

An extensive comparison of CB-SEM and PLS-SEM for reliability and validity

Asyraf Afthanorhan^a, Zainudin Awang^a and Nazim Aimran^{b*}

^aFaculty of Business and Management, UniSZA Gong Badak, 21300 Kuala Nerus, Terengganu, Malaysia

^bAdvanced Analytics Engineering Centre, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam., Malaysia

CHRONICLE

Article history:

Received: July 20, 2020

Received in revised format: August 29, 2020

Accepted: September 12, 2020

Available online: September 12, 2020

Keywords:

CBSEM

PLS-SEM

Standardized Loadings

Reliability and Validity

ABSTRACT

Structural Equation Modeling (SEM) includes measurement and structural model for hypothesis testing. The results yielded from structural model is unlikely to be valid if a poor loading of an indicator is selected. The impact of these erroneous result on standardized loading is disregard. Thus, knowing how poor loading can affect the validity of measurement model is a crucial issue. This paper attempts to compare the standardized loadings result between two prominent SEM methods (CBSEM and PLS-SEM) using three varied of simulation models (TRA, Loyalty and UTAUT model) to investigate their effects on reliability and validity of measurement model. The data for each model were generated using R software by setting the value of standardized loading and the construct correlations (N=50, 100, 200 and 500). The value of standardized loadings was set to 0.60 for each construct in the model while the construct correlations were set in the range between 0.45 to 0.65. Then, the AMOS 21.0 and ADANCO 2.0 were used to perform the statistical analysis. It shows that good standardized loading can increase the reliability and validity of construct representation. CBSEM is particularly yielded valid and unbiased estimation under confirmatory condition (established theory) compared with PLS-SEM. The results are illustrated with empirical examples. This paper provides updated evidence about CBSEM and PLS-SEM when assessing the measurement model.

© 2020 by the authors; licensee Growing Science, Canada.

1. Introduction

Many of the major concepts of interest in the standard social science and management research cannot be directly measured, but can only be assessed by hypothesizing the relationship between latent variable and other variables (item or manifest variable). As a consequence, measurement theory that relates latent variable and variables must be interrelated (Edwards & Bagozzi, 2000). This approach has been called ‘effect indicator’ or reflective construct which has gained popularity in many fields (Bollen, & Diamontopoulos, 2015). Major writings have focused on this type of modeling (Bollen & Lennox, 1991) in their specific research project due to less computational demand. By doing so, there are many more assessments needed to be satisfied before moving to hypothesize model. Previously, there are many literatures that showed the bias and inconsistency estimates frequently occurred with PLS-PM in research projects (Rönkko & Evermann, 2013; Rönkko et al., 2016; Antonakis et al., 2010; Aguirre-Urreta & Marakas, 2014; Marcoulides & Saunders, 2006; Goodhue, Lewis, & Thompson, 2012) from their simulation. However, their simulation results were unsuitable for social science scholars who use the SEM method for hypothesis testing. In additions, the results shown from their findings are very broad which makes it difficult to understand. Regarding on this matter, this study attempts to focus more specifically on evaluating the measurement model in terms of the construct reliability and validity. Apart from that, the three established model were used for simulation purpose with different

* Corresponding author.

E-mail address: ahmadnazim@uitm.edu.my (N. Aimran)

sample size to provide more comprehensive findings. Then, the two competitive methods were compared across sample size and population model. From this point of view, there are three objectives need to be conducted. First, the researchers intend to determine the biasedness of standardized loadings across sample size and model. Second, the reliability of each construct was compared between the CBSEM and PLS-SEM. Third, the convergent validity using Average Variance Extracted (AVE) approach were also compared. Lastly, the findings and recommendation are critically discussed.

2. Methodology

2.1 Standardized loading

One of the central assumptions of the application of SEM is that all the constructs must be included in the model which is based on the measurement theory (Aguirre-Urreta, Rönkko, & Marakas, 2016). This requirement is discussed not only in methodological literature (Kline, 2016; Westland, 2015; Mertens et al., 2017) but it also presents in guidelines and recommendations offered to researchers in the applied sciences (Aimran et al., 2017; Mohamad et al., 2019). For instance, the acceptable variable retained in the model must be above 0.60 of factor loading and sometimes it can be lower than that depending on their rationale. Though there is an extensive agreement that the omission of a relevant effect indicator is essential in SEM, the consequence of doing so have not been carefully considered. Since the consistent method is required to warrant this value. As noted by Jarvis et al. (2003), dropping effect indicator may change the meaning of the variables. Thus far, CBSEM is the best choice for this situation since this method is free from inconsistency estimates. The factor loading in PLS method is based on the weight vector as follows,

$$\bar{w} = \frac{\lambda_j}{\sqrt{\lambda_i \sum \lambda_{jj}}}, \tag{1}$$

which concludes that the weight vector is proportional to parameter value of factor loadings. This tradition is followed for the mode A that focus to reflective measurement model.

3. Research model for simulation

Fig. 1, Fig. 2, and Fig. 3 present the three population models that will be tested with two different methods of SEM. These population models were chosen due to their popularity in management and information research. As a matter of fact, these models were indeed classified as confirmatory model where the position and role of each construct and indicator are correctly specified. Thus, the data were generated for three population models using R software. The value of standardized loadings for each model is set to 0.60 as it is regarded as the minimum threshold in determining the strength of indicator to their respective construct. Also, the four items were embedded on each construct to pass the minimum requirement of construct nature. In terms of the sample size, four conditions were tested at 50, 100, 200 and 500.

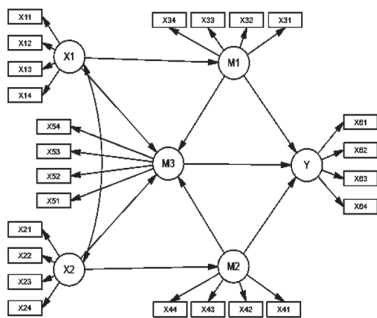


Fig. 1: TRA Model

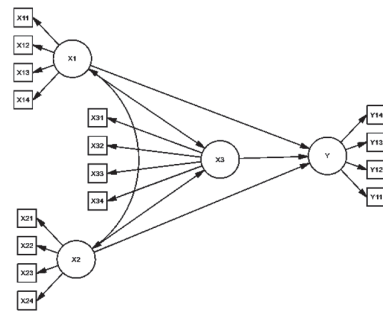


Fig. 2: Loyalty Model

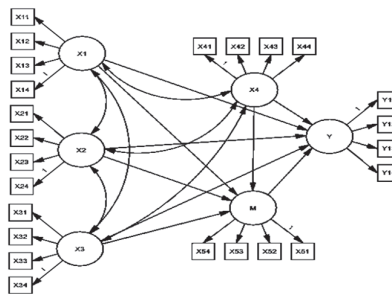


Fig. 3: UTAUT Model

be occurred in varying situations. However, since the underlying distribution of parameter estimates to CBSEM is normal and stringent assumptions, the chance of CBSEM exaggerated by the improper solutions is higher than the PLS-SEM. It can be marked by the finding appears in the Table 1 that showing the standardized loading of $Y11 = 1.033$ is higher than 1.0 at the small sample size ($N=50$). Based on this, it can be stated that the problem of Heywood cases under CBSEM occurred at the small sample size. Despite the fact that this problem may occur for the small samples, yet, the findings of factor loading become trustworthy at the samples of 100 and above. It means that, the reasonable of sample size for CBSEM is 100 validated as suggested from prior researches (Kline, 2016; Afthanorhan et al., 2019). In contrast, the PLS-SEM does not suffer from the presence of Heywood cases in TRA model. All the values of factor loadings for PLS-SEM are in the range of 0.0 to 1.0 which means the results for factor loadings are proper manner under assessment of measurement model. However, in terms of the parameter accuracy, the value of factor loading with PLS-SEM is highly different from the true value which means the biased estimate of factor loadings may exist in TRA model. These results are likely confirmed by the assessment of raw bias on factor loadings. Compared to CBSEM, the values of factor loadings are bias and thus the construct reliability for CBSEM may reliable than PLS-SEM. Also, it can be seen that how large the value of factor loadings with PLS-SEM deviates from the true value of population model.

In the simulation design, all the indicators of every construct were set using values of 0.60 which means the cut off for true value of factor loadings should be 0.60 per model. Since the findings are related to the comparison of factor loadings, the raw bias is calculated to assess the discrepancy values between the true values and estimated values. The smaller the values of raw bias of factor loadings, the more accurate of factor loadings per se. In order to approve the findings of factor loadings, the other two established models were tested repeatedly according to the same procedures that are Loyalty and UTAUT models.

For Loyalty and UTAUT models, the consistency of those estimation techniques is remained consistent at different levels. The Loyalty model is opted because the size model is slightly smaller than TRA and UTAUT model. With the researchable on different size model for adopting a strategy of Monte Carlo simulation, the potential of the estimation technique can be explained in detailed. With CBSEM on Loyalty and UTAUT model, the problem of Heywood cases does not exist in the model and the finding is similar for PLS-SEM.

4.2 Reliability and convergent validity results

Table 2 shows the results of reliability and validity and bias between CBSEM and PLS-SEM with varying of sample sizes. In addition, the values of reliability and validity constructs were compared with the true values of population model that is more efficient and informative in determining to which estimation techniques are reliable when assessing the measurement models which is one of the purposes for this analysis. Many studies believe the reliability and validity are essential features in the SEM context analysis in order to assure the path coefficients obtained can be generally trusted for managerial decision or for dynamics progress (Bagozzi & Philips, 1991; Brown, 2006; Bollen, 1989; Schumacker & Lomax, 2010; Kline, 2016; Raykov & Marcoulides, 2010; Harrington, 2009; Aziz, Afthanorhan & Awang, 2016).

However, previous studies have shown that the simulation approach is not common when SEM is applied to the assessment of reliability and validity such as those analyzed by Henseler et al. (2014) and Dijkstra & Henseler (2015). Paxton et al. (2001) assert that sample size is one of the most important variables in a simulation which means the study should adopt a strategy of different sample size for simulation addressed with the design factors. By addressing on their comment, the current study uses simulation for reliability and validity. For the true reliability of population model for TRA model, all constructs were set using a value of 0.692, meanwhile, the true validity for this model is set using a value of 0.360.

The analysis with CBSEM and PLS-SEM are conducted for every sample and then it is possible to interpret the comparative results on reliability and validity of measurement models. TRA model shows that the estimates of reliability and validity with CBSEM is much more accurate than those of PLS-SEM. To be sure, the bias of reliability and validity is depicted in Table 2 for verification. The bias for CBSEM is much smaller than PLS-SEM which means the estimation technique for CBSEM is lacking. In further illustrations, with the true reliability as 0.692, the estimates of reliability for CBSEM on TRA model is very close to the true value of population model. Unlike CBSEM, the distance reliability is estimated for PLS-SEM with true values which means their performance for assessing the constructs reliability cannot be dependable. The main reason of this problem is due to the effects of small indicator loadings. When small indicator loadings retained in the model, it means the indicator error is high.

Nevertheless, the assessment for PLS-SEM can be recognized by inclusion of the new reliability that is called Dijkstra-Henseler rho. Apparently, this new reliability estimates are mimic to the existence of the reliability estimates with CBSEM. Their estimations are slightly close to the true values and they are verified by the assessment of bias reliability. Subsequently, the other two models were determined with the same manner. Based on their performance, the authors confirmed that composite reliability is not appropriate when PLS-SEM is applied to common factor model. Other than that, it is much worthy of the PLS-SEM only concentrate their new reliability rho in determining the construct reliability. That is because their performance of the construct reliability is comparable with composite reliability with CBSEM.

In terms of the construct validity, the Average Variance Extracted (AVE) is tested with the true validity of population model. Surprisingly, the results show that the AVE estimates with CBSEM are more accurate than PLS-SEM across samples and models and the performance of CBSEM is not only exquisite for construct reliability. In this study, the authors set the value of true validity as 0.360 and it seems that all value estimates with CBSEM are in the acceptable range. Unfortunately for PLS-SEM, their AVE values are overestimated across samples and models.

Moreover, there is no new development to validity criterion for replacing the AVE criterion for assessing their construct validity. Instead, the PLS-SEM still relies on AVE criterion when assessing the measurement model whether in the form of composite model or common factor model. With these drawbacks appear in the terminology of PLS-SEM, the PLS users supposedly aware that research with this specified method for common factor model is not useful as it is remained incompletely for evaluating the measurement models.

5. Discussion and conclusion

One of the central assumptions of the application of structural equation modeling is that all the constructs must be included in the model which is based on the measurement theory (Aguirre-Urreta, Rönkko, & Marakas, 2016). The construct can be known as measurement model which is expressed how to measure the construct by means of a set of indicators (Jarvis, MacKenzie, & Podsakoff, 2003). Generally, there are two ways to conceptualize the measurement models: 1. Reflective measurement and 2. Formative measurement. The first approach is reflective measurement considered when there are some relationships between the construct with its indicators (Bollen, 1989). Different from reflective measurement perspectives, the formative measurement was illustrated when the relationships are going from the indicators to their constructs. Oftentimes, most of the studies prefer reflective measurement to operationalize the measurement theory. In a reflective measurement model, indicator loadings are the top priority for investigations.

To produce the indicator loadings, the CBSEM and PLS-SEM have their own way for estimating indicator loading when assessing the measurement model. In some cases, the indicator loadings can be known as factor loadings or pattern coefficients (Kline, 2016). The labeling of factor loading is not relevant for the composite method as PLS-SEM. In accordance to Rönkko & Evermann, (2013), the value of indicator loadings from composite method cannot be equated as CBSEM does. This indicator loading with composite method should not be acknowledged as factor loadings but known as composite loading. The composite based approach to structural equation modeling uses linearly combines indicators to form composite variables (Lohmoller, 1989; Sarstedt et al., 2016). This approach is different from what CBSEM does that consider construct as common factor that explain the covariation between their associated items (Jöreskog & Wold, 1982). Based on this reason, the construct with PLS-SEM fully depends on the value of associated items or indicators which implies that adding or omitting an indicator potentially modify the nature of the construct (Bollen, 2011; Aguirre-Urreta, Rönkko, & Marakas, 2016).

Therefore, adding or omitting the items from respective construct with composite method is more dangerous than the factor method. Other than that, the value of indicator loading for each item is always altered when some of the indicator from the same construct is removed, which implies the estimation with PLS-SEM is not consistent at all under assessment of measurement model. Indeed, the PLS-SEM is recently being improved by providing more assessment in statistical analysis but it is unjustifiable to claim it is appropriate for hypothesis testing (confirmatory) as it always biased.

With this approach, the researchers are difficult to confirm the value of their items for justifications and thus the meaning of the construct may have affected. Moreover, composite method generally inherits the measurement error of manifest variables. The measurement error is the difference between the true value of a variable and the value obtained by a measurement. There are many sources of measurement error, including poorly worded questions on a survey, misunderstanding of the scaling approach, and incorrect application of a statistical method (Hair et al., 2016). As such, there is no way composite method is affected by the presence of measurement error. It is indeed Consistent PLS (adjusted PLS-SEM) realize for this situation to reduce the measurement error by using MINRES approach (Dijkstra, 2010). Yet, this reduction does not clearly explain in any research insofar whether the estimates loadings are not affected by this detrimental effect of measurement error.

Unlike PLS-SEM, the CBSEM segregates the measurement error from the manifest variable to ensure the true estimate for factor loading can be obtained (McDonald, 1996). Another advantage with CBSEM is the tool of modification index has offered for this application to identify the multicollinearity problem in the model (McIntosh et al., 2014). The multicollinearity problem may appear when more than two indicators are involved for reflective measurement (Afthanorhan et al., 2019; Majid et al., 2019). This approach helps the researchers detect which item has potential to harm the construct meaning where PLS-SEM does not.

Based on the Table 1, the result clearly shows that PLS-SEM is much bias than those CBSEM under different situations. Their values of indicator loadings have been always shown overestimated under various situations which eventually may affect the construct meaning. In this sense, the bias of indicator loading with PLS-SEM exists due to the effect of measurement error. As mentioned earlier, the measurement error is not segregated from the variable but this effect is accumulated with the variable meanings to exhibit the value of indicator loadings.

Therefore, the value for indicator loading with PLS-SEM is always higher than those CBSEM. Strictly speaking, the indicator loading with PLS-SEM is not having 'conceptual unity' and sharing the similar theme (Bollen, 2011; Bollen & Pearl, 2013) which means that the construct must be measured by the corresponding variable where having the same concept or concept dimension (Bainter & Bollen, 2015). As such, the PLS-SEM does not have an ability to confirm this theme due the effect of measurement error. As to whether the PLS-SEM can produce the consistent indicator loading, it is useful to initially ensure the bias of this method is determined. Since the PLS-SEM is deductively not relevant for confirmatory mode due to the presence of bias (Rönkko & Evermann, 2013; Antonakis et al., 2010; McIntosh et al., 2014; Rönkko et al., 2016).

Finally, when we determine the indicator loadings under assessment of measurement model in the case for confirmatory purpose, we should consider to choose CBSEM over PLS-SEM which is did a better job of producing accuracy estimates for indicator loadings. Bentler (1982) stated that if the researchers interested in examining the relationships between variables and unobserved variables, the CBSEM is a method of choice in which latent variable cannot be reduced to weighted composite of observed variables. This, however, the solution for CBSEM is in question when using small sample size where the problem of Heywood cases occurred for this situation. This effect is not occurred with PLS-SEM even being tested with small samples. In this essence, the PLS-SEM seems 'always converged' and yielded proper solutions (no Heywood cases). Under these circumstances, the results showed that PLS-SEM cannot detect the mis-specified model (McIntosh et al., 2014; Evermann & Tate, 2010) where CBSEM can. It should be noted that the method with no guide for detecting the model misspecification is unjustifiable since the study in the sense of confirmatory needs warranted for evidence of measurement theory.

Implementing PLS-SEM for confirmatory modeling could produce erroneous conclusions (McDonald, 1996). The fact that PLS-SEM is appropriate for confirmatory modeling is not reasonable here where three population models being used. It is much appreciated if the PLS-SEM focus on bias property rather than ensuring the consistency property in all levels. For this reason, there are actually a few important exceptions where regression with composite is a consistent and unbiased estimator of latent variables such as model-implied instrumental variables (Bollen, 2019), Bartlett factor scores (Skrondal & Laake, 2001) and correlation-preserving factors scores (Grice, 2001). Among of these methods, the application of MIIV is available in R software (Gates, Fisher, & Bollen, 2019). With this presence, it is more sensible for those interested in composite method to structural equation modeling.

References

- Afthanorhan, A., Foziah, H., Rusli, R., & Khalid, S. (2019). The effect of service quality on customer satisfaction in three campuses of UniSZA. *International Journal of Innovation, Creativity and Change*, 7(10), 42-56.
- Aimran, A. N., Ahmad, S., Afthanorhan, A., & Awang, Z. (2017). The development of comparative bias index. *Paper presented at the AIP Conference Proceedings, 1870* doi:10.1063/1.4995935
- Aguirre-Urreta, M. I., & Marakas, G. M. (2014). A rejoinder to Rigdon et al.(2014). *Information Systems Research*, 25(4), 785-788.
- Aguirre-Urreta, M. I., Rönkkö, M., & Marakas, G. M. (2016). Omission of causal indicators: consequences and implications for measurement. *Measurement: Interdisciplinary Research and Perspectives*, 14(3), 75-97.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086-1120.
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly*, 421-458.
- Bainter, S. A., & Bollen, K. A. (2015). Moving forward in the debate on causal indicators: Rejoinder to comments. *Measurement: Interdisciplinary Research & Perspectives*, 13(1), 63-74.
- Bentler, P. M. (1982). Confirmatory factor analysis via noniterative estimation: A fast, inexpensive method. *Journal of Marketing Research*, 19(4), 417-424.
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research*, 17(3), 303-316.
- Bollen, K. A. (2011). Evaluating effect, composite, and causal indicators in structural equation models. *Mis Quarterly*, 35(2), 359-372.
- Bollen, K. A. (2019). Model Implied Instrumental Variables (MIIVs): An alternative orientation to structural equation modeling. *Multivariate Behavioral Research*, 54(1), 31-46.
- Bollen, K. A., & Diamantopoulos, A. (2015). In Defense of Causal-Formative Indicators: A Minority Report.
- Bollen, K. A., & Pearl, J. (2013). *Eight myths about causality and structural equation models*. In *Handbook of causal analysis for social research* (pp. 301-328). Springer Netherlands.
- Brown, G. T. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological Reports*, 99(1), 166-170.
- Dijkstra, T. K. (2010). *Latent variables and indices: Herman World's basic design and partial least squares*. In *Handbook of partial least squares* (pp. 23-46). Springer Berlin Heidelberg.
- Dijkstra, T. K., & Henseler, J. (2015). Consistent partial least squares path modeling. *MIS quarterly Management information systems quarterly*, 39(2), 297-316.

- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2), 155.
- Evermann, J., & Tate, M. (2010). Testing models or fitting models? Identifying model misspecification in PLS.
- Gates, K. M., Fisher, Z. F., & Bollen, K. A. (2019). Latent variable GIMME using model implied instrumental variables (MIIVs). *Psychological methods*.
- Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Comparing PLS to regression and LISREL: A response to Marcoulides, Chin, and Saunders. *MIS Quarterly*, 36(3), 703-716.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.
- Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press.
- Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., ... & Calantone, R. J. (2014). Common beliefs and reality about PLS comments on Rönkkö and Evermann (2013). *Organizational Research Methods*, 1094428114526928.
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218.
- Jöreskog, K. G., & Wold, H. O. (1982). *Systems under indirect observation: Causality, structure, prediction* (Vol. 139). North Holland.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford publications.
- Lohmöller, J. B. (1989). Predictive vs. structural modeling: Pls vs. ml. In *Latent variable path modeling with partial least squares* (pp. 199-226). Physica, Heidelberg.
- Majid, N. A., Zainol, F. A., Daud, W. N. W., & Afthanorhan, A. (2019). Cooperative entrepreneurship in Malaysian secondary schools: A review of current practices. *The Journal of Social Sciences Research*, 5(3), 812-818.
- Marcoulides, G. A., & Saunders, C. (2006). Editor's comments: PLS: a silver bullet? *MIS quarterly*, iii-ix.
- McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research*, 31(2), 239-270.
- McIntosh, C. N., Edwards, J. R., & Antonakis, J. (2014). Reflections on partial least squares path modeling. *Organizational Research Methods*, 1094428114529165.
- Mertens, W., Pugliese, A., & Recker, J. (2017). Models with Latent Concepts and Multiple Relationships: Structural Equation Modeling. In *Quantitative Data Analysis* (pp. 37-59). Springer, Cham.
- Mohamad, M., Afthanorhan, A., Awang, Z., & Mohammad, M. (2019). Comparison between CB-SEM and PLS-SEM: Testing and confirming the maqasid syariah quality of life measurement model. *Journal of Social Sciences Research*, 5(3), 608-614. doi:10.32861/jssr.53.608.614
- Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling*, 8(2), 287-312.
- Raykov, T., & Marcoulides, G. A. (2012). *A first course in structural equation modeling*. Routledge.
- Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, 26(4), 332-344.
- Rönkkö, M., & Evermann, J. (2013). A critical examination of common beliefs about partial least squares path modeling. *Organizational Research Methods*, 1094428112474693.
- Rönkkö, M., McIntosh, C. N., Antonakis, J., & Edwards, J. R. (2016). Partial least squares path modeling: Time for some serious second thoughts. *Journal of Operations Management*, 47, 9-27.
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies!. *Journal of Business Research*, 69(10), 3998-4010.
- Schumacker, R. E., & Lomax, R. G. (2010). *Structural equation modeling*. NY. Routledge.
- Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, 66(4), 563-575.
- Westland, J. C. (2015). *Structural equation modeling: from paths to networks* (pp. 47-60).

