New Theory Journal of
ISSN: 2149-1402

33 (2020) 56-75
Journal of New Theory
http://www.newtheory.org
Open Access

A Classical and Bayesian Approach for Parameter Estimation in Structural Equation Models

Naci Murat[1], Mehmet Ali Cengiz[2]

Article History
Received: 26.11.2020
Accepted: 15.12.2020
Published: 30.12.2020
Original Article

Abstract − Structural Equation Models (SEMs) with latent variables provide a general framework for modelling relationships in multivariate data. Although SEMs are most commonly used in studies involving intrinsically latent variables, such as happiness, quality of life, or stress, they also provide a parsimonious framework for covariance structure modelling. For this reason, they have become increasingly used outside of traditional social science applications. Frequentist inferences are based on point estimates and hypothesis tests for the measurement and latent variable parameters. Although most of the literature on SEMs is frequentist, Bayesian approaches have been proposed in the last years. This study aims to provide an easily accessible overview of a Classic and a Bayesian approach to SEMs. Due to the flexibility of the Bayesian approach, it is straightforward to apply the method in a comprehensive class of SEM-type modelling frameworks, allowing nonlinearity, interactions, missing data, mixed categorical, count, and continuous observed variables. The WinBUGS software package, which is freely available, can be used to implement Bayesian SEM analysis. Bayesian model fitting typically relies on MCMC, which involves simulating draws from the joint posterior distribution of the model unknowns (parameters and latent variables) through a computationally intensive procedure. The advantage of MCMC is that there is no need to rely on broad sample assumptions because exact posterior distributions can be estimated for any function of the model unknowns. In small to moderate samples, these exact posteriors can provide a more realistic measure of model uncertainty. Therefore, we use the MCMC method for the Bayesian approach in this study. All approaches given above are applied to the data obtained from Samsun Chamber of Commerce and Industry.

Keywords − *Structural equation models, Bayesian approach, MCMC, Bayesian structural equation models*

1. Introduction

The Structural Equation Model (SEM) is a multivariate statistical modelling technique that reveals the cause-effect relationship between measurable variables and non-measurable (implicit) variables. SEM consists of Observed/Measured Variables and unobservable or unmeasurable variables (Latent Variables) that can function as endogenous and exogenous. Since implicit variables cannot be measured directly, it is essential to define the measurable variables that the researcher wants to examine, and that is thought to represent the implicit variable. Measurable variables that describe implicit variables can be one or more. Therefore, the fact that makes the implicit variable measurable is assessing variables or variables that define the implicit variable.

In SEM, which is based on the causality relationship between implicit variables, each of the implicit variables is a linear function of the set of variables that were observed or measured. The parameters of these

[1] nacimurat@omu.edu.tr (Corresponding Author); [2] macengiz@omu.edu.tr
[1] Department of Industrial Engineering, Faculty of Engineering, Ondokuz Mayıs University, Samsun, Turkey
[2] Department of Statistics, Faculty of Art and Sciences, Ondokuz Mayıs University, Samsun, Turkey

linear functions are obtained using an analysis of covariance. It is tested that using the goodness of fit tests whether the researcher's model is compatible with the data’s variance-covariance structure. If the model predictions are accepted at the end of the test, the linear relationship between the implicit variables is assumed to be reasonable. SEM is a hybrid method that combines factor analysis and path analysis. Perhaps the main reason why SEM is so widely used today is that direct or indirect relationships between observable and unobservable variables can be analysed in a single model. SEM can also be considered as multiple regression analysis, and factor analysis performed simultaneously. Therefore, YEM; is also named with definitions such as causal analysis, causal modelling, concurrent structural modelling, covariance structure analysis, path analysis, or confirmatory factor analysis.

SEM is based on three basic analytical developments [1]. These are, respectively, path analysis, latent variable model, and general covariance estimation methods. Wright [2] started his first studies on road analysis and along with other studies, road analysis was developed and basic rules were established [2,3]. Today, SEM is widely used in many fields such as behavioral sciences, educational sciences, economics, marketing, health sciences and social sciences. In examining the structural relations of production practices, delivery time and productivity in Japan and Korea, the effect of time-based production on customer-specific production and adding value to the customer, the development of customer satisfaction index in the Turkish mobile phone industry, evaluation of customer satisfaction in the telephone industry with multi-level structural equation models, brand In the measurement of the value of the supply chain management, the effect of e-supply chain competence on competitive advantage and organizational performance, the effect of supplier development on purchasing performance, in modelling student success, In performing risk analysis in the coal mine construction project, The effects of total quality management practices applied in enterprises on employee performance, the examination of factors affecting individuals’ adoption of internet banking, in fraudulent financial reporting determination of auditor responsibility, investigation of the effect of critical control (success) factors in enterprise resource planning (ERP) applications, determinants of capital structure selection. Effects of depression and disease severity on quality of life, Use of SEM in decision tree models, Operation management, symptoms related to ecstatic; the role of fear of blood, injection and injury (KEY), estimation of post-stress traumas of child welfare institutions, processing speed, relationship between intelligence, creativity and school performance, use of incremental goodness of fit indexes in market research studies, structural equation for river water quality data model, role ambiguity, role conflict, relationships between job satisfaction and performance, structural equation technique and interactional stress and coping model.

Why not take advantage of our abilities, which we regularly use and call intuition, common sense, and sixth sense, for scientific purposes? Bayes Theory emerges as an alternative inference in the scientific use of such abilities. The classic inference is to conclude the population we do not have information about with sample data. Statistical operations such as confidence intervals and hypothesis tests are the basis of classical inference. However, “Life; It is the art of drawing sufficient conclusions from insufficient a priori.” Thomas Bayes, who has a similar opinion with Samuel Butler, has a different perspective on the audience’s inference based on the observed sample data. In general, he formed the chain of logic from causes to effects, from results to causes. In the last 30 years, using an approach different from other common basic approaches in statistical analysis has increased. This approach is Bayesian inference and is based on the well-known theorem put forward by Thomas Bayes in 1763. Thomas Bayes did not even predict that his simple probabilistic theorem would be a statistical method of inference. However, in the last 30 years, this theorem has influenced many statisticians and mathematicians, and Bayes statistics has been accepted as the primary method of statistical inference. Many researchers such as Jeffreys, de Finetti, Savage, and Lindley contributed to the development of Bayesian analysis. In recent years, the technical applicability of Bayes analysis has also rapidly developed using computers and has opened up new application areas. As a result of these developments, Bayesian analysis was expanded with researchers such as Berger [4] and Bernardo and Smith [5].

Today, Bayesian analysis is successfully applied in every discipline. Many applied studies have been revealed. A broad overview of these studies will be given below. It is often difficult to calculate the Bayesian factor. Different calculation methods have been proposed by Kass and Raftery [6]. A simple approach is used in the Bayesian Knowledge Criteria (BIC) YEM. For example, [7] on the LISREL model, Lee and Song [8] on the two-level SEM, Jedidi, Jagpal [9] applied to finite mixed SEM. Also, posterior simulation based approaches are used in Bayes Factor calculations. DiCiccio, Kass [10] detailed many methods from the Laplace approach to importance sampling. Gelman and Meng [11] developed a road sampling approach. Lee [12] gave the Bayesian approach with WinBUGS applications on linear, nonlinear and loss observation structural equation models. Wang and Fan [13] examined the factors affecting myopia disease with Bayesian structural equation models. Bayesian approach for semi-parametric structural equation models is given by Guo, Zhu [14]. In the first part of this study, firstly, literature information about structural equation models, secondly Bayesian approach and finally Bayesian structural equation models are given. In the second section under the title of basic information, basic concepts related to structural equation models and Bayesian approach are presented. Linear and nonlinear structural equation models are presented with a Bayesian approach in the material method section. In the fourth chapter, 2011 data of Samsun Chamber of Commerce and Industry member satisfaction are used. There are 4 factors that are thought to affect the overall satisfaction of the room. These 4 factors are guidance, solution, personnel and representation factors, respectively. Factors affecting the general satisfaction of the chamber were determined by both classical and Bayesian structural equation models.

2. Materials and Methods

2.1. Linear Structural Equation Modelling

Due to the nature of the problems and the design of the questionnaires in behaviour, education, medical and social sciences, data are usually obtained as sequential categorical variables. Examples of these variables; Scales such as attitude scales, likert scales, rating scales can be given. When questioned about some attitudes, the scale was “I do not approve”, “I have no idea”, “I approve”, while questioning about the effect of a drug was “worsened”, “did not change”, “got better” and when questioned about a political event, definitely “I don’t agree”, “I don’t agree”, “I have no idea”, “I agree”, “I absolutely agree”. Consider a five-point scale associated with responses to a political event. A common approach is to treat these integer values as continuous values drawn from the normal distribution. This approach does not cause serious problems if the histograms of the observation value are symmetrical and the frequencies of the central values are high. This will emerge in many cases when the “I have no idea” option is chosen. To claim that the observed variables are multivariate normally distributed, in most cases we have to choose the middle category. For example, “I have no idea” or “no change”. In many cases likert scales may have clutter at both ends. For example, such as “strongly agree (strongly disagree)” or “agree (disagree)”. Therefore, histograms are either skewed or bimodal as opposed to the variables involved. Treating such ordered categorical variables as normal may lead to erroneous results [15,16]. A better approach for evaluating discrete data is to consider these data as latent continuous variables from a specified threshold normal distribution. For a given data set, the ratios of 1, 2, 3, and 4 values are 0.05, 0.05, 0.40, and 0.5, respectively. From the histogram given in Figure 1, it is seen that the dashed data are skewed to the right.

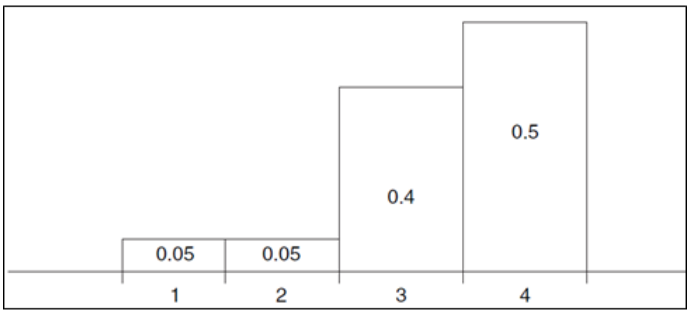

Fig. 1. Historical development of YEM [17]

The threshold approach for the analysis of this discrete data is to consider the discrete categorical data as normal variable y. There are no precise continuous measurements of y but they are related to the observed ordered categorical variable z. This relationship is expressed as follows:

$$z = k \qquad if \quad \propto_{k-1} < y < \propto_k \quad k = 1,2,3,4$$

Here, $-\infty < \propto_1 < \propto_2 < \propto_3 < \infty$ where $\propto_1, \propto_2$ and $\propto_3$ are threshold values. Then, the histogram of sequential categorical observations given in Figure 1 can be in a view with $N[0.1]$ distribution with appropriate threshold values as in Figure 2.

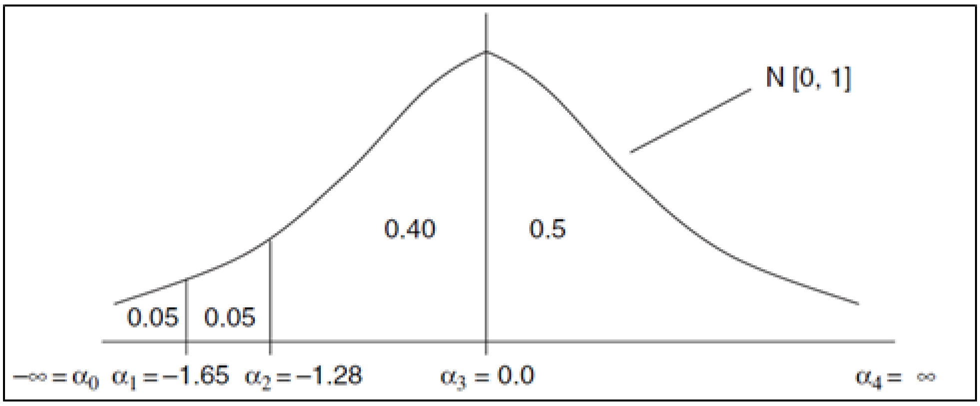

Fig. 2. Histogram chart of the scale

While the difference $\propto_2 - \propto_1$ may differ from the difference $\propto_3 - \propto_2$, unequal scales are allowed. Therefore, this threshold approach allows flexible modelling. As associated with a common normal distribution, it also allows parameters to be easily interpreted. It should be noted that temporary integer values ($k = 1,2,3,4$) are used only to represent the category; Only the frequencies of these values are important in statistical analysis. Structural equation modelling consisting of continuous and discrete data does not have a simple structure. Because it is necessary to calculate multiple integrals associated with cell probabilities determined by ordered categorical results [12].

Some multi-step methods have been introduced to reduce the computational difficulties of these integrals. The basic procedure of these multi-stage methods is polychromic and polyserial correlations, estimating the threshold value in the first stage, deriving the asymptotic distribution of the predictions in the second stage, and analysing the structural equation model with the generalized least-squares approach and covariance structural equation model in the last stage. There are different methods at the first stage to manage the different procedures given in PRELIS and LISREL. Different methods initially applied PRELIS and LISREL [18], LISCOMP and MPLUS [19], and Lee et al. [20] causes different operations. Multistep estimators, however, are not statistically optimal and need to invert a large matrix of size that increases very rapidly with the number of variables that can be observed at each stage of generalized least squares minimization. Besides multi-step operations, Reboussin and Liang [21] proposed an equality estimation approach and Shi and Lee [22] developed a Monte Carlo EM algorithm for the maximum likelihood analysis of a factor analysis model.

When dealing with sequential categorical variables in Bayesian analysis, the basic idea is to treat the expressed latent continuous measurements as hypothetically lost data and amplify them with observed data in posterior analysis. Using this data magnification strategy, the model based on the full data set becomes continuously variable. Sequences of observations of structural parameters, latent variables, and thresholds in infinitive analysis are simulated from the composite posterior distribution using a hybrid algorithm that is the combination of Gibbs sampling [23] and MH algorithm [24,25]. Combined Bayesian estimates of unknown thresholds, structural parameters, and latent variables are produced together with the standard error estimates of these estimates by using simulated observations. In addition to these point estimates, Bayesian model selection can be reviewed using the Bayes factor [12].

2.2. Application Material

The questions regarding the Samsun Chamber of Commerce and Industry member satisfaction survey used in this study. Table 1 is also given.

Table 1. Survey questions and related factors

Factors	Question	Factor Name	
a1	H1	General	General When you think about it in general, how satisfied you with are being a member of our chamber?
a1	H2		Generally speaking, how satisfied are you with the services of our room?
a2	H3	Guidance	Guidance Room responds to our requests in a timely manner
a2	H4		The efficiency of the chamber in strengthening the dialogue between the public authority and the industrialist is sufficient.
a2	H5		The efficiency of chamber services in the development of the sectors is sufficient.
a2	H6		Chamber efficiency is sufficient in terms of national and international expansion of the members.
a3	H7	Solution	Solution I find the room management successful in understanding our problems / needs related to the sector.
a3	H8		I can reach the management / concerned people when we need it
a3	H9		Room management has the ability to produce solutions to your sectoral problems
a3	H10		Our individual problems are taken into consideration by the room management.
a3	H11		I find room management successful in providing an environment and coordination that helps to solve individual problems.
a4	H12	Personal	Staff Attitudes and behaviours of personnel in business relations
a4	H13		Personnel being innovative and productive
a4	H14		Personnel bringing suggestions and guidance
a5	H15	Representation	Representation How satisfied are you with the representation level of the Chamber management?
a5	H16		How satisfied are you with the chamber management in terms of member relations?

3. Results

In this study, member satisfaction survey data with 616 companies randomly selected among Samsun Chamber of Commerce and Industry members in 2011 were used. The aim is to reveal the structural relationship between general satisfaction from Samsun Chamber of Commerce and Industry and guidance, solution, personnel and representation. The application consists of 2 steps. In the first step, the structural equation model was examined with confirmatory factor analysis using the LISREL package program. In the second step, the Bayesian structural equation model was analysed using the WinBUGS package program.

CLASSIC SOLUTION WITH LISREL

All relevant observed variables were associated with latent variables using a one-way road vehicle. The main reason why we draw the path diagram is that it allows easy visualization of the relationship between variables. After completing the figural representation of the relationship between variables, the solution phase was started. In the solution phase, the fit indices specified in the material and method section were examined and the final solution was obtained. Goodness of fit criteria are extremely important in obtaining the final solution. Although there is no comparison in the literature regarding the superiority of goodness of fit criteria, the LISREL program has highlighted the RMSEA value under the path diagram. Since the RMSEA value was not within the required fit criteria in the initial solution, the correction indices suggested by the program were examined and the solution process was repeated. The correction indices suggested as a result of the LISREL solution were used and associated as shown in Figure 3. The decrease in chi-square value is taken into account when using correction indices. The new solution is obtained by performing the correction process that provides the highest decrease in the chi-square value. Below, all the situations under the Estimates option in the LISREL program are shown on the path diagram.

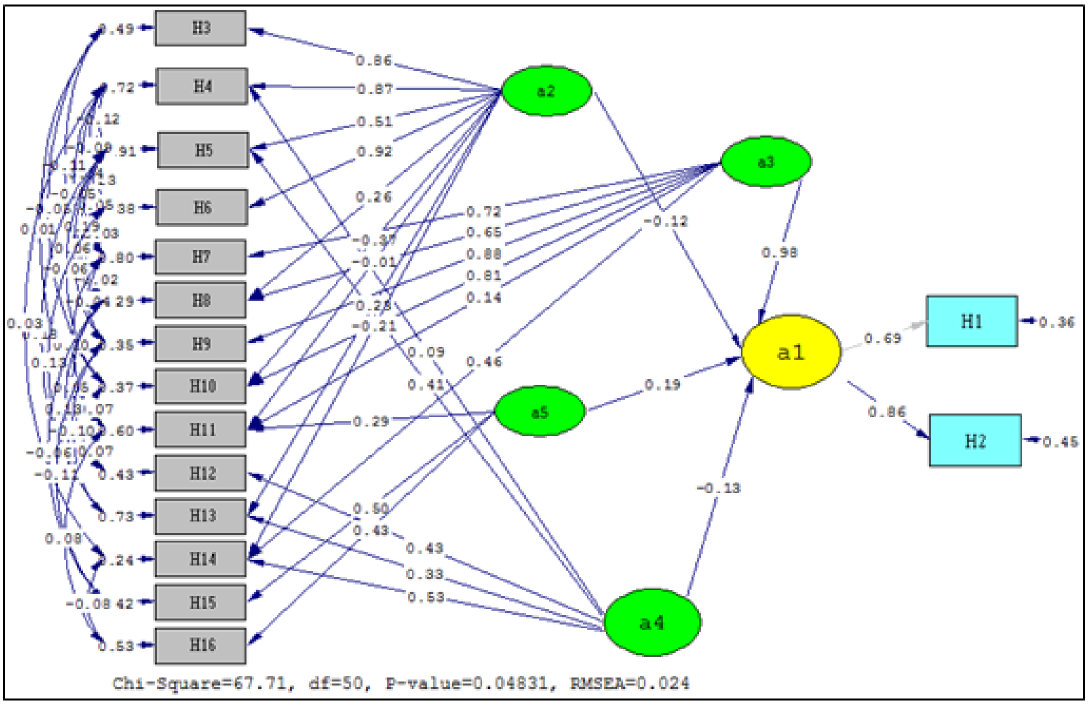

Fig. 3. Threshold values for the scale

The non-standardized coefficients obtained as a result of the LISREL solution are shown in Figure 3. The chi-square value was obtained as 67.71 at the program output. The ratio of the chi-square value to the degrees of freedom 50 is obtained as 1.35. This ratio shows us that the model established is a very powerful model. Another supporting indicator is the approximate root mean square error (RMSEA) value. The fact that this value is close to (0.024) 0 shows that the fit of the model is good. After checking the general fit of the model, the significance of the model parameters was examined.

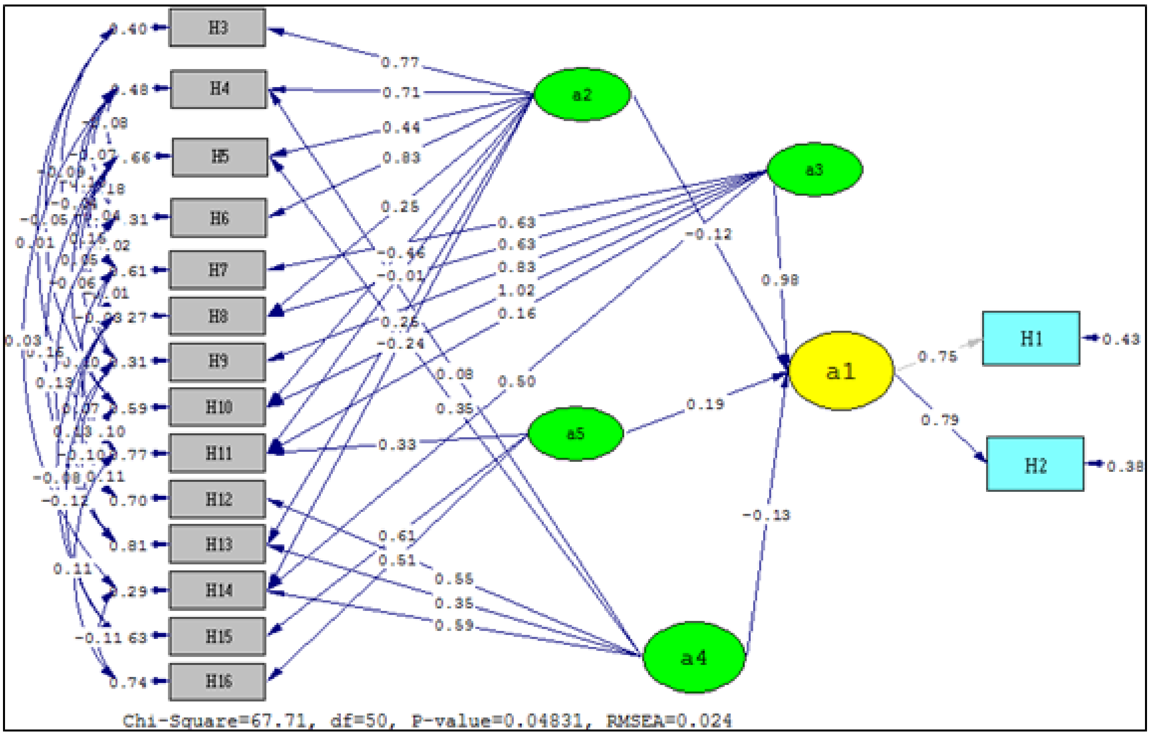

Fig 4. Unstandardized results

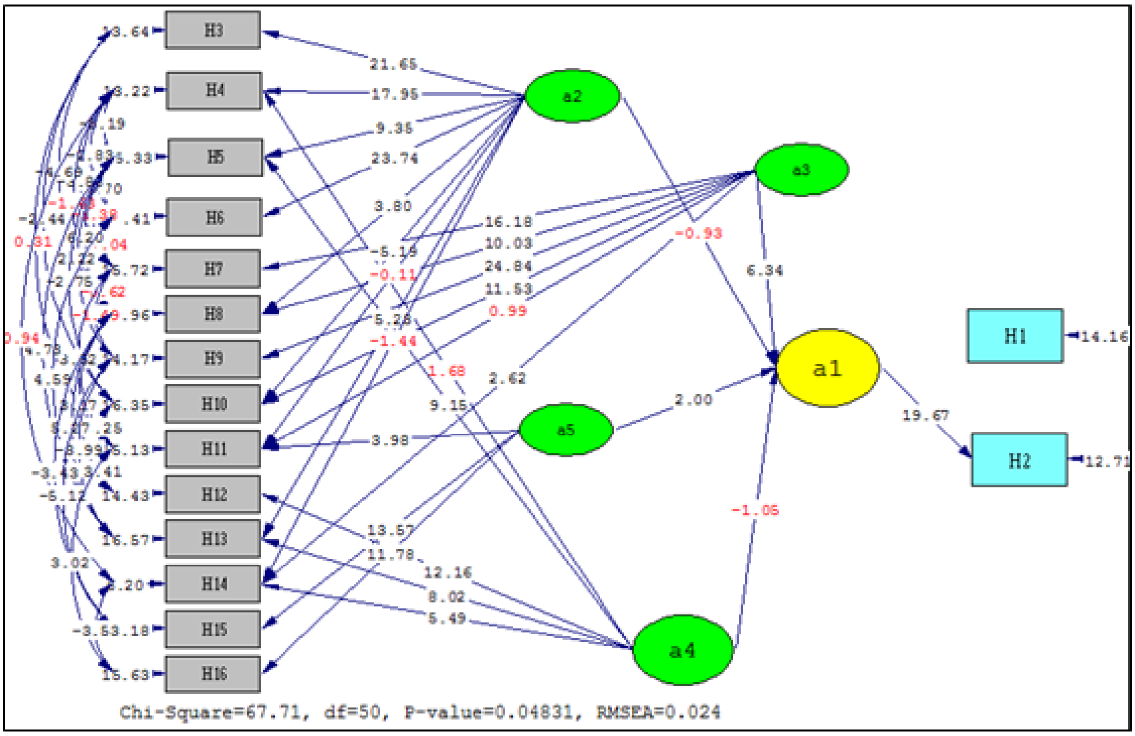

Fig. 5. Unstandardized results

Figure 4 contains the standardized path coefficients for the parameters, while the t-values of the non-statistically significant path coefficients on the path diagram are given in Figure 5. The t-values of the non-significant path coefficients are shown in red in the path diagram. When the t-values of the variables of guidance, solution, personnel and representation, which are thought to have an effect on general satisfaction, were examined, it was seen that the guidance ($t = -0.93$) and personnel ($t = -1.06$) variables were meaningless.

Interpreting Measurement Model Results

After the path diagram is obtained, the process of interpreting the analysis results is started. The results obtained are given below.

Table 2. Measurement model results

Factor / Expression	Standardized Loads	*t*-values	R^2
Factor a1			
H1	0.75		0.57
H2	0.79	19.67	0.62
Factor a2			
H3	0.77	21.65	0.60
H4	0.71	17.95	0.52
H5	0.44	9.35	0.34
H6	0.83	23.74	0.69
H8	0.25	3.80	0.73
H10	−0.46	**−5.19**	0.41
H11	−0.01	**−0.11**	0.23
H13	0.25	5.23	0.19
H14	−0.24	**−1.44**	0.71
Factor a3			
H7	0.63	16.18	0.39
H8	0.63	10.03	0.73
H9	0.83	24,84	0.69
H10	1.02	11.53	0.41
H11	0.16	**0.99**	0.23
H14	0.50	2.62	0.71
Factor a4			
H12	0.55	12.16	0.30
H13	0.35	8.02	0.19
H14	0.59	5.49	0.71
H4	0.08	**1.68**	0.52
H5	0.35	9.15	0.34
Factor a5			
H15	0.61	13.57	0.37
H16	0.51	11.78	0.26
H11	0.33	3.98	0.23

Measurement model results are given in Table 2. The standardized loads included in the measurement model results show the correlation between each observed variable and the implicit variable it is related to. Considering the first indicator of the implicit variable a1, H1, the correlation coefficient is 0.75. When the correlation coefficient is squared, R2 of H1 is 0.56. It is seen that the variability related to the implicit variable a1 is mostly explained by H2 (0.62). Fit criteria for the measurement model are given in Table 3.

Table 3. Fit criteria for the model

Fit Measurement	**Good Fit**	**Acceptable Fit**	**Results**
χ^2	$0 \leq \chi^2 \leq 2df$		67.71 $(P = 0.04831)$
P of Close Fit	≥ 0.05		1.00
$\frac{\chi^2}{df}$	$0 \leq \frac{\chi^2}{df} \leq 2$	$2 \leq \frac{\chi^2}{df} \leq 3$	1.35
RMSEA	$0 \leq$RMSEA≤ 0.05	$0.05 \leq$ RMSEA ≤ 0.08	0.024
NFI	$0.95 \leq$ NFI ≤ 1	$0.90 \leq$ NFI ≤ 0.95	0.99
NNFI	$0.97 \leq$ NNFI ≤ 1	$0.95 \leq$ NNFI ≤ 0.97	1.00
CFI	$0.97 \leq$ CFI ≤ 1	$0.95 \leq$ CFI ≤ 0.97	1.00
GFI	$0.95 \leq$ GFI ≤ 1	$0.90 \leq$ GFI ≤ 0.95	0.99
AGFI	$0.90 \leq$ AGFI ≤ 1	$0.85 \leq$ AGFI ≤ 0.90	0.96
PGFI	≥ 0.95		0.36
AIC			239.71
ECVI			0.39
IFI	≥ 0.95	$0.90 \leq$ IFI ≤ 0.94	1.00
RFI	≥ 0.90		0.99
Critical N			688.06

The adaptation criteria obtained for the model are presented in Table 3. When the results are examined, it is seen that the goodness of fit criteria are within the ranges recommended by the literature. It was seen that the value (1.35) obtained by dividing the Chi-square value by the degrees of freedom was within the acceptable range.

Table 4. Structural relationship coefficient values

Factor / Expression	**Standardized Loads**	***t*-values**	**R^2**
Factor a1			
a2	0.23	**1.61**	**0.93**
a3	0.59	3.20	
a4	0.32	4.28	
a5	−0.01	**−0.14**	

In Table 4, standardized loads and t-value are given regarding the structural relationship between the general satisfaction implicit variable and the implicit variables of guidance, solution, personnel and representation. According to the results, the path coefficients between the general satisfaction.

Table 5. Frequency disturbution from response to the questions

	Likert Scale					
	1	**2**	**3**	**4**	**5**	**Sum**
H1	2	67	119	294	134	616
H2	36	210	139	184	47	616
H3	16	127	177	168	128	616
H4	28	115	121	161	191	616
H5	27	97	197	130	165	616
H6	13	89	129	195	190	616
H7	45	106	184	183	98	616
H8	9	143	86	292	86	616
H9	30	159	157	217	53	616
H10	8	44	182	332	50	616
H11	8	51	220	249	88	616
H12	4	44	259	255	54	616
H13	3	42	206	190	175	616
H14	29	49	205	292	41	616
H15	4	57	122	361	72	616
H16	5	16	114	264	217	616

Implicit variable and the counselling and representation implicit variable were not found to be significant. Only the structural relationship between the general satisfaction implicit variable and the solution and personnel implicit variable was found to be significant. Bayesian solution has been implemented with WinBUGS package program. Before starting the Bayesian solution, frequency tables for 16 questions were prepared. The main purpose of extracting the frequency tables is to determine the percentage rates for each question of the 5-point Likert scale used in the questionnaire form and the threshold values required for analysis based on these rates. The frequency distribution of each question is given in the table below. The threshold values were started to be calculated by obtaining the percentages of the scale categories from the frequency table for each question. Threshold value calculation is made as one minus of the number of categories used in Likert scale. Threshold values in Table 6 were obtained from the reverse of the normal cumulative distribution by using the relevant frequency tables.

Table 6. Distribution percentage of response to the questions

	Likert Scale				
	1	**2**	**3**	**4**	**5**
H1	0.003247	0.108766	0.193182	0.477273	0.217532
H2	0.058442	0.340909	0.225649	0.298701	0.076299
H3	0.025974	0.206169	0.287338	0.272727	0.207792
H4	0.045455	0.186688	0.196429	0.261364	0.310065
H5	0.043831	0.157468	0.319805	0.211039	0.267857
H6	0.021104	0.144481	0.209416	0.316558	0.308442
H7	0.073052	0.172078	0.298701	0.297078	0.159091
H8	0.01461	0.232143	0.13961	0.474026	0.13961
H9	0.048701	0.258117	0.25487	0.352273	0.086039
H10	0.012987	0.071429	0.295455	0.538961	0.081169
H11	0.012987	0.082792	0.357143	0.404221	0.142857
H12	0.006494	0.071429	0.420455	0.413961	0.087662
H13	0.00487	0.068182	0.334416	0.308442	0.284091
H14	0.047078	0.079545	0.332792	0.474026	0.066558
H15	0.006494	0.092532	0.198052	0.586039	0.116883
H16	0.008117	0.025974	0.185065	0.428571	0.352273

The expressions to be used in the analysis in WinBUGS are given in the table below for both measurement models and structural equation model.

Structure of model in WinBUGS

Table 7. Symbolic representation of implicit and measurable variables

Factor	Questions	Node	Implicit	Variable	Node
a1	H1	1	a1	a2	gam[1]
a1	H2	lam[1]	a1	a3	gam[2]
a2	H3	1	a1	a4	gam[3]
a2	H4	lam[2]	a1	a5	gam[4]
a2	H5	lam[3]			
a2	H6	lam[4]			
a3	H7	1			
a3	H8	lam[5]			
a3	H9	lam[6]			
a3	H10	lam[7]			
a3	H11	lam[8]			
a4	H12	1			
a4	H13	lam[9]			
a4	H14	lam[10]			
a5	H15	1			
a5	H16	lam[11]			

Measurement of The Equality

```
for(j in 1:P){

y[i,j]~dnorm(mu[i,j],psi[j])I(thd[j,z[i,j]],thd[j,z[i,j]+1])
                ephat[i,j]<-y[i,j]-mu[i,j]
        }
        mu[i,1]<-eta[i]
        mu[i,2]<-lam[1]*eta[i]
        mu[i,3]<-xi[i,1]
        mu[i,4]<-lam[2]*xi[i,1]
        mu[i,5]<-lam[3]*xi[i,1]
        mu[i,6]<-lam[4]*xi[i,1]
        mu[i,7]<-xi[i,2]
        mu[i,8]<-lam[5]*xi[i,2]
        mu[i,9]<-lam[6]*xi[i,2]
        mu[i,10]<-lam[7]*xi[i,2]
        mu[i,11]<-lam[8]*xi[i,2]
        mu[i,12]<-xi[i,3]
        mu[i,13]<-lam[9]*xi[i,3]
        mu[i,14]<-lam[10]*xi[i,3]
        mu[i,15]<-xi[i,4]
        mu[i,16]<-lam[11]*xi[i,4]
```

Structural Equation

```
xi[i,1:4]~dmnorm(u[1:4],phi[1:4,1:4])
                eta[i]~dnorm(nu[i],psd)
                nu[i]<-gam[1]*xi[i,1]+gam[2]*xi[i,2]+gam[3]*xi[i,3]+gam[4]*xi[i,4]
                dthat[i]<-eta[i]-nu[i]
```

Threshold Values

```
thd=structure(
            .Data=c(-200.000.-2.722,-1.216,-0.510.
0.781,200.000.
-200.000.-1.568,-0.255, 0.319, 1.430.200.000.
-200.000.-1.944,-0.732, 0.049, 0.814,200.000.
-200.000.-1.691,-0.732,-0.180. 0.496,200.000.
-200.000.-1.708,-0.837, 0.053, 0.619,200.000.
-200.000.-2.031,-0.972,-0.319, 0.500.200.000.
-200.000.-1.453,-0.690. 0.110. 0.998,200.000.
-200.000.-2.180.-0.685,-0.289, 1.082,200.000.
-200.000.-1.658,-0.505, 0.155, 1.366,200.000.
-200.000.-2.227,-1.376,-0.306, 1.397,200.000.
-200.000.-2.227,-1.306,-0.118, 1.068,200.000.
-200.000.-2.484,-1.419, -0.004, 1.355,200.000.
-200.000.-2.585,-1.453,-0.234, 0.571,200.000.
-200.000.-1.674,-1.142,-0.102, 1.502,200.000.
-200.000.-2.484,-1.287,-0.533, 1.191,200.000.
-200.000.-2.404,-1.824,-0.775, 0.379,200.000),
            .Dim=c(16,6)),
```

Primarily, the point at which convergence was achieved was determined and this point was used as the burning period. Two methods were used to check whether convergence was achieved. The first of these is to examine the trace graphs for the related parameters. The trace graphs of the path coefficients for each measurement equations are given below. When the trace charts are examined, it is seen that the predicted values in the parameters gradually become stagnant and not take extreme values.

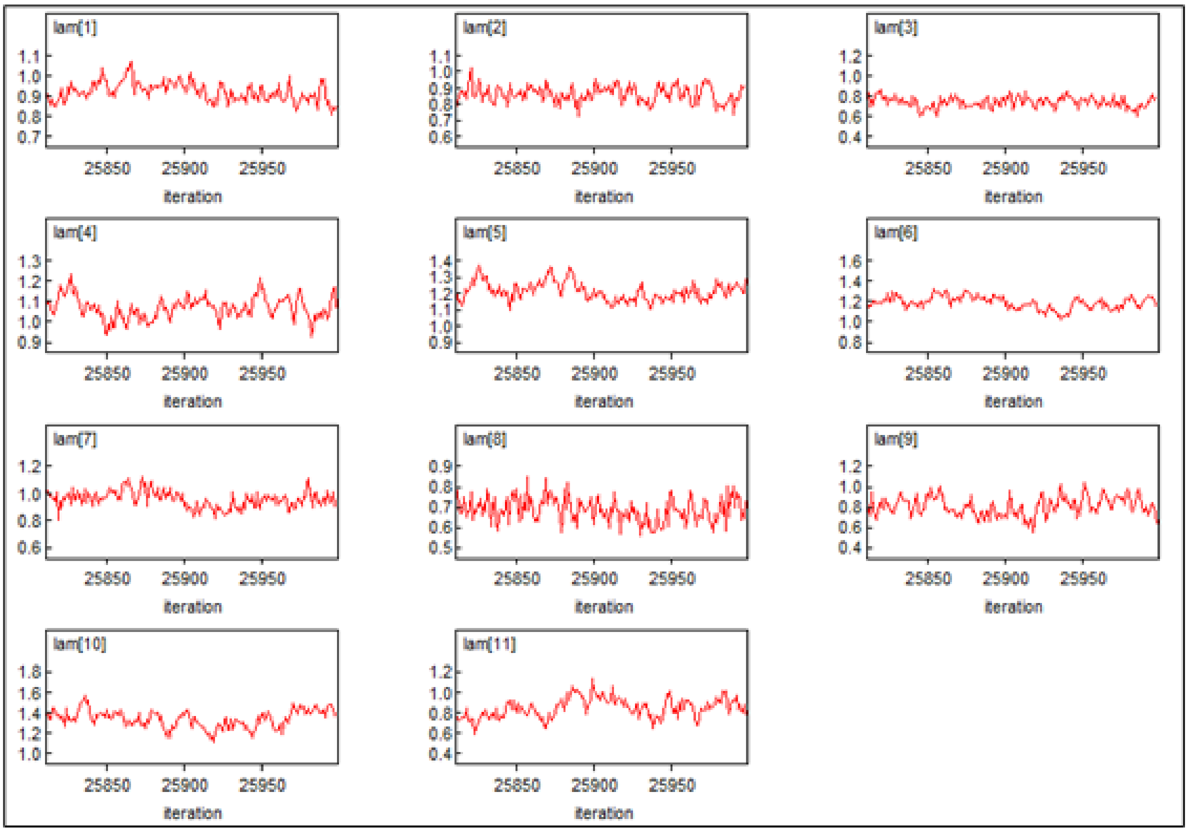

Fig. 6. Trace graphics related to measurement model parameters

When the trace graphs of the measurement model parameters in Figure 6 are examined, it is seen that there is no extreme value. After 11000 samples, it can be seen from the trace graphs that convergence is achieved.

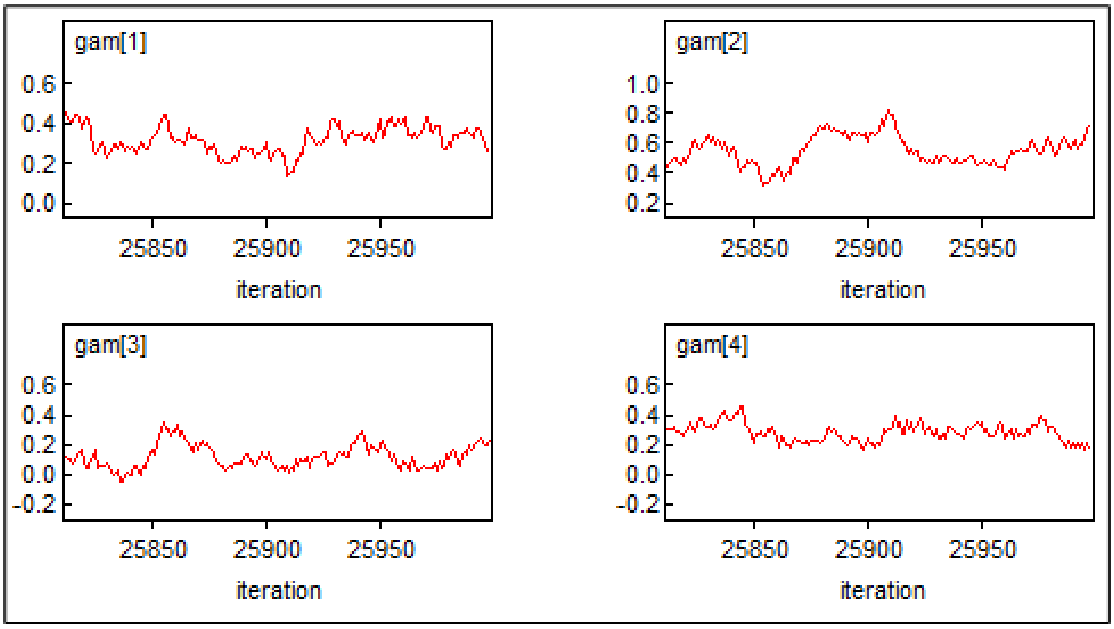

Fig. 7. Trace graphics related to measurement model parameters

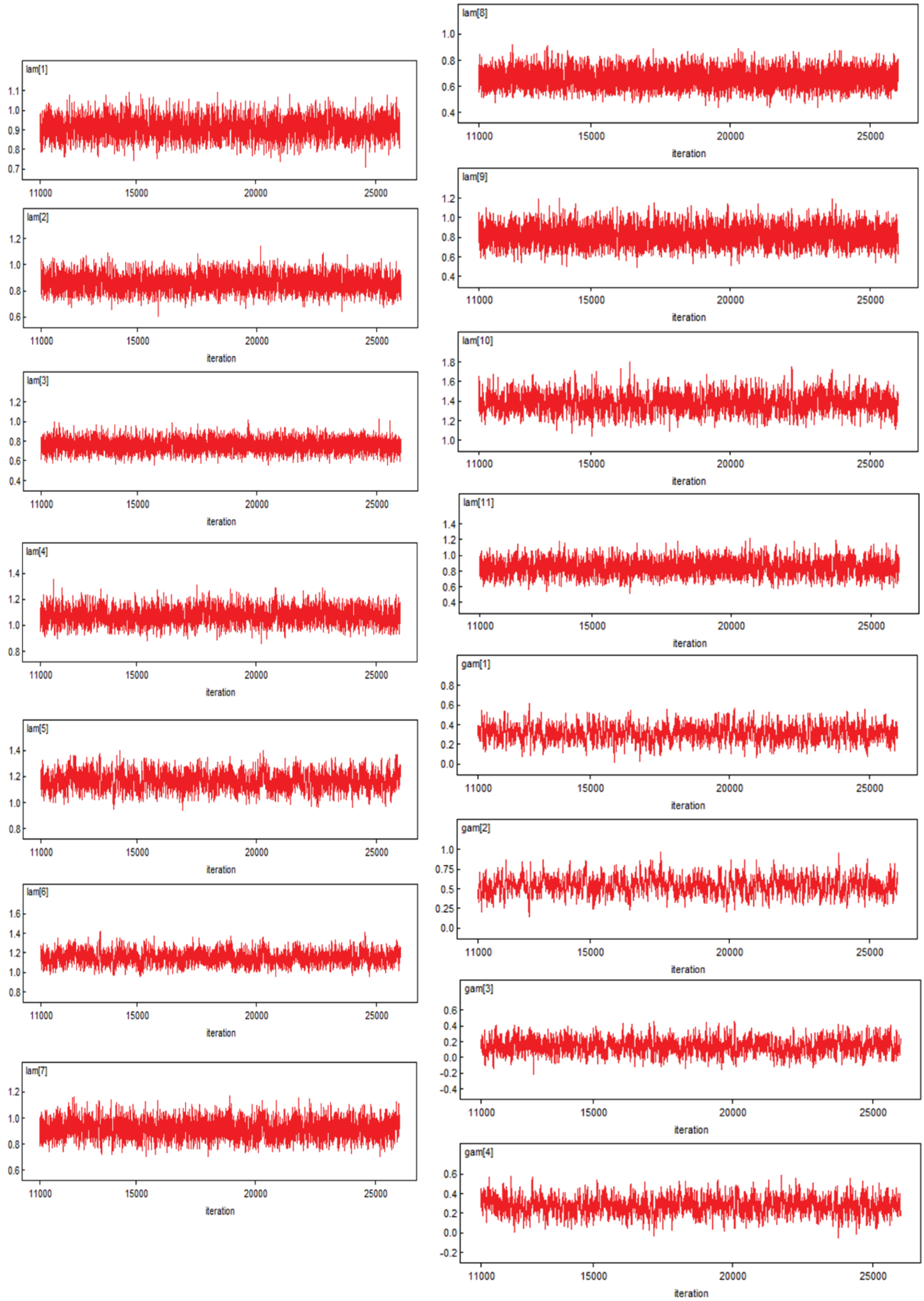

Fig. 8. Trace graphics related to structural model parameters

When the trace graphs of the structural equation model parameters in Figure 8 are examined, it is seen that there is no extreme value as in the measurement model parameters. After 11000 samples, it can be seen from the trace graphs that convergence is achieved. The interpretation of trace charts alone does not provide us with precise information on whether convergence is achieved. Secondly, the Thumb rule, which is a stronger method, is used to determine whether convergence is achieved. As a rule, MC errors for each parameter must be less than 5% of the standard deviation values. In the table below, it can be checked whether there is convergence for both the measurement and structural equation parameters.

Table 8. Convergence results of measurement model and structural equation model parameters

Node	Mean	Standard Deviation	Sd. (5%)	MC error	2.50%	Median	97.50%
gam[1]	0.3074	0.08184	0.004092	0.00328	0.1571	0.304	0.4626
gam[2]	0.5539	0.1152	0.00576	0.00549	0.348	0.5526	0.7694
gam[3]	0.1405	0.09059	0.00453	0.003564	−0.04055	0.1425	0.312
gam[4]	0.2714	0.08856	0.004428	0.003877	0.1012	0.2702	0.4458
lam[1]	0.9102	0.0487	0.002435	0.001313	0.8182	0.9095	1.005
lam[2]	0.8605	0.06729	0.003365	0.001976	0.7517	0.8582	0.9763
lam[3]	0.7595	0.06857	0.003429	0.002095	0.6467	0.757	0.8805
lam[4]	1.075	0.06825	0.003413	0.002101	0.9633	1.073	1.193
lam[5]	1.167	0.07395	0.003698	0.003261	1.045	1.166	1.296
lam[6]	1.154	0.07424	0.003712	0.003361	1.034	1.152	1.284
lam[7]	0.9194	0.07099	0.00355	0.00276	0.8021	0.9167	1.045
lam[8]	0.6654	0.06959	0.00348	0.002214	0.5485	0.6637	0.7891
lam[9]	0.8138	0.08949	0.004475	0.002305	0.65	0.811	0.9912
lam[10]	1.381	0.09121	0.004561	0.003211	1.217	1.378	1.565
lam[11]	0.8534	0.09682	0.004841	0.003385	0.6805	0.8497	1.04

As can be seen in Table 8, MC error values of all parameters related to measurement models and structural equation models are less than 5% of the standard deviation values. Buddha reveals the point at which convergence is achieved in the Bayesian solution more clearly than the trace graphs. The main purpose in finding the point at which convergence is achieved is to determine the burning period and to ensure that the estimates up to this period are not taken into account. After the burning period, 15000 updates were made, and parameter estimations were made for the model. The historical graphics for each parameter are represented in Figure 9.

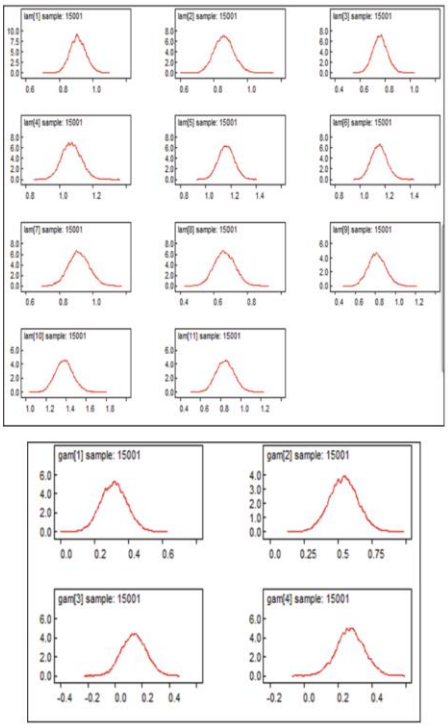

Fig. 9. History plots for parameters

Past graphs of the path coefficients related to both measurement models and structural equation model are given. From these graphs, it can be seen that there are no excessive fluctuations, and that each parameter converges. The Bayesian structural equation results obtained over 26000 samples, 11000 of which were taken using the burning period, are given in the table below.

Table 9. Bayesian prediction results

Node	Mean	Standard Deviation	Sd. (5%)	MC error	2.50%	Median	97.50%
gam[1]	0.3132	0.07689	0.003845	0.003078	0.1659	0.3129	0.4667
gam[2]	0.5436	0.1055	0.005275	0.004844	0.3391	0.5432	0.7545
gam[3]	**0.1401**	**0.08975**	**0.004488**	**0.003419**	**−0.03154**	**0.1398**	**0.3171**
gam[4]	0.2749	0.08222	0.004111	0.003308	0.1139	0.2741	0.4383
lam[1]	0.9111	0.04736	0.002368	0.001148	0.8211	0.9095	1.007
lam[2]	0.8605	0.05756	0.002878	0.001349	0.751	0.8595	0.9768
lam[3]	0.7586	0.05795	0.002898	0.001222	0.6478	0.7578	0.8758
lam[4]	1.073	0.05736	0.002868	0.001587	0.9642	1.071	1.188
lam[5]	1.166	0.0619	0.003095	0.002353	1.047	1.165	1.289
lam[6]	1.151	0.0609	0.003045	0.002406	1.036	1.15	1.275
lam[7]	0.9171	0.0623	0.003115	0.00192	0.7966	0.9156	1.042
lam[8]	0.6643	0.06102	0.003051	0.001395	0.5479	0.6627	0.7858
lam[9]	0.8176	0.0875	0.004375	0.002139	0.6521	0.815	0.9956
lam[10]	1.382	0.08879	0.00444	0.00332	1.216	1.381	1.563
lam[11]	0.8461	0.08819	0.00441	0.00268	0.6808	0.8449	1.023

Bayesian parameter estimation results are given in Table 9. Only the gamut [3] structural equation parameter is meaningless. Additionally, parameter estimates of classical and Bayesian measurement models were represented in the Table 10 and 11.

Table 10. Parameter estimates of classical and Bayesian measurement models

Factor/ Expression	LISREL	BAYES
Factor a1		
H1	0.75	1
H2	0.79	0.9111
Factor a2		
H3	0.77	1
H4	0.71	0.8605
H5	0.44	0.7586
H6	0.83	1.073
Factor a3		
H7	0.63	1
H8	0.63	1.166
H9	0.83	1.151
H10	1.02	0.9171
H11	0.16	0.6643
Factor a4		
H12	0.55	1
H13	0.35	0.8176
H14	0.59	1.382
Factor a5		
H15	0.61	1
H16	0.51	0.8461

Table 11. Classical and Bayesian structural model parameter estimates

Factor/ Expression	LISREL	BAYES
Factor a1		
a2	$\mathbf{0.23}$	0.31
a3	0.59	0.54
a4	0.32	$\mathbf{0.14}$
a5	$\mathbf{-0.01}$	0.27

With LISREL, it was seen that the implicit variables of counselling and representation did not have an effect on general satisfaction in the classical solution, only the solution and staff implicit variables were effective. In the Bayesian approach, contrary to the classical analysis, it was found that the counselling implicit variable and the representation implicit variable were significant, while the staff implicit variable was not significant.

4. Conclusion

In this study, classical structural equation models and Bayesian structural equation models are emphasized. Both approaches were applied to the survey data obtained from Samsun Chamber of Commerce and Industry. The data consists of 16 observed variables measuring 5 latent variables and 616 observations. Structural relationship and measurement models designed in classical analysis were created. First, the model fit indices were examined, and the analysis process was started. When the measurement models created in LISREL and the structural relationship were examined, it was seen that the model fit was not good and correction indices were needed. Correction indices were discussed in two parts. The correction indices in the first part show the relationships between observed variables and latent variables. The correction indices in this section show the decrease in the chi-square value calculated to evaluate model fit as a result of associating the observed variables under one latent variable with another latent variable. The correction indices in the second part are based on the independence of the errors for the observed variables. Since the model fit was not achieved in the initial solution, the analysis was performed using correction indices. In the initial solution, both the RMSEA value was higher than 0.08 and the division of the chi-square to the degrees of freedom was over 3. These values made it necessary to use correction indices in classical analysis. After the recommended corrections were made in the measurement model, it was determined that the model fit well according to all the criteria used in the assessment of goodness of fit. When the obtained fit indices were examined, the RMSEA value was found to be 0.024 and the division of the chi-square to the degrees of freedom was found to be 1.35. After examining the model fit indices, parameter estimation was started. Otherwise, it would not be reasonable to examine the parameter estimates unless the model fit is achieved. The significance of the parameter estimates related to the measurement models and the structural model was examined using t-values. Path coefficient with a t-value below 1.96 was considered to be insignificant. One of the most important advantages of the path diagram in classical analysis is that the meaningless relations are shown in shape and in different colours. This advantage allows for interpretation and viewing all relationships in a single photo. It was determined that the structural relationship between the counselling and representation implicit variables symbolized by a2 and a5 of the general satisfaction implicit variable was insignificant. While there was no significant relationship between the two latent variables and the general satisfaction implicit variable, a significant relationship was found between the other two latent variables (solution and staff). With this structural relationship obtained, 93% of the general satisfaction implicit variable is explained.

In the Bayesian structural equation model solution, model structures were defined in two stages: measurement models and structural model. By creating the frequency tables for the questions prepared using a 5-point Likert, (4) threshold values equal to one less than the number of Likert categories were calculated. The inverse of the cumulative normal distribution was used in calculating the threshold values. With the calculation of the threshold values, the model was started to be analysed. The most important step of the analysis phase is convergence. Unless convergence was achieved, model parameters were not estimated. It was seen that all parameters related to the model converged in 11000 iterations. Convergence has been examined in two stages. In the first stage, the trace graphs of each parameter were examined, and a stationary structure was observed. The interpretation of trace charts alone is not sufficient as a precise information. In the second stage, according to the Thumb rule; The condition that MC error values for each parameter should be less than 5% of the standard deviation value of the same parameter was examined. All parameters were found to meet this requirement and 11000 was used as the burning period. The purpose of using 11000 as the burning period is to ignore the parameter values in these iterations where convergence is not achieved. After the burning period, the parameters related to the model were obtained at the end of 15000 iterations. When the structural relationship between general satisfaction and the other 4 latent variables was examined, it was seen that the personnel implicit variable represented by a4 was meaningless, unlike the classical analysis.

Guidance, solution, personnel and representation factors affecting general satisfaction were examined in the data obtained from Samsun Chamber of Commerce and Industry, and it was revealed that classical and

Bayesian approaches give different results in terms of parameter estimates. While only the solution and personnel implicit variables were significant in the classical approach, the implicit variables of guidance, solution and representation were found to be significant in the Bayesian approach. This study was applied to the standard Samsun Chamber of Commerce and Industry questionnaire, whose data were prepared previously. The results can be obtained differently by redesigning and obtaining the questionnaire forms. Model comparisons were not emphasized in this thesis. In the classical approach, Akaike information criterion was calculated, and AIC, BIC and DIC calculations in Bayesian approach will be our future studies.

Although there are many a priori selection methods in the Bayesian approach, a priori selection in this study is limited to only conjugate a priori. Theoretical information is given about the use of other a priori such as Jeffrey's a priori. Therefore, a priori comparison and comparison of the adaptation criteria in the Bayesian approach has prepared a theoretical background for the studies to be carried out in the following years. In this study, the scale type is taken as a fixed 5-point Likert. There are no studies on the use of Bayesian structural equation when the types of scales are different. Studies in this field will provide new gains to the literature.

References

[1] K. A. Bollen, S*tructural Equations with Latent Variables*, in *Structural Equations with Latent Variables*. Wiley: New York, (1989) 1-9.

[2] S. Wright, *On The Nature of Size Factors.* Genetics 3(4) (1918) 367-74.

[3] L. Wolfle, *Sewall Wright on The Method of Path Coefficients: An Annotated Bibliography.* Structural Equation Modeling-a Multidisciplinary Journal - STRUCT EQU MODELING 6 (1999) 280-291.

[4] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistics, New York: Springer-Verlag, 1985.

[5] J. M. Bernardo, A. F. M. Smith, *Bayesian Theory*. Wiley Series in Probability and Statistics, 1994.

[6] R. E. Kass, A. E. Raftery, *Bayes Factors.* Journal of The American Statistical Association 90(430) (1995) 773-795.

[7] A. E. Raftery, *Bayesian Model Selection in Structural Equation Models.* Sage Focus Editions 154 (1993) 163-180.

[8] S. Y. Lee, X. Y. Song, *Hypothesis Testing and Model Comparison in Two-Level Structural Equation Models.* Multivariate Behavioral Research 36(4) (2001) 639-655.

[9] K. Jedidi, H. S. Jagpal, W. S. DeSarbo, *STEMM: A General Finite Mixture Structural Equation Model.* Journal of Classification 14(1) (1997) 23-50.

[10] T. J. DiCiccio, R. E. Kass, A. Raftery, L. Wasseman, *Computing Bayes Factors by Combining Simulation and Asymptotic Approximations.* Journal of the American Statistical Association 92(439) (1997) 903-915.

[11] A. Gelman, X. L. Meng, *Simulating Normalizing Constants: From Importance Sampling to Bridge Sampling to Path Sampling.* Statistical Science (1998) 163-185.

[12] S. Y. Lee, *Structural Equation Modeling: A Bayesian Approach*. Chichester: John Wiley & Sons, 2007.

[13] Y. F. Wang, T. H. Fan, *Bayesian Analysis of The Structural Equation Models with Application to a Longitudinal Myopia Trial.* Statistics in Medicine 31(2) (2012) 188-200.

[14] R. Guo, H. Zhu, S. M. Chow, J. G. Ibrahim, *Bayesian Lasso for Semiparametric Structural Equation Models.* Biometrics 68(2) (2012) 567-577.

[15] U. Olsson, *Maximum Likelihood Estimation of The Polychoric Correlation Coefficient.* Psychometrika 44(4) (1979) 443-460.

[16] S. Y. Lee, W. Y. Poon, P. Bentler, *Full Maximum Likelihood Analysis of Structural Equation Models with Polytomous Variables.* Statistics & Probability Letters 9(1) (1990) 91-97.

[17] J. C. Westland, *Lower Bounds on Sample Size in Structural Equation Modeling.* Electronic Commerce Research and Applications 9(6) (2010) 476-487.

[18] K. G. Jöreskog, D. Sörbom, *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, Ill.; Hillsdale, N.J.: SSI ; Erlbaum, Lawrence, Associates, 1996.

[19] L. K. Muthén, B. O. Muthén, *Mplus User’s Guide*. Los Angeles, CA: Muthén & Muthén, 2000.

[20] S. Y. Lee, W.Y. Poon, P. M. Bentler, *A Two-Stage Estimation of Structural Equation Models with Continuous and Polytomous Variables.* British Journal of Mathematical and Statistical Psychology 48(2) (1995) 339-358.

[21] B. A. Reboussin, K. Y. Liang, *An Estimating Equations Approach for the LISCOMP Model.* Psychometrika 63(2) (1998) 165-182.

[22] J. Q. Shi, S.Y. Lee, *Latent Variable Models with Mixed Continuous and Polytomous Data.* Journal of the Royal Statistical Society: Series B (Statistical Methodology) 62(1) (2000) 77-87.

[23] S. Geman, D. Geman, *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.* IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6(6) (1984) 721-741.

[24] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A H. Teller, E. Teller, *Equation of State Calculations by Fast Computing Machines.* The Journal of Chemical Physics 21(6) (1953) 1087-1092.

[25] W. K. Hastings, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications.* Biometrika 57(1) (1970) 97-109.