The comparison of estimation methods on the parameter estimates and fit indices in SEM model under 7-point Likert scale

Piotr Tarka

Abstract In this article, the author discusses the issues and problems associated with the influence of different estimation methods on the level of obtained parameters and goodness-of-fit of a Structural Equation Model (SEM) in the context of data measured on a 7-point Likert scale. Thus, the objective of the conducted analysis was to compare the selected methods of estimation such as *maximum likelihood* (ML), *maximum likelihood mean adjusted* (MLM), *maximum likelihood mean-variance adjusted* (MLMV), *weighted least squares* (WLS), *weighted least squares mean adjusted* (WLSM) and *weighted least squares mean-variance adjusted* (WLSMV) on the basis of respective parameter statistics, for which the quality of the SEM model fit was assessed. Eventually, among the presented methods, the best estimation procedure was selected. The area of empirical study and the subject of investigation refers to the opinion of consumers about the unethical behavior of companies in the area of marketing.

Poznan University of Economics and Business, Department of Market Research
ul. Niepodleglosci 10, 61-875 Poznan, Poland,
✉ piotr.tarka@ue.poznan.pl

ARCHIVES OF DATA SCIENCE (ONLINE FIRST) DOI 10.5445/KSP/1000058749/10
KIT SCIENTIFIC PUBLISHING ISSN 2363-9881
Vol. 2, No. 1, 2017

1 Introduction

In this article, various estimation methods are compared with regard to the SEM model, based on a preselected 7-point Likert scale. This technique (Likert, 1932) is commonly applied in many empirical studies, for it is fairly easy to develop and a highly reliable approach to measurement. However, the drawback of this technique lies in its specificity, especially in reference to indicators representing latent constructs, that are crude in nature. The crudeness arises from cutting the continuous scale of the theoretical construct into a set number of ordered categories. Another drawback is the error that is often brought into analysis due to imperfection of the scaling technique (Baker, Hardyck, and Petrinovich, 1966). Moreover, Likert measures do not have equally spaced intervals and the data obtained are typically considered as ordinal-scaled data or, at best, somewhere between ordinal and interval-scaled data, hence, only under special research conditions (e.g., with an extended range of response categories), the Likert scale provides quasi normally-distributed values and may yield satisfactory results in the context of normal estimators such as ML (Maximum Likelihood). If the goal of the researcher is to use item-level data and to diagnose the relations based on the covariance structure, the impact of data coupled with the choice of appropriate estimation technique must be also identified.

A main problem which is under discussion, refers to the methodological issues in the context of similarities/differences between various estimation methods: *Maximum likelihood* (ML), *maximum likelihood mean adjusted* (MLM), *maximum likelihood mean-variance adjusted* (MLMV), *weighted least squares* (WLS), *weighted least squares mean adjusted* (WLSM) and *weighted least squares mean-variance adjusted* (WLSMV) and the data collected on a 7-point Likert scale. Therefore, issues are not discussed which pertain to the selection of the optimal number of categories within the scale, as these issues have been profoundly described in the literature (Alwin 1992; Dawes 2008; Revilla, Saris, and Krosnick 2014; Tarka 2016). However, what one can infer from such studies is that the Likert scale with 7 categories ensures higher quality of information and plays the greatest advantage (as compared to the other variants as: 3, 4, 5, 6, 8, 9, 10) not only in the phase of data collection, but leads to better effects in the assessment of the CFA models which are responsible for the measurement of the respective latent variables. The empirical results, derived from simulations, indicate that scales with wider range (8, 9, 10) generate inconsistent scores, because respondents face problems with the appropriate differentiation of the particular point on the scale. In contrast, scales with a narrower range (e.g., 3, 4, 5, 6) generate insufficient information for the measurement models.

Given the above arguments, the author on the basis of empirical research (seeTarka (2016)) has focused on the 7-point option. After completing the comparative analysis of the estimation methods: ML, MLM, MLV, WLS, WLSM and WLSMV, an optimal solution was proposed in reference to SEM model which diagnosed relationships between two latent variables (see Sect. 3).

2 The Methods of Estimation of a SEM Model with Latent Variables

The Structural Equation Model (SEM) (see (Goldberger and Duncan, 1973; Bentler and Weeks, 1980; Muthén and Muthen, 2010)) in its most general form can be expressed by the following (as result of a modification of the LISREL notation):

$$\eta = \alpha_\eta + B\eta + \Gamma\xi + \zeta \tag{1}$$

$$Y = \alpha_Y + \Lambda_{Y\eta} + \varepsilon \tag{2}$$

$$X = \alpha_X + \Lambda_{X\xi} + \delta. \tag{3}$$

Equation 1 represents the latent variable model (SEM) where η is a vector of latent endogenous variables with B a matrix of regression coefficients for the impact of the latent endogenous variables on each other, ξ is the vector of latent exogenous variables with Γ a matrix of regression coefficients for the latent exogenous variable's impact on the latent endogenous variables, α_η is a vector of equation intercepts, and ζ is the vector of latent disturbances that have a mean of zero and are uncorrelated with ξ.

Equations 2 and 3 are the measurement models (CFA) in which Eq. 2 relates Y (a vector of observed variables) to η via a coefficient matrix of factor loadings, Λ_Y. The α_Y is a vector of equation intercepts, and ε is a vector of unique components that have a mean of zero and are uncorrelated with η, ξ, and ζ. Equation 3 is similarly defined as the indicators for the ξ latent variables. Since, each equation can be a factor model, the η and ξ may be classified as the respective latent variables constructed on the basis of the measurement models (Eqs. 2 and 3).

Also there are two general assumptions to be considered in the process of SEM construction. On the basis of **conceptual conditions**, an important question that should be answered affirmatively prior to engaging in SEM is whether the sample at hand comes from a *population* that is relevant to the theoretical ideas being evaluated. Another conceptual problem is whether the data are gathered under *appropriate conditions of measurement* in relation to the theory under investigation. The next issue involves whether a structural theory describes *cause-effect sequences* that occur over time, although the *instantaneous causation* (with simultaneous mutual influence of variables on each other) in specific situation also makes sense in a model (Strotz and Wold, 1960; Bentler and Freeman, 1983). Finally, an important condition of avalid evaluation of a theory via SEM is the *appropriate operationalization of latent variables*.

On the other hand, in case of **statistical conditions**, the researcher needs to assume that data originates from *independent observations* (cases, subjects, sampling units). The next aspect refers to the *selection of appropriate units of the sample*, as the existing methods in structural modeling are often based on the assumption that each of the units or cases in the population has an equal probability of being included in the sample to be studied. SEM also makes assumption about the *linearity of relations between the variables*, which must be conceptually and empirically appropriate to the theoretical questions being addressed. The final issue refers to the methodology for continuous variables in context of their *required distribution*. As such, if variables reflect strong abnormality in distributions then they fail in the estimation of the parameters and the model fit indices (Bentler and Lee, 1983)[1].

The above demands are specific for the successful design of a SEM model, however the optimal choice of the estimation method plays also a significant role. Considering this, one can find a few approaches to estimate a SEM model. For instance, in the light of the normal theory assumptions, one can mention Maximum Likelihood (ML) and Generalized Least Squares (GLS) estimators, which produce asymptotically unbiased, consistent estimates of parameters. In practice, a function of F_{ML} is most frequently used, which fits the covariance structure Σ to the sample covariance structure S by minimizing the discrepancy $F(S,\Sigma(\theta))$ as follows:

$$F_{ML}(\theta) = log|\Sigma(\theta)| + tr\{S\Sigma(\theta)^{-1}\} - log|S| - p. \tag{4}$$

[1] Only methods which are based on the distribution-free variables do not require it, since they enable application of the theory of Tyler (1983), Bentler (1983) and Browne (1984) to correct the normal theory statistics in order to obtain appropriate test statistics and standard errors.

where p is the number of indicators.

The problems with an ML estimator are that the ML is more prone to Heywood cases and is likely to produce markedly distorted solutions if minor misspecifications will be made to the model. Besides, in the social sciences the collected data are barely normal, and although the actual parameter estimates (e.g., factor loadings) may not be seriously affected, nonnormality causes biased standard errors as well as a poorly behaved χ^2-test of overall model fit. The alternative estimator, namely MLM with robust standard errors and χ^2 provides the same parameter estimates as ML, however the model χ^2 and standard errors of the parameter estimates are corrected for nonnormality in large samples. The similar yet alternative, MLMV assumes the correction based on the mean-and-variance.

Another solution, the Weighted Least Squares (WLS or ADF (see Hu, Bentler, and Kano (1992); Curran, West, and Finch (1996))), estimates a weight matrix based on the asymptotic variances and covariances of polychoric correlations that can be used in conjunction with a matrix of polychoric correlations in the estimation of SEM models (Muthén, 1984; Jöreskog, 1994). The WLS applies the following fitting function:

$$F_{WLS} = [S - \sigma(\theta)]^T W^{-1} [S - \sigma(\theta)]. \tag{5}$$

where S is the vector sample statistics (e.g. polychoric correlations), $\sigma(\theta)$ is the model-implied vector of population elements in $\Sigma(\theta)$, and W is a positive-definitive weight matrix.

Browne (1984) has proved that if a consistent estimator of the asymptotic covariance matrix of S will be selected for W (that is a positive-definitive weight matrix) then WLS will lead to asymptotically efficient parameter estimates and correct standard errors as well as a χ^2-distributed model test statistic. In short, he presented the solution for estimating the correct asymptotic covariance matrix in case of continuous but nonnormal distributions in data using observed fourth-order moments, whereas Muthén (1984) invented a methodology based on dichotomous, ordered categorical and continuous indicators. With this strategy, bivariate relationships between ordinal indicators are estimated with polychoric correlations and a SEM model is fit by WLS estimation (Muthén and Satorra, 1995).

Finally, Muthen, DuToit, and Spisic (1997) introduced a robust WLS estimator which profited from earlier works (Satorra and Bentler, 1990; Chou, Bentler, and Satorra, 1991). With this approach, parameter estimates are obtained by substituting a diagonal matrix, for W in the WLS function (Eq. 5), the elements of which are the asymptotic variances of the thresholds and polychoric correlation estimates (i.e., the diagonal elements of the original weight matrix). Once a vector of parameter estimates is obtained, a robust asymptotic covariance matrix can be used to obtain parameter standard errors. A typical matrix involves the full weight matrix W, however, it need not be inverted. In the end, a robust goodness-of-fit test statistic via calculation of a mean-adjusted (WLSM) and mean-and-variance-adjusted (WLSMV) χ^2 is obtained.

3 The Methodology and the Empirical Research

The subject of the empirical study referred to the opinion of consumers about the unethical behavior of companies within the area of marketing activities which influence their increased market consumption. The theoretical foundations were derived from the work of (Tarka, 2016). On their basis, two latent variables were formed (see Fig. 1): **Unfair Advertising Practices** (UAP) and **Lack of Social Responsibility** (LOSR). The relationship LOSR$\rightarrow$UAP was then examined through the agency of a SEM. Each latent variable has been loaded with the respective indicators measured on a 7-point scale, where only the marginal points were labeled, as: 1- *totally disagree*, 7 - *totally agree*:

1. (LOSR): *Firms often make attempts to get as much of their clients' wallets as possible* - (LOSR1); *Companies in pursuit of clients, have changed their marketing practices for worse* - (LOSR2).
2. (UAP): *Most of the advertising contents is misleading and far away from the truth* - (UAP1); *Advertisements prepared by companies can not be treated as a plausible source of information* - (UAP2); *Companies give false color to their products* - (UAP3).

The data were collected through a survey questionnaire (in 2014) among the academic community of students N = 200 (aged between 19-21) at five distinct universities in Poland (Adam Mickiewicz University, University of Technology, University of Economics, University of Life Sciences and University of Medical Sciences). Respondents were selected on the basis of the simple random method of units selection. The sampling frames with a complete list of units were provided by each university. The estimation of data was conducted in Mplus software (version 6.12) on the basis of six methods: ML, MLM, MLMV, WLS, WLSM and WLSMV. The obtained results (under the ML analysis of normally distributed continuous data) were compared to the results obtained from the

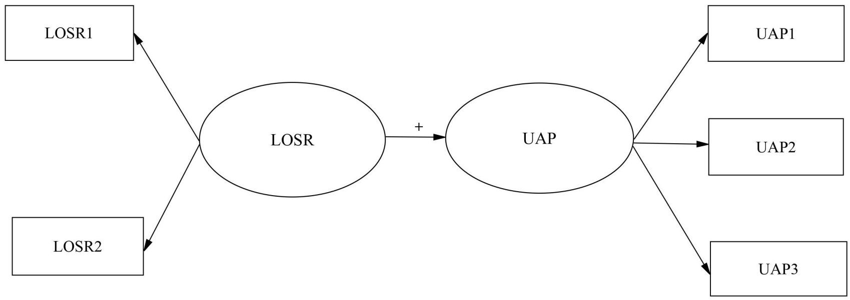

Fig. 1 The SEM model of the relationships between two latent variables: LOSR and UAP.

same data. However here, the author has assumed MLM and MLMV rescaling corrections. Finally, the same data were re-examined with the WLS, WLSM and WLSMV estimators. These methods were considered in the context of finding the best solution of the proposed SEM model assuming its quality is based on fit indices, parameter estimates and standard errors. The results are presented in the next two sections.

4 The Goodness of Fit Indices and the Estimation Methods

For the comparative analysis, selected goodness-of-fit indices were applied: χ^2, CFI, TLI, RMSEA and SRMR. The author's intention was to provide one fit statisticout of several different families of fit statistics rather than many fit statistics from the same family.

In the literature, one can distinguish four general types of measures (Marsh et al, 1988):

1. *Absolute Fit Indices - AFI* (χ^2, GFI, AGFI, Hoelter's CN, AIC, BIC, ECVI, RMR, SRMR): χ^2/df ratio, Goodness of Fit Index (GFI), the Adjusted Goodness of Fit Index (AGFI), Hoelter's CN (critical N), Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Expected Cross-validation Index (ECVI), the Root Mean Square Residual (RMR), and the Standardized Root Mean Square Residual (SRMR).

2. *Relative Fit Indices - RFI* (IFI, TLI, NFI): There are several relative fit indices in this family (which are not explicitly designed to provide penalties for less parsimonious models) including Bentler's Comparative Fit Index (CFI), Bollen's Incremental Fit Index (IFI, also called BL89 or Δ_2), the Tucker-Lewis Index (TLI, Bentler-Bonett Nonnormed Fit Index (NFI or BBNFI), or ρ_2), and the Bentler-Bonett Normed Fit Index (NFI).
3. *Parsimony Fit Indices - PFI* (PGFI, PNFI, PNFI2, PCFI): Parsimonious fit indices include PGFI (based on the GFI), PNFI (based on the NFI), PNFI2 (based on Bollen's IFI), PCFI (based on the CFI).
4. *Noncentrality Parameter - NP* (RMSEA, CFI, RNI, CI): The noncentrality-based indices include the Root Mean Square Error of Approximation (RMSEA) (not to be confused with RMR or SRMR), Bentler's Comparative Fit Index (CFI), McDonald and Marsh's Relative Noncentrality Index (RNI), and McDonald's Centrality Index (CI).

The first family *AFI* does not use an alternative model as a base for comparison, since they are derived from the fit of obtained and implied covariance matrices and the respective minimization function. In contrast, the *RFI* indices compare the χ^2 of the model tested with the null model. Third family of indices *PFI* imposes the adjustments on models within which simpler theoretical processes are favored over more complex ones. However in the literature, it is recommended to evaluate the SEM model independently of the parsimony considerations and to evaluate alternative theories favoring parsimony. With that approach, the researcher would not need to penalize models for having more parameters, but if simpler alternative models seemed to be as good as a more complex model, he might favor the simpler model. Finally, the last family of fit measures *NP* is grounded on the concept of the noncentrality parameter and the rationale for using it is that the standard χ^2 fit is based on a test holding that the null hypothesis is true.

Having presented the main types of goodness-of-fit indices we will now examine the hypothesized SEM model from the perspective of various estimation methods (Tab. 1). From the observation of the χ^2, CFI, TLI, RMSEA and SRMR indices we notice that the model estimated under ML, has obtained a value of the absolute fit index $\chi^2_{(4)} = 5.64$ at p = .23 thereby suggesting the fit of data to the hypothesized model is adequate. However, the WLS estimator is even better ($\chi^2_{(4)} = 5.10$, p-value = .28), also if we compare it to estimators as WLSM ($\chi^2_{(4)} = 5.17$, $p = .27$) and WLSMV ($\chi^2_{(4)} = 5.17$, $p = .27$). In contrast, the worst level of χ^2 can be observed for MLM ($\chi^2_{(4)} = 7.05$, $p = .14$) and

MLMV ($\chi^2_{(4)} = 6.81$, $p = .15$). Both solutions (MLM and MLMV) did not improve the final absolute fit of the model either, as compared to the ML estimator (see the Scaling Correction Factor (SCF = .80) for MLM). Also by reviewing the results of the SRMR index (Tab 1), we notice that particular estimation methods (ML, MLM, MLMV) generated values at the same level of .03.

Table 1 The summary of goodness of fit indices for the hypothesized SEM model.

Fit Indices							
Methods of estimation	χ^2	df(p)	BM χ^2 value/df(p)•	CFI	TLI	RMSEA	SRMR/WRMR
ML	5.64	4 (.23)	116.03/10(0.00)	.98	.96	.06	.03 (SRMR)
MLM	7.05 (.80*)	4 (.14)	119.74/10(0.00)	.97	.93	.08	.03 (SRMR)
MLMV	6.81	4 (.15)	104.48/10(0.00)	.97	.92	.08	.03 (SRMR)
WLS	5.10	4 (.28)	97.55/10(0.00)	.99	.97	.05	-
WLSM	5.17 (.49*)	4 (.27)	471.80/10(0.00)	.99	.99	.05	.26 (WRMR)
WLSMV	5.17 (.49*)	4 (.27)	471.80/10(0.00)	.99	.99	.05	.26 (WRMR)

Legend: * SCF - Scaling Correction Factor; • BM - Baseline Model and χ^2; WRMR - Weighted Root Mean Square Residual

In contrast to χ^2 which measures the extent to which a hypothesized model fits exactly the data, the RMSEA assesses the extent to which it fits reasonably well the data (Browne and Cudeck, 1989). Following the guidelines provided by MacCallum, Browne, and Sugawara (1996) who suggested .01, .05, and .08 levels as indicators of excellent, good, and mediocre fit of the RMSEA index, we find out that the estimators which didn't exceed the .05 level, were WLS, WLSM and WLSMV. This result suggests that the hypothesized SEM model has obtained a good fit. In case of other methods (especially the MLM and MLMV estimators, which obtained values of RMSEA = .08) this level appears to be slightly higher than expected.

By comparing the estimation methods on the basis of the CFI and TLI indices, we notice the proportionate improvement in model fit (starting from ML, through MLM, MLMV, WLS up to the WLSM and WLSMV methods). In particular, the WLS, WLSM and WLSMV estimators have produced substantial values (WLS: CFI = .99, TLI = .97; WLSM: CFI = .99, TLI = .99; WLSMV: CFI = .99, TLI = .99). Slightly lower values (i.e. CFI and TLI) were noted for: MLM (CFI = .97, TLI = .92), MLMV (CFI = .97, TLI = .92), and ML (CFI = .98, TLI = .96). However, their CFI and TLI levels were still fair enough to indicate a well-fitting model. Just to remind, if the value of CFI is close to

1.0, the index shows a perfect fit.[2] In contrast, the TLI denotes the alternative approach to the calculation of CFI. However, the differentiating aspect of the TLI is the inclusion of a penalty function for a model that is overly complex. In general, both indices indicated a well-fitting model.

At last, by inspecting the Weighted Root Mean Square Residual index (WRMR - which was calculated for the WLSM and WLSMV estimators with the assumption of categorical data) and accepting a cutoff .95 criterion as indicative of a good model, we can approve the estimated SEM model as appropriate.

5 The Estimation Methods on the Model Parameters

In the discussion over similarities and differences between estimation methods, one needs to pay attention to the fit of individual parameter estimates exhibiting the correct sign and size. Consequently, any estimates falling outside the admissible range, signal that either the SEM model is wrong or the input matrix has an insufficient level of information. Just to remind, parameter estimates (which are derived from covariance or correlation matrices) that are not positive definite and exhibit out-of-range values (such as correlations > 1.00) as well as negative variances (known as Heywood cases), exemplify unacceptable estimations of values. Also the presence of excessively large or small standard errors indicates poor quality of the model fit. For example, in case of standard errors approaching zero, the test statistic for their related parameters cannot be defined. Likewise, standard errors that are extremely large indicate that parameters cannot be determined. Standard errors are influenced by the units of measurement in observed and latent variables, as well as the magnitude of the parameter estimates, however no definitive criteria ("small" or "large) have been established so far in literature. Finally, the test z statistic (reflecting the parameter estimate divided by its standard error) helps in testing the estimate.

Now, the estimated information (which is presented in Tabs. 2-5, in columns 2-4, with the exception of the parameter symbols column 1) has been grouped according to the parameters model function. The initial block represents factor loadings ($\xi_1 \rightarrow \lambda_{LOSR1}$, $\xi_1 \rightarrow \lambda_{LOSR2}$, $\eta_1 \rightarrow \lambda_{UAP1}$, $\eta_1 \rightarrow \lambda_{UAP2}$, $\eta_1 \rightarrow \lambda_{UAP3}$), the next block represents the relationship between latent variables ($\xi_1 \rightarrow \eta_1$), and the next three blocks are the latent variable ($\xi_{1variance}$), the residual vari-

[2] Although a value $> .90$ was in literature originally considered of a well-fitting model, a revised cutoff value close to .95 has been strongly advised.

Table 2 Selected unstandardized parameters of the hypothesized SEM model - ML estimation.

Parameters	Estimate	Standard Error	Estimate/Standard Error	Two-tailed p-value
$\xi_1 \rightarrow \lambda_{LOSR1}$	1.00	0.00	999.00	999.00
$\xi_1 \rightarrow \lambda_{LOSR2}$	1.33	0.46	2.87	0.00
$\eta_1 \rightarrow \lambda_{UAP1}$	1.00	0.00	999.00	999.00
$\eta_1 \rightarrow \lambda_{UAP2}$	0.74	0.14	5.47	0.00
$\eta_1 \rightarrow \lambda_{UAP3}$	0.62	0.12	5.00	0.00
$\xi_1 \rightarrow \eta_1$	0.99	0.32	3.06	0.00
$\xi_{1variance}$	0.62	0.30	2.06	0.04
δ_{LOSR1}	1.48	0.29	4.95	0.00
δ_{LOSR2}	1.38	0.42	3.29	0.00
ε_{UAP1}	0.31	0.24	1.29	0.19
ε_{UAP2}	1.08	0.21	5.21	0.00
ε_{UAP3}	1.36	0.21	6.31	0.00
ζ_{η_1}	1.36	0.21	6.31	0.00

Table 3 Selected unstandardized parameters of the hypothesized SEM model - obtained in MLM and MLMV estimation.

Parameters	Estimate	Standard Error	Estimate/Standard Error	Two-tailed p-value
$\xi_1 \rightarrow \lambda_{LOSR1}$	1.00	0.00	999.00	999.00
$\xi_1 \rightarrow \lambda_{LOSR2}$	1.33	0.46	2.87	0.00
$\eta_1 \rightarrow \lambda_{UAP1}$	1.00	0.00	999.00	999.00
$\eta_1 \rightarrow \lambda_{UAP2}$	0.74	0.15	4.95	0.00
$\eta_1 \rightarrow \lambda_{UAP3}$	0.62	0.12	4.80	0.00
$\xi_1 \rightarrow \eta_1$	0.99	0.28	3.52	0.00
$\xi_{1variance}$	0.62	0.30	2.17	0.03
δ_{LOSR1}	1.48	0.30	4.95	0.00
δ_{LOSR2}	1.38	0.44	3.14	0.00
ε_{UAP1}	0.31	0.18	1.78	0.07
ε_{UAP2}	1.08	0.21	5.05	0.00
ε_{UAP3}	1.36	0.46	2.98	0.00
ζ_{η_1}	1.06	0.38	2.76	0.00

ances of indicators (δ_{LOSR1}, δ_{LOSR2}, ε_{UAP1}, ε_{UAP2}, ε_{UAP3} and the disturbance ζ_{η_1} pertaining to the latent variable)[3]. From their scores we infer that, in case of all estimation methods, all parameter estimates were reasonable and statistically significant (> 1.96). There were also no negative variances which might cause unacceptable estimated values. The only questionable parameter was noticed

[3] Two paths $\xi_1 \rightarrow \lambda_{LOSR1}$ and $\eta_1 \rightarrow \lambda_{UAP1}$ were fixed to 1.00 for purposes of identification.

Table 4 Selected unstandardized parameters of the hypothesized SEM model - WLS estimation.

Parameters	Estimate	Standard Error	Estimate/Standard Error	Two-tailed p-value
$\xi_1 \to \lambda_{LOSR1}$	1.00	0.00	999.00	999.00
$\xi_1 \to \lambda_{LOSR2}$	1.53	0.55	2.77	0.00
$\eta_1 \to \lambda_{UAP1}$	1.00	0.00	999.00	999.00
$\eta_1 \to \lambda_{UAP2}$	0.81	0.12	5.41	0.00
$\eta_1 \to \lambda_{UAP3}$	0.61	0.12	5.08	0.00
$\xi_1 \to \eta_1$	0.91	0.24	3.86	0.00
$\xi_{1variance}$	0.59	0.26	2.23	0.03
δ_{LOSR1}	1.51	0.30	5.11	0.00
δ_{LOSR2}	1.10	0.55	2.01	0.04
ε_{UAP1}	0.31	0.15	2.06	0.03
ε_{UAP2}	0.75	0.19	3.94	0.00
ε_{UAP3}	1.26	0.32	3.93	0.00
ζ_{η_1}	1.14	0.34	3.32	0.00

Table 5 Selected unstandardized parameters of the hypothesized SEM model - obtained in WLSM and WLSMV estimation.

Parameters	Estimate	Standard Error	Estimate/Standard Error	Two-tailed p-value
$\xi_1 \to \lambda_{LOSR1}$	1.00	0.00	999.00	999.00
$\xi_1 \to \lambda_{LOSR2}$	1.03	0.29	3.47	0.00
$\eta_1 \to \lambda_{UAP1}$	1.00	0.00	999.00	999.00
$\eta_1 \to \lambda_{UAP2}$	0.74	0.07	10.56	0.00
$\eta_1 \to \lambda_{UAP3}$	0.66	0.07	8.78	0.00
$\xi_1 \to \eta_1$	0.97	0.28	3.45	0.00
$\xi_{1variance}$	0.38	0.12	3.08	0.00
ζ_{η_1}	0.55	0.15	3.67	0.00

in the residual variance ε_{UAP1} which has obtained nonsignificant values, below 1.96 (according to information of column 4 for the respective three estimators: ML (with 1.29 estimate/standard error calculated at $p = 0.19$), MLM (with 1.78 estimate/standard error at $p = 0.07$), MLMV (with 1.78 estimate/standard error at $p = 0.07$). However, another finding indicates that when we use the estimator (WLS), the significance level of this parameter will be improved (see the WLS estimator where: 2.06 estimate/standard error was calculated at $p = 0.03$). In consequence, the WLS procedure eliminated the nonsignificance effect associated with the residual variance of UAP1.

Considering the size of the particular parameter estimates, we find it is more or less similar within the range of ML, MLM and MLMV estimators. Parameters did not considerably change in WLS estimation. For instance, factor loadings of the observed variables indicate only a slight modification within the obtained values: $\eta_1 \rightarrow \lambda_{UAP2}$ (ML = .74, MLM = .74, MLMV = .74, WLS = .81); $\eta_1 \rightarrow \lambda_{UAP3}$ (ML = .62, MLM = .62, MLMV = .62, WLS = .61). On the other hand, the size of standard errors has mostly decreased (with the exception of ε_{UAP3} in ML) in WLS as compared to ML, MLM and MLMV estimation methods: ε_{UAP1} (ML = .24, MLM = .18, MLMV = .18, WLS = .15); ε_{UAP2} (ML = .21, MLM = .21, MLMV = .21, WLS = .19); ε_{UAP3} (ML = .21, MLM = .46, MLMV = .46, WLS = .34).

Interestingly, in WLSM and WLSMV, the standard errors have shown even more descending values as compared to: ML, MLM, MLMV and WLS. For example, differences can be observed in case of the following parameters: $\xi_1 \rightarrow \lambda_{LOSR2}$ (ML = .46, MLM = .46, MLMV = .46, WLS = .55, WLSM = .29, WLSMV = .29); $\eta_2 \rightarrow \lambda_{UAP2}$ (ML = .14, MLM = .15, MLMV = .15, WLS = .12, WLSM = .07, WLSMV = .07); $\eta_1 \rightarrow \lambda_{UAP3}$ (ML = .12, MLM = .12, MLMV = .13, WLS = .12, WLSM = .07, WLSMV = .07); $\xi_{1variance}$ (ML = .30, MLM = .30, MLMV = .28, WLS = .26, WLSM = .12, WLSMV = .12); ζ_{η_1} (ML = .21, MLM = .38, MLMV = .38, WLS = .34, WLSM = .15, WLSMV = .15).

6 Conclusions

On the basis of the conducted analysis, we can eventually conclude that all estimators have produced approximately tolerable results. However, in searching for the optimal solution for the SEM model we find that its general fit as well as the parameter levels seemed to perform better for WLSM and WLSMV compared to the ML, MLM, MLMV and WLS estimation methods. Although, the range of response categories used within the 7-point Likert scale proves to be sufficient and approximates the conditions of the normal distribution, this type of scale is, in fact, still framed within the categorical answers. Hence, strict application of estimators based on normal theory, though this appears to be appropriate on the first sight, may not be enough from the perspective of obtaining the optimal solution for the SEM model. Just to remind, computing the correlation/covariance coefficients between indicators and application of the

ML estimator in a SEM model, needs the solid foundation of normality, at least in the context of a multivariate distribution for the respective latent variable. In other cases, such variables should be corrected with a MLM or MLMV estimator or should be restructured with order categorical scales and estimated with appropriate procedures, i.e. WLS, WLSM or WLSMV. Still, when the number of response categories on the scale is large (as it is in the case of a 7-point scale) and the data are approximately normal, failure to address the ordinality of the data can probably be negligible. One can even argue (given the normally distributed categorical variables), that continuous methods/estimators can be used with little worry when a variable has 7 categories. The problem however is, that in case of indicators which are measured on a 7-point Likert scale, they cannot completely produce the optimal solution in SEM as compared to the noncontinuous estimators. Therefore, a quick application of continuous estimators (ML, and its corrected alternative propositions: MLM and MLMV) seems to be dangerous, since they cause strong side effects on the χ^2 test of model fit, (including the CFI, TLI, RMSEA and SRMR fit indices). The factor loadings (as shown by the empirical results) can be also underestimated in ML, MLM and MLMV methods as compared to WLS, WLSM and WLSMV. Besides, the residual variance of estimates, more than other parameters, seems to be more sensitive to the continuous methods than the noncontinuous variants. Of these, we can infer that the WLSM and WLSMV estimation methods perform better in SEM modeling of a 7-point Likert-scaled data and yield more accurate test statistics, parameter estimates and standard errors.

References

Alwin DF (1992) Information transmission in the survey interview: Number of response categories and the reliability of attitude measurement. Sociological Methodology 22:83–118, DOI 10.2307/270993

Baker BO, Hardyck CD, Petrinovich LF (1966) Weak measurements vs. strong statistics - an empirical critique of s.s. stevens's proscriptions on statistics. Educational and Psychological Measurement 26(2):291–309, DOI 10.1177/001316446602600204

Bentler PM (1983) Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. Psychometrika 48(4):493–517, DOI 10.1007/BF02293875

Bentler PM, Freeman EH (1983) Tests for stability in linear structural equation systems. Psychometrika 48(1):143–145, DOI 10.1007/BF02314682

Bentler PM, Lee SY (1983) Covariance structures under polynomial constraints - applications to correlation and alpha-type structural models. Journal of Educational Statistics 8:207–222, DOI 10.2307/1164760, URL http://www.jstor.org/stable/1164760

Bentler PM, Weeks DG (1980) Linear structural equations with latent variables. Psychometrika 45(3):289–308, DOI 10.1007/BF02293905

Browne M, Cudeck R (1989) Single sample cross-validation indices for covariance structures. Multivariate Behavioral Research 24(4):445–455, DOI 10.1207/s15327906mbr2404_4

Browne MW (1984) Asymptotically distribution-free methods for the analysis of covariance structures. British Journal of Mathematical and Statistical Psychology 37(1):62–83, DOI 10.1111/j.2044-8317.1984.tb00789.x

Chou CP, Bentler PM, Satorra A (1991) Scaled test statistics and robust standard errors for nonnormal data in covariance structure analysis - a monte carlo study. British Journal of Mathematical and Statistical Psychology 44:347–357, DOI 10.1111/j.2044-8317.1991.tb00966.x

Curran PJ, West SG, Finch JF (1996) The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. Psychological Methods 1:16–29, DOI 10.1037//1082-989X.1.1.16

Dawes J (2008) Do data characteristics change according to the number of points used? an experiment using 5-point, 7-point and 10-point scales. International Journal of Market Research 50:61–77

Goldberger A, Duncan OD (1973) Structural equation models in the social sciences. Academic Press, New York

Hu L, Bentler PM, Kano Y (1992) Can test statistics in covariance structure analysis be trusted? Psychological Bulletin 112(2):351–362, DOI 10.1037//0033-2909.112.2.351

Jöreskog KG (1994) On the estimation of polychoric correlations and their asymptotic covariance matrix. Psychometrika 59(3):381–389, DOI 10.1007/BF02296131

Likert R (1932) A technique for the measurement of attitudes. Archives of Psychology 140:1–55

MacCallum RC, Browne MW, Sugawara HM (1996) Power analysis and determination of sample size for covariance structure modeling. Psychological Methods 1:130–149, DOI 10.1037//1082-989X.1.2.130

Marsh HW, Balla JR, McDonald RP (1988) Goodness of fit indexes in confirmatory factor analysis - the effect of sample size. Psychological Bulletin 103:391–410, DOI 10.1037//0033-2909.103.3.391

Muthén B (1984) A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. Psychometrika 49(1):115–132, DOI 10.1007/BF02294210

Muthen B, DuToit SHC, Spisic D (1997) Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Unpublished manuscript, Los Angeles

Muthén BO, Satorra A (1995) Technical aspects of muthén's liscomp approach to estimation of latent variable relations with a comprehensive measurement model. Psychometrika 60(4):489–503, DOI 10.1007/BF02294325

Muthén LL, Muthen B (2010) Mplus - statistical analysis with latent variables - user's guide. Muthén & Muthén, Los Angeles

Revilla AM, Saris WE, Krosnick JA (2014) Choosing the number of categories in agree-disagree scales. Sociological Methods and Research 43(1):73–97, DOI 10.1177/0049124113509605

Satorra A, Bentler PM (1990) Model conditions for asymptotic robustness in the analysis of linear relations. Computational Statistics and Data Analysis 10(3):235–249, DOI 10.1016/0167-9473(90)90004-2

Strotz RH, Wold HOA (1960) Recursive vs. nonrecursive systems - an attempt at synthesis. Econometrica 33(2):1–41, DOI 10.2307/1907731

Tarka P (2016) to appear in - studies in classification, data analysis and knowledge organization - analysis of large and complex data, Springer, Berlin, chap CFA MultiTrait-MultiMethod model in comparative analysis of 5, 7, 9 and 11 point scales

Tyler DE (1983) Robustness and efficiency properties of scatter matrices. Biometrika 70(2):411–420, DOI 10.1093/biomet/70.2.411