NAVIGATING THE FUTURE OF QUANTITATIVE RESEARCH : THE POWER OF STUCTURAL EQUATION MODELING

Sipnarong Kanchanawongpaisan *[1]

[*1] Suan Sunandha Rajabhat University

* Email: sipnarong.ka@gmail.com

Received 15 August 2024 Revised 27 August 2024 Accepted 29 August 2024

Abstract

This study delves into the advanced applications of Structural Equation Modeling (SEM) in modern quantitative research. SEM's versatility and power allow researchers to simultaneously examine multiple relationships and account for measurement errors, offering significant advantages over traditional regression models. This research highlights SEM's capacity to provide detailed and nuanced insights into complex constructs, particularly beneficial in social sciences, business administration, and psychology. A rigorous preparatory process is essential for the robustness and reliability of SEM models. This process includes defining the research problem, conducting a comprehensive literature review, developing a theoretical framework, identifying relevant variables, designing the study, and validating measurement instruments. Evaluating the measurement model fit using various indices, such as the Chi-Square Test, RMSEA, CFI, TLI, SRMR, GFI, and AGFI, ensures a comprehensive model accuracy assessment. The findings underscore the significant implications of SEM for advancing quantitative research methodologies. Researchers can enhance their studies' precision and explanatory power by leveraging SEM. This approach paves the way for exploring intricate relationships and contributes to developing sophisticated and reliable research techniques. This study provides an example process, valuable insights, and practical recommendations for researchers aiming to employ advanced statistical methods, ultimately leading to more robust and insightful findings in various research domains.

Keywords: Structural Equation Modeling (SEM), quantitative Research, measurement Model Fit

Introduction

Modern quantitative research requires sophisticated statistical methods to comprehend intricate correlations between variables. Structural Equation Modeling (SEM) is a technique that enables researchers to examine and estimate causal links using statistical data and theoretical assumptions (Little, 2023). The value of Structural Equation Modeling (SEM) rests in its capacity to effectively represent both observable and latent variables, providing a more nuanced examination of complex entities (Fu et al., 2024).

Structural Equation Modeling (SEM) is not just a statistical technique but a versatile tool that finds applications in various disciplines. Its effectiveness in solving complex research inquiries that classic regression models struggle with makes it essential in social sciences, business administration, and psychology. SEM's ability to comprehensively analyze the complex relationships among many components has been demonstrated in numerous studies, such as its use in investigating the influence of service quality on customer satisfaction across different sectors (Hair et al., 2014). This versatility of SEM will intrigue and pique the curiosity of researchers across various fields.

SEM, or Structural Equation Modeling, is a statistical technique that combines factor analysis and multiple regression analysis. It examines the structural correlation between observed variables and underlying constructs. Unlike traditional approaches, SEM can handle several dependent variables and evaluate both direct and indirect effects simultaneously (Byrne, 2016). One of the key benefits of using SEM is its ability to offer a comprehensive perspective by enabling the concurrent analysis of various relationships, thereby providing a holistic and enlightening view of the research data (Kurtaliqi et al., 2024).

Despite the significant advantages of SEM, more research is needed to comprehensively apply it in recent studies. Previous research has primarily focused on SEM's fundamental concepts and uses but has yet to explore its integration with other advanced statistical approaches. This study aims to fill this gap by examining the extensive application of SEM in contemporary research, providing a robust foundation for researchers to model complex interactions more accurately. By adopting these approaches, we can expand the scope of Structural Equation Modeling (SEM), particularly in disciplines that demand intricate depictions of relationships, such as psychometrics and econometrics. This integration promises growth and development in the field of SEM, offering hope for future research.

Preparatory Measures Before Building a Structural Equation Modeling (SEM) Model

Before creating a Structural Equation Modeling (SEM) model, researchers must engage in many essential preliminary measures to guarantee the model's validity and dependability. These stages are crucial for establishing a solid basis for the SEM analysis.

Step 1 Define the Research Problem:

Expose the research question or hypothesis. Gaining a comprehensive understanding of the problem you aim to tackle will direct the choice of variables and the design of the structural equation modeling (SEM) model. A precisely formulated research challenge guarantees that the structural equation modeling (SEM) model is concentrated and pertinent (Hair et al., 2014).

Step 2 Literature Review:

Researchers must examine the current literature extensively to uncover pertinent theories, models, and empirical evidence. This step facilitates the development of the theoretical framework and establishes a foundation for specifying the model (Vaithilingam et al., 2024). Conducting a literature review is essential to ensure that the model is based on existing information and fills in the gaps in the current research (Kline, 2011).

Step 3 Create a Theoretical Framework:

Using the literature review findings, construct a theoretical framework that delineates the anticipated associations among variables (Sharma et al., 2024). This framework functions as the architectural plan for the SEM model. A robust theoretical framework is necessary to direct the model design and guarantee that it is firmly based on known ideas (Byrne, 2016).

Step 4 Variable Identification:

Determine the observable (measured) and latent (unobserved) variables that should be incorporated into the model. Verify if these variables are theoretically justified and directly applicable to the study problem. Accurate and relevant variable identification ensures the model's precision and significance (Schumacker & Lomax, 2015)

Step 5 Develop the study:

Formulate the research design, encompassing the sampling technique, data collection protocols, and measuring tools. Ensure that the sample size is sufficient for SEM analysis, as SEM requires extensive samples to obtain dependable outcomes. A well-designed study improves the accuracy and consistency of the results (Westland, 2010).

Step 6 Design measurement instruments:

Develop or modify instruments (such as surveys or questionnaires) to capture the observed variables precisely. Verify the validity of these instruments by conducting pilot testing and reliability analysis. Accurate data can only be obtained using reliable and valid measurement instruments (Devellis, 2016).

Step 6 Data Collection:

Gather the data using the established process. Maximize data integrity by reducing biases and errors throughout the data collection process. The validity of the SEM analysis relies heavily on high-quality data (Hair et al., 2014).

Step 7 Initial Data Analysis:

Perform initial data analysis to identify any missing data or outliers and assess the normality of the data. Resolve any concerns related to data preparation for SEM analysis. Preliminary data analysis is crucial in ensuring the data is error-free and prepared for the modeling process (Tabachnick & Fidell, 2013).

Step 8 Conduct an Exploratory Factor Analysis (EFA)

Perform EFA to explore the underlying factor structure of the observed variables (Bollen et al., 2024). This step helps understand the data's dimensionality and refine the measurement model. EFA is useful for identifying potential latent constructs (Fabrigar & Wegener, 2011).

Step 9 Improve the Measurement Model:

Based on the Exploratory Factor Analysis (EFA) results, adjustments should be made to enhance the measurement model. Make sure that each hidden variable is accurately represented by its associated observable variables. Enhancing the measuring model enhances its validity and reliability (Brown, 2015).

Step 10 Confirmatory Factor Analysis (CFA):

Perform Confirmatory Factor Analysis (CFA) to assess the validity of the measurement model (Lesia et al., 2023). Confirmatory factor analysis (CFA) enables the evaluation of the extent to which the observed variables accurately reflect the underlying components. Validating the measurement model is an essential and crucial stage (Byrne, 2016).

Step 11 Evaluate Measurement Model Fit:

Assess the adequacy of the measurement model by examining different fit indices such as the Chi-Square Test, RMSEA, CFI, and TLI. Before advancing to the structural model, verifying that the model satisfies the established criteria for acceptable fit is imperative. Ensuring a good model fit is crucial to describing the data appropriately (Kline, 2015).

By rigorously adhering to these procedures, researchers may guarantee that their structural equation modeling (SEM) model is built upon a robust theoretical framework, precise measurement tools, and top-notch data. Properly preparing for the implementation of SEM in quantitative research is essential for success.

Table 1 Model Specification for Structural Equation Modelling

Fit Index	Acceptable Threshold	Description
Chi-Square Test	$p > 0.05$	A lower value indicates a better fit. A non-significant chi-square ($p > 0.05$) indicates a good fit.
Root Mean Square Error of Approximation (RMSEA)	< 0.06	RMSEA values less than 0.06 indicate a good fit between the model and the data.
Comparative Fit Index (CFI)	> 0.95	CFI values above 0.95 indicate a good fit. It compares the fit of a target model to an independent model.
Tucker-Lewis Index (TLI)	> 0.95	TLI values above 0.95 indicate a good fit. They are a comparison measure between a target and a baseline model.
Standardized Root Mean Square Residual (SRMR)	< 0.08	SRMR values less than 0.08 indicate a good fit. It represents the standardized difference between observed and predicted correlations.
Goodness of Fit Index (GFI)	> 0.90	GFI values above 0.90 indicate a good fit. It measures the proportion of variance accounted for by the estimated population covariance.
Adjusted Goodness of Fit Index (AGFI)	> 0.90	AGFI values above 0.90 indicate a good fit. The algorithm adjusts the GFI based on the model's degrees of freedom.

Assessing Measurement Model Fit in SEM

Chi-Square Test:

The Chi-Square Test evaluates the overall fit of the SEM model by assessing the discrepancy between the observed covariance matrix and the model-implied covariance matrix. A lower chi-square value indicates a better fit, suggesting the model-implied covariance matrix is close to the observed covariance matrix. A non-significant chi-square ($p > 0.05$) indicates no significant difference between the observed and model-implied matrices, suggesting a good fit (Zheng & Bentler, 2024). However, this test is sensitive to sample size, often leading to substantial results (indicating poor fit) in large samples even when the model fit is acceptable (Byrne, 2016).

Root Mean Square Error of Approximation (RMSEA):

RMSEA measures how well a model, with unknown but optimally chosen parameter estimates, would fit the population's covariance matrix. RMSEA values less than 0.06 indicate a good fit between the model and the data. It accounts for model complexity and is relatively insensitive to sample size, making it a reliable fit measure. Lower values of RMSEA indicate a better fit (Kline, 2015).

Comparative Fit Index (CFI):

CFI is an incremental fit index that compares the fit of the target model to the fit of an independent (null) model, which assumes that all variables are uncorrelated. CFI values above 0.95 indicate a good fit. It adjusts for sample size and model complexity, providing a comparison between the tested model and a baseline model. Higher CFI values signify that the model fits the data better than the null model, with values closer to 1 indicating a better fit (Hair et al., 2014).

Tucker-Lewis Index (TLI):

TLI, also known as the Non-Normed Fit Index (NNFI), is an incremental fit index that penalizes model complexity. TLI values above 0.95 indicate a good fit. Unlike CFI, TLI can sometimes exceed 1.0 or fall below 0.0. It compares the target model to a baseline model. It adjusts for the number of parameters in the model, making it helpful in evaluating the balance between model fit and complexity (Schumacker & Lomax, 2016).

Standardized Root Mean Square Residual (SRMR):

SRMR is an absolute fit index representing the standardized difference between observed and predicted correlations. SRMR values less than 0.08 indicate a good fit. It provides a straightforward measure of how well the model reproduces the sample data, with lower values indicating a better fit. SRMR is easy to interpret and helpful in assessing the overall fit of the model (Hu & Bentler, 1999).

Goodness of Fit Index (GFI):

GFI is an absolute fit index measuring the variance proportion accounted for by the estimated population covariance matrix. GFI values above 0.90 indicate a good fit. It assesses how the model-implied covariance matrix explains the observed covariance matrix. Higher GFI values suggest a better fit, with values above 0.90 desirable (Tabachnick & Fidell, 2013).

Adjusted Goodness of Fit Index (AGFI):

AGFI is a variant of the GFI that adjusts for the degrees of freedom in the model. AGFI values above 0.90 indicate a good fit. By adjusting for model complexity, AGFI provides a more conservative fit measure than GFI. Higher AGFI values suggest that the model accounts for a substantial proportion of the variance in the data, with values above 0.90 being preferred (Bollen, 1989).

The indices of the Measurement Model Fit in SEM provide several perspectives on the adequacy of the SEM model in representing the data, assisting researchers in verifying the accuracy of their model. By combining these indexes, it is possible to conduct a thorough evaluation of the model's suitability.

Sample SEM Model and Description

Title: Examining the Impact of Service Quality on Customer Satisfaction and Loyalty in the Airline Industry

Objective: To analyze the relationships between service quality, customer satisfaction, and customer loyalty in the airline industry using Structural Equation Modeling (SEM).

Variables:

Latent Variables:

Service Quality (SQ):

Five observed variables were measured (for easy drawing of the model, researchers should conduct abbreviation for variables): Tangibles (SQ1), Reliability (SQ2), Responsiveness (SQ3), Assurance (SQ4), and Empathy (SQ5).

Customer Satisfaction (CS):

Three observed variables were measured: Overall satisfaction (CS1), Satisfaction with services (CS2), and Satisfaction with staff (CS3).

Customer Loyalty (CL):

Three observed variables were measured: Repeat purchase intention (CL1), Willingness to recommend (CL2), and Loyalty program participation (CL3)

Hypotheses:

H1: Service quality positively influences customer satisfaction.

H2: Customer satisfaction positively influences customer loyalty.

H3: Service quality positively influences customer loyalty.

Model Framework

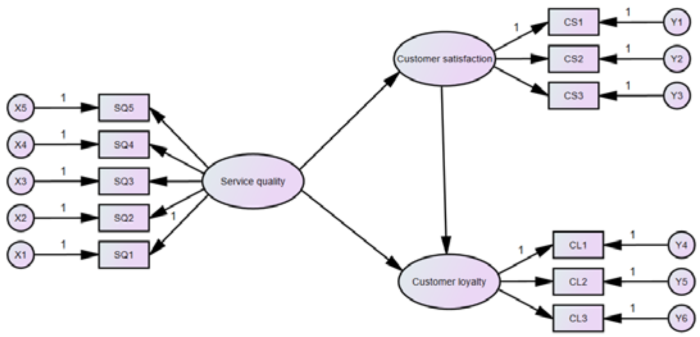

Figure 1: Path diagram Example model Illustrated by AMOS software program

SEM Analysis Steps

Step 1 Model Specification:

Define the relationships between latent variables (SQ, CS, CL) and their corresponding observed variables (SQ1-SQ5, CS1-CS3, CL1-CL3). Specify the hypothesized paths: SQ → CS, CS → CL, and SQ → CL.

Step 2 Model Identification:

Ensure that the model is identifiable, meaning there are sufficient data points to estimate the parameters. This involves setting constraints, such as fixing one of the loadings for each latent variable to 1.

Step 3 Data Collection:

Collect data from airline customers using a structured questionnaire to measure the observed variables. Ensure a sample size that is large enough to provide reliable SEM results.

Step 4 Preliminary Data Analysis:

Check for missing data, outliers, and normality. Conduct preliminary analyses to prepare the data for SEM.

Step 5 Model Estimation:

Estimate the model parameters using software like AMOS, LISREL, or Mplus. Estimation methods such as Maximum Likelihood Estimation (MLE) should be applied.

Step 6 Model Evaluation:

Evaluate the model fit using various indices, including the Chi-Square Test, RMSEA, CFI, TLI, SRMR, GFI, and AGFI. Ensure that the model meets acceptable fit criteria (table 1).

Step 7 Model Modification:

Modify the model based on fit indices and theoretical justification to improve fit if necessary. Then, re-estimate and re-evaluate it.

Step 8 Interpretation:

Interpret the estimated parameters and path coefficients to understand the relationships between service quality, customer satisfaction, and customer loyalty. Assess the direct and indirect effects and their significance.

Example Results:

H1: Service quality positively influences customer satisfaction (path coefficient = 0.75, $p < 0.001$), indicating that higher service quality leads to higher customer satisfaction.

H2: Customer satisfaction positively influences customer loyalty (path coefficient = 0.60, $p < 0.001$), suggesting that satisfied customers are more likely to remain loyal.

H3: Service quality positively influences customer loyalty (path coefficient = 0.40, $p < 0.01$), showing that high service quality directly enhances customer loyalty, even when controlling for customer satisfaction.

Conclusion

The SEM analysis of the relationships between service quality, customer satisfaction, and customer loyalty in the airline industry provides significant insights into how these constructs interact. The results confirm the hypothesized paths and underscore the importance of service quality in fostering customer satisfaction and loyalty.

Service Quality Positively Influences Customer Satisfaction (H1):

The analysis reveals a strong positive relationship between service quality and customer satisfaction (path coefficient = 0.75, $p < 0.001$). This finding indicates that higher levels of service quality significantly enhance customer satisfaction. Airlines that focus on improving tangibles, reliability, responsiveness, assurance, and empathy can expect a marked increase in customer satisfaction.

Customer Satisfaction Positively Influences Customer Loyalty (H2):

Customer satisfaction is shown to have a substantial positive effect on customer loyalty (path coefficient = 0.60, $p < 0.001$). Satisfied customers are more likely to engage in repeat purchases, recommend the airline to others, and participate in loyalty programs. This highlights the critical role of customer satisfaction in retaining customers and building long-term loyalty.

Service Quality Directly Influences Customer Loyalty (H3):

The direct relationship between service quality and customer loyalty is also significant (path coefficient = 0.40, $p < 0.01$). This finding suggests that improvements in service quality can directly lead to higher customer loyalty, even when accounting for the mediating effect of customer satisfaction. Airlines that invest in service quality enhancements can benefit from increased satisfaction and direct loyalty gains.

Conclusion

This study explores the potent and adaptable uses of Structural Equation Modeling (SEM) in contemporary quantitative research, emphasizing its crucial function in comprehending intricate connections among variables. Structural equation modeling (SEM) transcends typical regression models in its capacity to answer complex research issues by simultaneously analyzing many correlations and compensating for measurement errors. SEM is highly beneficial in social sciences, business administration, and psychology. Incorporating SEM with sophisticated analytical tools, such as vector analysis, signifies a notable progression in research methodology, enabling a more intricate and subtle examination of complicated constructions. An intensive preparatory process is essential to guarantee the strength and dependability of SEM models. The approach entails many key steps: defining the research problem, doing an extensive literature review, constructing a theoretical framework, identifying variables, designing the study, and validating measuring instruments. Assessing the adequacy of the measurement model fit through the utilization of diverse indices, including the Chi-Square Test, RMSEA, CFI, TLI, SRMR, GFI, and AGFI, offers distinct viewpoints on the model's appropriateness, guaranteeing that the model appropriately reflects the data.

The results of this study have significant consequences for academics who want to use sophisticated statistical methods in their quantitative studies. By using the advantages of SEM and incorporating vector analysis, researchers can improve their findings' accuracy and explanatory capability, facilitating the investigation of intricate associations and contributing to the advancement of sophisticated and dependable research procedures. Researcher are advised to utilize Structural Equation Modeling (SEM) in their investigations to represent complex interactions accurately. Subsequent investigations should continue to examine and enhance these approaches, tackle rising obstacles, and broaden their suitability across diverse fields. This work enhances the current knowledge by thoroughly analyzing SEM principles and showcasing their incorporation with vector analysis. This offers valuable insights for improving quantitative research procedures and achieving more reliable and insightful results.

References

Bollen, K. A. (1989). *Structural Equations with Latent Variables.* John Wiley & Sons.

Bollen, K. A., Gates, K. M., & Luo , L. (2024). A Model Implied Instrumental Variable Approach to Exploratory Factor Analysis (MIIV-EFA). *Psychometrika, 89*(1), 687–716.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research, 2nd ed.* Guilford Publications.

Byrne, B. M. (2016). *Structural Equation Modeling With AMOS 3rd Edition: Basic Concepts, Applications, and Programming, Third Edition.* New York: Routledge.

Devellis, R. F. (2016). *Scale Development: Theory and Applications.* Sage Publications.

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory Factor Analysis.* Oxford University Press.

Fu, C., Wang, J., Qu, Z., Skitmore, M., Yi, J., Sun, Z., & Chen, J. (2024). Structural Equation Modeling in Technology Adoption and Use in the Construction Industry: A Scientometric Analysis and Qualitative Review.

Sustainable Construction Engineering Processes, 16(9), 1–15. doi:https://doi.org/10.3390/su16093824

Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review, 26*(2), 106–121.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Kline, R. B. (2011). *Principles and practice of structural equation modeling (3rd ed.).* New York: Guilford Press.

Kurtaliqi, F., Miltgen, C. L., Viglia, G., & Pantin-Sohier, G. (2024). Using advanced mixed methods approaches: Combining PLS-SEM and qualitative studies. *Journal of Business Research, 172*(2024), 1–14. doi:https://doi.org/10.1016/j.jbusres.2023.114464

Lesia, M. P., Aigbavboa, C. O., & Thwala, W. D. (2023). Factors influencing residential location choice in South Africa: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). *Journal of Housing and the Built Environment, 39*(1), 133–160.

Little, T. D. (2023). *Longitudinal Structural Equation Modeling.* New York: The Guilford Press.

Schumacker, R. E., & Lomax, R. G. (2015). *A Beginner's Guide to Structural Equation Modeling 4th. Edition.* New York: Routledge. doi:https://doi.org/10.4324/9781315749105

Sharma, L., Bulsara, H. P., Bagdi, H., & Trivedi, M. (2024). Exploring sustainable entrepreneurial intentions through the lens of the theory of planned behavior: a PLS-SEM approach. *Journal of Advances in Management Research, 21*(1), 20–43. doi:https://doi.org/10.1108/JAMR-01-2023-0006

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics.* Pearson Education.

Vaithilingam, S., Ong, C. S., Moisescu, O. I., & Nair, M. S. (2024). Robustness checks in PLS-SEM: A review of recent practices and recommendations for future applications in business research. *Journal of Business Research, 172*(2024), 1–15. doi:https://doi.org/10.1016/j.jbusres.2023.114465

Westland, C. J. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications, 9*(6), 476–487. doi:https://doi.org/10.1016/j.elerap.2010.07.003

Zheng, B. Q., & Bentler, P. M. (2024). Enhancing Model Fit Evaluation in SEM: Practical Tips for Optimizing Chi-Square Tests. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–6. doi:https://doi.org/10.1080/10705511.2024.2354802