

# Bayesian Structural Equation Modeling: A Business Culture Application in Kenya

Mutitu Ephantus Mwangi, Antony Wanjoya

Jomo Kenyatta University of Agriculture and Technology, School of Mathematical Sciences, Nairobi, Kenya

## Email address:

e8mwass@gmail.com (M. E. Mwangi), awanjoya@gmail.com (A. Wanjoya)

## To cite this article:

Mutitu Ephantus Mwangi, Antony Wanjoya. Bayesian Structural Equation Modeling: A Business Culture Application in Kenya. *Science Journal of Applied Mathematics and Statistics*. Vol. 4, No. 2, 2016, pp. 37-42. doi: 10.11648/j.sjams.20160402.13

**Received:** October 12, 2015; **Accepted:** October 26, 2015; **Published:** March 16, 2016

---

**Abstract:** Structural equation modeling (SEM) is a multivariate method that incorporates regression, path-analysis and factor analysis. Classical SEM requires the assumption of multivariate normality to be met and large sample size, also choice is made either to ignore uncertainties or treat the latent variables as observed. National culture Data gathered in a study or survey may be inform of ordered categories and may not follow the assumptions of multivariate normality. This restricts the use of frequentist method of estimation. A Bayesian approach to SEM allows inclusion of this uncertainty and directly models the uncertainties in predictive models. In addition Bayesian SEM does not require constant variance normal disturbances and the sample size can be a small number. The development and application of Bayesian SEM has been relatively slow but it has been made possible by Gibbs sampler. The main purpose of the study was model National Culture in Kenya based on Hofstede model and business performance. Maximum likelihood Estimation was used to estimate the parameters in Classical SEM. Gibbs sampler algorithm was employed in Bayesian approach to SEM. This study used non-informative priors. The convergence of parameter was evaluated using proportional scale reduction procedure and trace and density plots. Data was gathered from employees in Nairobi through structured questionnaires. Bayesian SEM with non-informative prior gave the best estimates indicating that personal distance, individualism and long term orientation were significantly related to business performance. However, Uncertainty Avoidance had no significant relationship with business performance.

**Keywords:** Structural Equation Modeling, Maximum Likelihood Estimation, Markov Chain Monte Carlo, Proportional Scale Reduction

---

## 1. Introduction

### 1.1. Background of the Study

Concepts in social and other disciplines may be multidimensional, therefore univariate analysis which is the main form of analysis may result in estimates that are biased. Structural equation modelling (SEM) being a multivariate modelling technique can be used to examine hypothesized relationships between independent latent variables and dependent latent variables. Latent variables are quantities that are not measured directly or unobserved. Most of the models in statistics can be fit in SEM framework Skrondal and Rabe-Hesketh (2004). Majority of the literature in SEM uses frequentist estimation method where estimation is based on covariance matrix of all the outcomes and exposures that are observed Bollen (1989).

SEM tests the hypothesis that the matrix of variance of a set of measured variables is equal to the matrix of covariance implied by the hypothesized model. The relationship can be written as

$$\Sigma = \Sigma \theta$$

Where  $\Sigma$  is the population covariance matrix of a set of observed variables and  $\Sigma \theta$  is the population covariance matrix implied, a vector of model parameters. The vector defines the form of a particular SEM through the specification of means and intercepts, variances and covariance's, regression parameters, and factor loadings.

Fitting models in classical SEM requires certain assumptions to be fulfilled and a large sample size. If this condition is met then Maximum likelihood estimation is used to estimate parameters and their standard errors.

In most scenarios, data obtained in a study may violate this assumptions, for example Natural culture data may be in ordered categories, and clearly such data does not follow normal distribution. In this case the use of MLE estimation is restricted because parameter estimates and their standard errors tend to be biased, U. H. Olsson, T. Foss, S. T. & Howell. (2000).

Contrast to classical SEM where computation is based on matrix of sample covariance, Bayesian approach to SEM uses raw data observations and latent measurements are treated as missing data. Here the model is analysed on the basis of complete data set, R. Scheines, H. Hoijtink & Boomsma. (1999).

Culture refers to the way of life of a people, their values, and customs and how this influence individual behaviour in a society Peretomode (2012). Though there are a number of cultural models, Hofstede model is the most famous and methodologically supported theory of national culture, Tzeremes and Halkos (2008). This study was based on Hofstede model which identifies cultural dimensions namely power distance, uncertainty avoidance, individualism and long term orientation.

### 1.2. Literature Review

Majority of literature on SEM uses frequentist method of estimation, Laird and Ware (1982). A Bayesian SEM permits direct latent variables estimation and posterior distribution multi-modality to be seen, which is unnoticeable in frequentist estimation method, R. Scheines, H. Hoijtink & Boomsma. (1999). In contrast to traditional approach to SEM, Bayesian SEM does not depend on asymptotic assumptions or validation of routines for estimation or data procedures of model fit Levy (2011).

Previous research has shown that Bayesian approach to SEM give better estimation results for parameters than traditional approach to SEM in small samples, Lee, S. Y., & Song (2004). Inclusion of Prior information in Bayesian approach to SEM permits the model to mirror elementary beliefs about the circumstance and provision of knowledge for under identified parameters in traditional approach to SEM, R. Scheines, H. Hoijtink & Boomsma. (1999).

Lee and Song (2004) used Bayesian procedures to model data with ignorable and non-ignorable missingness. Jackman (2000) concluded that Bayesian approach to SEM offers a joining procedure to missing data and latent variables. Lee (2007) highlighted that the major gain of Bayesian approach lies in the ability of MCMC to estimate non-standard models which fail to be estimated by traditional approaches.

Little Studies directly relating Bayesian approach and traditional approach to SEM have been conducted. Price (2012) conducted a study and recommended that Bayesian SEM performs well mainly in circumstances where traditional theory fails such as in small samples and non-linear relationships. Levy, R., Mislevy, R. J., & Sinharay (2009) discusses how frequentist approach to SEM brings misunderstanding about discrete data models and modelling item responses, and how Bayesian approach to SEM helps to clear these misconceptions.

Bayesian approach to SEM perspective is based on probability and distributional reasoning thus poses some difficulties, similarly MCMC have been criticized being tough to implement, However Yuan, Y., & MacKinnon (2009) argues that it is easier to set up the procedure of estimation for MCMC than to follow the required procedures in classical SEM.

More than forty years have passed since the introduction of Hofstede cultural model that has brought national culture understanding of different countries. Tzeremes and Halkos (2008) highlights that the impact of National culture and the performance of the organization is a major concern and therefore Organizations or institutions have to synchronize their operations with National culture to gain competitive advantage, Newman and Nollen (1996). Little research has been done on Hofstede national culture dimensions in examining the business performance in different cultural groups in Kenya McCrae, R.R., Terraciano, A., Realo, A., & Allik (2008). This issue constitutes a gap in knowledge.

## 2. Methodology

### 2.1. Data Set

The study targeted employees in organization around Nairobi. Primary data was gathered from 109 employees in various organizations in Nairobi through a structured questionnaire. The study was based on national culture in Kenya.

The data in this study consisted of 35 indicator variables with four exogenous latent variables and one endogenous latent variable namely one outcome latent variable for Business performance (BP) which had 6 indicator variables, and four explanatory latent variables for National Business culture namely Power Distance (PD) which had 6 indicator variables and measured the layers of management between an individual employee and the highest level of management. The second was Uncertainty Avoidance (UA) which had 4 indicators and measured employee's comfort with unstructured environments that is unknown situations where surprising events may occur.

The third was Individualism (IND) which had 10 indicator variables and measured not only the degree to which an employee maintained her/his unique attributes, but also the degree to which she/he became integrated into the collective group. The fourth was Long-Term Orientation (LTO) which had 9 indicator variables and measured long-term values, such as perseverance and thrift, against short-term values such as respect for tradition, fulfilment of social obligations and avoiding personal embarrassment. All indicator variables were on 5-point Likert score summative regarding the importance of each item to the employee.

### 2.2. Classical SEM

Basic SEM model consists of measurement and structural equation. The Measurement equation in SEM approach is given by:

$$x_i = \Lambda \omega_i + \varepsilon_i, i = 1, \dots, n$$

Considering our model  $x_i$  is a  $35 \times 1$  vector of indicator variables describing the  $5 \times 1$  random vector of latent variables  $\omega_i$ .

$\Lambda$  Is  $35 \times 5$  matrix of the loading coefficients as obtained from the regression of  $x_i$  on  $\omega_i$

Where

$$\omega_i = (BP, PA, UA, IND, LTO)^T$$

And  $\varepsilon_i$  is  $35 \times 1$  random vectors of the measurement errors which follow  $N(0, \psi_\varepsilon)$ . It is assumed that for  $i = 1, \dots, n$ ,  $\omega_i$  is independent and follows a normal distribution  $N(0, \Phi)$  and uncorrelated with the random vector  $\varepsilon_i$

Matrix notation of measurement equation is as follows

$$X_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 & 0 \\ \lambda_{31} & 0 & 0 & 0 & 0 \\ \lambda_{41} & 0 & 0 & 0 & 0 \\ \lambda_{51} & 0 & 0 & 0 & 0 \\ \lambda_{61} & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & \lambda_{82} & 0 & 0 & 0 \\ 0 & \lambda_{92} & 0 & 0 & 0 \\ 0 & \lambda_{102} & 1 & 0 & 0 \\ 0 & 0 & \lambda_{123} & 0 & 0 \\ 0 & 0 & \lambda_{133} & 0 & 0 \\ 0 & 0 & \lambda_{143} & 0 & 0 \\ 0 & 0 & \lambda_{153} & 0 & 0 \\ 0 & 0 & \lambda_{163} & 0 & 0 \\ 0 & 0 & \lambda_{173} & 0 & 0 \\ 0 & 0 & \lambda_{183} & 0 & 0 \\ 0 & 0 & \lambda_{193} & 0 & 0 \\ 0 & 0 & \lambda_{203} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda_{224} & 0 \\ 0 & 0 & 0 & \lambda_{234} & 0 \\ 0 & 0 & 0 & \lambda_{244} & 0 \\ 0 & 0 & 0 & \lambda_{254} & 0 \\ 0 & 0 & 0 & \lambda_{264} & 0 \\ 0 & 0 & 0 & \lambda_{274} & 0 \\ 0 & 0 & 0 & \lambda_{284} & 0 \\ 0 & 0 & 0 & \lambda_{294} & 1 \\ 0 & 0 & 0 & 0 & \lambda_{315} \\ 0 & 0 & 0 & 0 & \lambda_{325} \\ 0 & 0 & 0 & 0 & \lambda_{335} \\ 0 & 0 & 0 & 0 & \lambda_{345} \\ 0 & 0 & 0 & 0 & \lambda_{355} \end{bmatrix} \begin{bmatrix} PD \\ UA \\ IND \\ LTO \\ BP \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \\ e_{10} \\ e_{11} \\ e_{12} \\ e_{13} \\ e_{14} \\ e_{15} \\ e_{16} \\ e_{17} \\ e_{18} \\ e_{19} \\ e_{20} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{24} \\ e_{25} \\ e_{26} \\ e_{27} \\ e_{28} \\ e_{29} \\ e_{30} \\ e_{31} \\ e_{32} \\ e_{33} \\ e_{34} \\ e_{35} \end{bmatrix}$$

To obtain the structural equation the latent variable  $\omega_i$  is partitioned into  $(\eta_i, \xi_i)$  where  $\eta_i$  and  $\xi_i$  are  $m \times 1$  and  $n \times 1$  vectors of latent variables respectively. The structural equation is then expressed as

$$\eta_i = \beta \eta_i + \Gamma \xi_i + \delta_i, i = 1, \dots, n$$

Where  $\beta$  is  $m \times m$  matrix of structural parameters depicting the relationship between the endogenous latent variables which is assumed to have zeros in the diagonal,  $\Gamma$  is  $m \times n$  matrix of regression parameter governing the relationship of endogenous latent variables and exogenous latent variables.

$\delta_i$  is  $m \times 1$  vector of disturbances which is assumed  $N(0, \psi_\delta)$  where  $\psi_\delta$  is a diagonal covariance matrix.  $\delta_i$  is also assumed to be uncorrelated with  $\xi_i$ .

Our study contains one endogenous latent variable and therefore  $\eta_i = 0$ , the resulting equation becomes

$$\eta_i = \Gamma \xi_i + \delta_i$$

Finally the structural equation was modelled as follows:

$$BP = \gamma_1 PD + \gamma_2 UA + \gamma_3 IND + \gamma_4 LTO + \delta_i$$

Where  $(PA, UA, IND, LTO)^T$  follows  $N(0, \Phi)$  distribution and is independent with  $\delta_i$ .

### 2.3. Bayesian SEM

Under Bayesian SEM data matrix are let to be

$$Y = (y_1, y_2, \dots, y_n)$$

and

$$X = (x_1, x_2, \dots, x_n)$$

And the matrix of latent variables is let to be

$$\Omega = (\omega_1, \omega_2, \dots, \omega_n)$$

The structural parameter  $\theta$  contains

$$\Phi, \psi_\delta, \psi_\varepsilon, \Lambda, \text{ and } \Lambda_\omega$$

Bayesian estimates  $\theta$  and  $\Omega$  are obtained by applying Markov chain Monte Carlo (MCMC). This is done through Gibbs sampler which generates a sequence of random observations from the joint distribution  $\theta, \Omega | Y$ .

### 2.4. Gibbs Sampler

Considering lee (2007), Gibbs sampler generates a sequence of random observation and is carried out as follows:

Gibbs sampler algorithm starts with initializing

$$(\theta^{(0)}, \Omega^{(0)}, Y^{(0)})$$

Then simulation is done for

$$(\theta^{(1)}, \Omega^{(1)}, Y^{(1)})$$

The  $r^{th}$  iteration becomes,

$$(\theta^{(r)}, \Omega^{(r)}, Y^{(r)})$$

The Gibbs sampler is implemented as follows:

- Generate  $\Omega^{(r+1)}$  from  $(\Omega | \theta^{(r)}, Y^{(r)})$
- Generate  $\theta^{(r+1)}$  from  $(\theta | \Omega^{(r+1)}, Y^{(r)})$
- Generate  $Y^{(r+1)}$  from  $p(Y | \Omega^{(r+1)}, \theta^{(r+1)})$

Samples converge to the desired posterior distribution

under regulatory conditions. Determination of posterior distribution involves selection of prior distribution for  $(\Lambda, \psi_\varepsilon)$  and  $\Phi$ .

The prior distribution for  $(\Lambda, \psi_\varepsilon)$  and  $\Phi$  can be gotten from conjugate type distribution, this is by letting  $\psi_{\varepsilon k}$  be the  $k^{th}$  diagonal element of  $\psi_\varepsilon$  and  $\Lambda_k$  be the  $k^{th}$  row of  $\Lambda$ .

This study considered

$$(\Lambda_k | \psi_{\varepsilon k}^{-1}) \sim N(\Lambda_{0k}, \psi_{\varepsilon k}, H_{0yk})$$

$H_{0yk}$  is a hyper parameter assumed to be described by uninformative prior distribution.

### 2.5. Parameter Convergence

Intra –chain and inter-chain convergence of the sampling chains were assessed. The coda package in R was used to plot sampling trace and a sampling density for every parameter that was sampled and was present in the MCMC samples generated. The model was estimated with 10,000 total iterations, 5,000 burn-in and 5,000 post-burn-in, thin steps were 500. The number of chains were 2, number of iteration per chain was 100000.

Proportional scale reduction (PSR) by Gelman, A. & Rubin (1992) was also used to determine the convergence of parameters. PSR is a ratio that examines the variation between chains  $B$  and the variation within chains.

The proportional scale reduction is computed as follows

$$PSR = \sqrt{\frac{\hat{W} + \hat{B}}{\hat{W}}}$$

Where

$$\hat{B} = \frac{1}{m-1} \sum_{j=1}^m (\bar{\theta}_{.j} - \bar{\theta}_{..})^2$$

$$\hat{W} = \frac{1}{m} \sum_{j=1}^m \frac{1}{n} \sum_{i=1}^n (\bar{\theta}_{ij} - \bar{\theta}_{.j})^2$$

The number of chains is, the number of iterations is  $n$ ,  $\bar{\theta}_{ij}$  is the parameter value estimated at  $i^{th}$  iteration and  $j^{th}$  chain. Good parameter convergence is seen when PSR ratio is close to 1.

### 2.6. Bayesian Prior

Prior distributions specified for the parameters are updated through the data likelihood to form posterior distributions under Bayesian approach to SEM.

The study assumed that there was no prior knowledge thus a non-informative prior of a normal distribution with mean  $\mu = 0$  and very variance  $\delta^2 = 10^{10}$  was used to provide no useful information about the parameters that it describes. Here the assumption was that there was no prior beliefs about the regression parameters of national business culture and Business performance.

## 3. Results and Discussion

### 3.1. Classical SEM

Table 1. Fitted model for classical SEM.

| Regression | Estimate | Std.err | Z-value | P(> Z ) |
|------------|----------|---------|---------|---------|
| BP~        |          |         |         |         |
| PD         | 0.616    | 0.461   | 1.336   | 0.182   |
| UA         | -0.061   | 0.130   | -0.469  | 0.639   |
| IND        | -0.431   | 0.441   | -0.978  | 0.328   |
| LTO        | -0.283   | 0.237   | -1.195  | 0.232   |

RMSEA = 0.086, CFI = 0.828 and TLI = 0.898.

$$BP_{classSEM} = 0.616PD - 0.061UA - 0.431IND - 0.283LTO$$

This study found out that though model for classical SEM fit the data reasonably well as indicated by model fits, all culture dimensions (Power Distance (PD), Uncertainty Avoidance (UA), Individualism (IND) and Long-Term Orientation (LTO)) did not significantly predict business performance (BP) at 5% level of significance since the p-values were greater than 0.05.

### 3.2. Bayesian SEM

Table 2. Bayesian estimates of structural model using non informative prior.

| Parameters | Estimate | Std.err |
|------------|----------|---------|
| $\gamma_1$ | 0.720*   | 0.170   |
| $\gamma_2$ | -0.092   | 0.069   |
| $\gamma_3$ | -0.549*  | 0.078   |
| $\gamma_4$ | -0.386*  | 0.084   |

\* significant at 5% level

$$BP_{BaySEM} = 0.720PD - 0.092UA - 0.549IND - 0.386LTO$$

Table 2 indicates that personal distance has the greatest effect on Business performance then followed by individualism and then long term orientation.

The relationship between personal distance and business performance was significant in Bayesian SEM. Personal distance which reflect the level of perceived power distance felt by respondent positively and significantly influenced Business performance with an estimate of 0.720.

This study also found that individualism had a significant but negative influence on business performance with an estimate of -0.549. An increase in individualism reduces team work and consequently poor performance.

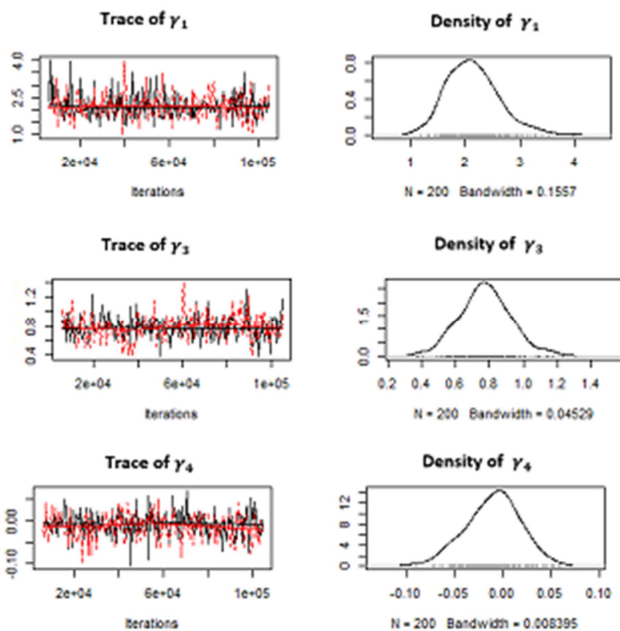
Long term orientation was also significant and had a negative influence on organizational performance with an estimate of -0.386. Uncertainty Avoidance had no significant relationship with Business performance. Compared to ML estimation, Bayesian estimation tends to result better estimates and smaller errors.

### 3.3. Parameter Convergence

#### 3.3.1. Trace and Density Plots

Trace plots and density plots were used to examine intra-

chain and inter-chain convergence.



**Figure 1.** Trace plot and density plot for three parameters of the Bayesian model using non-informative prior showing a good solution.

The plots on the left of the figure shows the trace plots for sampled values for the two chains while to the right is density plot the which presents the sampled values overall frequency for the two chains. The algorithm convergence as indicated by all plots was in less than 100000 iteration. The convergence plots exhibit a tight, horizontal band for both of the parameters presented. This tight band indicates the parameters likely converged properly. The plots shows that the two sampling chains had stable average each and they converged on the same values.

This was confirmed by the density plots on the right that showed a fairly normal distribution.

### 3.3.2. Gelman's Potential Scale Reduction Factor (PSRF)

**Table 3.** Potential scale reduction factor (PSRF).

| Parameters               | Point Estimate | Upper C.I |
|--------------------------|----------------|-----------|
| $\gamma_1$               | 0.997          | 1.00      |
| $\gamma_2$               | 0.998          | 1.00      |
| $\gamma_3$               | 1.010          | 1.05      |
| $\gamma_4$               | 1.038          | 1.17      |
| Multivariate psrf = 1.01 |                |           |

Table 3 indicates PSRF that is less than 1.1 and by [Gelman et al., 2004] this is good convergence, which is true for all parameters.

## 4. Conclusions and Recommendations

The purpose of this study was to construct a business culture model using data gathered from different

organizations in Nairobi (the Capital city in Kenya). Bayesian SEM was implemented using OpenBUGS and R. To compare the results Classical SEM was also implemented using lavaan package in R.

This study concluded, that the Bayesian approach to SEM performed better than the classical SEM for small samples, the classical estimation method gave almost similar estimates to Bayesian SEM but the error of the estimates were much higher in classical SEM as compared to Bayesian SEM. Under Bayesian SEM, this study used uninformative priors. Since the study used small samples. Out of the two models obtained estimates of Bayesian SEM with non-informative prior were the best.

The model indicated that personal distance has the greatest effect on the Business performance then followed by individualism and then long term orientation. Uncertainty Avoidance had no significant relationship at 5% level of significance.

The study found out power distance positively and significantly affected business performance in that the more the power distance the higher the business performance this might be due to the influence of managers to their subordinates through sanctions and rewards.

The study found out that individualism negatively and significantly affected the business performance. An increase in individualism leads to reduction in performance. This is because when people become more individualistic they seek independence and power and therefore focus on personal goals than their work and consequently perform poorly.

Long term Orientation had a negative and significant relationship with performance of the organization. This study shows that an increase in Long -term Orientations may result in poor performance, this may be due majority of the organizations heads focus much on short term results and generally employees order their lives based on abilities. In other words the negative influence of LTO indicates that the participants in the sample value more immediate gratification of their needs and have a much stronger focus on spending, analytic thinking and therefore perform poorly. Uncertainty Avoidance had no significant relationship with business performance.

This research presents proof that national cultural values affects the performance of organization. This study therefore recommends that the organizations in Kenya should be attentive and concerned about the negative influence of performance attributed by national cultural values.

Since Bayesian approach to SEM performed better than the classical SEM, this the study recommends the use of Bayesian approach to SEM in cases where frequentist theory breaks down such as with small samples.

## References

- [1] Bollen, K. (1989), 'Structural equations with latent variables. wiley & sons, new york'.

- [2] U. H. Olsson, T. Foss, S. T. & Howell. R. (2000), 'The performance of ml, gls, and wls estimation in sem under conditions of misspecification and non-normality.', *Structural Equation Modeling* 7, 557–595.
- [3] Skrondal, A. and S. Rabe-Hesketh (2004), 'Generalized latent variable modeling: multilevel longitudinal, and structural equation models'.
- [4] R. Scheines, H. Hoijtink & Boomsma., A. (1999), 'Bayesian estimation and testing of structural equation models.' *Psychometrika* 64, 37–52.
- [5] Lee, S. Y., & 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.
- [6] Lee, S. Y. (2007), 'Bayesian estimation of structural equation models. in s. y. lee(ed.)', *Structural Equation Modeling: A Bayesian Approach* pp. 67–109.
- [7] Yuan, Y., & MacKinnon D. P. (2009), 'Bayesian mediation analysis', *Psychological Methods* 14(4), 301–322.
- [8] Tzeremes, N. G. and G. E. Halkos (2008), 'Does the home country's national culture affect mncs performance? empirical evidence of the worlds top 100 east west non financial mnc's', *Global Economic Review* 37(4), 405–427.
- [9] McCrae, R. R., Terraciano A. Realo A. & Allik J. (2008), 'Interpreting globe societal practices scales', *Journal of Cross-Cultural Psychology* 39(6), 805–810.
- [10] Newman, K. L and S. D Nollen (1996), 'Culture and congruence: The fit between management practices and national culture', *Journal of International Business Studies* 27(4), 753–778.
- [11] Gelman, A., Carlin J. B. Stern H. S. and D. B. Rubin (2004), *Bayesian Data Analysis*, 2nd ed edn, CRC Press., London.