach (Lee and Song, 2004; Dunson, 2000) that is dedicated to the explanation of the model. This approach is based on the prior and posterior distributions in the estimation phase, and it has several advantages such as: use of prior information that is available in the data, can give more reliable results for small samples and also for complex models (Lee and Song, 2004). However, this method requires the fixation of some parameters to achieve the goal of model identification. Furthermore, the latent variables are assumed to be independent and normally distributed.

For the non parametric methods, we refer to the Partial Least Squares (PLS) approach (Wold, 1982, 1985) which has been considered as an approach oriented toward prediction (Jöreskog and Wold, 1982). Unlike the parametric methods, PLS requires no assumption on the distribution of data and on model identification. Additionally, the PLS can be a powerful estimation method in case of small samples and models with high complexity (Sarstedt et al., 2014). A review of applications of PLS approach is given by Hair et al. (2012) and Nitzl (2016).

To overcome some assumptions required in parametric methods, different semiparametric approaches have been presented in the literature. We cite for example the Generalized Maximum Entropy (GME) for the SEM proposed by Ciavolino and Al-Nasser (2009), this method is oriented to maximize the expected self information using the maximum entropy. For more details on the GME method and its applications, see Papalia and Ciavolino (2011), Ciavolino et al. (2015), Ciavolino and Dahlgaard (2009) and Ciavolino and Carpita (2015). Another approach, in this context, proposed by Yang and Dunson (2010) namely Bayesian semiparametric SEM, based on the Dirichlet Process. The aim advantage of this method is that it allows the latent variables to have unknown distributions.

In this paper, we will present another solution to the normality problem of latent variables by constructing a semiparametric approach based on combination of the PLS and the Bayesian approaches.

In the following, we will first present the basic concepts of SEM. Then, we recall briefly the PLS and Bayesian methods and we present a new hybrid method to estimate the SEMs. We also highlight its advantages compared with the classical Bayesian one. Finally, an application on simulated and on real data will be discussed. The proofs of posterior distributions used in the proposed method are reported in the Appendix.

2 Basic concepts of SEM

Let J blocks of manifest variables (MVs) decomposed into $\boldsymbol{X_k^i} = (x_{k1}^i, ..., x_{kp_k}^i), k = 1, ..., q$ and $\boldsymbol{Y_j^i} = (y_{j1}^i, ..., y_{jp_j}^i), j = q+1, ..., J$ with $i = 1, ..., n$ and n is the sample size, where $\boldsymbol{X_k^i}$ constitutes the observable expression of a standardized exogenous latent variable (LV) $\boldsymbol{\xi_k} = (\xi_k^1, \xi_k^2, ..., \xi_k^n), k = 1, ..., q$ and $\boldsymbol{Y_j^i}$ constitutes the observable expression of a standardized endogenous LV $\boldsymbol{\eta_j} = (\eta_j^1, \eta_j^2, ..., \eta_j^n), j = q+1, ..., J$.

In the following, we denote by:

- $p = p_1 + p_2 + ... + p_q + p_{q+1} + ... + p_J$ the number of MVs,
- $\boldsymbol{Z^i = [X_1^i, X_2^i, ..., X_q^i, Y_{q+1}^i, Y_{q+2}^i, ..., Y_J^i]'}$, $i = 1, ..., n$ the matrix of observed data,
- $\begin{aligned}\boldsymbol{\Omega} &= \boldsymbol{[\xi_1', \xi_2', ..., \xi_q', \eta_{q+1}', \eta_{q+2}', ..., \eta_J']'} \\ &= \boldsymbol{[\omega_1, \omega_2, ..., \omega_n]}\end{aligned}$ the set of LVs in the SEM,

where $\boldsymbol{\omega_i} = (\xi_1^i, \xi_2^i, ..., \xi_q^i, \eta_{q+1}^i, ..., \eta_J^i)'$, $i = 1, ..., n$.

Commonly, the SEM is presented by two submodels, the measurement (outer) and the structural (inner) models.

2.1 Measurement model

The measurement model is the part of the SEM that examines relationship between the LVs and their MVs.

We will adopt two different formulas for this model, (1) and (2); one will be used for the PLS approach and the other for the Bayesian one.

Each block of MVs is related to its LV by a simple regression equation as following,

$$\begin{cases} \boldsymbol{(X_k^i)'} = \boldsymbol{\Lambda_k^x}\xi_k^i + \boldsymbol{\varepsilon_k^i} \\ k = 1, \cdots, q \ , \ \ i = 1, \cdots, n \end{cases} \text{and} \begin{cases} \boldsymbol{(Y_j^i)'} = \boldsymbol{\Lambda_j^y}\eta_j^i + \boldsymbol{\varepsilon_j^i} \\ j = q+1, \cdots, J \ , \ \ i = 1, \cdots, n \end{cases} \tag{1}$$

where

- $\boldsymbol{\Lambda_k^x} = (\lambda_{k1}^x, \lambda_{k2}^x, ..., \lambda_{kp_k}^x)'$, $k = 1, \cdots, q$ and $\boldsymbol{\Lambda_j^y} = (\lambda_{j1}^y, \lambda_{j2}^y, ..., \lambda_{jp_j}^y)'$, $j = q+1, \cdots, J$ are the vectors of regression parameters (loadings),
- $\boldsymbol{\varepsilon_k^i}$ and $\boldsymbol{\varepsilon_j^i}$ are the error terms assumed to be uncorrelated with $\boldsymbol{\xi_k}$ and $\boldsymbol{\eta_j}$ respectively, and are normally distributed.

The formulas (1) can be presented as follows,

$$\boldsymbol{Z^i} = \boldsymbol{\Lambda\omega_i} + \boldsymbol{\varepsilon_i} \ \ , i = 1, ..., n \ , \tag{2}$$

where

- $\boldsymbol{Z^i}$ is a $(p \times 1)$ observed random vector,
- $\boldsymbol{\Lambda}$ is a $(p \times J)$ factor loading matrix,
- $\boldsymbol{\omega_i}$ is a $(J \times 1)$ random vector of LVs,
- $\boldsymbol{\varepsilon_i}$ is a $(p \times 1)$ random vector of the measurement errors which is independent of $\boldsymbol{\omega_i}$,
- It is assumed that $\boldsymbol{\varepsilon_i} \sim \mathcal{N}(0, \boldsymbol{\Psi_\varepsilon})$.

In the measurement model, we should assess the reliability of any constructs. Three reliability measures are used in this paper: Cronbach's alpha, Composite reliability and Average variance extracted; for more details see, for example, Fornell and Larcker (1981) and Chin (1998).

The formulas of these measures are presented in the following only for the MVs associated with exogenous LVs. The same procedure is applied to calculate these coefficients for the MVs associated with endogenous LVs.

- ***Cronbach's alpha***

For each block of standardized MVs, the formula of Cronbach's alpha can be written as,

$$\begin{cases} \alpha_k = \dfrac{\sum\limits_{h \neq h'} cor(\boldsymbol{x_{kh}},\ \boldsymbol{x_{kh'}})}{p_k + \sum\limits_{h \neq h'} cor(\boldsymbol{x_{kh}},\ \boldsymbol{x_{kh'}})} \dfrac{p_k}{p_k - 1} \\ k = 1, \cdots, q\ ,\ \ h, h' \in \{1, 2, ..., p_k\} \end{cases} \tag{3}$$

- ***Composite Reliability***

Let us suppose that all the MVs and the LVs are standardized, the Composite Reliability CR for each block is then defined by,

$$\begin{cases} CR_k = \dfrac{(\sum\limits_{h=1}^{p_k} \lambda_{kh}^x)^2}{(\sum\limits_{h=1}^{p_k} \lambda_{kh}^x)^2 + \sum\limits_{h=1}^{p_k} var(\boldsymbol{\varepsilon_{kh}})} \\ k = 1, \cdots, q \end{cases} \tag{4}$$

- ***Average Variance Extracted***

Assuming standardized manifest and latent variables, the Average Variance Extracted AVE for each block is calculated as follows,

$$\begin{cases} AVE_k = \dfrac{\sum\limits_{h=1}^{p_k} (\lambda_{kh}^x)^2}{\sum\limits_{h=1}^{p_k} (\lambda_{kh}^x)^2 + \sum\limits_{h=1}^{p_k} var(\boldsymbol{\varepsilon_{kh}})} \\ k = 1, \cdots, q \end{cases} \tag{5}$$

2.2 Structural model

The structural model allows to examine the relationships between endogenous and exogenous LVs by the following equation,

$$\boldsymbol{\eta^i} = \boldsymbol{B\eta^i} + \boldsymbol{\Gamma\xi^i} + \boldsymbol{\delta^i} \quad \text{or} \quad \boldsymbol{\eta^i} = \boldsymbol{\Lambda_\omega \omega_i} + \boldsymbol{\delta^i} \quad , \quad i = 1, ..., n \; , \tag{6}$$

where

- $\boldsymbol{\xi^i} = (\xi_1^i, ..., \xi_q^i)^{'}$, $\boldsymbol{\eta^i} = (\eta_{q+1}^i, ..., \eta_J^i)^{'}$, $i = 1, ..., n$,
- $\boldsymbol{\Lambda_\omega} = \boldsymbol{(B, \Gamma)}$,
- $\boldsymbol{\omega_i} = \boldsymbol{(\xi^i, \eta^i)'} = (\xi_1^i, \xi_2^i, ..., \xi_q^i, \eta_{q+1}^i, ..., \eta_J^i)^{'}$,
- $\boldsymbol{B}$ is a $((J-q) \times (J-q))$ matrix of structural parameters governing the relationship among the endogenous LVs,
- $\boldsymbol{\Gamma}$ is $((J-q) \times q)$ a regression parameter matrix relating the endogenous LVs and exogenous LVs,
- $\boldsymbol{\delta^i} \sim \mathcal{N}(0, \boldsymbol{\Psi_\delta})$ is a vector of the measurement errors.

In the following sections, we are interested in the estimation of LVs $\boldsymbol{\Omega}$ and the unknowns parameters $\boldsymbol{\theta} = \boldsymbol{(\Lambda, \Lambda_\omega, \Psi_\varepsilon, \Psi_\delta)}$.

3 PLS approch

PLS approach is one of the tools used to estimate the parameters of Structural Equation Models, which was proposed by Wold (1982, 1985). A detailed theoretical presentation of this approach is given by Tenenhaus et al. (2005).

The PLS model is decomposed into two submodels, measurement and structural models as defined in (1) and (6) respectively. The estimation of LVs $\boldsymbol{\Omega}$ contained in these models is based on the following iterative PLS algorithm composed by two steps:

•***Step 1: outer estimate of LVs.***

The standardized LVs $\boldsymbol{\xi_k}$ and $\boldsymbol{\eta_j}$ are estimated as linear combinations of their centered MVs, as follows:

$$\begin{cases} \widehat{\xi_k^i}^{out} = \boldsymbol{X_k^i \pi_k} \\ k = 1, \cdots, q \;\; , \;\; i = 1, \cdots, n \end{cases} \quad \text{and} \quad \begin{cases} \widehat{\eta_j^i}^{out} = \boldsymbol{Y_j^i \pi_j} \\ j = q+1, \cdots, J \;\; , \;\; i = 1, \cdots, n \end{cases} \tag{7}$$

where $\boldsymbol{\pi_k} = (\pi_{k1}, \pi_{k2}, ..., \pi_{kp_k})^{'}$, $k = 1, \cdots, q$ and $\boldsymbol{\pi_j} = (\pi_{j1}, \pi_{j2}, ..., \pi_{jp_j})^{'}$, $j = q+1, \cdots, J$ are the external weights.

The external weights $\boldsymbol{\pi_k}$ and $\boldsymbol{\pi_j}$ require initial values and to update these weights we use either Mode A or Mode B as described follows, Mode A:

$$\begin{cases} \pi_{kl} = Cov(\boldsymbol{x_{kl}}, \widehat{\boldsymbol{\xi_k}}^{int}) \\ k = 1, \cdots, q\ , l = 1, ..., p_k \end{cases} \quad \text{and} \quad \begin{cases} \pi_{jh} = Cov(\boldsymbol{y_{jh}}, \widehat{\boldsymbol{\eta_j}}^{int}) \\ j = q+1, \cdots, J\ , h = 1, ..., p_j \end{cases} \tag{8}$$

Mode B:

$$\begin{cases} \boldsymbol{\pi_k = (X_k'X_k)^{-1}X_k'}\widehat{\boldsymbol{\xi_k}}^{int} \\ k = 1, \cdots, q \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{\pi_j = (Y_j'Y_j)^{-1}Y_j'}\widehat{\boldsymbol{\eta_j}}^{int} \\ j = q+1, \cdots, J \end{cases} \tag{9}$$

where $\widehat{\boldsymbol{\xi_k}}^{int}$ and $\widehat{\boldsymbol{\eta_j}}^{int}$ are the inner estimates of LVs given below.

Note that the Mode A is appropriate for a block with a reflective measurement model and mode B for a formative one.

- ***Step 2: inner estimate of LVs.***

\- The inner estimate of the standardized LV $\boldsymbol{\xi_k}$ is defined by:

$$\begin{cases} \widehat{\xi_k^i}^{int} = \sum_{j=q+1}^{J} c_{kj} e_{kj} \widehat{\eta_j^i}^{out} \\ k = 1, \cdots, q\ , \quad i = 1, \cdots, n \end{cases} \tag{10}$$

where $c_{kj} = 1$ if $\boldsymbol{\xi_k}$ is connected with $\boldsymbol{\eta_j}$ and $c_{kj} = 0$ otherwise. And e_{kj} is an internal weight.
\- The inner estimate of the standardized LV $\boldsymbol{\eta_j}$ is given by:

$$\begin{cases} \widehat{\eta_j^i}^{int} = \sum_{k=1}^{q} c_{jk} e_{jk} \widehat{\xi_k^i}^{out} + \sum_{h \neq j} c_{jh} e_{jh} \widehat{\eta_h^i}^{out} \\ j, h \in \{q+1, \cdots, J\}\ , \quad i = 1, \cdots, n \end{cases} \tag{11}$$

where $c_{jk} = 1$ if $\boldsymbol{\eta_j}$ is connected with $\boldsymbol{\xi_k}$ and $c_{jk} = 0$ otherwise; and $c_{jh} = 1$ if $\boldsymbol{\eta_j}$ is connected with $\boldsymbol{\eta_h}$ and $c_{jh} = 0$ otherwise. e_{jk} and e_{jh} are internal weights.
The internal weights can be selected using one of the schemes as defined in Tenenhaus et al. (2005), namely centroid scheme, factorial scheme and structural scheme.

Afterwards, the parameters $\boldsymbol{\theta}$ are estimated by simple and multiple regressions.

3.1 Model validation

The SEM, in case of PLS approach, can be validated at three levels: the quality of the measurement model, the quality of the structural model, and the quality of the global model. Usually, these qualities are quantified with the following coefficients.

Note that these coefficients can be calculated also for Bayesian and hybrid approaches presented in the next sections.

- ***Communality***

The communality H^2 (Vinzi et al., 2010) evaluates the quality of the measurement model for each block. For each block, it is defined by the following formula:

$$\begin{cases} H_k^2 = \frac{1}{p_k}\sum_{l=1}^{p_k} cor^2(\boldsymbol{x_{kl}}, \widehat{\boldsymbol{\xi_k}}^{out}) \\ k = 1, \cdots, q \end{cases} \text{ and } \begin{cases} H_j^2 = \frac{1}{p_j}\sum_{h=1}^{p_j} cor^2(\boldsymbol{y_{jh}}, \widehat{\boldsymbol{\eta_j}}^{out}) \\ j = q+1, \cdots, J \end{cases} \tag{12}$$

where $\widehat{\boldsymbol{\xi_k}}^{out} = (\widehat{\xi_k^1}^{out}, \widehat{\xi_k^2}^{out}, ..., \widehat{\xi_k^n}^{out})$ and $\widehat{\boldsymbol{\eta_j}}^{out} = (\widehat{\eta_j^1}^{out}, \widehat{\eta_j^2}^{out}, ..., \widehat{\eta_j^n}^{out})$.

This index represents the variance proportion of MVs explained by their associated LVs.

- ***Redundancy***

The redundancy F^2 (Vinzi et al., 2010) evaluates the quality of the structural model for each endogenous block, taking into account the measurement model. It is defined as:

$$F_j^2 = H_j^2 \times R_j^2 \ , \ \ j = q+1, ..., J \ , \tag{13}$$

where R_j^2 is the coefficient of determination between the endogenous variable $\widehat{\boldsymbol{\eta_j}}^{out}$ and all LVs that explain $\widehat{\boldsymbol{\eta_j}}^{out}$.

- ***GoF index***

A global index of Goodness-of-Fit (GoF)(Tenenhaus et al., 2004), can be proposed as the geometric mean of the average communality ($\overline{H^2}$) and the average of the determination coefficient ($\overline{R^2}$) associated to the endogenous LVs :

$$GoF = \sqrt{\overline{H^2} \times \overline{R^2}} \tag{14}$$

The GoF designed as an index for validating the PLS model globally. According to Wetzels et al. (2009), a PLS model is considered as valid when the GoF is greater than 0.36 .

4 Bayesian approach of SEM

The Bayesian approach of SEM has been proposed , especially using a Markov Chain Monte Carlo (MCMC), by a number of authors; for example, Lee and Song (2003, 2004), Ansari et al. (2000), Dunson (2000), Jedidi and Ansari (2001). In this approach, the set of unknown parameters in the SEM model are regarded as random, and the Bayesian solution is obtained via the analysis of the posterior density of these parameters.

Similarly to the other methods, the Bayesian approach is decomposed into two submodels, measurement and structural models as defined in (2) and (6) respectively. The main hypothesis of the Bayesian approach is to assume that the LVs $\boldsymbol{\omega_i}$ are mutually independent and normally distributed $\mathcal{N}(0, \boldsymbol{\Phi})$. The expression of $\boldsymbol{\theta}$ becomes $\boldsymbol{\theta} = (\boldsymbol{\Lambda}, \boldsymbol{\Lambda_\omega}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Psi_\delta}, \boldsymbol{\Phi})$.

In the Bayesian approach procedure, all LVs $\boldsymbol{\Omega}$ in the model are treated as hypothetical missing data, then using the posterior analysis on the basis of the augmented complete dataset $(\boldsymbol{Z}, \boldsymbol{\Omega})$ (Lee and Song, 2004), where
$\boldsymbol{Z} = (\boldsymbol{Z^1}, \boldsymbol{Z^2}, ..., \boldsymbol{Z^n})'$.

To obtain the estimation of $\boldsymbol{\theta}$ from Bayesian approach, we apply firstly the Gibbs sampler then we check the convergence of this algorithm, afterward the Bayesian estimates are given by the mean of the posterior density of $\boldsymbol{\theta}$ given $\boldsymbol{Z}$.

4.1 Gibbs sampler

The Gibbs sampler is a Markov Chain algorithm used to generate a sequence of random observations from the joint posterior distribution $p(\boldsymbol{\theta}, \boldsymbol{\Omega}/\boldsymbol{Z})$ of $\boldsymbol{\theta}$ and $\boldsymbol{\Omega}$ given $\boldsymbol{Z}$. It is implemented as follows,

- *Initialization*:
  The algorithm begins with arbitrary initial values $\boldsymbol{\Omega}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$,
- At the $(j+1)^{th}$ iteration, by using the current values $\boldsymbol{\Omega}^{(j)}$ and $\boldsymbol{\theta}^{(j)}$,
  - *Step 1*: Generate a random variable $\boldsymbol{\Omega}^{(j+1)}$ from the conditional distribution $p(\boldsymbol{\Omega}/\boldsymbol{\theta}^{(j)}, \boldsymbol{Z})$
  - *Step 2*: Generate a random variable $\boldsymbol{\theta}^{(j+1)}$ from the conditional distribution $p(\boldsymbol{\theta}/\boldsymbol{\Omega}^{(j+1)}, \boldsymbol{Z})$

In this algorithm, we need to select the prior distribution of $\boldsymbol{\theta}$. In this article, we choose the prior distribution for $\boldsymbol{\theta}$ via the following conjugate type distribution; more details for the prior information are given by Lee (2007). We denote by,

- $\psi_{\varepsilon k}$: the k^{th} diagonal elements of $\boldsymbol{\Psi}_{\boldsymbol{\varepsilon}}$,
- $\boldsymbol{\Lambda}_k'$: the k^{th} row of $\boldsymbol{\Lambda}$,
- $\psi_{\delta k}$: the k^{th} diagonal elements of $\boldsymbol{\Psi}_\delta$ and $\boldsymbol{\Lambda}_{\omega k}'$ be the k^{th} row of $\boldsymbol{\Lambda}_\omega$.

The prior distributions for all parameters in $\boldsymbol{\theta}$ are given by:

$$\begin{cases} \psi_{\varepsilon k}^{-1} \sim \Gamma(\alpha_{0\varepsilon k}, \beta_{0\varepsilon k}) \ , \\ \psi_{\delta k}^{-1} \sim \Gamma(\alpha_{0\omega k}, \beta_{0\omega k}) \ , \\ (\boldsymbol{\Lambda}_k/\psi_{\varepsilon k}) \sim \mathcal{N}(\boldsymbol{\Lambda}_{0k}, \psi_{\varepsilon k}\boldsymbol{H}_{0zk}) \ , \\ (\boldsymbol{\Lambda}_{\omega k}/\psi_{\delta k}) \sim \mathcal{N}(\boldsymbol{\Lambda}_{0\omega k}, \psi_{\delta k}\boldsymbol{H}_{0\omega k}) \ , \\ \boldsymbol{\Phi} \sim \mathcal{W}_q(\boldsymbol{R_0}^{-1}, \rho_0) \ , \end{cases} \tag{15}$$

where $\Gamma(.,.)$, $\mathcal{N}(.,.)$ and $\mathcal{W}(.,.)$ denote respectively Gamma, Normal and Wishart distributions, and $\alpha_{0\varepsilon k}$, $\alpha_{0\omega k}$, $\beta_{0\varepsilon k}$, $\beta_{0\omega k}$, $\boldsymbol{H}_{0zk}$, $\boldsymbol{H}_{0\omega k}$, $\boldsymbol{R}_0$ and ρ_0 are the associated hyperparameters.

The derivation of the conditional distribution that is required in the Gibbs sampler process is discussed by Lee (2007) and Song and Lee (2012). And these posterior distributions are given by:

$$\begin{cases} (\boldsymbol{\omega_i}/\boldsymbol{Z^i}, \boldsymbol{\Lambda}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Phi}) \sim \mathcal{N}((\boldsymbol{\Phi^{-1}} + \boldsymbol{\Lambda}'\boldsymbol{\Psi_\varepsilon^{-1}}\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}'\boldsymbol{\Psi_\varepsilon^{-1}}\boldsymbol{Z^i}, (\boldsymbol{\Phi^{-1}} + \boldsymbol{\Lambda}'\boldsymbol{\Psi_\varepsilon^{-1}}\boldsymbol{\Lambda})^{-1}) , \\ (\boldsymbol{\Phi}/\boldsymbol{Z}, \boldsymbol{\Omega}) \sim \mathcal{IW}(\boldsymbol{\Omega}\boldsymbol{\Omega}' + \boldsymbol{R}_0^{-1}, \rho_0 + n) , \\ (\psi_{\varepsilon k}^{-1}/\boldsymbol{Z}, \boldsymbol{\Omega}) \sim \Gamma(\frac{n}{2} + \alpha_{0\varepsilon k}, \beta_{\varepsilon k}) , \\ (\boldsymbol{\Lambda}_k/\boldsymbol{Z}, \boldsymbol{\Omega}, \psi_{\varepsilon k}^{-1}) \sim \mathcal{N}(\boldsymbol{a}_k, \psi_{\varepsilon k}\boldsymbol{A}_k) , \\ (\psi_{\delta k}^{-1}/\boldsymbol{\Omega}) \sim \Gamma(\frac{n}{2} + \alpha_{0\delta k}, \beta_{\delta k}) , \\ (\boldsymbol{\Lambda}_{\omega k}/\boldsymbol{\Omega}, \psi_{\delta k}^{-1}) \sim \mathcal{N}(\boldsymbol{a}_{\omega k}, \psi_{\delta k}\boldsymbol{A}_{\omega k}) , \end{cases} \tag{16}$$

where

- $\mathcal{IW}$ denotes the inverse Wishart distribution,

- $\boldsymbol{A}_k = (\boldsymbol{H}_{0zk}^{-1} + \boldsymbol{\Omega}\boldsymbol{\Omega}')^{-1}$,

- $\boldsymbol{a}_k = A_k(\boldsymbol{H}_{0zk}^{-1}\boldsymbol{\Lambda}_{0k} + \boldsymbol{\Omega}\boldsymbol{Z}_k)$; (with $\boldsymbol{Z}_k'$ is the k^{th} row of $\boldsymbol{Z}$) ,

- $\beta_{\varepsilon k} = \beta_{0\varepsilon k} + \frac{1}{2}[\boldsymbol{Z}_k'\boldsymbol{Z}_k - \boldsymbol{a}_k'\boldsymbol{A}_k^{-1}\boldsymbol{a}_k + \boldsymbol{\Lambda}_{0k}'\boldsymbol{H}_{0zk}^{-1}\boldsymbol{\Lambda}_{0k}]$,

- $\boldsymbol{A}_{\omega k} = (\boldsymbol{H}_{0\omega k}^{-1} + \boldsymbol{\Omega}\boldsymbol{\Omega}')^{-1}$,

- $\boldsymbol{a}_{\omega k} = \boldsymbol{A}_{\omega k}(\boldsymbol{H}_{0\omega k}^{-1}\boldsymbol{\Lambda}_{0\omega k} + \boldsymbol{\Omega}\boldsymbol{\Omega}_{1k})$,

- $\beta_{\delta k} = \beta_{0\delta k} + \frac{1}{2}[\boldsymbol{\Omega}_{1k}'\boldsymbol{\Omega}_{1k} - \boldsymbol{a}_{\omega k}'\boldsymbol{A}_{\omega k}^{-1}\boldsymbol{a}_{\omega k} + \boldsymbol{\Lambda}_{0\omega k}'\boldsymbol{H}_{0\omega k}^{-1}\boldsymbol{\Lambda}_{0\omega k}]$.

4.2 Convergence of Gibbs sampler

Under mild regularity conditions (Song and Lee, 2012), the joint distribution $p(\boldsymbol{\Omega}^{(j)}, \boldsymbol{\theta}^{(j)}/\boldsymbol{Z})$ of $\boldsymbol{\Omega}^{(j)}$ and $\boldsymbol{\theta}^{(j)}$ given $\boldsymbol{Z}$ converges to the desired posterior distribution $p(\boldsymbol{\theta}, \boldsymbol{\Omega}/\boldsymbol{Z})$ after a sufficiently large number of iterations J (Song and Lee, 2012). This number J can be determined by:

1) Plots of the simulated sequences of the individual parameters. At convergence, parallel sequences generated with different starting values should mix well together(Song and Lee, 2012).

2) Generate several parallel sequences of observations with different starting values. Then, computing the "Estimated Potential Scale Reduction (EPSR)" values. The whole simulation procedure is considered to be converged if all the EPSR values are less than 1.2. For more details on this index, see Lee and Song (2004) and Lee (2007).

4.3 Bayesian estimate of θ

The Bayesian estimate of $\boldsymbol{\theta}$ (or each unknown parameter in $\boldsymbol{\theta}$), denoted by $\widehat{\boldsymbol{\theta}}_B$, is obtained by the mean of the posterior density $p(\boldsymbol{\theta}|\boldsymbol{Z})$ of $\boldsymbol{\theta}$ given $\boldsymbol{Z}$. Theoretically, this estimate could be obtained through integration. However, it is often difficult to do it explicitly. To avoid this problem, monte carlo simulation methods should be appealed. In fact, a sufficiently large number of draws from $p(\boldsymbol{\theta}|\boldsymbol{Z})$ are simulated using Gibbs sampler. After that, the mean and/or other useful statistics can be approximated through the simulated observations (Song and Lee, 2012).
The convergence of the Gibbs sampler allows us to collect T observations $(\boldsymbol{\theta}^{(1)}, \cdots, \boldsymbol{\theta}^{(T)})$. Thereby, the Bayesian estimate of $\boldsymbol{\theta}$ is obtained by calculating the average of these T simulated observations. The same procedure is applied to estimate the set of LVs $\boldsymbol{\Omega}$.

5 A hybrid method for SEM

We present now our proposed hybrid method which combine the two methods previously presented (PLS and Bayesian). The aim advantage is to overcome the assumptions on normality and independence of LVs required in classical Bayesian approach. For this purpose, the PLS method is an efficient alternative especially because it also overcomes the problem of model identification. Otherwise, we take advantage of the fact that the Bayesian method allows the use of prior information. As a consequence of combination of the two methods, our approach provides reliable results for small samples and complex models. Finally, using the Gibbs sampler only for the estimation of unknown parameters provides a gain in terms of algorithm complexity.

The process of this technique can be decomposed into two steps:
- The first step consists to obtain the scores of LVs by PLS technique, so the SEM can be treated as familiar regression models (or simultaneous equation models) without LVs.
- The second step allows to obtain the estimates of unknown parameters by using Bayesian approach.

In the following,

- $\boldsymbol{\xi_k}^{pls} = (\xi_k^{1\,pls}, \xi_k^{2\,pls}, ..., \xi_k^{n\,pls})$, $k = 1, ..., q$ denote the exogenous LVs scores obtained by the PLS approach.
- $\boldsymbol{\eta_j}^{pls} = (\eta_j^{1\,pls}, \eta_j^{2\,pls}, ..., \eta_j^{n\,pls})$, $j = q+1, ..., J$ denote the endogenous LVs scores obtained by the PLS approach.
  We note that $\boldsymbol{\xi_k}^{pls}$ and $\boldsymbol{\eta_j}^{pls}$ are treated as observed variables in this case.
- $\boldsymbol{\Omega}^{pls} = [(\boldsymbol{\xi_1}^{pls})', (\boldsymbol{\xi_2}^{pls})', ..., (\boldsymbol{\xi_q}^{pls})', (\boldsymbol{\eta_{q+1}}^{pls})', (\boldsymbol{\eta_{q+2}}^{pls})', ..., (\boldsymbol{\eta_J}^{pls})']'$
  $= [\boldsymbol{\omega_1}^{pls}, \boldsymbol{\omega_2}^{pls}, ..., \boldsymbol{\omega_n}^{pls}]$ the set of LVs scores,
  where $\boldsymbol{\omega_i}^{pls} = (\xi_1^{i\,pls}, \xi_2^{i\,pls}, ..., \xi_q^{i\,pls}, \eta_{q+1}^{i\,pls}, ..., \eta_J^{i\,pls})'$, $i = 1, ..., n$.

As previous approaches, SEM in the case of the hybrid method can be presented by two submodels:

5.1 Measurement model

The regression equation linking the scores of LVs and their observations is given as follows:

$$\boldsymbol{Z^i} = \boldsymbol{\Lambda\omega_i^{pls}} + \boldsymbol{\varepsilon_i} \ , \ i = 1, ..., n \ , \tag{17}$$

or

$$\boldsymbol{Z} = \boldsymbol{\Sigma\beta} + \boldsymbol{\varepsilon} \ , \tag{18}$$

where

- $\boldsymbol{Z = (Z^1, Z^2, ..., Z^n)'}$ is a $(n \times p)$ matrix of observed data,
- $\boldsymbol{\Sigma} = (\boldsymbol{\Omega}^{pls})'$ is a $(n \times J)$ matrix of LVs scores,
- $\boldsymbol{\beta = \Lambda'}$ is a $(J \times p)$ matrix of regression parameters,
- $\boldsymbol{\varepsilon = (\varepsilon_1, ..., \varepsilon_n)'}$ is a $(n \times p)$ matrix of measurement error, assumed to be $N(0, \boldsymbol{\Psi_\varepsilon})$.

5.2 Structural model

The scores of LVs are linked between them by the following equations:

$$\boldsymbol{\eta^i}^{\ pls} = \boldsymbol{B\eta^i}^{\ pls} + \boldsymbol{\Gamma\xi^i}^{\ pls} + \boldsymbol{\delta^i} \ \text{ or } \ \boldsymbol{\eta^i}^{\ pls} = \boldsymbol{\Lambda_\omega\omega_i}^{pls} + \boldsymbol{\delta^i} \quad , \quad i = 1, ..., n \ , \tag{19}$$

or

$$\boldsymbol{\eta}^{pls} = \boldsymbol{\Sigma\beta_\omega} + \boldsymbol{\delta} \ , \tag{20}$$

where

- $\boldsymbol{\xi^i}^{\ pls} = (\xi_1^{i \ pls}, \xi_2^{i \ pls}, ..., \xi_q^{i \ pls})' \quad , \quad \boldsymbol{\eta^i}^{\ pls} = (\eta_{q+1}^{i \ pls}, ..., \eta_J^{i \ pls})'$,
- $\boldsymbol{\eta}^{pls} = (\boldsymbol{\eta^1}^{\ pls}, \boldsymbol{\eta^2}^{\ pls}, ..., \boldsymbol{\eta^n}^{\ pls})'$ is a $(n \times (J-q))$ matrix of endogenous LVs scores,
- $\boldsymbol{\beta_\omega = \Lambda'_\omega = (B, \Gamma)'}$ is a $(J \times (J-q))$ matrix of regression parameters,
- $\boldsymbol{\delta}$ is a $(n \times (J-q))$ matrix of measurement error, assumed to be $N(0, \boldsymbol{\Psi_\delta})$.

We note, in the following, $\boldsymbol{\theta = (\beta, \beta_\omega, \Psi_\varepsilon, \Psi_\delta)}$ the unknown parameters contained in the SEM. We will estimate these parameters $\boldsymbol{\theta}$ by Bayesian method using Gibbs sampler. In order to simplify the estimation of $\boldsymbol{\theta}$, the equations (18) and (20) will be separately estimated; therefore $\boldsymbol{\theta}$ will be decomposed into: $\boldsymbol{\theta_1 = (\beta, \Psi_\varepsilon)}$ and $\boldsymbol{\theta_2 = (\beta_\omega, \Psi_\delta)}$.

5.3 Estimation of the parameter $\boldsymbol{\theta_1}$ in the regression model (18)

The Bayesian estimate of $\boldsymbol{\theta_1}$ can be obtained by applying the "Gibbs sampler" algorithm, we generate a sufficiently large sample of $\boldsymbol{\theta_1}$ noted $(\boldsymbol{\theta_1}^{(1)}, \boldsymbol{\theta_1}^{(2)}, \cdots, \boldsymbol{\theta_1}^{(T)})$, from the posterior density $p(\boldsymbol{\theta_1}/\boldsymbol{Z}, \boldsymbol{\Sigma})$ of $\boldsymbol{\theta_1}$ given $\boldsymbol{Z}$ and $\boldsymbol{\Sigma}$, as follows,

- *Initialization*:
  The algorithm begins with arbitrary initial values $\boldsymbol{\theta_1}^{(0)} = (\boldsymbol{\beta}^{(0)}, \boldsymbol{\Psi_\varepsilon}^{(0)})$,

- At the $(j+1)^{th}$ iteration, with current values $\boldsymbol{\theta_1}^{(j)} = (\boldsymbol{\beta}^{(j)}, \boldsymbol{\Psi_\varepsilon}^{(j)})$,

  - *Step 1*: Generate a random variable $\boldsymbol{\Psi_\varepsilon}^{(j+1)}$ from the conditional distribution $p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma})$
  - *Step 2*: Generate a random variable $\boldsymbol{\beta}^{(j+1)}$ from the conditional distribution $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}^{(j+1)}, \boldsymbol{Z}, \boldsymbol{\Sigma})$

The conditional distributions $p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma})$ and $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma})$ are required in the implementation of Gibbs sampler.

Note that the distribution of $\boldsymbol{\Sigma}$ is independent of $\boldsymbol{\theta_1}$, i.e $p(\boldsymbol{\Sigma}/\boldsymbol{\theta_1}) = p(\boldsymbol{\Sigma})$. Then, the expression of the posterior distribution $p(\boldsymbol{\theta_1}/\boldsymbol{Z}, \boldsymbol{\Sigma})$ is written as:

$$\begin{aligned} p(\boldsymbol{\theta_1}/\boldsymbol{Z}, \boldsymbol{\Sigma}) &= p(\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma}) \\ &\propto p(\boldsymbol{Z}/\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Sigma}) p(\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}) \\ &= p(\boldsymbol{Z}/\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Sigma}) p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}) p(\boldsymbol{\Psi_\varepsilon}) . \end{aligned} \tag{21}$$

The symbol "$\propto$" denotes proportional, and the density $p(\boldsymbol{Z}/\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Sigma})$ represents the likelihood of $\boldsymbol{Z}$. So,

$$p(\boldsymbol{Z}/\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{n}{2}} exp\{-\frac{1}{2} Trace[(\boldsymbol{Z} - \boldsymbol{\Sigma}\boldsymbol{\beta})'(\boldsymbol{Z} - \boldsymbol{\Sigma}\boldsymbol{\beta})\boldsymbol{\Psi_\varepsilon}^{-1}]\} . \tag{22}$$

Therefore, to derive $p(\boldsymbol{\theta_1}/\boldsymbol{Z}, \boldsymbol{\Sigma})$, we need the prior distribution $p(\boldsymbol{\Psi_\varepsilon})$ and $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon})$. For that reason, we suggest the following conjugate prior distribution for $\boldsymbol{\Psi_\varepsilon}$ and $\boldsymbol{\beta}$:

$$\begin{cases} \boldsymbol{\Psi_\varepsilon} \sim \mathcal{IW}(\boldsymbol{T}_0, \rho_0) , \\ \text{vect}(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}) \sim \mathcal{N}(\text{vect}(\boldsymbol{\beta}_0), \boldsymbol{\Psi_\varepsilon} \otimes \boldsymbol{P}_0^{-1}) , \end{cases} \tag{23}$$

where

- "vect" denotes vectorization and "$\otimes$" is a symbol of the Kronecker product,
- $\mathcal{IW}(.,.)$ and $\mathcal{N}(.,.)$ are respectively the inverse Wishart and the normal distributions,
- $\boldsymbol{T}_0$, ρ_0, $\boldsymbol{\beta}_0$ and $\boldsymbol{P}_0$ are the associated hyperparameters.

We prove in Appendix that the posterior distributions of $p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z},\boldsymbol{\Sigma})$ and $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon},\boldsymbol{Z},\boldsymbol{\Sigma})$ are written as follows:

$$\begin{cases} (\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z},\boldsymbol{\Sigma}) \sim \mathcal{IW}(\boldsymbol{M}, n+\rho_0)\ , \\ \mathrm{vect}(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon},\boldsymbol{Z},\boldsymbol{\Sigma}) \sim \mathcal{N}(\mathrm{vect}(\boldsymbol{\beta}_n), \boldsymbol{\Psi_\varepsilon}\otimes \boldsymbol{U})\ , \end{cases} \tag{24}$$

where

- $\boldsymbol{M} = \boldsymbol{Z}'\boldsymbol{Z} - \boldsymbol{\beta}_n'(\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{P}_0)\boldsymbol{\beta}_n + \boldsymbol{\beta}_0'\boldsymbol{P}_0\boldsymbol{\beta}_0 + \boldsymbol{T}_0$,
- $\boldsymbol{\beta}_n = (\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{P}_0)^{-1}(\boldsymbol{\Sigma}'\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}} + \boldsymbol{P}_0\boldsymbol{\beta}_0)$,
- $\boldsymbol{U} = (\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{P}_0)^{-1}$,
- $\widehat{\boldsymbol{\beta}}$ is the estimate of $\boldsymbol{\beta}$ obtained by the Ordinary Least Squares.

So, we can conclude that each element in the matrix $\boldsymbol{\beta}$ is normally distributed.

Finally, to obtain the estimate $\widehat{\boldsymbol{\theta}}_1$ of $\boldsymbol{\theta_1}$, we ensure firstly the convergence of the Gibbs sampler. Then we collect T observations $(\boldsymbol{\theta}_{\mathbf{1}}^{(1)}, \boldsymbol{\theta}_{\mathbf{1}}^{(2)}, \cdots, \boldsymbol{\theta}_{\mathbf{1}}^{(T)})$ obtained by this algorithm. Hence, the estimate of each unknown parameter is obtained by calculating the average of these T simulated observations.

We follow the same reasoning for the other regression model (20), and we suggest the following conjugate prior distribution of $\boldsymbol{\theta_2}$.

- **Prior distributions of $\boldsymbol{\theta_2} = (\boldsymbol{\beta}_\omega, \boldsymbol{\Psi_\delta})$:**

$$\begin{cases} \boldsymbol{\Psi_\delta} \sim \mathcal{IW}(\boldsymbol{S}_0, \sigma_0)\ , \\ \mathrm{vect}(\boldsymbol{\beta}_\omega/\boldsymbol{\Psi_\delta}) \sim \mathcal{N}(\mathrm{vect}(\boldsymbol{\beta}_{\omega 0}), \boldsymbol{\Psi_\delta}\otimes \boldsymbol{R}_0^{-1})\ , \end{cases} \tag{25}$$

where $\boldsymbol{S}_0$, σ_0, $\boldsymbol{\beta}_{\omega 0}$, and $\boldsymbol{R}_0$ are the associated hyperparameters.

In the same way, the posterior distributions of $\boldsymbol{\theta_2}$ is given by:

- **Posterior distributions of $\boldsymbol{\theta_2}$:**

$$\begin{cases} (\boldsymbol{\Psi_\delta}/\boldsymbol{\eta}^{pls}, \boldsymbol{\Sigma}) \sim \mathcal{IW}(\boldsymbol{F}, n+\sigma_0)\ , \\ \mathrm{vect}(\boldsymbol{\beta}_\omega/\boldsymbol{\Psi_\delta}, \boldsymbol{\eta}^{pls}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathrm{vect}(\boldsymbol{\beta}_{\omega n}), \boldsymbol{\Psi_\delta}\otimes \boldsymbol{D})\ , \end{cases} \tag{26}$$

where

- $\boldsymbol{F} = (\boldsymbol{\eta}^{pls})'\boldsymbol{\eta}^{pls} - \boldsymbol{\Lambda}_n'(\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{R}_0)\boldsymbol{\beta}_{\omega n} + \boldsymbol{\beta}_{\omega 0}'\boldsymbol{R}_0\boldsymbol{A}_0 + \boldsymbol{S}_0$,
- $\boldsymbol{\beta}_{\omega n} = (\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{R}_0)^{-1}(\boldsymbol{\Sigma}'\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}_\omega + \boldsymbol{R}_0\boldsymbol{\Lambda}_0)$,
- $\boldsymbol{D} = (\boldsymbol{\Sigma}'\boldsymbol{\Sigma}+\boldsymbol{R}_0)^{-1}$,
- $\widehat{\boldsymbol{\beta}}_\omega$ is the estimate of $\boldsymbol{\beta}_\omega$ obtained by the ordinary least squares.

6 Application

6.1 Simulation study

In this section, we present a simulation study to compare the performance of the hybrid method with the PLS approach and the Bayesian one. We consider a SEM with three LVs ($\boldsymbol{\xi_1}, \boldsymbol{\eta_2}, \boldsymbol{\eta_3}$) and nine observed variables ($\boldsymbol{x_{11}}, \boldsymbol{x_{12}}, \boldsymbol{x_{13}}, \boldsymbol{y_{24}}, \boldsymbol{y_{25}}, \cdots, \boldsymbol{y_{39}}$), as shown in the Figure 1a. The data set is obtained on the basis of the models defined by (1) and (6) with $q = 1$, $J = 3$, $p_1 = p_2 = p_3 = 3$, $p = 9$ and the sample size $n = 50$. Hence, the measurement and structural equations associated to this model are defined as follows:

$$\begin{cases} \boldsymbol{x_{1l}} = \lambda_l^x \, \boldsymbol{\xi_1} + \boldsymbol{\varepsilon_l} \quad , \quad l = 1,2,3 \\ \boldsymbol{y_{2h}} = \lambda_h^y \, \boldsymbol{\eta_2} + \boldsymbol{\varepsilon_h} \quad , \quad h = 4,5,6 \\ \boldsymbol{y_{3h}} = \lambda_h^y \, \boldsymbol{\eta_3} + \boldsymbol{\varepsilon_h} \quad , \quad h = 7,8,9 \end{cases} \tag{27}$$

$$\begin{cases} \boldsymbol{\eta_2} = \gamma_1 \, \boldsymbol{\xi_1} + \boldsymbol{\delta_1} \\ \boldsymbol{\eta_3} = \gamma_2 \, \boldsymbol{\xi_1} + \gamma_3 * \boldsymbol{\eta_2} + \boldsymbol{\delta_2} \end{cases} \tag{28}$$

where the true values of $\lambda_l^x, l = 1,2,3$; $\lambda_h^x, h = 4,5,...,9$ and $\gamma_j, j = 1,2,3$ are given in the Table 1a. The error terms $\boldsymbol{\varepsilon_l}, l = 1,2,3$ and $\boldsymbol{\varepsilon_h}, h = 4,5,...,9$ are drawn from $\mathcal{N}(0,1)$.

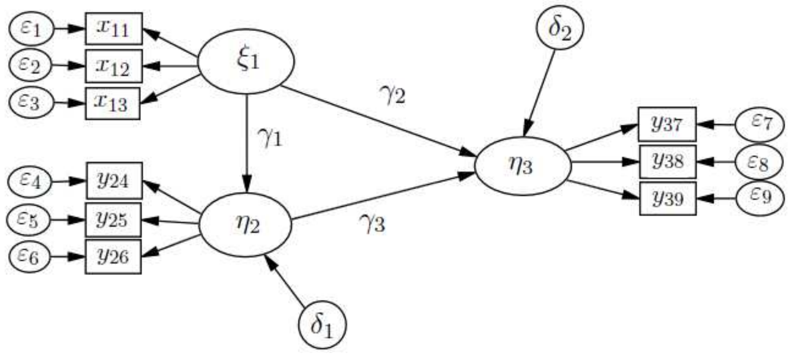

Figure 1a: Path diagram corresponding to the simulation model

As underlined, one of contributions of our method is the absence of the normality assumption contrary to the Bayesian approach; the data are simulated in a manner to have non normal distribution for the LVs scores. For this reason, the LV $\boldsymbol{\xi_1}$ is drawn from the continuous uniform distribution $\mathcal{U}[-3,3]$. Then $\boldsymbol{\eta_2}$ and $\boldsymbol{\eta_3}$ are obtained directly from the equations defined in (28), where $\boldsymbol{\delta_1}$ and $\boldsymbol{\delta_2}$ are drawn from $\mathcal{N}(0,1)$.

The stability of the proposed method is proved by varying the values of the parameters. For this purpose, we consider three simulation cases as presented in the Table 1a.

Table 1a: Simulation plan

	True values											
	γ_1	γ_2	γ_3	λ_1^x	λ_2^x	λ_3^x	λ_4^y	λ_5^y	λ_6^y	λ_7^y	λ_8^y	λ_9^y
Simulation 1	0.7	0.5	0.3	0.9	0.9	0.9	0.7	0.7	0.7	0.8	0.8	0.8
Simulation 2	0.6	0.5	0.3	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Simulation 3	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

We use the R software with **plspm** package to realize the PLS, whereas the WinBUGS software (Lunn et al., 2000) is used to execute the Bayesian one and the hybrid method. Note that in the Bayesian approach, we fixed some parameters to one ($\lambda_1^x = \lambda_4^y = \lambda_7^y = 1$) to identify the model.

To ensure that the scores obtained from the PLS approach are not normally distributed, we have plotted the densities of these scores for each simulation as shown in the Figures 1, 2 and 3. From which we conclude that the LVs scores are not necessarily all normally distributed.

Per each simulation the following outcomes are measured:

• Reliability measures.

• The validation indices.

• The loadings $\lambda_l^x, l = 1, 2, 3$; $\lambda_h^y, h = 4, 5, ..., 9$ and the inner coefficients $\gamma_j, j = 1, 2, 3$; with the standard errors SE, using the three approaches.

• The correlations between the estimated LVs scores obtained from hybrid method and the true LVs.

• **Reliability measures**

Before comparing the parameters estimation and the performance of the three methods, we should check the reliability of any constructs using Cronbach's alpha, Composite reliability CR and AVE (note that these measures are the same for the three methods). The values of these indices are presented in the Table 11 for simulation 1, in the Table 21 for simulation 2 and in the Table 31 for simulation 3. These tables show that all LVs in the model are reliable since Cronbach's alpha, CR and AVE coefficients are above the conventionally accepted threshold (Cronbach's alpha and CR are larger than 0.7, and AVE is larger than 0.5).

• **The validation indices**

In order to compare the empirical performance of the three approaches, we compute the communality, the Redundancy, the R-square and GoF index for all the three simulations. The measures of these indices are reported in the Table 12 for simulation 1, in the Table 22 for simulation 2 and in the Table 32 for simulation 3 (note that the Redundancy and the R^2 may not be computed for the exogenous LV $\boldsymbol{\xi_1}$ and the values of communality are common for the three approaches, see column 2 of the Tables 12, 22 and 32). According to the results in these tables, we see that the communalities are always greater than the conventionally accepted threshold (all communalities are larger than 0.5), these results reflect the validity of the measurement models in each simulation.

Moreover, from the values of F^2, R^2 and GoF index, we conclude that the both approaches hybrid and PLS provide globally better performance of structural and global models compared to the Bayesian approach.

Additionally, when the values of loadings are not large, particularly in the simulations 2 and 3, the hybrid method outperforms both PLS and Bayesian approaches in term of the validity of the global model (see the Tables 22 and 32).

• **Estimation of model parameters**

The Tables 13, 23 and 33 summarize the simulation results about the coefficients $\lambda_l^x, l = 1, 2, 3$; $\lambda_h^y, h = 4, 5, ..., 9$; $\gamma_j, j = 1, 2, 3$ and the standard errors (SE) for each simulation. From these tables, it seems that always the estimates obtained by the hybrid method are very close to the true values. Moreover, these estimates seem more significant in terms of standard errors.

In addition, the estimates derived from our method are more accurate and more significant when the values of the loadings are small ($\lambda_l^x = \lambda_h^y = 0.5$ or 0.6) (see the Tables 23 and 33).

• **LVs prediction accuracy**

Another way to show the performance of the proposed approach is to evaluate the LVs prediction accuracy (Ciavolino and Nitti, 2013). For this goal, the correlations between the true LVs and the estimated scores (noted $\widehat{\boldsymbol{\xi}}_1^{hyb}$, $\widehat{\boldsymbol{\eta}}_2^{hyb}$ and $\widehat{\boldsymbol{\eta}}_3^{hyb}$) are reported in the Tables 14, 24 and 34. It can be seen that all the correlation coefficients are larger than 0.838. These results bring us to the conclusion that, the proposed method provides nice accuracy in the estimation of LVs.

Table 11: Reliability measures for simulation 1

LV	Cronbach's alpha	Composite reliability	AVE
ξ_1	0.902	0.939	0.837
η_2	0.840	0.904	0.759
η_3	0.825	0.896	0.742

Table 12: Communality H^2, Redundancy F^2, R-square R^2 and GoF for simulation 1

		Hybrid Method		Bayesian Approach		PLS	
LV	H^2	F^2	R^2	F^2	R^2	F^2	R^2
ξ_1	0.837	-	-	-	-	-	-
η_2	0.759	0.390	0.514	0.380	0.512	0.392	0.516
η_3	0.742	0.406	0.547	0.401	0.541	0.414	0.558
GoF index		0.639		0.640		0.641	

Table 13: The estimates obtained by the three approaches and their SE for simulation 1

		Hybrid Method		Bayesian Approach		PLS	
Parameter	True value	Estimate	SE	Estimate	SE	Estimate	SE
γ_1	0.7	0.71	0.07	0.66	0.13	0.71	0.10
γ_2	0.5	0.50	0.11	0.44	0.18	0.52	0.13
γ_3	0.3	0.29	0.14	0.38	0.18	0.28	0.13
λ_1^x	0.9	0.95	0.05	1	-	0.90	0.06
λ_2^x	0.9	0.84	0.04	0.93	0.14	0.90	0.06
λ_3^x	0.9	0.50	0.03	0.98	0.13	0.93	0.05
λ_4^y	0.7	0.66	0.05	1	-	0.82	0.08
λ_5^y	0.7	0.69	0.04	0.78	0.13	0.85	0.07
λ_6^y	0.7	0.79	0.03	0.99	0.14	0.93	0.05
λ_7^y	0.8	0.74	0.06	1	-	0.77	0.09
λ_8^y	0.8	0.75	0.03	0.91	0.14	0.91	0.06
λ_9^y	0.8	0.77	0.04	0.88	0.14	0.88	0.07

Table 14: Correlation between the estimated LV scores obtained from hybrid method and the true LV for simulation 1

$Corr(\widehat{\boldsymbol{\xi}_1}^{hyb}, \boldsymbol{\xi_1})$	$Corr(\widehat{\boldsymbol{\eta}_2}^{hyb}, \boldsymbol{\eta_2})$	$Corr(\widehat{\boldsymbol{\eta}_3}^{hyb}, \boldsymbol{\eta_3})$
0.952	0.902	0.942

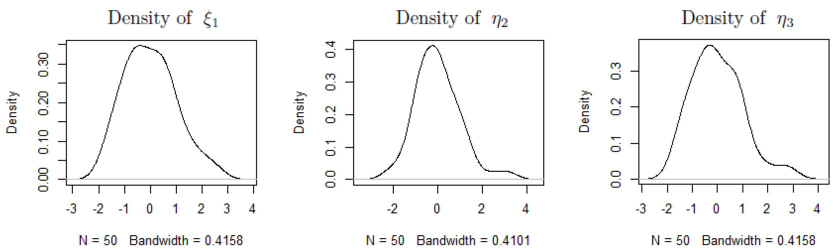

Figure 1: The densities of the three LVs scores for simulation 1

Table 21: Reliability measures for simulation 2

LV	Cronbach's alpha	Composite reliability	AVE
ξ_1	0.820	0.893	0.736
η_2	0.816	0.891	0.730
η_3	0.746	0.856	0.665

Table 22: Communality H^2, Redundancy F^2, R-square R^2 and GoF for simulation 2

		Hybrid Method		Bayesian Approach		PLS	
LV	H^2	F^2	R^2	F^2	R^2	F^2	R^2
ξ_1	0.736	-	-	-	-	-	-
η_2	0.730	0.240	0.318	0.233	0.319	0.235	0.322
η_3	0.665	0.288	0.434	0.287	0.434	0.293	0.442
GoF index		0.601		0.517		0.520	

Table 23: The estimates obtained by the three approaches and their SE for simulation 2

		Hybrid Method		Bayesian Approach		PLS	
Parameter	True value	Estimate	SE	Estimate	SE	Estimate	SE
γ_1	0.6	0.59	0.08	0.61	0.17	0.56	0.11
γ_2	0.5	0.45	0.10	0.34	0.18	0.35	0.13
γ_3	0.3	0.29	0.10	0.43	0.19	0.39	0.13
λ_1^x	0.6	0.70	0.04	1	-	0.89	0.06
λ_2^x	0.6	0.68	0.05	0.95	0.17	0.84	0.07
λ_3^x	0.6	0.62	0.04	0.82	0.16	0.82	0.08
λ_4^y	0.6	0.64	0.04	1	-	0.83	0.08
λ_5^y	0.6	0.59	0.04	0.76	0.14	0.86	0.07
λ_6^y	0.6	0.54	0.03	0.72	0.14	0.86	0.07
λ_7^y	0.6	0.56	0.05	1	-	0.75	0.09
λ_8^y	0.6	0.66	0.04	0.87	0.18	0.84	0.07
λ_9^y	0.6	0.54	0.03	0.66	0.14	0.84	0.07

Table 24: Correlation between the estimated LV scores obtained from hybrid method and the true LV for simulation 2

$Corr(\widehat{\boldsymbol{\xi}_1}^{hyb}, \boldsymbol{\xi_1})$	$Corr(\widehat{\boldsymbol{\eta}_2}^{hyb}, \boldsymbol{\eta_2})$	$Corr(\widehat{\boldsymbol{\eta}_3}^{hyb}, \boldsymbol{\eta_3})$
0.896	0.877	0.873

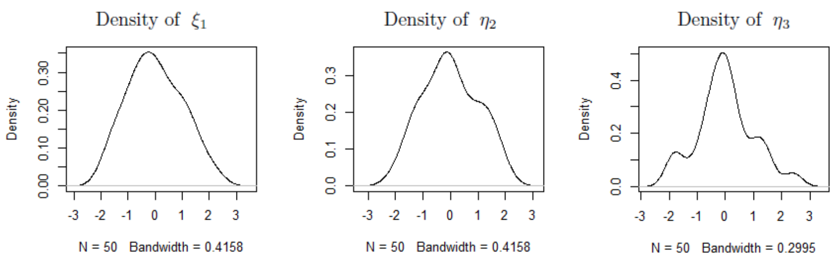

Figure 2: The densities of the three LVs scores for simulation 2

Table 31: Reliability measures for simulation 3

LV	Cronbach's alpha	Composite reliability	AVE
ξ_1	0.807	0.886	0.722
η_2	0.701	0.808	0.590
η_3	0.705	0.836	0.629

Table 32: Communality H^2, Redundancy F^2, R-square R^2 and GoF for simulation 3

		Hybrid Method		Bayesian Approach		PLS	
LV	H^2	F^2	R^2	F^2	R^2	F^2	R^2
ξ_1	0.722	-	-	-	-	-	-
η_2	0.580	0.211	0.362	0.210	0.360	0.232	0.390
η_3	0.629	0.324	0.515	0.326	0.510	0.332	0.534
GoF index		0.573		0.533		0.550	

Table 33: The estimates obtained by the three approaches and their SE for simulation 3

		Hybrid Method		Bayesian Approach		PLS	
Parameter	True value	Estimate	SE	Estimate	SE	Estimate	SE
γ_1	0.5	0.59	0.07	0.53	0.16	0.63	0.11
γ_2	0.5	0.41	0.10	0.42	0.18	0.43	0.12
γ_3	0.5	0.50	0.15	0.46	0.20	0.38	0.12
λ_1^x	0.5	0.61	0.04	1	-	0.84	0.07
λ_2^x	0.5	0.75	0.05	1.17	0.20	0.84	0.07
λ_3^x	0.5	0.59	0.04	0.87	0.16	0.85	0.07
λ_4^y	0.5	0.43	0.04	1	-	0.72	0.09
λ_5^y	0.5	0.48	0.04	0.67	0.18	0.79	0.09
λ_6^y	0.5	0.50	0.04	0.76	0.19	0.76	0.09
λ_7^y	0.5	0.53	0.03	1	-	0.87	0.07
λ_8^y	0.5	0.48	0.05	0.62	0.16	0.72	0.09
λ_9^y	0.5	0.48	0.04	0.63	0.15	0.77	0.09

Table 34: Correlation between the estimated LV scores obtained from hybrid method and the true LV for simulation 3

$Corr(\widehat{\boldsymbol{\xi}_1}^{hyb}, \boldsymbol{\xi_1})$	$Corr(\widehat{\boldsymbol{\eta}_2}^{hyb}, \boldsymbol{\eta_2})$	$Corr(\widehat{\boldsymbol{\eta}_3}^{hyb}, \boldsymbol{\eta_3})$
0.886	0.838	0.879

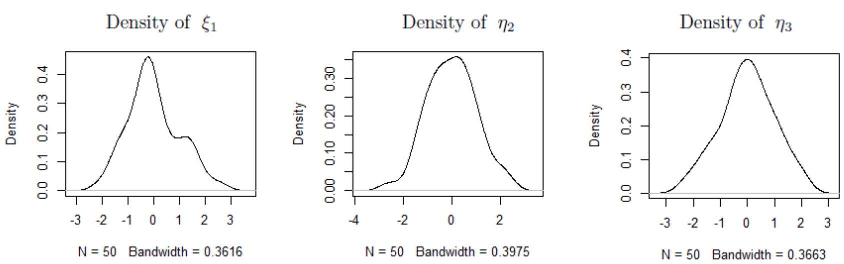

Figure 3: The densities of the three LVs scores for simulation 3

6.2 Application on real data

In order to illustrate the performance of the estimation obtained by the hybrid method in case of complex model and small sample size, we test a theoretical model describing the effect of perceived justice on recovery satisfaction and relationship quality in the case of airline overbooking problem (Zarrouk, 2008). The path diagram of this model is given in Figure 4. Then, we apply the three previous approaches to estimate this model.

In this application, we use only a real small sample size ($n = 50$ passengers) picked randomly from the complete database ($n = 420$ passengers).

Similar to the simulation study, we have firstly plotted the densities of all the LVs scores (Figure 5), from which we conclude that the LVs are not necessarily all normally distributed.

Then, we calculated the reliability measures (Table 41), from where we observe that all LVs are reliable (Cronbach's alpha ≥ 0.702, Composite reliability ≥ 0.819 and AVE ≥ 0.557).

Moreover, the validation indices of measurement, structural and global models are presented in the Table 42. According to this table, it can be seen that the three approaches provide almost similar performances.

The two last tables (Tables 43 and 44) summarize the estimates of the path coefficients γ_j and the loadings λ_k with the standard errors SE. Globally, the estimates obtained from the hybrid method are more significant since they have the smallest SE. These results confirm those obtained by simulation study.

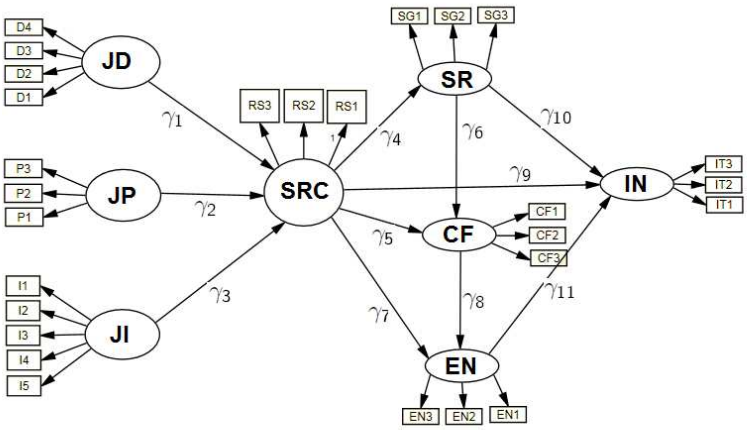

Figure 4: Path diagram corresponding to recovery satisfaction model

where,

JD=Justice_Distributive; JP=Justice_Procédurale; JI=Justice_Interactionnelle; SRC=Satisfaction_Recovery; SR=Satisfaction_Relationnelle; CF=Confiance; EN=Engagement; IN=Intention.

The measurement and structural equations associated to this model are defined as follows:

$$\begin{cases} D_i = \lambda_{1i} * JD + \varepsilon_{1i} \ ; \ i = 1, \cdots, 4 \\ P_i = \lambda_{2i} * JP + \varepsilon_{2i} \ ; \ i = 1, \cdots, 3 \\ I_i = \lambda_{3i} * JI + \varepsilon_{3i} \ ; \ i = 1, \cdots, 5 \\ RS_i = \lambda_{4i} * SRC + \varepsilon_{4i} \ ; \ i = 1, \cdots, 3 \\ SG_i = \lambda_{5i} * SR + \varepsilon_{5i} \ ; \ i = 1, \cdots, 3 \\ CF_i = \lambda_{6i} * CF + \varepsilon_{6i} \ ; \ i = 1, \cdots, 3 \\ EN_i = \lambda_{7i} * EN + \varepsilon_{7i} \ ; \ i = 1, \cdots, 3 \\ IT_i = \lambda_{8i} * IN + \varepsilon_{8i} \ ; \ i = 1, \cdots, 3 \end{cases} \text{and} \begin{cases} SRC = \gamma_1 * JD + \gamma_2 * JP \\ \qquad + \gamma_3 * JI + \delta_1 \\ SR = \gamma_4 * SRC + \delta_2 \\ CF = \gamma_5 * SRC + \gamma_6 * SR + \delta_3 \\ EN = \gamma_7 * SRC + \gamma_8 * CF + \delta_4 \\ IN = \gamma_9 * SRC + \gamma_{10} * SR \\ \qquad + \gamma_{11} * EN + \delta_5 \end{cases}$$

Table 41: Reliability measures for real data

LV	Cronbach's alpha	Composite reliability	AVE
JD	0.702	0.819	0.557
JP	0.792	0.879	0.708
JI	0.917	0.938	0.751
SRC	0.933	0.957	0.882
SR	0.931	0.956	0.878
CF	0.896	0.936	0.829
EN	0.957	0.972	0.920
IN	0.875	0.924	0.802

Table 42: Communality H^2, Redundancy F^2, R-square R^2 and GoF for real data

		Hybrid Method		Bayesian Approach		PLS	
LV	H^2	F^2	R^2	F^2	R^2	F^2	R^2
JD	0.557	0	0	0	0	0	0
JP	0.708	0	0	0	0	0	0
JI	0.751	0	0	0	0	0	0
SRC	0.882	0.511	0.580	0.487	0.553	0.564	0.590
SR	0.878	0.152	0.174	0.151	0.172	0.154	0.175
CF	0.829	0.351	0.453	0.356	0.429	0.387	0.467
EN	0.920	0.644	0.700	0.635	0.690	0.657	0.714
IN	0.802	0.563	0.697	0.539	0.671	0.563	0.701
GoF index		0.640		0.630		0.648	

Table 43: The estimates of γ_j obtained by the three approaches and their SE for real data

	Hybrid Method		Bayesian Approach		PLS	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE
γ_1	0.439	0.031	0.517	0.151	0.451	0.098
γ_2	0.380	0.038	0.435	0.212	0.374	0.113
γ_3	0.176	0.040	0.231	0.176	0.169	0.114
γ_4	0.445	0.055	0.429	0.138	0.419	0.131
γ_5	0.504	0.115	0.436	0.137	0.409	0.117
γ_6	0.447	0.238	0.520	0.143	0.402	0.117
γ_7	0.193	0.107	0.143	0.133	0.181	0.095
γ_8	0.706	0.145	0.836	0.140	0.727	0.095
γ_9	0.010	0.137	-0.064	0.136	-0.085	0.100
γ_{10}	0.549	0.204	0.679	0.166	0.539	0.113
γ_{11}	0.370	0.168	0.401	0.154	0.416	0.128

Table 44: The estimates of λ_k obtained by the three approaches and their SE for real data

	Hybrid Method		Bayesian Approach		PLS	
Parameter	Estimate	SE	Estimate	SE	Estimate	SE
λ_{11}	0.913	0.024	1	-	0.904	0.061
λ_{12}	0.933	0.019	1.054	0.138	0.933	0.051
λ_{13}	0.678	0.038	0.619	0.169	0.680	0.104
λ_{14}	0.244	0.052	0.194	0.173	0.274	0.137
λ_{21}	0.705	0.034	1	-	0.800	0.085
λ_{22}	0.889	0.025	1.021	0.194	0.896	0.063
λ_{23}	0.928	0.034	0.893	0.204	0.825	0.080
λ_{31}	0.868	0.028	1	-	0.886	0.066
λ_{32}	0.899	0.025	1.020	0.131	0.902	0.061
λ_{33}	0.875	0.030	1.008	0.136	0.887	0.066
λ_{34}	0.801	0.033	0.903	0.144	0.829	0.079
λ_{35}	0.877	0.034	0.890	0.142	0.828	0.080
λ_{41}	0.905	0.019	1	-	0.937	0.050
λ_{42}	0.955	0.018	0.969	0.103	0.954	0.042
λ_{43}	0.953	0.022	0.923	0.109	0.926	0.053
λ_{51}	0.984	0.025	1	-	0.925	0.054
λ_{52}	0.933	0.024	1.055	0.124	0.936	0.050
λ_{53}	0.899	0.019	1.050	0.122	0.950	0.044
λ_{61}	0.954	0.024	1	-	0.938	0.049
λ_{62}	0.947	0.019	1.011	0.102	0.965	0.037
λ_{63}	0.829	0.040	0.771	0.127	0.823	0.081
λ_{71}	0.934	0.023	1	-	0.942	0.047
λ_{72}	1.009	0.017	0.944	0.095	0.968	0.035
λ_{73}	0.934	0.018	0.954	0.096	0.967	0.036
λ_{81}	0.911	0.020	1	-	0.956	0.041
λ_{82}	0.860	0.039	0.691	0.119	0.780	0.089
λ_{83}	0.929	0.022	0.914	0.099	0.939	0.049

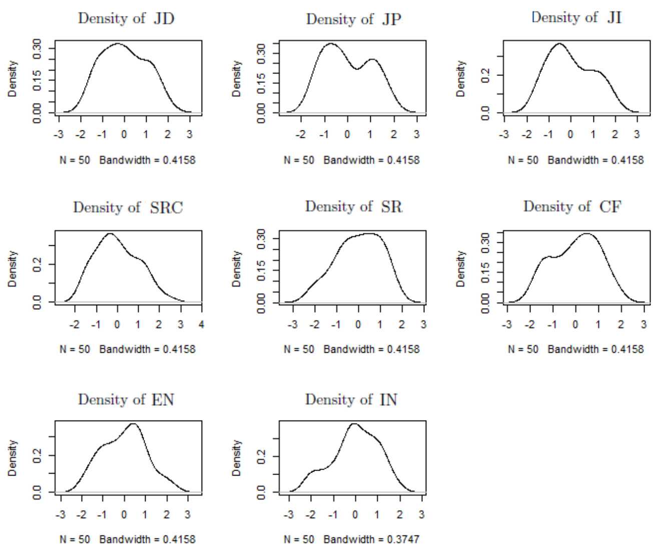

Figure 5: The densities of the three LVs scores in case of real data

7 Conclusion

This paper introduces a semiparametric approach based on combination of the PLS and the Bayesian approaches. This new method presents several advantages such as, the absence of identification problem and the assumption of normality of the LVs is not necessary. It provides reliable results for small samples and complex models. Furthermore, this technique uses a simpler Gibbs sampler. The results obtained from application on simulated and on real data show the great performance of the estimates obtained in case of hybrid method even when the LVs are not normally distributed. However, the proposed method doesn't take into account the non linearity that can exist in the measurement and structural models. Also, the interactions between explanatory LVs are not considered in our method. Further research is required to compare the performance of different semiparametric approaches: hybrid method, Generalized Maximum Entropy and Bayesian semiparametric SEM.

Appendix:

Proof of the posterior distribution of $p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z},\boldsymbol{\Sigma})$ and $p(\boldsymbol{A}/\boldsymbol{\Psi_\varepsilon},\boldsymbol{Z},\boldsymbol{\Sigma})$

First, we recall the expression of model (18): $\boldsymbol{Z}=\boldsymbol{\Sigma\beta}+\boldsymbol{\varepsilon}$ with $\boldsymbol{\theta_1}=(\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon})$. Then we write,

$$\begin{aligned} p(\boldsymbol{\theta_1}/\boldsymbol{Z},\boldsymbol{\Sigma}) &= p(\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z},\boldsymbol{\Sigma}) \\ &\propto p(\boldsymbol{Z}/\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon},\boldsymbol{\Sigma})p(\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon}) \\ &= p(\boldsymbol{Z}/\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon},\boldsymbol{\Sigma})p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon})p(\boldsymbol{\Psi_\varepsilon})\ . \end{aligned}$$

Now, we replace the densities $p(\boldsymbol{Z}/\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon},\boldsymbol{\Sigma})$, $p(\boldsymbol{\Psi_\varepsilon})$ and $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon})$ by their expressions (22) and (23) given in section 5, then we obtain,

$$\begin{aligned} p(\boldsymbol{\theta_1}/\boldsymbol{Z},\boldsymbol{\Sigma}) &\propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{n}{2}}exp\{-\frac{1}{2}Trace[(\boldsymbol{Z}-\boldsymbol{\Sigma\beta})^{'}(\boldsymbol{Z}-\boldsymbol{\Sigma\beta})\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}}exp\{-\frac{1}{2}Trace[(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^{'}\boldsymbol{P}_0(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+\rho_0}{2}}exp\{-\frac{1}{2}Trace[\boldsymbol{T}_0\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &= |\boldsymbol{\Psi_\varepsilon}|^{-\frac{n}{2}}|\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}}exp\{-\frac{1}{2}Trace[((\boldsymbol{Z}-\boldsymbol{\Sigma\beta})^{'}(\boldsymbol{Z}-\boldsymbol{\Sigma\beta}) \\ &\quad +(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^{'}\boldsymbol{P}_0(\boldsymbol{\beta}-\boldsymbol{\beta}_0))\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+\rho_0}{2}}exp\{-\frac{1}{2}Trace[\boldsymbol{T}_0\boldsymbol{\Psi_\varepsilon^{-1}}]\}\ . \end{aligned}$$

Let $\boldsymbol{\beta}_n=(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)^{-1}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}\widehat{\boldsymbol{\beta}}+\boldsymbol{P}_0\boldsymbol{\beta}_0)$,
where $\widehat{\boldsymbol{\beta}}$ is the estimate of $\boldsymbol{\beta}$ obtained by the Ordinary Least Squares method.

So,

$$\begin{aligned} (\boldsymbol{Z}-\boldsymbol{\Sigma\beta})^{'}(\boldsymbol{Z}-\boldsymbol{\Sigma\beta})+(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^{'}\boldsymbol{P}_0(\boldsymbol{\beta}-\boldsymbol{\beta}_0) &= (\boldsymbol{\beta}-\boldsymbol{\beta}_n)^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)(\boldsymbol{\beta}-\boldsymbol{\beta}_n) \\ &\quad +\boldsymbol{Z}^{'}\boldsymbol{Z}-\boldsymbol{\beta}_n^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)\boldsymbol{\beta}_n \\ &\quad +\boldsymbol{\beta}_0^{'}\boldsymbol{P}_0\boldsymbol{\beta}_0\ . \end{aligned}$$

Therefore,

$$\begin{aligned} p(\boldsymbol{\beta},\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z},\boldsymbol{\Sigma}) &\propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{n}{2}}|\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}}exp\{-\tfrac{1}{2}Tr[((\boldsymbol{\beta}-\boldsymbol{\beta}_n)^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)(\boldsymbol{\beta}-\boldsymbol{\beta}_n))\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times exp\{-\tfrac{1}{2}Tr[(\boldsymbol{Z}^{'}\boldsymbol{Z}-\boldsymbol{\beta}_n^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)\boldsymbol{\beta}_n+\boldsymbol{\beta}_0^{'}\boldsymbol{P}_0\boldsymbol{\beta}_0)\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+\rho_0}{2}}exp\{-\tfrac{1}{2}Tr[\boldsymbol{T}_0\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &= |\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}}exp\{-\tfrac{1}{2}Tr[((\boldsymbol{\beta}-\boldsymbol{\beta}_n)^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)(\boldsymbol{\beta}-\boldsymbol{\beta}_n))\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+n+\rho_0}{2}}exp\{-\tfrac{1}{2}Tr[(\boldsymbol{Z}^{'}\boldsymbol{Z}-\boldsymbol{\beta}_n^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)\boldsymbol{\beta}_n \\ &\quad +\boldsymbol{\beta}_0^{'}\boldsymbol{P}_0\boldsymbol{\beta}_0+\boldsymbol{T}_0)\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &= |\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}}exp\{-\tfrac{1}{2}Tr[((\boldsymbol{\beta}-\boldsymbol{\beta}_n)^{'}(\boldsymbol{\Sigma}^{'}\boldsymbol{\Sigma}+\boldsymbol{P}_0)(\boldsymbol{\beta}-\boldsymbol{\beta}_n))\boldsymbol{\Psi_\varepsilon^{-1}}]\} \qquad \textbf{(A.1)} \\ &\quad \times|\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+n+\rho_0}{2}}exp\{-\tfrac{1}{2}Trace[\boldsymbol{M}\boldsymbol{\Psi_\varepsilon^{-1}}]\}\ , \qquad \textbf{(A.2)} \end{aligned}$$

with, $\boldsymbol{M} = \boldsymbol{Z}'\boldsymbol{Z} - \boldsymbol{\beta}_n'(\boldsymbol{\Sigma}'\boldsymbol{\Sigma} + \boldsymbol{P}_0)\boldsymbol{\beta}_n + \boldsymbol{\beta}_0'\boldsymbol{P}_0\boldsymbol{\beta}_0 + \boldsymbol{T}_0$.
And we know that:
$p(\boldsymbol{\beta}, \boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma}) = p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma})p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma})$.

From **(A.1)** and **(A.2)**, we can write,
$p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{p+1+n+\rho_0}{2}} exp\{-\frac{1}{2}Trace[\boldsymbol{M}\boldsymbol{\Psi_\varepsilon^{-1}}]\}$,
And
$p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma}) \propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}} exp\{-\frac{1}{2}Trace[(\boldsymbol{\beta} - \boldsymbol{\beta}_n)'(\boldsymbol{\Sigma}'\boldsymbol{\Sigma} + \boldsymbol{P}_0)(\boldsymbol{\beta} - \boldsymbol{\beta}_n)\boldsymbol{\Psi_\varepsilon^{-1}}]\}$.
Let $\boldsymbol{U^{-1}} = (\boldsymbol{\Sigma}'\boldsymbol{\Sigma} + \boldsymbol{P}_0)$, then the expression of $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma})$ becomes:

$$\begin{aligned} p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma}) &\propto |\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}} exp\{-\frac{1}{2}Trace[(\boldsymbol{\beta} - \boldsymbol{\beta}_n)'\boldsymbol{U}^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}_n)\boldsymbol{\Psi_\varepsilon^{-1}}]\} \\ &= |\boldsymbol{\Psi_\varepsilon}|^{-\frac{J}{2}} exp\{-\frac{1}{2}Trace[\boldsymbol{U^{-1}}(\boldsymbol{\beta} - \boldsymbol{\beta}_n)\boldsymbol{\Psi_\varepsilon^{-1}}(\boldsymbol{\beta} - \boldsymbol{\beta}_n)']\} . \end{aligned}$$

Consequently, the posterior distributions of $p(\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma})$ and $p(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma})$ are given as follows:

$$\begin{cases} (\boldsymbol{\Psi_\varepsilon}/\boldsymbol{Z}, \boldsymbol{\Sigma}) \sim \mathcal{IW}(\boldsymbol{M}, n + \rho_0) , \\ (\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma}) \sim \mathcal{MN}_{Jp}(\boldsymbol{\beta}_n, \boldsymbol{U}, \boldsymbol{\Psi_\varepsilon}) , \end{cases}$$

where $\mathcal{MN}_{Jp}(.,.,.)$ denotes multivariate normal distribution, and J is the order of the matrix $\boldsymbol{U}$ and p is the order of the matrix $\boldsymbol{\Psi_\varepsilon}$.

From where, we conclude that, $\mathbf{vect}(\boldsymbol{\beta}/\boldsymbol{\Psi_\varepsilon}, \boldsymbol{Z}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathbf{vect}(\boldsymbol{\beta}_n), \boldsymbol{\Psi_\varepsilon} \otimes \boldsymbol{U})$. Then each element in $\boldsymbol{\beta}$ is normally distributed.

We follow the same proof for the model (20).

References

Ansari, A., Jedidi, K., and Jagpal, S. (2000). A hierarchical bayesian methodology for treating heterogeneity in structural equation models. *Marketing Science*, 19(4):328–347.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern methods for business research*, 295(2):295–336.

Ciavolino, E. and Al-Nasser, A. D. (2009). Comparing generalised maximum entropy and partial least squares methods for structural equation models. *Journal of nonparametric statistics*, 21(8):1017–1036.

Ciavolino, E. and Carpita, M. (2015). The gme estimator for the regression model with a composite indicator as explanatory variable. *Quality & Quantity*, 49(3):955–965.

Ciavolino, E., Carpita, M., and Al-Nasser, A. (2015). Modelling the quality of work in the italian social co-operatives combining npca-rsm and sem-gme approaches. *Journal of Applied Statistics*, 42(1):161–179.

Ciavolino, E. and Dahlgaard, J. J. (2009). Simultaneous equation model based on the generalized maximum entropy for studying the effect of management factors on enterprise performance. *Journal of applied statistics*, 36(7):801–815.

Ciavolino, E. and Nitti, M. (2013). Simulation study for pls path modelling with high-order construct: A job satisfaction model evidence. In *Advanced Dynamic Modeling of Economic and Social Systems*, pages 185–207. Springer.

Dunson, D. B. (2000). Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(2):355–366.

Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, pages 39–50.

Hair, J. F., Sarstedt, M., Pieper, T. M., and Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long range planning*, 45(5-6):320–340.

Hoyle, R. H. (2012). *Handbook of structural equation modeling.* Guilford Press.

Jedidi, K. and Ansari, A. (2001). Bayesian structural equation models for multilevel data. *New developments and techniques in structural equation modeling*, pages 129–157.

Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika*, 57(2):239–251.

Jöreskog, K. G. and Wold, H. (1982). The `ML` and `PLS` techniques for modeling with latent variables: historical and comparative aspects. In Jöreskog, K. G. and Wold, H., editors, *Systems Under Indirect Observation, Part 1*, pages 263–270, North-Holland, Amsterdam.

Lee, S.-Y. (2007). *Structural equation modeling: A Bayesian approach*, volume 711. John Wiley & Sons.

Lee, S.-Y. and Song, X.-Y. (2003). Model comparison of nonlinear structural equation models with fixed covariates. *Psychometrika*, 68(1):27–47.

Lee, S.-Y. and Song, X.-Y. (2004). Evaluation of the bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research*, 39(4):653–686.

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). Winbugs-a bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*, 10(4):325–337.

Nitzl, C. (2016). The use of partial least squares structural equation modelling (pls-sem) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, 37:19–35.

Papalia, R. B. and Ciavolino, E. (2011). Gme estimation of spatial structural equations models. *Journal of classification*, 28(1):126–141.

Reinartz, W., Haenlein, M., and Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based sem. *International Journal of research in Marketing*, 26(4):332–344.

Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., and Hair, J. F. (2014). Partial least squares structural equation modeling (pls-sem): A useful tool for family business researchers. *Journal of Family Business Strategy*, 5(1):105–115.

Song, X.-Y. and Lee, S.-Y. (2012). A tutorial on the bayesian approach for analyzing structural equation models. *Journal of Mathematical Psychology*, 56(3):135–148.

Tenenhaus, M., Amato, S., and Esposito Vinzi, V. (2004). A global goodness-of-fit index for pls structural equation modelling. In *Proceedings of the XLII SIS scientific meeting*, volume 1, pages 739–742.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., and Lauro, C. (2005). Pls path modeling. *Computational statistics & data analysis*, 48(1):159–205.

Vinzi, V. E., Trinchera, L., and Amato, S. (2010). Pls path modeling: from foundations to recent developments and open issues for model assessment and improvement. In *Handbook of partial least squares*, pages 47–82. Springer.

Wetzels, M., Odekerken-Schröder, G., and Van Oppen, C. (2009). Using pls path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS quarterly*, pages 177–195.

Wold, H. (1982). Soft modelling: the basic design and some extensions. *Systems under indirect observation, Part II*, pages 36–37.

Wold, H. (1985). Partial least squares. *Encyclopedia of statistical sciences.*

Yang, M. and Dunson, D. B. (2010). Bayesian semiparametric structural equation models with latent variables. *Psychometrika*, 75(4):675–693.

Yanuar, F. (2014). The estimation process in bayesian structural equation modeling approach. In *Journal of Physics: Conference Series*, volume 495, page 012047. IOP Publishing.

Zarrouk, Z. (2008). *Injustice, compensation et qualité de la relation: le cas d'une défaillance volontaire - le surbooking.* PhD thesis, Pau.