American Journal of Applied Mathematics
2024, Vol. 12, No. 5, pp. 133-140
https://doi.org/10.11648/j.ajam.20241205.13

Review Article

Theoretical Approaches Review on Covariance Based Sem Using Lisrel, Partial Least Based Sem Using Smart PLS and Component Based Sem Using Gesca

Umi Narimawati[1], Jonathan Sarwono[2,*]

[1]Economic Department, Computer University of Indonesia, Bandung, Indonesia
[2]Science and Technology Department, International Women University, Bandung, Indonesia

Abstract

The aim of the research is to review theories underlying the Structural Equation Modeling (SEM) procedure based on covariance (CBSEM), partial least square (PLSSEM) and component (GESCA SEM). The methods used are meta-analysis and systematic secondary data search. Results of the study are: First, theories underlying the CBSEM, PLSSEM and GESCA SEM procedures produce different characteristics in each SEM model. CBSEM models consist of two sub models, namely 1) Factor Analysis Model consisting of a) Exploratory Factor Analysis (EFA) which is designed for a situation where the relationship between indicators and latent variables is unknown or unclear; b) Confirmatory Factor Analysis (CFA) which is used for research where the researcher already has knowledge about the structure of the underlying latent variable (construct) and c) Full Latent Variable Model (LV). 2) PLSSEM consists of two sub model, namely reflective and formative models. GESCA SEM consists of structural / inner model and measurement / outer model. Second, the primary characteristics of CBSEM, PLSSEM and GESCA SEM are requirements of the amount of data sample; the sample data origin; and the software used to calculate the data due to the different statistical formulation, namely LISREL, SmartPLS and GSCA Pro Windows. Third, the main differences among the CBSEM, PLS SEM and GESCA SEM are in the uses of the unstandardized regression coefficients (b) versus the standardized regression coefficients (β). Thus, the researchers that are going to use those procedures must consider those three important findings.

Keywords

Structural Equation Modeling (SEM), CBSEM, PLSSEM, GESCA SEM

1. Introduction

There are several definitions that explain what structural equation modeling (SEM) is. First, SEM uses various types of models to describe the relationship between observed variables, with the same basic goal of providing a quantitative test of the theoretical model hypothesized by the researcher. More specifically, various theoretical models can be tested in SEM that hypothesize how a set of variables defines a construct (latent variables) and how these constructs relate to each other [22]. Second, Schumacker, R.E & Whitettaker, TA [23] mentions that the specific characteristic of SEM lies in the presence of several specific things including: observed and unobserved variables; standardized and normalized residuals;

*Corresponding author: jonathan@iwu.ac.id (Jonathan Sarwono), jsarwono007@gmail.com (Jonathan Sarwono)

Received: 7 August 2024; **Accepted:** 2 September 2024; **Published:** 20 September 2024

direct, indirect and total influence; the existence of 4 conditions concerning the causal relationship: 1) temporal sequence in the relationship of variables, 2) covariation and correlation between variables, 3) other causes that are controlled statistically and using research design, 4) independent variable (X) which is manipulated so as to cause a change in the value of the dependent variable (Y). In other words, a change in the first variable where there is an arrow in the direction of the next variable will change the second variable where the arrow heads towards it [15]. Third, Structural Equation Modeling (SEM) is a group of statistical procedures and methods for modeling relationships between variables. The model can include unobserved variables (latent variables) and observed variables (indicators). For this reason, SEM has been referred to as latent variable modeling. The main data of the SEM procedure is covariance so that SEM is also referred to as covariance structure modeling. The purpose of using SEM is to estimate causal relationships between variables, which is why SEM is also referred to as a causal model [12].

Furthermore, SEM is also defined as a statistical modeling technique that is linear and general. The techniques included in this SEM are factor analysis, path analysis and linear regression. In this regard, SEM can be referred to as a general multivariate analysis technique. SEM is also a statistical technique used to build and test statistical models which are generally in the form of causal models. That is why SEM is called a hybrid technique which includes confirmatory aspects of factor analysis, path analysis and regression that can be considered as a special case in SEM. In relation to that, SEM also has a function similar to multiple regression, even though SEM is a more accurate analytical technique because it considers interaction modeling, nonlinearity, correlated independent variables, measurement errors, interference with correlated errors, error terms, several latent independent variables, each of which is measured using many indicators, and one or two latent dependent variables, each of which is measured by several indicators. Accordingly, SEM can be used another alternative that is stronger than using multiple regression, path analysis, factor analysis, time series analysis, and covariance analysis.

SEM is also called as a statistical methodology that uses a confirmatory approach in hypothesis testing (confirmatory approach) for multivariate analysis of a structural theory based on certain symptoms. Typically, this theory presents a causal process that results in observations of several variables. According to Byrne, the term SEM contains two important procedural aspects, namely: 1) a causal process in a study is represented by a group of structural equations, namely regression equations. 2) This structural relationship can be described in a model to clarify the conceptualization of the theory being studied (Bentler, 1988) as cited by [1]. The hypothesized model can be tested statistically in a simultaneous analysis of an overall system of variables to determine the extent to which the model fits the data. If the model fit is met, then the model must be able to show the feasibility of the relationship between variables as suggested. Furthermore, Byrne [1] made his own definition of SEM as follows: "SEM is a popular analytical method that allows to examine various models that can explain data structures. SEM is a statistical procedure used to explain the relationship among several variables. In explaining the relationship, SEM examines the structure of the relationship expressed by several equations that are similar to the equations in multiple linear regression. The equation describes the relationship of constructs (latent variables) used in the analysis [14].

From those definitions, it can be concluded that SEM has the characteristics of an analytical technique that functions as confirmation rather than explanatory. That is, a researcher is more likely to use SEM to determine whether a particular model is valid or not than to use it to find a particular model suitable or not, although SEM analysis often includes elements used to explain. Furthermore, to study SEM correctly, it is necessary to first understand linear regression and path analysis where the model of the relationship between the most basic cause and effect variables occurs in both procedures

2. Literature Review

The structural equation modeling (SEM) analysis procedure has the following chronological order: it starts with the discovery of linear regression, then path analysis, factor analysis, then structural equation models [22]. Nevertheless, the discovery of SEM cannot be separated from the services of Pearson who found a correlation procedure between two variables that resulted in the value of the correlation coefficient which, later on, is known as the Pearson Product Moment Correlation. In relation to that, development of the Pearson correlation coefficient value (r_{xy}) is, then, used as the basis for the linear regression procedure. In the linear regression model, the value of the correlation coefficient is used as the basis for calculating the coefficient of determination (R^2) in which the value of the regression coefficient is the square of the value of the Pearson correlation coefficient. In other words, the coefficient of determination can be calculated by squaring the value of the Pearson correlation coefficient in the linear regression model. Even though another theory develops which states that this R^2 comes from the results of calculations with the formula as follows.

$$R^2 = \frac{SSreg}{SStotal} = 1 - \frac{SSres}{SStotal}$$

In relation to the modeling problems, linear regression is the first procedure used as a model that uses two main values, namely the Pearson correlation coefficient and the least squares criterion to calculate the regression weights (regression coefficients, especially the unstandardized regression coefficients). The main function of linear regression used to predict the value of the dependent variable (Y) based on the value of the independent variable (X) which is based on a linear relationship between the variables X and Y will minimize the number of squares of residual errors using the mathematical equation Y = a + b x. In its development, SEM uses the value of the unstandardized regression coefficient (b) derived from this equation. Accordingly, SEM modeling is much inspired by the regression equation model.

Furthermore, another procedure that inspires the development of SEM is factor analysis invented by Charles Spearman. Spearman in his experiment used the correlation coefficient to determine the relationship between variables which was then made as a factor model. In this experiment Spearman uses a set of correlated items where the responses to that set of items can be summed into a certain value that is used to measure, define or make inferences about a construct which later this construct in SEM is called a latent variable (unobserved variable). Some of these correlated items are known as indicators that make up certain constructs. In his experiment Spearman finds two construct factors that make up intelligence. It is the first time name it as factor analysis. Spearman's findings are reinforced by the findings of Thrustone who developed a factor model application and proposed an instrument to generate observation values that can be used to conclude a particular construct. Factor analysis can be concluded as a second model that precedes modeling in SEM. Furthermore, the term Confirmatory Factor Analysis (CFA) is first used by Howe, Anderson and Rubin, and Lawley. CFA was seriously developed by Karl Joreskog in 1960 to test a set of items can be defined as a construct.

The next important procedure that contributes to SEM is the path model developed by Sewell Wright which later became known as path analysis which is used to model the relationship between observed variables (indicators in the sense of SEM) sequentially with the aim of parsing the model of correlation relationship into direct and indirect relationship models. Where in path analysis the correlation coefficient is used to measure the relationship between independent variables which in the context of path analysis is called an exogenous variable and the standardized regression coefficient (β) is used as the regression weight used to measure the relationship between the independent variable (exogenous) and the variable. dependent (endogenous) in a certain path diagram which is then referred to as the path coefficient from exogenous to endogenous variables. In path analysis, the model of the relationship between exogenous and endogenous variables uses variables that can be observed directly or which are then referred to as manifest variables or indicators in the context of SEM.

Thus SEM is basically a combination of regression models, factor analysis and path analysis as discussed above by adding variables that cannot be observed directly or constructs (latent) or factors in the context of factor analysis, or latent variables in the context of SEM. To measure the relationship between latent variables and the relationship between latent variables and their indicators, SEM uses regression weights or unstandardized regression coefficients (b). SEM takes the path diagram from the path analysis procedure, uses latent variables from the factor analysis model, uses the regression coefficient value from the regression analysis procedure. From the three previous models, SEM finally becomes a model of the relationship between latent variables known as the structural model and the relationship between latent variables and their indicators known as the measurement model. In the next development the SEM procedure is called as the Covariance Based SEM which is then followed by the emerging new SEM procedures, namely Partial Least Square SEM (PLS SEM) and Generalized structured component analysis SEM (Gesca SEM).

From those backgrounds mentioned above, the research questions arising in this study are: 1) What theories underlying the SEM procedure? 2) What are striking features of CBSEM, PLSSEM and GESCA SEM 3) What makes different among the CBSEM, PLS SEM and GESCA SEM?

3. Methodology

The method used in this study is meta- analysis. The meta-analysis is used to review the similar literature [17]. The main requirement for using meta-analysis is the study of the results of the same research (Glass, 1981, as quoted by Narimawati, Umi & Sarwono, Jonathan, 2020). Besides, the present writer also uses a systematic secondary data search through Google Scholar relating with similar topics that are discussed by the present writer.

4. Results and Discussion

4.1. Research Result

4.1.1. Covariance Based SEM Using LISREL

The followings are the main striking features of Covariance Based SEM using LISREL:

Uses of Theory: Theories relating to the subject under study in SEM play an important role because theory will serve as the basic foundation in building models in research regarding the specification of measurement and structural models. There are 3 roles of theory in SEM, namely: 1) creating a relationship specification that serves to define the model; 2) to establish a causal relationship; and 3) to develop a specific modeling strategy.

Because the SEM procedure is a confirmatory analytical procedure; the role of theory is very important. Theory is used as a means of testing and confirming the model in research. Theories also mainly play a role in specifying the relationship among the variables under study where the relationship among these variables will reflect which variables are independent variables (exogenous) and which variables are dependent (endogenous) variables and which variables act as latent variables and manifest variables or indicators.

The theory of the field of science in SEM is also used as a determinant of causal relationships on the variables studied. Therefore the causal relationship in SEM arises because of the confirmation of the theory or in other words the causal findings in the model are inherent or derived in theory not from the results of statistical estimates. The causal relationship in SEM is the same as the causal relationship in Path Analysis, which is independent of statistics. This means that if we conclude that certain exogenous variables affect certain endogenous variables, it is not because the results of statistical calculations that are based on data alone, but because the theory has said that the two variables have a causal relationship.

Basic Terms: In general, the understanding of the relationship between variables in SEM is the same as that in statistics, namely the existence of an independent variable (X) and a dependent variable (Y). This term applies to directly observable variables (indicators). Furthermore, in SEM the variables are generally divided into two, namely latent variables and variables that can be observed directly. Latent variables are theoretical constructs that cannot be observed directly. This theoretical construct is an abstract phenomenon that cannot be measured directly; that is why this construct is called latent variable or factor. Since latent variables cannot be measured directly, they must be defined operationally in terms of the behavior they represent. Under these conditions, the latent variable must be associated with at least one directly observable variable/indicator. Thus the measurement of the latent variable is possible because of these indicators. The logic is that if an assessment is made of the behavior; then the assessment will build a direct measurement of the variables observed directly and indirectly on these latent variables which incidentally is the underlying construct. Measurement of directly observed variables can be done using measurement instruments such as tests, questions in questionnaires, interviews or other instruments that are suitable for data collection. The result of this measurement is called the measurement value. That is why the values that represent these variables are called observables which have another name as manifest variables. In the context of SEM this variable serves as an indicator of the underlying construct. That is why directly observable variables are also called indicators or manifest variables; while the latent variable is also known as a construct or factor. The indicators in the SEM are reflective, meaning that the indicators on a particular latent variable are a reflection of the latent variable. In other words, latent variables underlie these indicators.

Functions: The SEM procedure functions as: First, it allows for more flexible assumptions. Second, the use of confirmatory factor analysis is to reduce measurement error by having many indicators in one latent variable. Third, the attractiveness of the graphical modeling interface to make it easier for users to read the output of the analysis results. Fourth, the possibility of testing the model as a whole rather than the coefficients individually. Fifth, the ability to test models using several dependent variables. Sixth, the ability to model the intermediate variables. Seventh, the ability to model the error term. Eighth, the ability to test the outer coefficients between several groups of subjects. Ninth is the ability to deal with difficult data, such as time series data with autocorrelation errors, abnormal data, and incomplete data.

Main Applications: The main applications of structural equation modeling include: 1) Causal modeling, also known as path analysis, which hypothesizes causal relationships between variables and tests causal models by using a system of linear equations. Cause and effect models may include manifest variables or also known as indicators and latent variables (construct) or both; 2) Confirmatory factor analysis, a continuation technique of factor analysis in which hypotheses are tested for factor loadings and their intercorrelation; 3) Second order factor analysis, a variation of the factor analysis technique in which the correlation matrix of the common factors is analyzed on its own factors to create second order factors; 4) Regression models, an advanced technique of linear regression analysis in which the regression weights are constrained to be equal to each other, or are specified for their numerical values; 5) Covariance structure models, in which the model hypothesizes that the covariance matrix has a certain shape; 6) Correlation structure models, in which the model hypothesizes that the correlation matrix has a certain shape.

Model Fit Index: The model fit index in SEM is used to determine whether the model made is based on observational data in accordance with the theoretical model or not. There are two categories of model fit index used in SEM, namely absolute and complementary model fit index. The absolute fit index consists of: 1) Chi Square, 2) Goodness of Fit Index (GFI), 3) Root Mean Square Error of Approximation (RMSEA), 4) Root Mean Square Residual (RMR) and Standardized Root Mean Residual (SRMR), and Normed Chi Square. The complementary model fit index consists of 1) critical ratio (CR), 2) Standard Coefficients, 3) measurement error, 4) regression weight, 5) model specification, 6) Maximum Likehood Estimation (MLE), 7) The significance (probability) value, 8) construct reliability, 9) Extract Variant, 10) Adjusted Goodness of Fit Index (AGFI), 11) The Minimum Sample Discrepancy Function (CMNF), 12) Tucker Lewis Index (Tucker Lewis Index (TLI), 13) Comparative Fit Index (CFI), 14) Parsimony Fit Index, 15) Reliability Test, 16) Parsimony Based Indexes of Fit (PGFI), 17) Normed Fit Index (NFI), 18) Relative Fit Index (RFI), 19) First Fit Index (PRATIO), 20) Noncentrality Parameter (NCP), 21) The Expected Cross Validation Index (ECVI), 22) Hoelter's Critical N (CN), 23) Residual (Jöreskog, K.G. & Sörbom, D., 2022) [14] and (Narimawati, Umi & Sarwono, Jonathan, 2023) [18].

Partial Least Square SEM Using Smart PLS

Definition: Partial Least Square SEM contains two components of understanding that can be used as a basis for SEM using PLS. The first understanding is about PLS and the second is about SEM. The term PLS (partial least square) is found in procedures that are included in the family of regression procedures [2]. If we study these procedures, it will be found a sub-procedure called Partial Least Squares (PLS) regression. The Partial Least Squares (PLS) regression procedure is used to estimate partial least squares regression models or known as projections to the latent structure. PLS is a predictive technique which is an alternative to ordinary least squares regression (OLS), canonical correlation, or structural equation modeling (SEM), besides that PLS is very useful when several independent variables / predictors are highly correlated with each other, or when the number of predictors exceeds the number of cases. PLS combines features from principal component analysis and multiple regression. The procedure for using PLS is carried out in two stages, namely first, by removing a series of latent factors that explain as much as possible the covariance between the independent and dependent variables; second, predict the value of the dependent variable using independent variable decomposition [4].

One definition of SEM that is in accordance with the definition above states that Structural equation modeling (SEM) is a statistical technique used to build and test statistical models which are usually in the form of causal models. SEM is actually a hybrid technique that includes confirmatory aspects of factor analysis, path analysis and regression which can be considered a special case of SEM. Slightly different from previous definitions, it says that structural equation modeling (SEM) develops and has a function similar to multiple regression, however, it seems that SEM is a stronger analysis technique because it considers interaction modeling, nonlinearity, correlated independent variables, measurement error, correlated error terms, several latent independent variables, each of which is measured using many indicators, and one or two latent dependent variables, and is also measured using several indicators. Accordingly, according to this definition, SEM can be used as an alternative that is more powerful than using multiple regression, path analysis, factor analysis, time series analysis, and covariance analysis [21]. SmartPLS. http://smartpls.de SEM PLS is an alternative procedure where the data is not normally distributed. Therefore SEM PLS is also known as a soft modeling technique where the requirements are not as strict as those in Covariance Based SEM, for example in terms of measurement scale, sample size and residual distribution.

Main Features: PLS SEM should be seen as a more general form of SEM that supports composite as well as general factor models. Composite versus general factors, are factors that are not the same. This means that the PLS-SEM model is the same as the traditional SEM model. The coefficient of the first variable does not have to be closely related to the last.

PLS-SEM is based on a component approach that uses a type of principal components analysis to construct latent variables. In PLS SEM covariation can also be explained by the relationship between indicators. On a path, a covariance arrow can connect each indicator to each other indicator. PLS-SEM approach gives more weight to indicators with higher predictive validity. PLS-SEM assumes that the indicators vary in degrees that are respectively related to those measured in the latent variable.

The path model is reflective if in the path diagram the causal arrow leads from the latent variable (factor) to the indicator variable being measured. On the other hand, the path model is formative if the direction of the arrow from the indicator variable is towards the latent variable. The reflective model is also called the “Model A” model and the formative model is called “Model B”.

In reflective models, indicators are a representative set of items that reflect the latent variable they measure. The reflective model assumes that the factors (latent variables) are "reality" and the measured variables are samples of all indicators of that reality. This implies that omitting one indicator may not matter much if the other indicators are also representative. The latent variable will still have the same meaning even if one indicator is removed.

In the formative model, each indicator represents a dimension of the meaning of the latent variable. A set of indicators collectively represents all dimensions of a latent variable. The formative model assumes that indicators are reality and all of these indicators are dimensions of factors (latent variables). Eliminating indicators in a formative model is the same as eliminating dimensions of meaning. This can cause the meaning of the latent variable to change. To the extent that the omitted dimensions are important, then the meaning of the latent variable will change. Combining multiple indicators is another problem that arises in formative models. Because the indicator items in the formative model represent different dimensions, it is possible that one item may be negatively correlated with another.

Some assumptions in PLS SEM include:

1. The main assumption in using PLS SEM is that it does not require following the normality assumption because PLS SEM does not treat data as in covariance-based SEM where in SEM the data is required to have a normal distribution. This allowance allows us to use data that is not normally distributed.
2. The next assumption is that PLS SEM can use a small sample size unlike covariance-based SEM which requires researchers to use a large sample size because SEM is a procedure that is categorized into a multivariate procedure where almost all multivariate procedures require a large amount of data, for example at least -at least 400. On the other hand, PLS SEM does not require researchers to use large amounts of data. Accordingly, this procedure provides benefits for users when they have difficulty searching for large amounts of data [3].
3. Does not require randomization of samples, so samples selected using non-probability approaches, such as ac-

cidental sampling, purposive sampling and the like can be used in PLS SEM.
4. Allows formative indicators to measure latent variables other than reflective indicators. This is not permitted in covariance-based SEM which uses reflective indicators only.
5. PLS SEM allows for dichotomous latent variables.
6. PLS SEM makes allowances for the necessity of having an interval measurement scale. Thus researchers can use measurement scales other than intervals.
7. The residual distribution in PLS SEM is not required as in covariance-based SEM where in SEM the residual distribution must be as small as possible as in linear regression.
8. PLS SEM is suitable as a procedure used to develop theory at an early stage. This is different from covariance-based SEM which uses theory to be confirmed using sample data.
9. The regression approach in PLS SEM is more suitable than in covariance-based SEM.
10. In PLS SEM only recursive (cause - effect) models are allowed and do not allow non-recursive (reciprocal) models as in covariance-based SEM.
11. PLS SEM allows very complex models with many latent and indicator variables.

4.1.2. Model Fit Measurement

The model fit includes the following matters: 1) Cronbach's Alpha reflects the reliability of all indicators in the model, 2) Convergent validity using the extracted variance (AVE), 3) discriminant validity uses the criteria presented by Fornell - Larcker and 'cross-loadings', 4) Heterotrait-monotrait (HTMT). SRMR: Is an abbreviation for Standardized Root Mean Residual, 5) Chi Square, 6) Normed Fit Index, 7) d_ULS: Squared Euclidean distance value, 8) d_G1: Geodesic1 distance value, 9) d_G2: Geodesic2 distance value 10) Rho_A. [10]

4.1.3. Component Based SEM Using GESCA

Definition: Generalized structured component analysis (GSCA) is a component-based approach to structural equation modeling (SEM), in which the latent variables are approximated by weighted composites of indicators (Hwang, H., et.al., 2017). This method can be alternative to partial least squares called as generalized structured component analysis. Moreover, GSCA based SEM consists of three sub-models: 1) the weighted relation, 2) component measurement, and 3) structural models [13]. The weighted relation model explicitly defines a component as a weighted sum of indicators, where the weights assigned to a set of indicators to shape a component of parameters. The component measurement model is used to specify the regression relationships of a set of indicators on its component, in which the regression coefficients which is called component loadings constitute parameters. The structural model is to specify the path analytic relationships between components, consisting of the component path coefficients as parameters [7].

Furthermore, generalized structured component analysis is suggested as an alternative to partial least squares structural equation modeling (PLS SEM). SEM based on Generalized structured component analysis (GSCA) is used to compute scores on small samples. It proposes a universal least squares optimization criterion. That is why, SEM based on GSCA provides measure of model fit and can manage more diverse path analyses. GSCA SEM consists of three models: 1) weighted relation model, 2) measurement model, and 3) structural model [8].

4.1.4. Main Characteristics

First, Model Specification: The model specification consists of 3 (three) sub models and the sub-models are joint together into one.

Second, Parameter Estimation: The parameter estimation employs a single least square optimization function and alternates the least square algorithm.

Third, the indicator is formative.

4.1.5. Model Fit Index

The overall model fit in measures in SEM based on GSCA are namely FIT and AFIT, GFI (Goodness of Fit Index), SRMR (Standardized Root Mean Residual), Cronbach Alpha, Ave, HTMT, OPE, PVE, VIF and R square [7].

4.2. Discussion

From the research result above, there are some important matters to be discussed as follows:

First, covariance based structural equation modeling (CBSEM) is development of the previous statistic analysis procedures, namely correlation, regression, path analysis and factor analysis. That is why the path coefficient of CB SEM uses an unstandardized regression coefficient (b) whose score range is > 0 up to infinite depending on score unit. Furthermore, the CBSEM makes use of the R square which is originated from the squared Pearson correlation coefficient as well as the path diagram taken from the path analysis procedure. This is in line with the opinion stating that the SEM method is development of path analysis and multiple linear regression, which are included in multivariate statistical analysis models (Chang 1981 in Hidayat, Rachmat & Wulandari, Patricia, 2022). Moreover, SEM procedure gives many benefits for researchers in relation to building research models that employs many variables, assessing constructs (latent variables) which cannot be observed and measured directly [11].

Second, covariance Based SEM (CBSEM) requires some classical assumptions, such as normal distribution meaning that the sample data must originate from the population that has normal distribution; linearity meaning that the relationship between the independent (exogenous) variable and the dependent (endogenous) variable shapes the straight line from the below left side of the XY curve into the above right side;

collinearity meaning that there is no very high or low correlation between the independent variables when there are more the independent variables in the model under study; and heteroscedasticity meaning that there is an equal variance.

Third, the CBSEM requires amount of large data around 200 using the 0.05 error tolerance and around 400 using the 0.01 error tolerance. This means that only large data that can make a better chance to get the significant result in the relationship between the independent and dependent variable.

Fourth, the CBSEM uses a universal compulsory goodness of fit index and complementary model fit index. This implies that rigid analysis procedure needs more requirement meaning that the more the requirements the better the result is. Even though good results in calculation do not merely depends on this goodness of fit index.

Fifth, CBSEM employs a reflective indicators only. This means that the indicators reflect its respective construct. In other words, the construct or latent variable underlies its indicators.

Sixth, Partial Least Square SEM (PLS SEM) employee less rigid requirements, such as enabling to use small data, the data do not have normal distribution. This finding is strengthened by the ideas stating by Memon, Muntaz Ali, et.al. [16] saying that "small sample size and non-normal data can be used in PLS SEM". Moreover similar opinion also stating that PLS SEM is an alternative procedure of CB SEM when the model is very complex, normal data assumption is not fulfilled, and the sample size is small [5].

Seventh, PLS SEM employs both reflective and formative indicators. This finding is convenient with the explanation stating that one of the advantages using the PLS SEM is that this procedure can employ the reflective and formative indicators [6]. In reflective models, indicators are a representative set of items that reflect the latent variable they measure. The reflective model assumes that the factors (latent variables) are "reality" and the measured variables are samples of all indicators of that reality. This implies that omitting one indicator may not matter much if the other indicators are also representative. The latent variable will still have the same meaning even if one indicator is removed. In the formative model, each indicator represents a dimension of the meaning of the latent variable. A set of indicators collectively represents all dimensions of a latent variable. The formative model assumes that indicators are reality and all of these indicators are dimensions of factors (latent variables). Eliminating indicators in a formative model is the same as eliminating dimensions of meaning. This can cause the meaning of the latent variable to change. To the extent that the omitted dimensions are important, then the meaning of the latent variable will change. Combining multiple indicators is another problem that arises in formative models. Because the indicator items in the formative model represent different dimensions, it is possible that one item may be negatively correlated with another [19].

Eighth, the path coefficient of PLS SEM uses the standardized regression coefficient (β). Standardized regression coefficient's values ranges from >0 up to 1. This value can be changed into the unstandardized regression coefficient when it is needed.

Ninth, GESCA based SEM tries to compensate some weaknesses of the PLS SEM relating to its algorithm. This is in line with the opinion stating that GESCA SEM is a method that is developed in order to fulfil the weakness existing in PLS SEM, namely the overall goodness of fit [20].

Tenth, the most striking feature of GESCA SEM is that it employs three sub-models, namely 1) the weighted relation, 2) component measurement, and 3) structural models. Moreover, partial least squares structural equation modeling (PLSSEM) and generalized structural component analysis SEM (GESCA) build composite-based structural equation modeling (SEM) methods, that have attracted considerable interest among methodological and applied researchers [9].

5. Conclusion

The conclusions of this study are as follows:

Theories underlying the CBSEM, PLSSEM and GESCA SEM procedures produce different characteristics in each SEM model which finally distinguish the ways the researchers use each SEM procedure.

Striking features of CBSEM, PLSSEM and GESCA SEM are requirements of the amount of data sample; the sample data origin; and the software used to calculate the data due to the different statistical formulation.

Primary Differences among the CBSEM, PLS SEM and GESCA SEM lies in the uses of the unstandardized regression coefficients (b) versus the standardized regression coefficients (β).

Author Contributions

Umi Narimawati: Conceptualization, Funding acquisition, Supervision

Jonathan Sarwono: Data curation, Methodology, Writing – original draft, Writing – review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Byrne, B. M. 2001. Structural Equation Modeling With LISREL, Basic Concepts, Applications, and Programming. New Jersey: Lawrence Erlbaum Associates Publishers.

[2] Garson, David. (2016) Partial Least Square: Regression and Structural Equation Model. School of Public and International Affairs, North Carolina State University. www.statisticalassociates.com

[3] Garson, D.G. 2006. Structural Equation Model. World Wide Web: http://hcl.chass.ncsu.edu/ssl/ssl.htm

[4] Hair, J. F. Ringle, C. M & Sarstedt, M. (2011) PLS-SEM: indeed a silver bullet. *Journal of Marketing Theory and Practice*, vol. 19, no. 2 (spring 2011), pp. 139–151. © 2011 M.E. Sharpe, In.

[5] Hair, Jo & Alamer, Abdullah, (2022) Partial Least Squares Structural Equation Modelling (PLS-SEM) in Second Language and Education Research: Guidelines Using an Applied Example. PLS SEM in Second Language Research.

[6] Hanafiah, M. Hafiz, (2020) Formative vs. Reflective Measurement Model: Guidelines for structural equation modeling research. International Journal of Analysis and Applications Volume 18, Number 5.

[7] Hwang, Heungsun (2023) GSCA Pro—Free Stand-Alone Software for Structural Equation Modeling. Structural Equation Modeling: A Multidisciplinary Journal. Routledge. Taylor & Francis Group.

[8] Hwang, Heungsun,. (2013) Generalized Structured Component Analysis: A Component-based Approach to Structural Equation Modeling. Department of Psychology. McGill University.

[9] Hwang, Heungsun, et, al., (2019) A concept analysis of methodological research on composite-based structural equation modeling: bridging PLSPM and GSCA. Behaviormetrika.

[10] Henseler, J. Ringle, C. M. & Sinkovicks, R. R. (2009). The use of partial least square modeling in international marketing. *New Challenges to International Marketing Advances in International Marketing*, Volume 20, 277-319.

[11] Hidayat, Rachmat & Wulandari, Patricia, (2022) Structural Equation Modelling (SEM) in Research: Narrative Literature Review. Open Access Journal of Social Sciences Volume 5, Issue 6.

[12] Hoyle, R. H, (2012). Handbook of Structural Equation Modeling. London: The Guilford Press.

[13] Hwang, H., et. al., (2017) Generalized Structured Component Analysis with Uniqueness Terms for Accommodating Measurement Error. Frontier in Psychology.

[14] Jöreskog, K. G. & Sörbom, D. (2022). LISREL 11: Examples Guide. Chapel Hill, NC: Scientific Software International, Inc.

[15] Loehlin, J. C & Beaujean, A. A, 2017. Latent Variable Models. New York: Routlegde Taylor & Francis Group.

[16] Memon, Muntaz Ali, et. al. (2017) A Review of the Methodological Misconceptions and Guidelines related to the Application of Structural Equation Modeling: a Malaysian scenario. Journal of Applied Structural Equation Modeling. 1 (1) June 2017.

[17] Narimawati, Umi & Sarwono, Jonathan, (2020) Ragam Analisis. Yogyakarta: Penerbit Andi.

[18] Narimawati, Umi & Sarwono, Jonathan, (2023) PLS SEM untuk Riset Skripsi, Tesis dan Disertasi. Yogyakarta: Penerbit Andi.

[19] Narimawati, Umi & Sarwono, Jonathan, (2023) Mengenal Pemodelan Persamaan Struktural dengan LISREL. Yogyakarta: Penerbit Gava Media & Penerbit Graha Ilmu.

[20] Purwanto, Agus., et. al, (2021) Education Research Quantitative Analysis for Little Respondents: Comparing of Lisrel, Tetrad, GSCA, Amos, SmartPLS, WarpPLS, and SPSS. Jurnal Studi Guru dan Pembelajaran, Vol. 4, No. 2, August.

[21] Ringle, C. M., Wende, S. & Will, A. (2005). SmartPLS. http://smartpls.de

[22] Schumacker, R. E. & Lomax, R. G, 2010. A Beginner's Guide to Structural Equation Modeling. New York: Taylor & Francis Group.

[23] Schumacker, R. E & Whitettaker, T. A. 2022. A Beginner Guide to Structural Equation Modeling Edisi 5. Routledge Taylor & Francis.