

PARTIAL LEAST SQUARES APPROACH TO STRUCTURAL EQUATION MODELING FOR TOURISM RESEARCH

Mehmet Mehmetoglu

ABSTRACT

Tourism research contains a large share of consumer behavior-orientated studies using multidimensional constructs (exogenous/endogenous). Accordingly, scholars have mainly made use of a two-step approach that can be referred to as PCA-MLR (principal component analysis and then ordinary least squares multiple linear regression analysis) to examine the relationships among exogenous and endogenous constructs in a statistical model. Although this two-step approach has contributed to the advancement of tourism research, it still suffers from a number of drawbacks which can readily be overcome by a so-called second-generation statistical tool, namely, partial least squares approach to structural equation modeling (PLS-SEM). The current chapter explains and illustrates (with an application to tourism data) the advantages (e.g., several layers of estimations, suiting small sample sizes, robustness to multicollinearity, model-based clustering, etc.) of PLS-SEM both from a statistical and practical point of view. Finally, an elucidation is also provided for

Advances in Hospitality and Leisure, Volume 8, 43–61

Copyright © 2012 by Emerald Group Publishing Limited

All rights of reproduction in any form reserved

ISSN: 1745-3542/doi:10.1108/S1745-3542(2012)0000008007

suggesting PLS-SEM as an alternative to PCA-MLR instead of COV-SEM (covariance-based structural equation modeling). The chapter concludes by proposing that PLS-SEM is a reliable and flexible statistical approach that is of high value, in particular, for applied research.

Keywords: Partial least squares; structural equation modeling; latent variable; PLS; SEM

INTRODUCTION

Tourism research is known to be dominated by use of quantitative methods (Mehmetoglu, 2004). One particular quantitative approach that is often utilized in tourism studies is sequentially employing principal component analysis (PCA) and ordinary least squares multiple linear regressions (hereafter referred to as MLR) in tandem. In the first stage of the data analysis, researchers typically employ PCA to linearly transform an original set of variables into a substantially smaller set of uncorrelated variables (components/factors) that represents most of the information in the original set of variables (Dunteman, 1989). In the second stage, researchers go on and usually use the resulting components (through composite or factor score) as a set of independent (predictor) variables endeavoring to explain the variation of a dependent (outcome) variable in an MLR model.

The combinational use of PCA and MLR can be abbreviated as PCA-MLR belonging to the family of first-generation data analysis techniques (Gefen, Straub, & Boudreau, 2000). Although PCA-MLR may have contributed to the advancement of tourism research, it, however, suffers from a number of drawbacks which easily are overcome by its second-generation counterpart (i.e., structural equation modeling, SEM), making accordingly the latter a more reliable statistical approach (Sosik, Kahai, & Piovoso, 2009). One of the second-generation statistical tools representing a serious alternative to PCA-MLR is the PLS approach to SEM, which can also be abbreviated as PLS-SEM.

As not many applications of PLS-SEM (e.g., Loureiro & Gonzales, 2008; Mendes, Do Valle, & Guerreiro, 2011; Tiedemann, van Birgele, & Semeijn, 2009) are existent in tourism research, in line with Chin's (1998) suggestion this chapter aims to remind or make (tourism) researchers aware of this approach. As such, the chapter conceptually explains the methodological characteristics of PLS-SEM and illustrates the PLS approach to SEM with

an application to tourism data. In so doing, the advantages (seen from statistical and practical perspectives) of PLS-SEM (as opposed to PCA-MLR) will also be explicated. Finally, an explanation is also provided as to why PLS-SEM is proposed as a serious and sometimes more sound alternative to PCA-MLR instead of covariance-based SEM (COV-SEM) approaches.

PLS Approach to SEM

PLS-SEM is a variance-based structural equation modeling that has its origins in the early work of Herman Wold (1975) on partial least squares for principal component analysis (Esposito Vinzi, Trinchera, & Amato, 2010). PLS-SEM resembles a system of interdependent MLR equations estimating a network of links among latent variables (expression of manifest variables) as well as the links between the manifest (observed) variables and their respective latent constructs (Esposito Vinzi et al., 2010). The equations estimating the links between the manifest variables and their latent constructs represent the outer (measurement) model, while the equations estimating the relationships among latent variables constitute the inner (structural) model (Henseler, Ringle, & Sinkovics, 2009). This is visualized in Fig. 1 in which the equations above the ellipses represent the measurement model and the equations below the ellipses depict the structural model part of a PLS-SEM analysis. In the measurement part, latent variable scores (e.g., Y_1 , Y_2 , etc.) for each subject are computed (e.g., $Y_1 = \beta_0 + \beta X_1 + \beta X_2 + \beta X_3$) and then the resulting scores are used to ascertain coefficients of the paths between latent variables and to explain the variance of specific latent variables (e.g., $Y_1 = \beta_0 + \beta Y_1 + \beta Y_2$) in the structural part (see Tenenhaus, 2008).

Although this procedure at first sight may appear to be identical to PCA-MLR, unlike in PCA-MLR though, first, the latent variable scores in PLS-SEM are computed based on weightings (without assuming equal weights) of their respective manifest variables or indicators (Haenlein & Kaplan, 2004). As such, indicators exhibiting weaker relationships with associated indicators and latent variables are given lower weightings (Chin, Marcolin, & Newsted, 2003). Second, the estimations (relations between latent variables) taking place in the structural part are also taken into account in the computation of latent variable scores in PLS-SEM, that is, the whole model is estimated simultaneously. This procedure contributes certainly to increasing the predictive power of a model (Sosik et al., 2009).

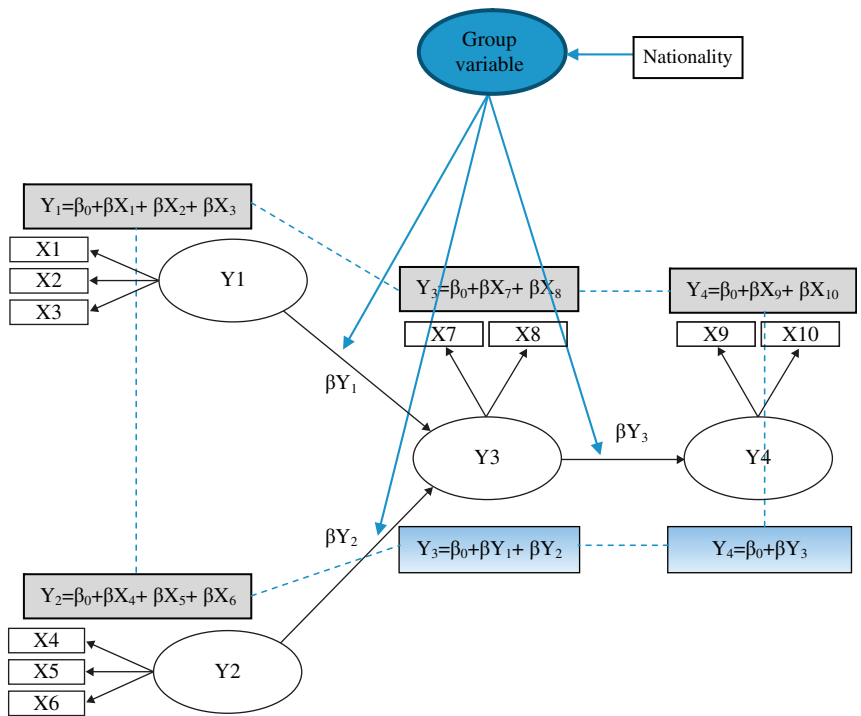


Fig. 1. A PLS-SEM Analysis.

Relatedly, in PLS-SEM latent variable scores can readily be computed not only for latent variables with reflective indicators but also for those with formative indicators (Haenlein & Kaplan, 2004), a possibility which clearly has been absent in tourism studies adopting PCA-MLR. In these studies there has, in other words, not even been made any distinction between reflective and formative indicators. Treating formative indicators as reflective ones may lead to incorrect conclusions about the relations between latent variables in the structural part (Law & Wong, 1999). Further, PLS-SEM enables estimation of a measurement model (i.e., MIMIC) with latent variables expressed by a combination of reflective and formative indicators, an option that may be appropriate to use in some cases (e.g., tourist motivations and activities reflecting and forming a joint latent variable) in tourism research (see Fig. 2).

Another advantage that PLS-SEM has is that it, unlike PCA-MLR, allows more than one layer of linkages between independent and dependent

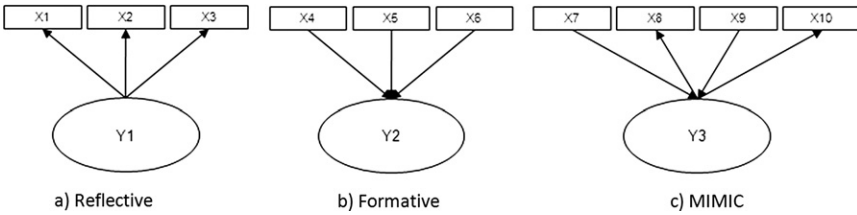


Fig. 2. Three Ways of Modeling a Latent Variable in PLS-SEM.

variables at the same time (Gefen et al., 2000), a feature that makes it possible for the analyst to define one/several variables as both independent and dependent variables in the same model (Haenlein & Kaplan, 2004). This particular feature enables the analyst to perform mediation analyses (and particularly complex ones) in a more accurate and efficient manner as compared to PCA-MLR. For instance, as depicted in Fig. 1, PLS-SEM can easily examine the mediator role of Y_3 by determining magnitudes of the indirect effects of Y_1 and Y_2 on Y_4 via Y_3 .

In addition to mediation analysis, PLS-SEM enables the analyst to examine moderating/interaction effects in a model. First, the moderating effects in PLS-SEM can be examined in the same manner as in traditional MLR approach by creating the interaction term with product indicator, whereas the second way of assessing moderating effects in PLS-SEM can be facilitated through multigroup comparison approach (Henseler & Fassott, 2010). Here, groups (two or more) can be compared based on path coefficients, outer weights/loadings, and goodness-of-fit (GoF) quality indexes as well as means of latent variable scores (Chin, 2003). Fig. 1 shows a graphical presentation of a moderator analysis (domestic vs. foreign tourists) through multigroup comparison approach. To determine whether there exist significant differences between the model parameters across the groups, t -tests (requires normality) or permutation (assumption-free) tests are utilized (Chin & Dibbern, 2010).

While the multigroup comparison approach is one of the techniques used for a priori type of segmentation studies, the factor and cluster analysis (run sequentially) has by far been the most common way of conducting a posteriori type of segmentation works in tourism research. This approach suffers, however, from a number of considerable drawbacks. The main problem with the factor and cluster analysis approach is that it totally is a subjective procedure, in that segmentations are conducted by examining patterns in a set of variables, and not model based in which this set of

variables can be related to a criterion (response) variable. This is exactly another invaluable feature that PLS-SEM offers, namely, a model-based segmentation procedure. There are five such procedures one of which is the response-based unit segmentation in partial least squares (REBUS-PLS) modeling (Esposito Vinzi et al., 2010). Briefly, this procedure identifies clusters that are homogeneous in terms of signs and magnitudes of path coefficients between the independent and dependent variables in a PLS-SEM model. In so doing, the clustering procedure takes into consideration not only the structural but also the measurement part of the model (Esposito Vinzi et al., 2010). For instance, in a simple model with two independent variables (price and quality attributes) and a dependent variable (satisfaction), one can obtain a cluster which is price sensitive and another cluster which is quality oriented. (Fig. 3)

According to Sosik et al. (2009), PLS-SEM, as opposed to MLR, (a) allows for an easier modeling of both continuous and categorical variables (observed/latent), which are modeled through dummy variables; (b) deals with issues of unreliability and heteroscedasticity effectively; (c) requires no normality of data distributions (Squillacciotti, 2010); (d) requires no observation independence as the linear combinations of variables/scores are themselves statistically independent; (e) requires no variable metric uniformity, placing minimal demands on measurement scale (i.e., can be nominal/interval/ordinal); and finally (f) is robust to multicollinearity (Kristensen & Eskilden, 2010), in that it uses a model for both independent and dependent data which accounts for the correlation structure of the data in an appropriate manner. Incidentally, a latent variable can also be expressed by a set of indicators with different scales (0–1, 1–5, etc.) when these indicators are standardized prior to the analysis (Esposito Vinzi, 2009).

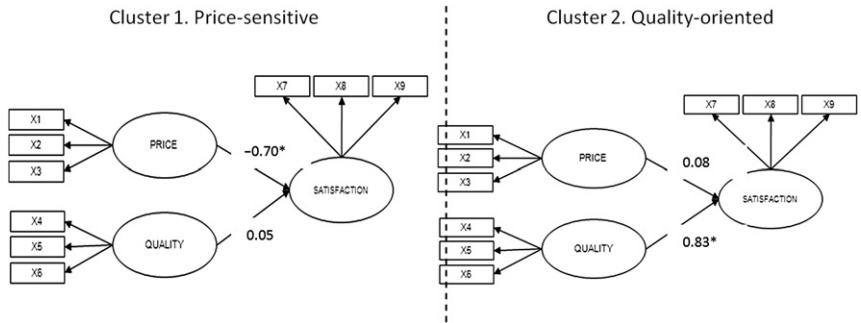


Fig. 3. Illustration of Two Model-Based Clusters through PLS-SEM.

A further benefit that PLS-SEM offers is that, unlike PCA-MLR, PLS-SEM enables the analyst to construct higher-order models, also known as multidimensional factor models (second, third, and even fourth) (see for instance [Wetzels, Odekerken-Schröder, & van Oppen, 2009](#)), which can be either estimated on its own as a measurement model or in tandem with the structural part of a model depending on the purpose of the analysis. This feature can prove to be very useful when estimating complex models including overarching theoretical concepts such as push travel motivations that the analyst can easily construct through higher-order factor modeling. In PLS-SEM, the manifest variables are used twice: for the first-order latent construct (i.e., primary loadings) and for the second-order latent construct (i.e., secondary loadings) (see [Wetzels et al., 2009](#)).

Another advantage of PLS-SEM is a practical one, in that all of the so far presented analytical procedures can readily be performed with PLS-based software (e.g., XLSTAT-PLSPM, SmartPLS, SPAD-PLS, and PLSGraph) depending on the features available in the software. In the following section, XLSTAT-PLSPM is incidentally the software chosen to explicitly illustrate how a complex PLS-SEM model can be estimated and evaluated. For this purpose, a conceptual model (see [Fig. 4](#)) is constructed in such a way to

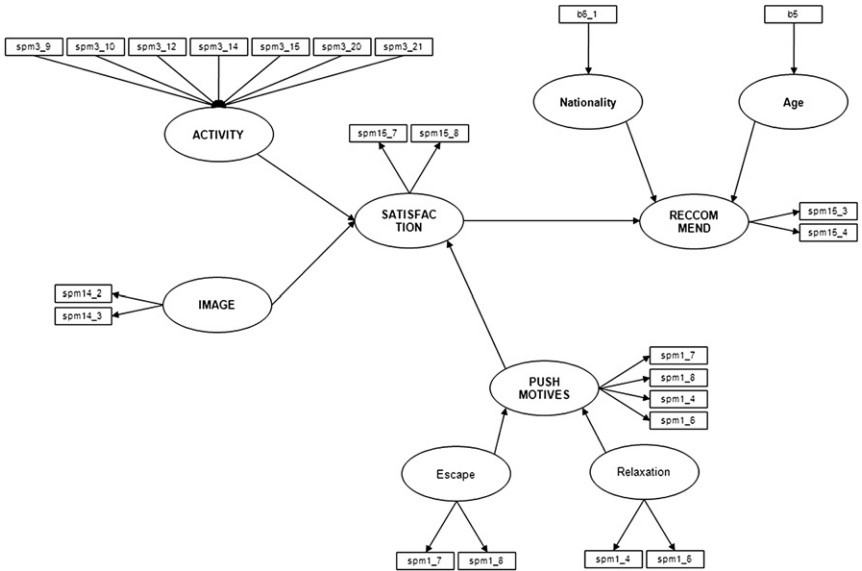


Fig. 4. A Complex Conceptual PLS-SEM Model.

contain several elements that are related to the features (e.g., reflective/formative, higher-order factors, categorical variables) explained earlier in this section.

Application of PLS-SEM to Tourism Data

The aim of this section is to apply PLS-SEM to a set of tourism data based on which the conceptual model in Fig. 4 is constructed. The data used comprises responses provided by 996 respondents who spent their vacation in the Lillehammer region of Norway in the summer of 2010. The data was collected through a four-page long questionnaire consisting of questions related to travel motives (push and pull), travel activities, image, trip characteristics, as well as sociodemographic data. For the purpose of this chapter, a set of activities, a set of travel motives, image, satisfaction level, willingness to recommend, nationality, and age were used as the study's variables.

Measurement of the Study's Variables

Travel activities were dichotomous variables (0 = no; 1 = yes) based upon which the respondents were asked to indicate whether or not they had participated or planned to participate in the following activities: hiking, cycling, fishing, golfing, visiting farms, seeing plants and animals, and kayaking/canoeing/rafting. Travel motives were measured using an ordinal scale (1 = not important; 5 = very important), asking the respondents to reveal how important the following motives were for choosing to travel on their present vacation: to get away from the everyday activities, to escape boredom, to be able to relax mentally, and to invigorate myself. Image attributes were also measured employing an ordinal scale (1 = not at all; 5 = to a great degree), asking the respondents to indicate to what degree they perceived the destination as exciting and stimulating.

Satisfaction level and willingness to recommend were also measured using an ordinal scale (1 = completely disagree; 5 = completely agree), asking the respondents to indicate to what degree they agreed with the following statements: overall, I am satisfied with my vacation in the Lillehammer region, and my vacation experience in the Lillehammer region exceeded my expectations (satisfaction level), I am going to recommend to my friends to go on vacation in the Lillehammer region, and I am going to say positive

things about Lillehammer as a destination (willingness to recommend). Additionally, two other variables were added to the model, namely, nationality and age. Nationality was a dichotomous variable (1 = Norwegian; 2 = foreign) and age was a continuous variable.

Model Construction

To be able to illustrate the diversity of analytical capabilities of PLS-SEM, as seen in Fig. 4, a formative latent variable (activity) expressed by a set of categorical variables, three reflective latent variables (image, satisfaction level, and willingness to recommend) expressed by ordinal variables, and a second-order reflective latent variable (push motives) were constructed. Further, a categorical single item (nationality) and a continuous single item (age) were included in the model. Further, two layers of linkages were constructed to facilitate an illustration of a mediation analysis.

ASSESSMENT AND RESULTS OF A PLS-SEM MODEL

Although PLS estimates both the measurement model and the structural model simultaneously, the assessment of the estimated full model should involve a two-step process encompassing (1) the examination of the measurement model and (2) the assessment of the structural model (Henseler et al., 2009). Whereas the measurement model allows us to examine whether the constructs are measured with satisfactory accuracy, the structural model lets us assess the explanatory power of the model (see Cool, Dierickx, & Jemison, 1989). The former is, however, a prerequisite for the latter as it only makes sense to evaluate the structural model when the measurement model exhibits evidence of reliability and validity (Henseler et al., 2009).

Assessment of the Measurement Model

Since the measurement model contained both reflective and formative and second-order constructs, different criteria are needed to assess psychometric properties of each of these constructs. As far as reflective constructs are concerned, one should examine their composite reliabilities (CR), average variances extracted (AVE), item loadings' size and significance, and discriminant validity (see Liang, Saraf, Hu, & Xue, 2007).

Reliability resembles the internal consistency (i.e., homogeneity and unidimensionality) of a reflective latent construct, a condition which is absolutely necessary as changes in one indicator imply changes in others (Esposito Vinzi et al., 2010). The most common index used to check the reliability of a latent construct is Cronbach's α (used often in PCA-MLR studies), which provides an estimate for the reliability of a latent construct based on the indicator intercorrelations. Cronbach's α does assume tau equivalence (i.e., all indicators are equally important in expressing the latent variables) (Esposito Vinzi et al., 2010), which may not always be the case in reality. Thus, Chin (1998) suggests that Dillon–Goldstein's rho (D.G. rho) is a better measurement of reliability than Cronbach's α as D.G. rho does not assume tau equivalence as it is based on the results from the model (i.e., loadings directly) rather than the intercorrelations among the indicators in the raw data (Esposito Vinzi et al., 2010). Regardless of the index used, the reliability of a latent construct should, as a rule of thumb, be larger than 0.7 (Henseler et al., 2009).

Average variance extracted (average communality) measures the amount of variance (total variance minus unique variance) that a latent variable captures from its corresponding indicators (Chin, 1998). An AVE value of 0.73 for a latent variable would suggest that the latent variable encompasses 73 percent of the variance of its indicators. This figure is simply arrived at by taking the average of the individual communalities of the corresponding indicators (e.g., $(0.609 + 0.854)/2$). Communalities for individual indicators are obtained by squaring the standardized loadings (0.780 and 0.924) of these indicators. As it is generally recommended by scholars (e.g., Chin, 1998) that a latent variable should capture at least 50 percent of the variance of its indicators, the suggested minimum AVE value for a latent variable is 0.50, exhibiting the convergent validity of a latent construct (Fornell & Larcker, 1981). AVE can in fact also be used as an alternative to the existing reliability indexes (α , D.G. rho, etc.) despite the fact that it is a more conservative approach (Fornell & Larcker, 1981). Incidentally, 1 minus AVE (communality) or 1 minus reliability index would yield the unique variance (i.e., unexplained/uncaptured variance) of a latent variable.

The fact that a latent variable should capture at least 50 percent of its indicators sets also a standard (i.e., minimum 0.7) for the size of the standardized loading for an indicator, in that this figure, as explained above, corresponds to a communality of 0.50. Needless to say, this means that 50 percent of the indicator's variance is accounted for by its underlying component/factor. As PCA does not factor out the unique variance of the indicators (Brown, 2006) the standardized loadings should preferably be

much larger than 0.7, a condition which would reduce the amount of unique variance included in the factor extraction in PCA. This then may contribute to estimation of more consistent path coefficients in the structural part (see [Chin, 1998](#)). Finally, it is further expected that the loadings of the indicators are statistically significant as well.

Discriminant validity can be demonstrated when the average variance extracted values of latent variables are larger than the squared correlations among the latent variables in PLS-SEM model, showing that each latent variable shares more variance with its associated indicators than with any other latent variable expressed by different sets of indicators in the model ([Fornell & Larcker, 1981](#)).

As far as the examination of second-order reflective constructs is concerned, a two-step procedure must be followed. First, the measures in the first order are subjected to the same criteria (item loading's size and significance, AVE, etc.) as those that apply to reflective constructs explained just above. Second, the measures in the second order should be assessed based on the level of both CR and AVE (see [Wetzels et al., 2009](#)). In other words, reliability and average variance extracted values of second-order constructs should, respectively, as guidelines, be larger than 0.7 and 0.5.

When it comes to the assessment of formative measures, the only definite criterion one has is to check for the variance inflation factor (VIF) in order to confirm that there is no multicollinearity between the formative indicators forming the construct ([Moreno & Casillas, 2008](#)). As such, VIF values above 2.5 can be considered indicative of multicollinearity ([Allison, 1999](#)) and the concerned indicators should be removed from the equation. As PLS-SEM produces latent variable scores, the concerned latent variable score can be regressed on the indicators forming the latent variable in a separate regression analysis to examine if there is a problem of multicollinearity. Another criterion suggested by [Henseler et al. \(2009\)](#) that can be considered is to check for significant formative indicators; the authors, however, recommend inclusion of the nonsignificant formative indicators as long as this can conceptually be justified. The most important criterion is, however, a strong theoretical rationale the analyst can provide for constructing a latent variable as formative (see [Fornell & Bookstein, 1982](#)).

When it comes to single-item constructs, they are constructed in the same manner as formative constructs (at least in XLSTAT-PLSPM), in that the single item alone forms a latent construct regardless of the conceptual nature (reflective/formative) of the indicator. As such all of the single item's variance is included in its corresponding latent construct (i.e., $h^2 = 1$).

Table 1 provides the details of the measurement model of the current study so that it can be assessed based on the just-mentioned criteria. As observed, all of the reflective measures show satisfactory levels of reliabilities (D.G. $\rho > 0.70$) and average variance extracted ($AVE \geq 0.50$). Further, all of the loadings were above the guideline of 0.70 and statistically significant. Finally, the AVE values were higher than the squared correlations between

Table 1. Results of Measurement Model.

Constructs/Indicators	Loadings	D.G. ρ	AVE
Image		0.90	0.82
Exciting	0.90		
Stimulating	0.91		
Satisfaction		0.86	0.73
I am satisfied with my vacation in the Lillehammer Region	0.77		
My vacation experience in the LR exceeded my expectations	0.93		
Recommendation		0.94	0.89
I am going to recommend to my friends to vacation in the LR	1.00		
I am going to say positive things about LR as a destination			
Nationality		1.00	1.00
Nationality	1.00		
Age		1.00	1.00
Age	1.00		
Escape		0.82	0.67
To get away from the everyday	0.74		
To escape boredom	0.90		
Relaxation		0.85	0.73
To be able to relax mentally	0.82		
To invigorate myself	0.89		
Push motives (second-order)		0.80	0.50
Escape (first-order)			
Relaxation (first-order)			
Activity	Weights	VIF	
Hiking	0.056	1.078	
Cycling	0.036	1.092	
Fishing	1.680	1.108	
Golfing	1.068	1.049	
Visiting farms	1.111	1.039	
Seeing plants and animals	1.432	1.079	
Kayaking/canoeing/rafting	-0.259	1.054	

the latent variables in the model. As such, the reflective part of the measurement model demonstrates both convergent and discriminant validity. When it comes to the formative part, it is here seen that there was no issue of multicollinearity as VIF values were all lower than the suggested figure of 2.5. Although not all of the indicator weights were statistically significant, they were all kept in the model as it was theoretically justifiable asking people what activities they had taken part in to capture the diversity of these activities.

Assessment of the Structural Model

As the measurement model was shown to be an appropriate one (i.e., reliable and valid) the next step is to assess the structural part of the model. As far as the structural model is concerned, one should first examine the multiple correlation coefficients squared (R^2) for each dependent (endogenous) variable in the structural part (Chin, 1998), showing the amount of variance in the dependent variable explained by its corresponding model. The interpretation of R^2 in PLS-SEM is identical to that in traditional MLR (Chin, 1998), and unsurprisingly, there are no definite guidelines for the evaluation of a model as “good or bad” only based upon the magnitude of R^2 as this will be dependent on several factors. This is why scholars are encouraged to do a substantial consideration regarding the nature of the dependent variable, the number and relevance of the independent variables and assumptions etc. (see Allison, 1999; Keith, 2006). Having done so, a model found to explain 25 percent of the variance in a dependent variable can be considered reasonable or having a large effect seen in relation to ANOVA (Keith, 2006). Incidentally, according to Falk and Miller (1992, cited in Duarte & Raposo, 2010), the variance explained for each dependent variable should, as a rule of thumb, be greater than 10 percent.

As R^2 alone cannot be used to evaluate the quality of the whole structural model (i.e., several equations) (Esposito Vinzi et al., 2010), an additional global criterion of GoF has been proposed by Tenenhaus et al. (2004, cited in Esposito Vinzi et al., 2010), whose intent is to account for the performance of the model at both the measurement and the structural parts (all of the R^2 values are taken into consideration) with a focus on overall prediction performance of the model (Chin, 2010). Thus, GoF can be used as an index for validating the model globally (Tenenhaus, Esposito Vinzi, Chatelin, & Lauro, 2005). The GoF index is bounded between 0 and

1; models exhibiting *relative GoF* values equal to or higher than 0.90 can be considered good (Esposito Vinzi et al., 2010).

Another aspect to examine in the structural part of a PLS-SEM model is 3S (sign, size, and significance) of the path coefficients as done in PCA-MLR studies. The standardized path coefficients in PLS-SEM are interpreted in the same manner as in traditional MLR (Chin, 1998). According to Chin (2010), Cohen's (1988) well-established f^2 -values of 0.02, 0.15, and 0.35 can be used to determine whether a predictor (exogenous variable) has, respectively, a small, medium, or large effect. Here, too, scholars are encouraged to do a substantial consideration (i.e., examine the unstandardized coefficients in relation to the measurement scales) to avoid possible incorrect conclusions concerning the importance of the predictors. Further, if the analyst is examining sets of predictors for a dependent latent variable, where the baseline model is compared with adding two or more additional variables, an *F*-test can be performed (Chin, 2010).

Finally, when examining the 3S of path coefficients in a PLS-SEM model with several layers, the indirect effects (i.e., amount of mediation) and their significance should also be identified and assessed. The magnitudes of the indirect effects are interpreted in the same manner as the direct effects. As far as the significance testing of indirect effects is concerned, one should first investigate whether the links between the independent (initial) variable and the dependent variable, between the independent variable and the mediator, and between the mediator and the dependent variable are statistically significant (Baron & Kenny, 1986). Indirect effects with large (≥ 30) standardized coefficients can be considered significant (Esposito Vinzi, 2009). The significance of the indirect effects can, however, be tested by using the well-known Sobel's *Z*-statistic (see Helm, Eggert, & Garnefeld, 2010) or Bootstrap procedure (see Streukens, Wetzels, Daryanto, & de Ruyter, 2010).

Table 2 provides the details of the structural model of the current study so that it can be assessed based on the just-mentioned criteria. As seen, the R^2 values exceeded the required minimum level of 0.10, and further the R^2 value of the main dependent variable (i.e., recommendation) was 0.45 which is much larger than 0.25, considered to be large effect, indicating that the current model was a good one. Further, the *relative GoF* value was 0.96, clearly above the threshold of 0.90, demonstrating a good predictive performance of the global model. As can be observed in Table 2, the three exogenous variables (activity, image, and push motives) had significant positive effects on the first endogenous variable (satisfaction). While the effects of activity and push motives on satisfaction were relatively small

Table 2. Results of Structural Model.

Relations	Direct Effect	Indirect Effect	R^2	GoF
Activity $\rightarrow$ Satisfaction	0.074*	—	—	—
Activity $\rightarrow$ Satisfaction $\rightarrow$ Recommendation	—	0.049	—	—
Image $\rightarrow$ Satisfaction	0.481*	—	—	—
Image $\rightarrow$ Satisfaction $\rightarrow$ Recommendation	—	0.317	—	—
Push Motives $\rightarrow$ Satisfaction	0.115*	—	—	—
Push Motives $\rightarrow$ Satisfaction $\rightarrow$ Recommendation	—	0.076	—	—
Satisfaction model	—	—	0.29	—
Satisfaction $\rightarrow$ Recommendation	0.660*	—	—	—
Nationality $\rightarrow$ Recommendation	-0.094*	—	—	—
Age $\rightarrow$ Recommendation	0.080*	—	—	—
Recommendation model	—	—	0.45	—
The whole model	—	—	—	0.96

* $p < 0.01$.

($\beta < 0.12$), the effect of image on satisfaction was large ($\beta > 0.45$). Moreover, satisfaction and age had significant positive effects on the main endogenous variable (recommendation), whereas nationality had significant negative effect on recommendation (i.e., Norwegian tourists were on average more likely to recommend the destination). The effects of nationality and age on recommendation were low ($\beta < 0.10$), whereas the effect of satisfaction on recommendation was, as expected, considerably large ($\beta > 0.60$). When it comes to the indirect effects, image had relatively large ($\beta > 0.30$) and accordingly significant effect on recommendation via satisfaction, whereas activity and push motives had small ($\beta < 0.08$) and accordingly nonsignificant indirect effect on recommendation via satisfaction.

PLS-SEM versus COV-SEM

PLS-SEM has so far been treated as a typically better alternative to PCA-MLR as the former offers several statistical and practical advantages and benefits over the latter. One can, however, raise the question, “Why use PLS-SEM instead of COV-SEM?” which is a legitimate issue that must be addressed here. Although PLS-SEM (variance-based) and COV-SEM (covariance-based) allow for most of the features that have so far been presented, the former is often preferable to the latter on the basis of the following reasons.

First and most importantly, PLS-SEM is a more appropriate alternative to PCA-MLR than COV-SEM as the purpose of most PCA-MLR studies in tourism research is indeed to explain as much variance as possible in the concerned dependent variable. This claim is based upon the fact that covariance-based SEM analysis approaches involve parameter estimation procedures whose purpose is to reproduce as closely as possible the observed covariance matrix, whereas PLS aims to maximize the variance explained in all dependent variables (Hulland, 1999). One consequence of this difference is that GoF measures (i.e., root mean square error of approximation [RMSEA]) should be used for evaluating COV-SEM works, whereas R^2 should be used for evaluating the performance of PLS-SEM studies (Hulland, 1999) as is the case for PCA-MLR studies.

Secondly and related to the above point, COV-SEM assumes strong theoretical underpinning, whereas PLS-SEM does not make such an assumption, and as such COV-SEM should in fact be considered complementary rather than alternative to each other (Chin, 1998). This implies that studies of mainly empirically oriented nature (i.e., early-stage studies, applied research, etc.) should most certainly adopt a PLS-SEM approach to data analysis, whereas studies whose purpose is to test a theory are recommended to employ a COV-SEM approach. In this respect, studies adopting PLS-SEM approach can be seen as exploratory whereas COV-SEM studies can be viewed as confirmatory.

Despite the fact that PLS-SEM and COV-SEM are complementary approaches, there are, however, some cases in which PLS-SEM is suggested to be considered as an alternative to COV-SEM itself. These are when (1) sample sizes are small (Smith & Barclay, 1997; Tenenhaus, 2008), (2) multicollinearity is severe (Cassel, Hackl, & Westlund, 2000; Kristensen & Eskilden, 2010), (3) multivariate non-normality is the case (Cassel et al., 2000; Squillacciotti, 2010), (4) improper and nonconvergence solutions occur in COV-SEM (Bagozzi & Yi, 1994), (5) measures are not well established (Chin, 2010; Smith & Barclay, 1997), and (6) the model is highly complex with a large number of indicators and/or latent variables (Chin, 2010).

CONCLUSION

The current review and illustration of PLS-SEM show that tourism scholars using PCA-MLR for prediction purposes would be better off with adopting PLS-SEM for exactly the same purpose as PLS-SEM proves to be a more

reliable data analytical approach taking both the measurement and structural models into consideration. From a practical perspective, PLS-SEM is also a superior technique due to its huge flexibility (i.e., complex and various modeling opportunities). It appears further that many studies in tourism research have mistakenly adopted COV-SEM mainly due to the modeling flexibilities (e.g., several/various relations) provided by SEM software even in cases where the purpose of the study is prediction. In such cases, PLS-SEM should be the standard procedure. COV-SEM is primarily confirmatory and should be utilized accordingly. However, PLS-SEM is also considered to be a good alternative to COV-SEM particularly when the sample size is small. Finally, PLS-SEM is indeed an invaluable data analysis approach for research conducted for the tourism industry as such research is often data driven and prediction oriented.

ACKNOWLEDGMENTS

I would like to thank Professor Gudbrand Lien of Lillehammer University College for valuable comments on an earlier version of this chapter.

REFERENCES

- Allison, P. D. (Ed.). (1999). *Multiple regression – A primer*. London: Pine Forge Press.
- Bagozzi, R. P., & Yi, Y. (1994). Advanced topics in structural equation models. In R. P. Bagozzi (Ed.), *Advanced methods of marketing research* (pp. 1–51). Oxford: Basil Blackwell Ltd.
- Baron, R. M., & Kenny, D. A. (1986). The Moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Cassel, C. M., Hackl, P., & Westlund, A. H. (2000). On measurement of intangible assets: A study of robustness of partial least squares. *Total Quality Management*, 11(7), 897–907.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). London: Psychology Press.
- Chin, W. W. (2003). A permutation procedure for multi-group comparison of PLS models. Paper presented at the PLS'03 International symposium – Focus on Customers, September 15–17, Lisbon.
- Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 655–690). Heidelberg: Springer.

- Chin, W. W., & Dibbern, J. (2010). An introduction to a permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 171–193). Heidelberg: Springer.
- Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), 21–41.
- Cohen, E. (1988). Traditions in the qualitative sociology of tourism. *Annals of Tourism Research*, 15(1), 29–46.
- Cool, K., Dierickx, I., & Jemison, D. (1989). Business strategy, market structure and risk-return relationships: A structural approach. *Strategic Management Journal*, 10(6), 507–522.
- Duarte, P. A. O., & Raposo, M. L. B. (2010). A PLS model to study brand preference: An application to the mobile phone market. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 449–485). Heidelberg: Springer.
- Dunteman, G. H. (1989). *Principal components analysis*. London: Sage.
- Esposito Vinzi, V. (2009, October 12–13). Introduction to PLS (partial least squares) world and PLS path modeling: Model and basic algorithm. Lillehammer, Norway.
- Esposito Vinzi, V., Trinchera, L., & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 47–82). Heidelberg: Springer.
- Fornell, C., & Bookstein, F. (1982). A comparative analysis of two structural equation models: LISREL and PLS applied to market data. In C. Fornell (Ed.), *A second generation of multivariate analysis* (pp. 289–323). New York, NY: Praeger.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement errors. *Journal of Marketing Research*, 18, 39–50.
- Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural equation modelling and regression: Guidelines for research practice. *Communications of the Association for Information Systems*, 4, 1–79.
- Haenlein, M., & Kaplan, A. M. (2004). A Beginner's guide to partial least squares analysis. *Understanding Statistics*, 3(4), 283–297.
- Helm, S., Eggert, A., & Garnefeld, I. (2010). Modeling the impact of corporate reputation on customer satisfaction and loyalty using partial least squares. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 515–534). Heidelberg: Springer.
- Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 713–735). Heidelberg: Springer.
- Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, 20, 277–319.
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, 20, 195–204.
- Keith, T. Z. (2006). *Multiple regression and beyond*. London: Pearson Education, Inc.

- Kristensen, K., & Eskilden, J. (2010). Design of PLS-based satisfaction studies. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 247–277). Heidelberg: Springer.
- Law, K. S., & Wong, C.-S. (1999). Multidimensional constructs M structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management*, 25(2), 143–160.
- Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly*, 31(1), 59–87.
- Loureiro, S. M. C., & Gonzales, F. J. M. (2008). The importance of quality, satisfaction, trust, and image in relation to rural tourist loyalty. *Journal of Travel & Tourism Marketing*, 25(2), 117–136.
- Mehmetoglu, M. (2004). Quantitative or qualitative? A content analysis of Nordic research in tourism and hospitality. *Scandinavian Journal of Hospitality and Tourism*, 4(3), 176–190.
- Mendes, J., Do Valle, P. O., & Guerreiro, M. (2011). Destination image and events: A structural model for the Algarve case. *Journal of Hospitality Marketing & Management*, 20(3/4), 366–384.
- Moreno, A. M., & Casillas, J. C. (2008). Entrepreneurial orientation and growth of SMEs: A causal model. *Entrepreneurship Theory and Practice*, May, 507–528.
- Smith, J. B., & Barclay, D. W. (1997). The effects of organizational differences and trust on the effectiveness of selling partner relationships. *Journal of Marketing*, 61(1), 3–21.
- Sosik, J. J., Kahai, S. S., & Piovoso, M. J. (2009). Silver bullet or voodoo statistics? A primer for using the partial least squares data analytic technique in group and organization research. *Group and Organization Management*, 34(1), 5–36.
- Squillacciotti, S. (2010). Prediction oriented classification in PLS path modeling. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: concepts, methods and applications* (pp. 219–223). Heidelberg: Springer.
- Streukens, S., Wetzels, M., Daryanto, A., & de Ruyter, K. (2010). Analyzing factorial data using PLS: Application in an online complaining context. In V. E. Vinzi, W. W. Chin, J. Henseler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 567–587). Heidelberg: Springer.
- Tenenhaus, M. (2008). Component-based structural equation modelling. *Total quality management*, 19(7–8), 871–886.
- Tenenhaus, M., Esposito Vinzi, V., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics and Data Analysis*, 48, 159–205.
- Tiedemann, N., van Birgele, M., & Semeijn, J. (2009). Increasing hotel responsiveness to customers through information sharing. *Tourism Review*, 64(4), 12–26.
- Wetzels, M., Odekerken-Schröder, G., & van Oppen, C. (2009). Using PLS Path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly*, 33(1), 177–195.
- Wold, H. O. A. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon & V. Cappecchi (Eds.), *Quantitative sociology* (pp. 307–359). New York, NY: Academic Press.