

# Tourism sector, Travel agencies, and Transport Suppliers: Comparison of Different Estimators in the Structural Equation Modeling

Nataša Kovačić<sup>1</sup>, Darja TOPOLŠEK<sup>2</sup> and Dejan DRAGAN<sup>2</sup>

<sup>1</sup> University of Rijeka / Faculty of Tourism and Hospitality Management, Opatija, Croatia

<sup>2</sup> University of Maribor/Faculty of Logistics, Celje, Slovenia

**Abstract** — *The paper addresses the effect of external integration (EI) with transport suppliers on the efficiency of travel agencies in the tourism sector supply chains. The main aim is the comparison of different estimation methods used in the structural equation modeling (SEM), applied to discover possible relationships between EI-s and efficiencies. The latter are calculated by the means of data envelopment analysis (DEA). While designing the structural equation model, the exploratory and confirmatory factor analyses are also used as preliminary statistical procedures. For the estimation of parameters of SEM model, three different methods are explained, analyzed and compared: maximum likelihood (ML) method, Bayesian Markov Chain Monte Carlo (BMCMC) method, and unweighted least squares (ULS) method. The study reveals that all estimation methods calculate comparable estimated parameters. The results also give an evidence of good model fit performance. Besides, the research confirms that the amplified external integration with transport providers leads to increased efficiency of travel agencies, which might be a very interesting finding for the operational management.*

**Key words**—Tourism Sector, Structural Equation Modeling, Estimation methods, External Integration, Efficiency of Travel Agencies.

## I. INTRODUCTION

Structural equation modeling (SEM) is a family of advanced statistical tools for modeling the relationships between different types of variables. It can deal with an enormous number of exogenous and endogenous variables, as well as with unobserved latent variables (factors - constructs) expressed by linear combinations of the measured indicator (manifest - item) variables [1, 2]. Since some of the variables involved in SEM are latent, structural equation modeling is sometimes also called as latent variable modeling. One of the primary goals of SEM is an estimation of causal effects between addressed variables. For this reason, SEM modeling has been also referred as causal path modeling [1]. The latter represents the extension of multiple regression analysis and enables efficient background for modeling the complex causal relationships among the multiple variables [3].

Another name for SEM modeling is covariance structures modeling, since the investigation of particular covariance and correlation patterns among the treated variables is engaged here and the covariance analysis methods are used for SEM estimation [1, 2, 3]. Different names for SEM modeling are all consistent with Bollen's definition from 1989, who proposed that SEM be based on three main analytical methodologies: (1) path analysis, (2) latent variable analysis and modeling, and (3) covariance estimation methods [2, 4].

In statistical manner, SEM can be treated as an integration, generalization, and extension of familiar general linear statistical models such as multiple regression modeling, analysis of variance (ANOVA), and factor analysis [1]. Due to a broad spectrum of covariance analysis methods, which can provide accurate estimates, SEM can be conducted for different types of data, such as continuous, ordinal, longitudinal, cross-sectional, and so [1, 2, 3, 5].

Since SEM is a confirmatory type of modeling, we can simplistically say that it combines the multiple regression analysis, simultaneous equations models, and confirmatory factor analysis (CFA) into the comprehensive statistical modeling framework [1, 2, 3]. SEM model comprises two main submodels, measurement and structural submodel, which can be estimated simultaneously [2]. The

latter means that the relations between observed indicators and latent variables (measurement part) and the causal relations between the latent variables among themselves (structural part) can be evaluated in a single model [1, 3]. The only condition for simultaneous estimation of both submodels is the ensuring of the full-information estimation methods [1, 3].

Applications of SEM that concentrate exclusively on the relationships between latent variables and their observed indicators are usually referred to as CFA analysis [1]. In this case, the SEM model comprises only one part, that is measurement submodel. Contrary to exploratory factor analysis (EFA), the CFA investigates how well the measured indicators characterize the unobserved latent variables. For this purpose, individual statistical tests are applied while the factor structures are hypothesized in advance and then verified empirically. Although EFA does not test a certain theory, but only derive particular factor structure from the data, it can be a still useful preliminary guideline for subsequent CFA. This way, the nature of the latent variables can be initially inspected, and a preliminary insight of the relationships between factors and their measured indicators can be initially provided [3, 6].

One of the main aims of SEM modeling is to explore whether the hypothesized theoretical model consistently reflects the measured data [1]. For this purpose, the different Goodness-of-fit (GOF) indices are used to verify if a model defined by the researcher is consistent with variance-covariance patterns in the data [2]. By other words, this means that the GOF indices support us to identify the level of plausibility and adequacy of assumed relationships between the variables addressed in the SEM model [1, 3].

According to [1, 2], the most frequently used estimation methods in SEM modeling are: maximum likelihood (ML) method, generalized least squares (GLS) method, unweighted least squares (ULS) method, weighted least squares (WLS) methods, also called asymptotically distribution-free (ADF) methods, and Bayesian Markov Chain Monte Carlo (BMCMC) methods. All mentioned estimators have their advantages and weaknesses, such as issues about normality violations, appropriateness of sample size, and so.

The ML estimator is probably the most popular among researchers because it is justly robust against violations of normality conditions [2]. Its solution maximizes the probability that the observed covariances belong to the population, which has its variances and covariances produced by the process implied by the model, where a multivariate normal distribution is assumed [2].

Surprisingly, despite the extensive use of SEM modeling in many areas, a relatively little research has been reported in the scholarly literature on comparison of achieved results with several different estimation methods applied to the same real cases [7]. This is particularly true for the tourism sector and tourism supply chains, which are addressed in this study. More precisely, the paper addresses an investigation of possible impacts of external integration (EI) with transport providers on the efficiency (EFF) of travel agencies. For this purpose, the SEM model is constructed on the basis of questionnaires' data collected in the survey, which was proceeded among the chosen Croatian agencies.

The main aim of the SEM model is to identify how the differences in the integration level regarding the different kinds of suppliers (water, air, bus, and rail) influence on the efficiency of the agencies. From existing literature, it is evident that practically none of similar research has been done in the field. But as noted in some previous studies [8-14], the amplified integration of members in the tourism supply chains definitely leads to improved performance and bigger quality of services. Since there is a quite big gap detected in the existing literature about similar kind of research as it was ours, we believe that the findings of this study might serve as a major contribution of this paper.

The efficiencies of the travel agencies are calculated by the means of data envelopment analysis (DEA). During the construction of SEM model, the EFA and CFA analyses are also employed as preliminary stages. Namely, it is recommended to conduct the CFA alone before the estimation of the entire SEM model since its measurement part must be firstly separately statistically evaluated. The reason is to verify independently if hypothesized factor model, reflected in its indicators, adequately fits the real data.

For the estimation of parameters of given SEM model, three different estimators, ML estimator, BMCMC estimator, and ULS estimator, are applied. For these estimators, their characteristics and different properties are shortly stressed. Afterward, the comparison of achieved estimation results by these estimators is performed and discussed.

When a SEM model is finally estimated and appropriately evaluated, it can be used to identify the relationships between the measured item indicators and the latent factors EI-s. Also, the causal directional paths from these factors to the latent factor EFF can be investigated. While doing these calculations, the program package IBM SPSS V21 and its extension AMOS were applied.

## II. THE MAIN STEPS OF SEM MODELING

The main steps of SEM modeling procedure are depicted in Fig. 1 [3, 10, 15]. In the first step, the data (usually gained by the means of the survey) are collected, and all statistical measures are selected. Herein, possible assumptions are also applied, such as those about the sample size, normality level of the data, the outliers, and so. In the second step, the suitable model must be specified, and possible causal paths between the variables must be defined according to the certain theory. Afterward, identification and estimation of the model are the next steps of the modeling process.

While proceeding model identification, the caution must be taken about so-called over-identification [1, 15]. In the process of model estimation, proper estimation method must be chosen due to satisfactory performance of GOF indices. The model verification follows in the SEM modeling development. At this step, the model fit quality is evaluated via the calculation of different fit indices. If they indicate a poor model performance, the model must be adjusted and re-specified. Otherwise, the constructed model is prepared for suitable interpretation and the report of computed results.

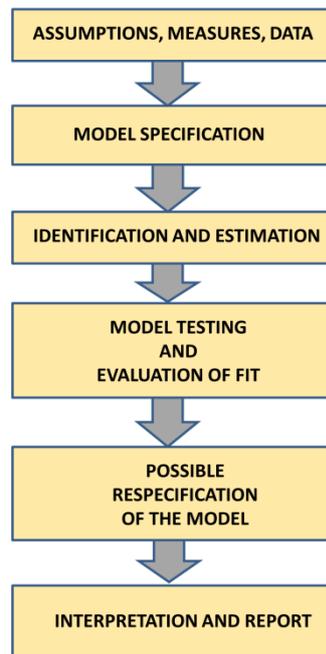


Figure 1. The basic steps of SEM modeling process

## III. ESTIMATION METHODS IN SEM MODELING PROCESS

Model estimation is one of the most important tasks of any SEM application. The quality of estimated parameters, their accompanying standard error estimates, and model fit statistics rely on the selection of appropriate estimation methods [1]. As generally in statistics, necessary properties of estimators comprise asymptotic unbiasedness, consistency, and efficiency.

The basic components of data for SEM modeling and analyses are sample variances and covariances of measured variables [1]. When a hypothesized SEM model is set, individual measured variables can be expressed as a function of unknown parameters (path coefficients or factor loadings) and other measured or unobserved variables in the model [1]. These functions refer to structural relations among the variables and are also called as "structural equations". From these equations, variances and covariances of measured variables can be expressed in relationship to unknown parameters in the model (i.e. causal path weights, factor loadings, and variances and covariances of latent factors). Since these variances and covariances are specific for a given model (model-specific), they are called model-implied variances and covariances [1].

Afterward, during the estimation process, the main issue is to find such model parameters that the model-implied variances and covariances are as close as possible to the observed sample variances and covariances [2].

Fig. 2 shows most frequently used estimation methods, used in SEM modeling process. All the details about those methods can be investigated in the scholarly literature [1, 3, 6, 15]. In the sequel, the main properties of corresponding estimators will be explained.

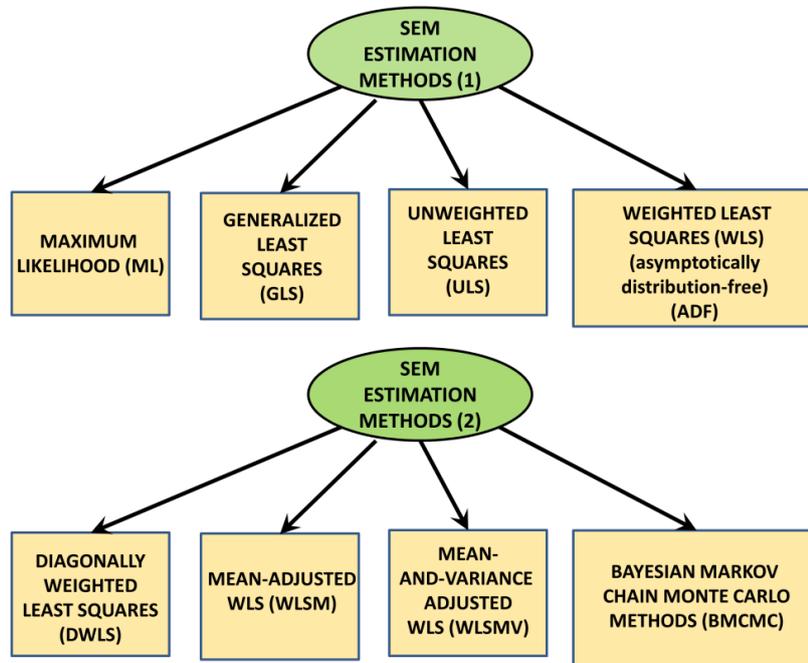


Figure 2. SEM estimation methods

**ML estimator** is most commonly used estimator in SEM modeling. In general, it requires certain statistical assumptions about the normality and provides quite accurate estimates in the case of continuous variables with normal distribution [15]. The possible non-normality of the data is usually detected with the computation of skewness index (SI) and kurtosis index (KI) of the measured variables. If SI and KI indices provide the evidence of at least approximate normality, then the application of ML estimator is not problematic. But in the case of ordinal data, the decision about the suitability of this estimator can become more complex. In general, most of the researchers agree that the ML estimator can be applied in the case of ordinal variables, if they have at least 5 categories and do not significantly depart from the normality conditions [3, 5, 15].

For the ML estimator, the following conditions must be (at least approximately) fulfilled [1, 3]:

- the normality of univariate distributions of the addressed variables;
- the bivariate normality of the joint distribution of any pair of addressed variables;
- the homoscedasticity and linearity of bivariate scatterplots.

This estimator associates ML method with the observed covariance matrix. During the ML estimation procedure, the parameters, which minimize the difference between data sample covariance matrix and model implied covariance matrix, are estimated. In the case of multivariate normality of observed variables and correct specification of the model, the ML estimator is characterized as asymptotically consistent, efficient, unbiased, and normally distributed [1]. Then the model fit statistic is chi-square  $\chi^2$  asymptotically distributed, which enables us to apply an inspection of the overall fit of the model to the data by the means of  $\chi^2$  test. And even more, when this case happens, the asymptotic variance-covariance matrix of the ML estimator also provides the calculation of standard error estimates, which enables us to conduct the significance tests [1]. On the contrary, when the measured data are substantially non-normal, the calculated parameters are still relatively unbiased while the model chi-square  $\chi^2$  statistic can be overestimated, and standard error estimates can be deflated.

As an alternative to the ML estimator, many other estimators have been designed (see Fig. 2). They possess different statistical properties and various estimation mechanisms, which are adopted for the computation of estimated parameters. Also, they can have different GOF indices and different assumptions about the observed variables. Some of them can effectively face with the non-normal character of the variables while the others are even specially developed for the categorical variables. The main properties of these estimators can be shortly stressed as follows.

**GLS estimator** minimizes the so-called weighted residual function by the means of different iterative algorithms. This estimator assumes the multivariate normality of the data with no excessive kurtosis. It is also characterized as asymptotically unbiased, consistent, efficient, and normally distributed full-information estimator.

**ULS estimator** does not have any assumptions about the distribution of measured variables. In general, it is less efficient than maximum likelihood estimator. As it turns out, it has one specific requirement, which demands that all indicators must be observed on the same scale.

**WLS (ADF) estimators** are also insensitive to the distributional properties of the measured variables. When asymptotic covariance matrix is applied here, these estimators also involve forth-order moments around the mean, which are additionally included in estimation besides the second order moments. For the adequate estimation, these methods rigorously require the large sample size. Since the full-weight matrix must be inverted here, they are computationally very expensive estimators.

**DWLS estimators** are very useful in the case of significantly non-normal ordinal variables when we are also dealing with so-called polychoric correlations between the categorical variables. To avoid the computational wastefulness of WLS estimators, DWLS estimators might usually be a better choice [1].

**WLSMV and WLSMV estimators** are extraordinary estimators, which were specially designed for the variables with categorical nature. Herein, the corrected  $\chi^2$  test statistics is also available [6]. As it turns out, among the estimation methods, based on polychoric correlations, the WLSMV method has been indicated to yields better results than the WLS and WLSM estimator in type I error control [1].

**BMCMC estimators** are very appropriate for the noticeably non-normal categorical variables. These estimators do not require any assumption about the asymptotic normality of the estimated parameters. The reason is that the Bayesian credibility intervals only rely on percentiles of the posterior distribution, not limited to any fixed form [1, 3].

The BMCMC estimators are modern and up-to-date alternative to other SEM estimation methods [1, 5, 16]. They use the Markov Chain Monte Carlo (MCMC) procedures for gradual reduction of the uncertainty in the parameter estimates [5]. Since these methods are insensitive to the normality issues, they can be competently used to examine the correctness of results of other, more classical SEM estimation methods, particularly in the case of ordinal and slightly non-normal data [5].

Besides, it was reported in several studies that the BMCMC estimators can provide more accurate estimates for smaller sample sizes than some other estimation methods, such as, for example, the ML method [17, 18]. In general, the main property of these methods is the capability of combining the prior knowledge about the parameters with the fact that the modeling process does not depend on the asymptotic theoretical baseline [19]. This property becomes particularly essential in the case of small sample size and ordinal or markedly non-normal data.

The main philosophy of Bayesian estimation is the fact that every parameter can be addressed as a random variable with associated probability distribution. Then the assumed prior probability distribution can be combined with the empirical information carried in the sample data by the means of the Bayes' theorem, which gives us the posterior distribution [16].

The uncertainty in the estimated parameters is afterward progressively reduced by the generation of new data, which are produced from the original sample by using the MCMC procedure [5, 16]. The latter picks up the repeated samples from the given dataset and generates a big number of the estimates for each model parameter. This way, posterior probabilities of those parameters can be also derived, and the mean values of posterior distributions can be used for the parameter estimates [16].

Since the maximum likelihood estimator provides quite a big number of GOF indices, it is usually desired to use it in the estimation process. As Olsson and his colleagues suggest [20], it is convenient to employ several estimators while doing estimation (for example ML, WLS, and GLS estimators, etc.), and then investigate whether all of them provide similar estimation results. If so, we have an additional confirmation that the model structure is correctly identified, as well as the parameter estimates accurate enough. Such logic was for example used in work [21], where authors applied two estimators, ML estimator, and ULS estimator. Similarly, following the suggestions of Byrne [5],

researchers in study [22] compared the results of ML and BMCMC estimators, where Bayesian method was used to reaffirm the results of ML estimator.

In our case, we have used three estimators, ML estimator, ULS estimator, and BMCMC estimator, by which it was possible to compare the calculated results for parameter estimates of the SEM model.

#### IV. CONCEPTUAL FRAMEWORK, HYPOTHESIZED MODEL, AND SURVEY

A conceptual framework of the hypothesized model is depicted in Fig. 3. Data collection was carried out by a conduction of a survey among 671 travel agencies, located alongside the North-East coast in the Adriatic Sea. The questionnaire was divided into two parts. The first one was interrelated with the external integration indicators for each type of transport supplier. It consisted of 11 ordinal variables, as follows:  $W_i = Q_i^W, i = 1, \dots, 11$  (water suppliers),  $A_i = Q_i^A, i = 1, \dots, 11$  (air suppliers),  $B_i = Q_i^B, i = 1, \dots, 11$  (bus suppliers), and  $R_i = Q_i^R, i = 1, \dots, 11$  (rail suppliers). These measures were needed to evaluate the behavioral magnitudes of external integration with transport suppliers and encompassed the following crucial EI dimensions: interaction, consultation, and collaboration [10].

Creation of ordinal variables  $W_i, A_i, B_i, R_i$  was based on interviewing the managers of travel agencies. They were asked to estimate the level of relationships with transport suppliers on the ordinal scale from 1 (zero cooperation) to 5 (total cooperation). The structure of survey questions of the first part of the questionnaire is shown in Fig. 3.

The second part of the questionnaire consisted of five indicators. They can be called as an agencies' inner variables, denoted by:  $x_i, i = 1, 2, 3, y_j, j = 1, 2$ , where  $x_i$  refer to 'input variables', while  $y_j$  refer to 'output variables'. The meaning of these variables can be seen in Fig. 3. These variables were needed to calculate the efficiencies of the travel agencies, similarly as it was reported in study [23].

From Fig. 3 can be seen that we conducted four main hypotheses  $H_i, i = 1, 2, 3, 4$ , which indicate that external integrations with transport suppliers, denoted by  $EI_i, i = 1, 2, 3, 4$ , do have a certain influence on the efficiency (EFF) of the agencies. Additionally, six sub-hypotheses have been applied, by which it was supposed that external integrations are interrelated among themselves as well (their connections are not shown in Fig. 3).

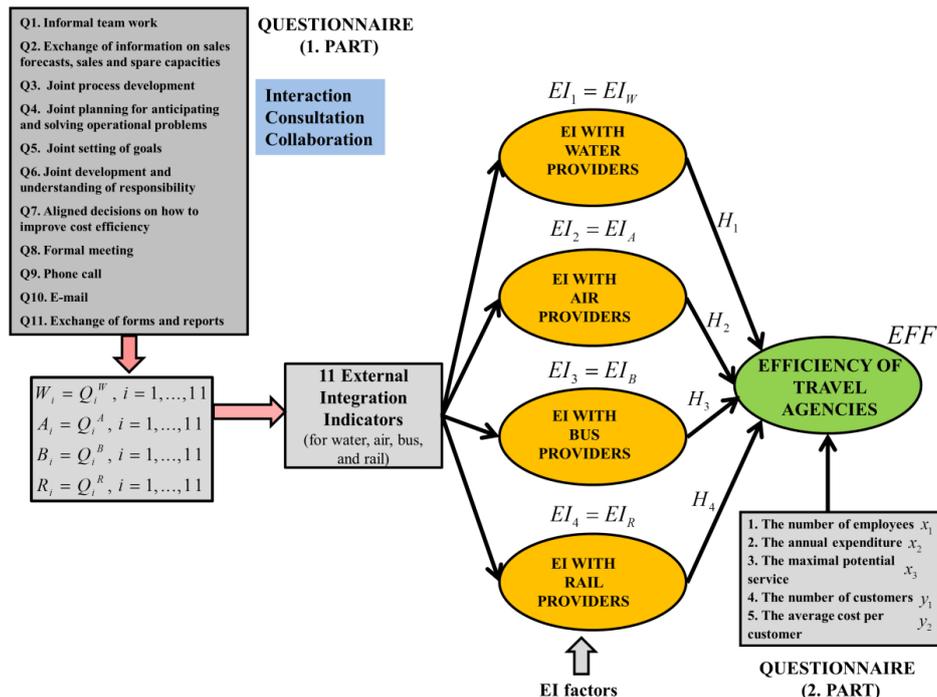


Figure 3. Conceptual framework of hypothesized model

V. SEM MODELING PROCESS FOR THE CASE OF AGENCIES

Fig. 4 depicts the main methodologies used in the SEM modeling process. In the first stage, the DEA analysis was applied to estimate the efficiencies on the basis of agencies' inner variables. This method is based on linear programming principles and is designed for measuring the efficiency of given decision-making unit (DMU) [23, 24]. Herein, the most efficient DMU (i.e. company or some other organization) represents the "frontier", with which the relative performance of all other DMUs in the sample must be compared. Then any deviation from the frontier is reflected as inefficiency [24].

In the next stage of the modeling process, the EFA analysis was employed. It investigated the nature of external integration based latent factors  $EI_i, i = 1, 2, 3, 4$ , assessed via the observation of measured indicators  $W_i, A_i, B_i, R_i$ . Afterward, the CFA analysis assisted us to evaluate statistically how well the measured indicators represent the corresponding factors. This way, the confirmation of hypothesized factor structure was executed by the means of certain statistic tests and calculation of GOF indices. So, the recommendations of many authors [1, 6, 15, 16] were taken into the consideration, who claim that it is needed to apply the CFA individually before the estimation of the overall SEM model. This enabled us to carry out the statistical evaluation of measurement theory completely independently to be completely convinced of the adequacy of the measurement part of SEM model.

After the completion of CFA analysis, the SEM modeling procedure was the final stage of our model design process. At this stage, the structural part of the SEM model was also constructed, and both sub-models, measurement and structural, were linked together and transformed into the overall SEM model. Naturally, before the total completion of SEM modeling process, the adequate validation of the model has been also processed by the means of GOF indices. The final SEM model revealed the causal path relations between the external integrations  $EI_i, i = 1, 2, 3, 4$  and the efficiency  $EFF$  of the agencies, as well as the interrelations among  $EI_i$  themselves. The weights of all paths and interrelations were also estimated in this process by the means of ML, ULS, and BMCMC estimator.

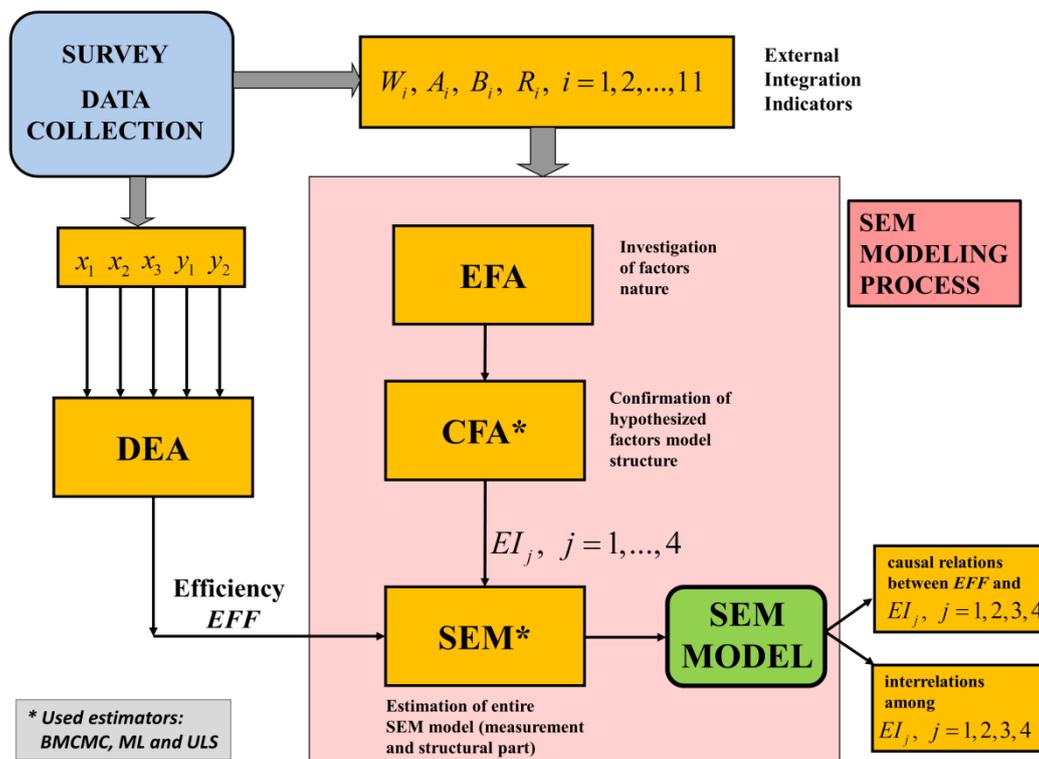


Figure 4. The main methodologies used in the SEM modeling process for the case of agencies

## VI. PRACTICAL NUMERICAL RESULTS

### A. Statistical properties of the data

It turned out that only 61 travel agencies were prepared to cooperate fully with the survey since the others were unresponsive to reveal enough information about their economic activities. The conditions for the normality of collected data were not severely violated, but only slightly, since the range of SI was  $(-1.6, 2.69)$ , while the range of KI was  $(-1.516, 6.49)$ . According to several authors [25-30], these ranges are acceptable for further analysis and do not represent any serious non-normality problem, even in the case of relatively small sample size.

### B. Exploratory factor analysis

Correctness of conducting the factor analysis was inspected by the execution of two tests, Bartlett's test of sphericity (BTS) and the Kaiser-Meyer-Olkin (KMO) test [1, 6, 15]. The value for BTS value was highly significant ( $\chi^2 = 1998.664$  with  $df = 300$  and  $p < 0.001$ ), while the KMO value was  $0.763 > 0.5$ . According to recommendations of some authors [31, 32], the calculated KMO, and BTS values confirmed that the factor analysis can be consistently applied for the further research.

For the extraction of factors and estimation of their loadings, the principle axis factoring (PAF) algorithm was employed. Since the so-called communalities  $h_i^2$  [3, 6] of some indicators were not acceptable, and the cross-loadings were quite significant, 19 ill-fitting indicators were excluded from the further analysis. Afterward, the factors' extraction was conducted again on the basis of PAF algorithm and PROMAX with Kaiser Normalization rotation method [5].

When the rotation was completed, the significant cross-loadings were completely eliminated while the communalities of retained indicators became sufficiently large, which means the range:  $(0.574 \leq h_i^2 \leq 0.891)$ . Also, the final indicators' loadings  $\lambda_{ij}$  on addressed factors reached the adequate value ( $\lambda_{ij} \geq 0.689$ ), as it is recommended in the literature [6, 33].

Table 1 shows the derived Cronbach alpha coefficients (CAC), eigenvalues, the percent of the variance, and the cumulative variance percent of the extracted factors. Naturally, these results are presented for the case of retained indicators. CAC values are all bigger than 0.7, as suggested by Hair and his colleagues [6]. According to these authors, the derived results presented in this section had satisfied all the crucial requirements for the convergent and discriminant validity, as well as the reliability and internal consistency. Additionally it turned out that according to our expectations, the different types of transport suppliers' indicators were significantly loaded on their own factors. Or, by other words, the water transport indicators  $W_i = Q_i^W, i = 1, \dots, 11$  were significantly loaded on 'water' factor  $EI_W$ , the air transport indicators  $A_i = Q_i^A, i = 1, \dots, 11$  were significantly loaded on 'air' factor  $EI_A$ , and so.

Table 1. The derived CAC coefficients, eigenvalues, the percent of the variance, and the cumulative variance percent of the extracted factors.

Measure	Factor			
	$EI_W$	$EI_A$	$EI_B$	$EI_R$
Cronbach Alpha	0,960	0.972	0.921	0.947
Eigenvalues	9,951	5.495	2.725	2.346

<b>% of Variance</b>	39.805	21.980	10.902	9.383
<b>Cumulative %</b>	39.805	61.785	72.687	82.069

### C. Confirmatory factor analysis

In CFA, the indicators were firstly allocated to the corresponding factors according to the preliminary recommended guidelines of the EFA analysis, and the estimation of measurement sub-model was executed by the means of all three estimators, ML, BMCMC, and ULS. Since calculated GOF indices indicated the un-adequate results, 6 more ill-fitting indicators have been additionally excluded from the further analysis. Afterward, the estimation was repeated again, and GOF indices have become adequate. The estimation results have implied that all three estimators achieved similar results for estimated parameters.

After the successful estimation and GOF indices based evaluation, the process of further confirmation of our measurement theory was the next step of CFA analysis. This means that the overall measurement sub-model was assessed for its validity and reliability. In this context, the convergent and discriminant validities are particularly important. The convergent validity encompasses the properly high values of factor loadings, the adequate composite reliability (CR), and the appropriate average variance extracted (AVE) [6]. According to Hair et al. [6], the threshold of CR is 0.7 while the threshold for AVE is 0.5.

The discriminant validity demands that the AVE of each factor should be larger than the squared correlation  $CORR^2$  between this factor and any other factor, which means:  $\sqrt{AVE} > CORR$  [6, 34].

In our case, the conditions about convergent and discriminant validities have been confirmed for all three estimators. For example, when ML estimator was used, the range of estimated standardized factor loadings was:  $0.741 \leq \lambda_{ij} \leq 0.984$ , which indicates the appropriate convergent validity. Standardized factor loadings have taken similarly high range in the case of BMCMC and ULS estimators. The convergent validity was also confirmed through the calculation of CR and AVE values. Table 2 shows results for CR and AVE regarding each addressed factor when the ML estimation was applied. Obviously, the calculated CR values are all bigger than 0.7 while computed AVE values are all greater than 0.5. CR and AVE have taken similar values for addressed factors when other two estimators, BMCMC, and ULS, were conducted.

As it turned out for all three employed estimators, the discriminant validity of the measurement sub-model was also similarly confirmed. So, the condition  $\sqrt{AVE} > CORR$  was always fulfilled, irrespective of the type of used estimator. After the confirmation of all necessary tests in CFA analysis, we became positively convinced that the independently observed measurement sub-model is adequate. This so, we can reliably say that the transport indicators are indeed significantly loaded on their own transport factors, as it was already implied in the case of EFA analysis.

Table 2. The convergent validity of the measurement sub-model (case of ML estimator)

<b>Factor</b>	<b>Measure</b>	
	<b>CR</b>	<b>AVE</b>
$EI_A$	0.969	0.819
$EI_W$	0.949	0.791
$EI_B$	0.883	0.717
$EI_R$	0.947	0.819

### D. Structural equation model

Following the methodological framework in Fig. 4, the derivation of the structural sub-model was the final stage of SEM modeling process. Here, both sub-models, measurement and structural, were also connected and converted into the entire SEM model. Besides all the involved variables introduced in the previous sections, the single-indicator factor EFF was also added to the entire SEM model structure. This is logical since our primary interest in this study was to reveal and estimate the causal paths between the integration factors  $EI_W, EI_A, EI_B, EI_R$  on one side, and the efficiency factor EFF on the other side.

Table 3 shows the comparison of achieved estimation results for all used estimators, ML, ULS, and BMCMC estimator. These results refer to the standardized weights of causal paths between integration factors and efficiency factor, as well as to the standardized weights of correlations among the integration factors themselves.

Table 3. Comparison of achieved estimation results (Maximum likelihood, Bayesian estimation, Unweighted least squares)

Standardized Weight on:	Type of relation	ML estimator	BMCMC estimator	ULS estimator	Significance
$EI_W \rightarrow EFF$	Causal path	<b>0.421</b>	<b>0.41</b>	<b>0.408</b>	yes
$EI_A \rightarrow EFF$	Causal path	<b>0.499</b>	<b>0.477</b>	<b>0.395</b>	yes
$EI_B \rightarrow EFF$	Causal path	-0.148	-0.13	-0.037	no
$EI_R \rightarrow EFF$	Causal path	-0.009	-0.024	-0.05	no
$EI_W \leftrightarrow EI_B$	Correlation	<b>0.538</b>	<b>0.537</b>	<b>0.429</b>	yes
$EI_W \leftrightarrow EI_R$	Correlation	<b>0.484</b>	<b>0.486</b>	<b>0.412</b>	yes
$EI_A \leftrightarrow EI_B$	Correlation	<b>0.502</b>	<b>0.53</b>	<b>0.464</b>	yes
$EI_B \leftrightarrow EI_R$	Correlation	<b>0.39</b>	<b>0.348</b>	<b>0.292</b>	yes
$EI_W \leftrightarrow EI_A$	Correlation	0.166	0.153	0.177	no
$EI_A \leftrightarrow EI_R$	Correlation	0.165	0.158	0.225	no

Careful observation of results in Table 3 leads us to the following conclusions for all three used estimators:

1. The weights of two causal paths and four correlations have been estimated as positive and statistically significant;
2. Roughly speaking, all estimators have provided more or less similar results for those weights, which were significant. This is particularly true for the ML and BMCMC estimators, which gave quite comparable results. The results achieved by the ULS estimator slightly diverge from results of other two estimators but are still comparable to them.
3. Since the results of all three estimators are sufficiently close to each other, we can derive similar conclusions on their basis. Additionally, we can simultaneously apply the GOF indices of all addressed estimators in order to carry out an as much reliable model fit evaluation as possible.

Table 4 shows the GOF indices obtained for the case of ML estimator. Their values were compared with the required threshold values, given in the literature [1, 5, 6, 15]. Since the comparison gave adequate results, it was concluded that the overall SEM model provides a reasonably good fit to the

data. As it turned out, the same conclusion was derived on the basis of observing of GOF indices related to the ULS and BMCMC estimators.

Table 4. GOF indices for developed SEM model (Maximum likelihood case)

Fit Index	Description	Value for ML estimator
$\chi^2$	Chi-Square $\chi^2$ of the discrepancy between the sample and the fitted covariances' matrices	152.329
$\left(\chi^2/df\right)$	Relative Chi-Square $\chi^2$ of the discrepancy	1.058
<b>RMSEA</b>	Root Mean Square Error of Approximation	0.031
<b>NFI</b>	Normed-fit Index	0.911
<b>NNFI (TLI)</b>	Non-Normed-fit Index (Tucker-Lewis Index)	0.993
<b>CFI</b>	Comparative Fit Index	0.995
<b>IFI</b>	Bollen's Incremental Fit Index	0.995
<b>SRMR</b>	Standardized Root Mean Square Residual	0.0602

Based on estimation results presented in table 3, Fig. 5 can be created, which corresponds to the conceptual framework shown in Fig. 3. Fig. 5 represents the final estimated SEM model, where the retained indicator items are also depicted. Dashed lines refer to the causal paths or correlations with statistically insignificant weights (c.f. Table 3).

Achieved results imply that the effect of factors  $EI_B$  and  $EI_R$  on efficiency  $EFF$  cannot be supported, so the hypotheses  $H_3, H_4$  are rejected. However, on the other side, the first two hypotheses,  $H_1$  and  $H_2$ , evidently can be accepted, which implies that the effect of factors  $EI_W$  and  $EI_A$  on efficiency  $EFF$  is confirmed in this study.

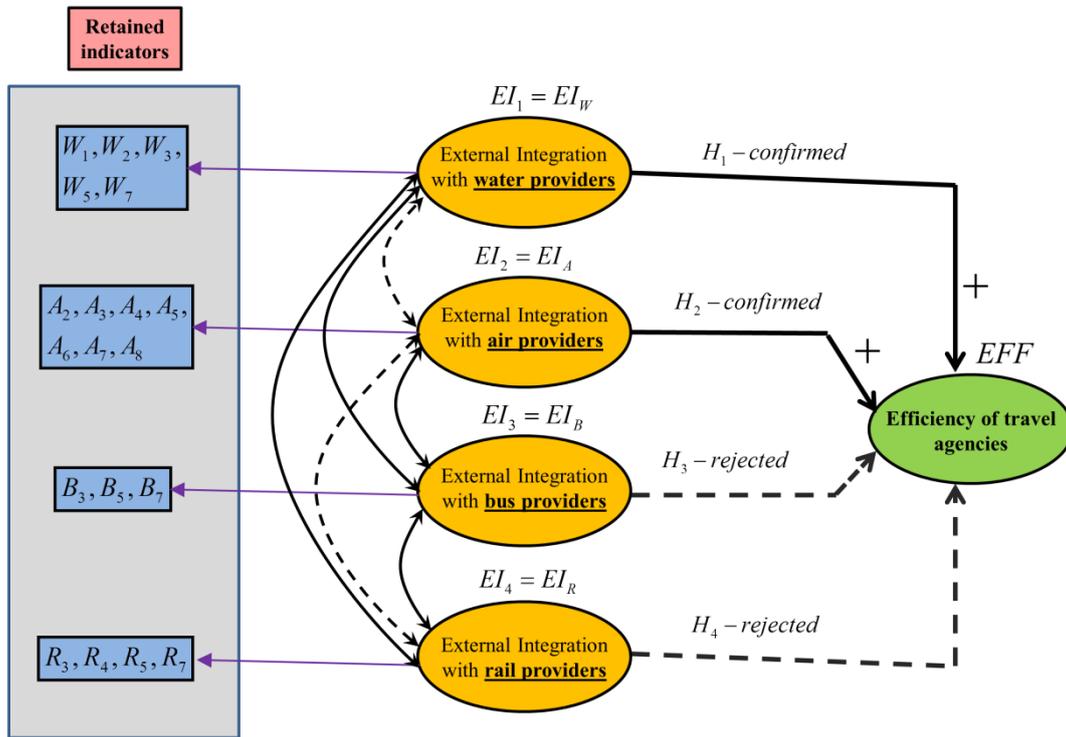


Figure 5. The main properties of the final estimated SEM model

#### E. Discussion about achieved results

On the basis of obtained results, we can say that the external integration with water and air suppliers has some positive impact on the efficiency of the agencies. This so, the amplified level of integration would probably also lead to their higher efficiency. In other words, if the addressed agencies increase the level of interaction, collaboration, and consultation with the water and air suppliers as much as possible, the result would most likely be a higher financial profit and bigger number of satisfied customers, as well as a lower level of operational costs.

The reason why the hypotheses  $H_3, H_4$  have not been supported should be deeper inspected.

Perhaps the sample size was too small, or the addressed agencies have quite distinct characteristics. Namely, they possess their own transport vehicles, so it is probably counter-productive to hire the latter from elsewhere. Additionally, since the railway infrastructure in agencies' neighborhood does not provide perfect driving conditions, the agencies' leaders expressed particular doubt about deepening the collaboration with the railway authorities.

## VI. CONCLUSION

The paper addressed the relationships between external integration with different transport suppliers on one side and efficiency of travel agencies located in Croatia on the other side. For this purpose, the SEM model was developed on the basis of the data collected by the means of a survey. While conducting the SEM modeling process, the EFA and CFA analyses have been also applied. For the calculation of efficiencies, the DEA analysis was performed.

The primary focus of the paper was the comparison of the performance of three different estimators, used in the SEM modeling procedure. Within the estimation framework, the following estimators were studied: Maximum likelihood estimator, Bayesian estimator, and Unweighted least squares estimator. The achieved standardized results show that all estimators have computed similar significant weights of causal paths between EI factors and EFF factor, as well as similar significant weights of correlations among EI factors themselves.

The investigation of GOF indices provided an evidence of good fit to the data performed by the derived SEM model. The latter has revealed that enlarged integration with air and water suppliers would most likely lead to increased efficiency. This might represent a very interesting finding for the management from the financial profit and operational costs' point of view.

## REFERENCES

1. R. H. Hoyle, *Handbook of Structural Equation Modeling*, The Guilford Press: New York, 2012.
2. T. F. Golob, *Structural Equation Modeling for Travel Behavior Research, Transportation Research Part B: Methodological*, vol. 37, pp. 1-25, 2003.
3. D. Dragan, D. and D. Topolšek, "Introduction to Structural Equation Modeling: Review, Methodology and Practical Applications", 11th International conference on Logistics & Sustainable Transport, Celje, June 2014.
4. K. A. Bollen, *Structural Equations with Latent Variables*, John Wiley & Sons: New York, 1989.
5. B. M. Byrne, *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming*, 2<sup>nd</sup> ed., Routledge: London, 2009.
6. J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, 7<sup>th</sup> ed., Prentice Hall: New Jersey, 2010.
7. T. W. Andreassen, B. G. Lorentzen, and U. H. Olsson, "The Impact of Non-Normality and Estimation Methods in SEM on Satisfaction Research in Marketing", *Quality & Quantity*, vol. 40, pp. 39–58, 2006.
8. Y. Yilmaz, and U.S. Bititci, "Performance measurement in tourism: a value chain model", *International Journal of Contemporary Hospitality Management*, vol. 18, pp 341-349, 2006.
9. F. Alamdari, "Regional development in airlines and travel agents relationship", *Journal of Air Transport Management*, vol. 8, pp 339-348, 2002.
10. H. Atkinson, and J.B. Brown, "Rethinking performance measures: assessing progress in UK hotels", *International Journal of Contemporary Hospitality Management*, vol. 13, pp. 128-136, 2001.
11. D. Topolšek, E. Mrnjavac, and N. Kovačić, "Integration of travel agencies with transport providers", *Tourism Management Perspectives*, vol. 9, pp. 14–23, 2014.
12. L. Mia, and A. Patiar, "The use of management accounting systems in hotels: an exploratory study", *Hospitality Management*, vol. 20, pp. 111-128, 2001.
13. P.A. Phillips, "Hotel performance and competitive advantage: a contingency approach", *International Journal of Hospitality Management*, vol. 11, pp. 359-65, 1999.
14. L. Enz, L. Canina, and K. Walsh, "Hotel industry average: an inaccurate tool for measuring performance", *The Cornell Hotel and Restaurant Administration Quarterly*, vol. 42 (6), pp. 22-32, 2001.
15. R. B. Kline, *Principles and Practice of Structural Equation Modeling*, 3rd ed., The Guilford Press: New York, 2011.
16. J. L. Arbuckle, *IBM SPSS Amos 20 User's Guide*, IBM, Amos Development Corporation: Chicago, b.d.
17. J. Evermann, and M. Tate, "Bayesian Structural Equation Models for Cumulative Theory Building in Information Systems—A Brief Tutorial Using BUGS and R", *Communications of the Association for Information Systems*, vol. 34, pp. 1481–1515, 2014.
18. T. Asparouhov, and B. Muthén, *Bayesian analysis using Mplus*, Technical appendix. Los Angeles, Muthén & Muthén, 2010.
19. P. Congdon, *Applied Bayesian Modeling*, West Sussex, England: John Wiley & Sons, 2003.
20. U. H. Olsson, T. Foss, S. V. Troye, and R. D. Howell, "The Performance of ML, GLS, and WLS Estimation in Structural Equation Modeling Under Conditions of Misspecification and Nonnormality", *Structural Equation Modeling*, vol.7, no. 4, pp. 557–595, 2000.
21. E. Penelo, C. Viladrich, and J. M. Domènech, "Perceived parental rearing style in childhood: internal structure and concurrent validity on the Egna Minnen Beträffande Uppfostran-Child Version in clinical settings", *Comprehensive Psychiatry*, vol. 51, pp. 434–442, 2010.
22. M. R. A. Hamid, Z. Mustafa, F. Idris, M. Abdullah, and N. R. M. Suradi, "Measuring Value-Based Productivity: A Confirmatory Factor Analytic (CFA) Approach", *International Journal of Business and Social Science*, vol. 2, no. 6, April 2011.
23. R. Fuentes, "Efficiency of travel agencies: A cases study of Alicante, Spain", *Tourism management*, vol. 32, no. 2, pp. 75-87, 2011.
24. W.W. Cooper, L.M. Seiford, and K. Tone, K., *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Massachusetts, USA: Kluwer Academic Publishers, 2006.
25. J.B. Ullman, "Structural equation modeling: Reviewing the basics and moving forward", *Journal of Personality Assessment*, vol. 87, no. 1, pp. 35–50, 2006.
26. R. Weston, and P.A. Gore, "A Brief Guide to Structural Equation Modeling", *The Counseling Psychologist*, vol. 34, no. 5, pp. 719-751, 2006.
27. C.P. Chou, and P.M. Bentler, *Estimates and tests in structural equation modelling*, In R. Hoyle, *Structural equation modeling: Concepts, issues, and applications*, Thousand Oaks, CA: Sage, pp. 37–55. 1995.
28. X. Zhai, A.M.M. Liu, and R. Fellows, "Human Resource Practices in Chinese Construction Organizations: Development of a Measurement Scale", *International Journal of Architecture, Engineering and Construction*, vol. 2, no. 3, pp. 170-183, 2013.

29. M. Lei, and R. G. Lomax, "The effect of varying degrees of nonnormality in structural equation modelling," *Struct. Equ. Modeling*, vol. 12, pp. 1–27, 2005.
30. P. J. Curran, S. G. West, and J. F. Finch, "The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis," *Psychol. Methods*, vol. 1, pp. 16–29, 1996.
31. M. Sahin, A. Todiras, P. Nijkamp, B. Neuts, and C. Behrens(2013), "A Structural Equations Model for Assessing the Economic Performance of High-Tech Entrepreneurs", In R. Capello, and T.P. Dentinho, *Globalization Trends and Regional Development* , vol. 48, 2013.
32. M.T. Frohlich, and R. Westbrook, R., "Arcs of integration: an international study of supply chain strategies", *Journal of Operations Management*, vol. 19, pp. 185–200, 2001.
33. D.A. Kenny, "Measuring Model Fit", Retrieved on February 6, 2014, from <http://davidakenny.net/cm/fit.htm>
34. C. Fornell, D.F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, vol. 18, pp. 39-50, 1981.

## AUTHORS

**A. Nataša Kovačić**, Phd, is the Higher Assistant at the Faculty of Tourism and Hospitality Management, University of Rijeka, Opatija, Croatia (e-mail: [natasak@ftthm.hr](mailto:natasak@ftthm.hr)).

**B. Darja Topolšek**, PhD, is the Assistant Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: [darja.topolsek@um.si](mailto:darja.topolsek@um.si)).

**C. Dejan Dragan**, Phd, is the Assistant Professor at the Faculty of Logistics, University of Maribor, Celje, Slovenia (e-mail: [dejan.dragan@um.si](mailto:dejan.dragan@um.si)).