

# Comparison of Covariance-Based Structural Equation Model and Partial Least Squares Equality Models

Duygu Vargör<sup>1\*</sup> and Tuncay Öğretmen<sup>2</sup>

<sup>1</sup>Ministry of Education, Turkey

<sup>2</sup>Ege University, Faculty of Education, Department of Assessment and Evaluation in Education, İzmir-Turkey

\*Corresponding Author

Duygu Vargör, Dr., Ministry of Education, Turkey.

Submitted: 2025, Aug 18; Accepted: 2025, Oct 17; Published: 2025, Dec 08

**Citation:** Vargör, D., Öğretmen, T. (2025). Comparison of Covariance-Based Structural Equation Model and Partial Least Squares Equality Models. *Gen Surgery Clin Med*, 3(2), 01-09.

## Abstract

The aim of this research is to compare the differences between the objectives, distribution assumptions, sample sizes, parameters, fit indices, and measurement models of the covariance-based structural equation model (CB-SEM) and the partial and consistent partial least squares structural equation models (PLS-SEM and PLSc-SEM) to contribute to future studies. Data from Turkey's Information and Communication (ICT) scale of the Program for International Student Assessment (PISA) for the year 2018 was used. Exploratory factor analysis (EFA) was initially conducted on the data from a sample of 5963 individuals, followed by confirmatory factor Analysis (CFA) using CB-SEM, PLS-SEM, and PLSc-SEM. CFA was performed by obtaining normal and non-normal distributions from the same sample data. The structure validity and reliability, goodness-of-fit indices, item parameters, and latent variable parameters obtained using CB-SEM, PLS-SEM, and PLSc-SEM were compared. The CB-SEM model fit indices provide a better method for explaining how well a hypothetical model fits the experimental data. PLS-SEM and PLSc-SEM, on the other hand, have sufficient reliability and validity parameters for the weight of the items, while the confidence intervals, estimations, and variances of the items are insufficient. This study concludes that it is not appropriate to claim that PLS-SEM is a preferred method when the sample size is small, and the data distributions are non-normal. It is essential for the observed data to be consistent with the hypothesis and theory; otherwise, the analysis results may lead to errors and misconceptions.

**Keywords:** PLS-SEM, PLSc-SEM, CB-SEM, Structural Equation

## 1. Introduction

Within the branch of statistics, different analysis methods are used based on the characteristics of the data and the relationships between them. Multivariate analysis encompasses methods that involve the simultaneous analysis of two or more variables [1]. Hair et al. divided these analysis techniques into two groups (exploratory and confirmatory), and these groups were further classified as first and second generation [2]. Exploratory first-generation techniques include cluster analysis, exploratory factor analysis (EFA), and multidimensional scaling, while second-generation techniques include partial least squares structural equation modeling (PLS-SEM). Confirmatory first-generation techniques consist of analysis of variance (ANOVA), logistic regression, multiple regression, and confirmatory factor analysis

(CFA), whereas second-generation techniques include covariance-based structural equation modeling (CB-SEM).

Multivariate analysis techniques vary based on the number of variables in the data, the distribution of the variables, the relationships between the variables, the ranking and grouping characteristics of the data, and the hypotheses and estimations determined.

Although first-generation multivariate data analysis techniques are widely used by many researchers, they have three common limitations:

- 1) The assumption of a regression-based simple model structure
- 2) The assumption that all variables are observable
- 3) The assumption that all variables are measured without error[3].

---

Considering the first limitation, it is misleading to examine each structure as a simple model of cause-and-effect relationships. For example, multiple regression analysis should be examined as a complex structure that includes both dependent and independent variables and can be predicted in fragments. Regarding the second limitation, when factor analysis is performed, unobservable qualities of an abstract feature may emerge. The third limitation involves a systematic or random error in each measurement. Owing to the three limitations mentioned above, first-generation techniques are considered inadequate in social sciences and many scientific research fields, as they include unobservable variables, such as attitudes, intentions, and self-efficacy [4]. To overcome these limitations, researchers have turned to second-generation methods, referred to as structural equation modeling (SEM). The fundamental reason researchers are drawn to SEM is its ability to model the relationships between multiple dependent and independent variables simultaneously.

The relationship between the hypothesis and the variables established while constructing the SEM is crucial. Because when determining the model, the relationships between dependent (endogenous) and independent (exogenous) variables should be examined. Another factor to be considered when establishing a model in SEM is determining whether the measurement model is reflective or formative. The structure of the relationship between the latent variable and observed variables should be determined. If the latent variable is the cause of the measurement statements, it results in a reflective measurement model, whereas if the latent variable is a consequence of the measurement statements, it leads to a formative measurement model.

### 1.1. Factor Analysis

One of the reasons researchers prefer SEM is its efficiency in studies involving latent variables. Factor analysis is a useful statistical method that can explain the covariance between observed variables and latent variables. In general, factor analysis assumes that the common variance is due to the factors, and the number of factors is assumed to be less than the number of variables. Total variance is the sum of the common variance with other variables and the variance of the variable itself or the variance arising from measurement errors [5]. In factor analysis, the dataset is reduced in size to reduce the dimensions of the model. With the help of the computer program used for EFA, researchers can produce as many factors as they want, or they can also produce a single factor structure. The factors generated in EFA are specified in CFA. CFA is the part of SEM that deals with measurement models of relationships between latent and observed variables. Therefore, EFA is also referred to as an unrestricted measurement model, while CFA is referred to as restricted measurement models. Different rotation techniques are used in EFA, while rotation cannot be performed in CFA. The purpose of rotation in EFA is to reduce unexplained variance by shifting factor axes. The most significant difference between EFA and CFA is that, in EFA, it is assumed that the variance of each variable is not shared with other variables, while CFA allows the calculation of shared variance among some variables.

In 1973, after Karl Jöreskog developed the CB-SEM and, later, Wold developed the variance-based structural equation model, both methods began to be used [6]. While the main purpose of both methods is to estimate relationships between structures and indicators, they differ fundamentally in their statistical understandings and especially in how they handle the measurement models of structures [7]. CB-SEM aims to maximize the explained variance using the maximum likelihood (ML) method to minimize the difference between the observed and predicted covariance matrices. CB-SEM initially calculates the covariances of variables, and only this variance is included in the derived solutions. CB-SEM attempts to explain the variance of variables with a latent variable (common factor) and the presence of individual errors for each variable. Therefore, CB-SEM is a highly suitable method for finding common factors—that is, reflective models [8]. CB-SEM formative measurements can also be used, but certain restrictions need to be applied. PLS-SEM, which is used in different disciplines, was developed by Lohmöller as an alternative method to CB-SEM [9]. PLS-SEM estimates the parameters of a set of equations in a structural equation model by combining principal component analysis with regression-based path analysis. In PLS-SEM, the model is examined as a composite structure. PLS-SEM has three basic features.

The first is to examine the model as an internal and external measurement model. In the internal measurement model, the linear relationship between the latent variable and its indicators, as well as the covariances between the latent variables, are calculated. On the other hand, in the external measurement model, the linear relationship between latent variables and observed variables is examined. The second feature of PLS-SEM is that, while examining external measurement models, a causal and predictive approach is used by separating the latent variables into formative and reflective. The third feature is the examination of the relationship between the formative and reflective latent variables and indicators of the created structure. For this reason, both a causal and predictive approach are generally adopted. Unlike CB-SEM, PLS-SEM does not examine variances separately. In PLS-SEM, not only the correlations between the indicators but also the total variance in the observed variables (indicators) are taken into account. Therefore, when estimating model relationships in the PLS-SEM approach, all variances where external variables are common with internal variables should be included [9].

Bentler and Huang, Dijkstra, and Dijkstra and Henseler discovered the "consistent PLS-SEM (PLSc-SEM)" method by using the PLS-SEM approach and some features of the CB-SEM approach [10,11]. PLS-SEM is used in studies for exploring and predicting the structure, while CB-SEM is used in studies that aim to test structures used in testing the structure.

The fit indices obtained in SEM are crucial for the evaluation of the analysis. Absolute fit indices determine how well the covariance matrix data obtained from the tested model fit the existing sample covariance matrix data [12]. However, since there is no single statistical significance test that defines the correct model,

it is necessary to consider multiple criteria and evaluate model fit simultaneously based on various criteria [13]. The fit indices vary depending on the calculation method used in the data analysis and the type of data used. In general, the compliance criteria show the extent to which the specified SEM fits the data used. Chi-square ( $\chi^2$ ), RMSEA, GFI, SRMR, CFI, TLI, etc. fit indices are used in data analysis. Chi-square ( $\chi^2$ ): The  $\chi^2$  test statistic is used for hypothesis testing to assess the goodness-of-fit of a structural equation model. The  $\chi^2$  value tests whether the covariance matrix ( $\Sigma$ ) obtained from the model is equal to the sample covariance matrix  $\Sigma(\theta)$  of the data. Root Mean Square Error of Approximation (RMSEA): RMSEA fit index determines how well the covariance matrix data obtained from the tested model, using error squares, fit the data in the existing sample covariance matrix. The RMSEA value ranges from 0 to 1. An RMSEA value less than or equal to 0.08 indicates that it is acceptable [14]. Standardized Root Mean Square Residual (SRMR): When calculating SRMR, the inconsistency between the model and the data is measured. For a good fit, the SRMR value should be less than or equal to 0.08 [12].

Normed Fit Index and Tucker Lewis Index (NFI and TLI): NFI compares the chi-square value of the zero model, which suggests that the measured variables are not related to the chi-square value obtained from the prediction model [12]. In fact, it compares the values of the model that is assumed to exist among the variables with the values of the model that is claimed to be unrelated. Therefore, the NFI value ranges from 0 to 1, and a value greater than 0.90 indicates a good fit. When the degree of freedom is taken into account while calculating the NFI value, the non-NFI (NNFI) value is obtained. NNFI is also known as the Tucker Lewis Index (TLI) [5]. Comparative Fit Index (CFI): The CFI value assumes that all latent variables are uncorrelated and compares the sample covariance matrix with the value produced by this unrelated assumed model [15]. The CFI value lies between 0 and 1, and a value greater than 0.90 indicates a good fit [12]. In cases where the fit indices are not acceptable in CFA, it may be difficult to redefine the model or create a new model. Therefore, it may be necessary to make adjustments (modifications) based on the results obtained because of CFA. These indices suggest detailed modifications to the researcher by examining the covariances between variables and latent variables [16].

### 1.2. Relevant Publications and Research

PLS-SEM has been widely used as a statistical modeling technique in recent years. In the past few decades, numerous studies have been conducted across different disciplines, comparing CB-SEM, PLS-SEM, and PLSc-SEM methods. When these studies are examined, it is evident that they primarily focus on method comparisons, the separate application of methods, and the determination of the reliability and validity of scales. In their study, Hair et al. examined the CCFA approach, which they referred to as the composite CFA [17]. They used the CFA method to validate both reflective and formative measurement models while employing the PLS-SEM approach. In this study, they stated that the CCFA method consists of linear composite structures and has advantages in validating the measurement model.

In their study on customer satisfaction and purchase intention, Dash, and Paul investigated the differences between the two methods using experimental data, without considering the problems related to sample size and distribution [18]. They concluded that PLS-SEM should be referred in composite-based models, while CB-SEM or PLSc-SEM should be used in factor-based models. Many studies have been published on the use of the PLS-SEM method and the discovery of its infrastructure. Hair et al. also examined different SEMs related to the method in their book on PLS-SEM [4,17]. When the field literature in Turkey was examined, it was found that Civelek conducted separate analyzes using the CB-SEM and PLS-SEM methods on a small and non-normal data group [19]. The analyses performed in this study were compared and, although PLS-SEM is the most preferred method for small sample sizes and non-normal data distribution, the data obtained through both methods were found to be compatible. Polat also investigated the validity and reliability of the scale applied to teachers using PLS-SEM [20]. In the study, the 29-item scale was applied to 98 individuals, and the reliability and validity of the scale were examined using the PLS-SEM approach. The study concluded that, in the field of Educational Sciences, PLS-SEM can be an alternative method in cases where the assumption of normal distribution is not fulfilled.

### 1.3. Purpose of the Study

In today's world, advancements in technology necessitate the adaptation of education policies to this change. International exams and practices conducted globally have become important in order for countries to assess their own levels and guide their education policies. One of these practices is the Program for International Student Assessment (PISA), implemented by the Organization for Economic Co-operation and Development (OECD), which aims to measure the ability of 15-year-old students to use their reading comprehension, mathematics, and science knowledge to cope with real-life problems. Therefore, the structural validity and qualifications of the scales used in the PISA application are crucial for comparing the data. In these applications, there are inconsistencies regarding whether there is a relationship between student achievement and achievements in information and communication technologies (ICT). Two fundamental issues were identified in these studies. The first issue relates to the difficulty of observing student achievement and the lack of a clear definition of the same. The second issue is that ICT are rapidly evolving, making it challenging to separate their effects from the environment [21]. In light of this information, it is believed that conducting factor analyses using two different approaches, EFA (Exploratory Factor Analysis) and CFA (Confirmatory Factor Analysis), on the data from the PISA 2018 ICT scale for the Turkish sample will contribute to the field literature.

The aim of this study is to determine the structural validity of the exams in which the validity and reliability of the internationally used scales—such as the PISA exam—are determined by using PLS-SEM, PLSc-SEM and CB-SEM. Upon examining the literature on the field, new features related to approaches have been discovered in recent years in studies conducted using PLS-SEM and PLSc-

SEM. In their study, Hair et al. examined the CCFA approach, which they called combined CFA [17]. The CCFA method is used to confirm both reflective and formative measurement models when using PLS-SEM. In the study, they stated that the CCFA method consists of linear composite structures and has advantages in confirming the measurement model. However, it is necessary to expand and update the developments in the field of PLS-SEM in light of new studies and research. The purpose of this study is to contribute to the literature by comparing the three approaches using the same sample in sub-data groups with different distribution characteristics.

In this study, three problems have been presented:

- Do fit indices, item parameters, and latent variable parameters differ when CB-SEM, PLS-SEM, and PLSc-SEM are applied to the PISA 2018 ICT Scale Turkey sample in different sample sizes?
- Does the validity and reliability of the scale differ when CB-SEM, PLS-SEM, and PLSc-SEM are applied to the Turkish sample of the PISA 2018 ICT Scale?
- Do fit indices, item parameters, and latent variable parameters differ when CB-SEM, PLS-SEM, and PLSc-SEM are applied to the PISA 2018 ICT Scale Turkey sample under the assumption of normal distribution and right or left skewed distribution?

In addition, since each of the three different methods has its limitations, it is believed that investigating the advantages and disadvantages of using them under which distribution and in solving which problems will contribute to the literature.

#### 1.4. Limitations of the Study

In this study, the ICT014 and ICT015 coded subtests of the application of the PISA 2018 ICT scale were examined. Both sub-scales use a 4-point Likert-type scale with responses being "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree." Both scales consisted of five items. While the normal and non-normal distributions were obtained in the same sample, a new variable consisting of the total score was added. For different distribution assumptions, data were selected using the R program. All variables of the selected data were organized using Microsoft Excel. One of the most important limitations of this study is that, although the results obtained are similar to that of previous studies, comparing both methods might have been affected by the characteristics of the dataset used. In fact, the results obtained for another SEM with a small sample size or different distribution

characteristics may differ. The ICT014 and ICT015 coded tests of the application of the PISA 2018 ICT scale used in this study were examined using the SmartPLS program. The analysis revealed the reflective nature of SEM and the relatively low number of latent variables as another limitation. Having different fit indices associated with the software programs used in this study posed some challenges in comparing the two approaches. Additionally, changing the calculation methods used may cause differences in the obtained results.

#### 2. Methodology

The aim of this research is to investigate an existing situation in detail. Additionally, the purpose of the study is to present the data obtained through interviews and observations to the reader in an organized and interpreted way. In descriptive analyses, a cause-effect relationship is established between the findings and, if necessary, comparisons are made between the cases [22]. Therefore, the descriptive analysis method was used in this study. In PISA applications, the school sample is determined using the two-layer random sampling method. One of the innovations made in the PISA 2018 application is the addition of the ICT survey. Research data were obtained from the official website of the OECD, which prepared the PISA 2018 assessment. In the PISA 2018 application, the ICT questionnaire evaluates students' thoughts regarding the use of digital media and digital tools with 20 items rated on a 4-point Likert-type scale. The data of students who did not answer any of the items of the questionnaire were excluded from the dataset to avoid the formation of normal, right, and left skewed sub datasets within the dataset. In this study, the data of 5963 students who answered the 10 items in the Turkey ICT014 and ICT015 scales were used. Descriptive statistics for both tests were calculated. It was observed that the group is heterogeneous, the distribution has a kurtosis (0.502), and the skewness (-0.371) is close to zero. Based on this information, it can be inferred that the distribution is normal.

Before conducting EFA using IBM SPSS25, the Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity were performed. With a KMO value of 0.908, the sample size can be considered excellent, and factor analysis can be conducted. According to Bartlett's Test of Sphericity, it can be concluded that the test data is derived from a multivariate normal distribution. EFA was then performed using IBM SPSS25.

Item	Load Value After Rotation		
	Factor Common Variance	Factor 1	Factor 2
I.1	0,538	<b>0,805</b>	0,296
I.2	0,711	<b>0,804</b>	0,295
I.3	0,683	<b>0,803</b>	0,257
I.4	0,736	<b>0,798</b>	0,215
I.5	0,733	<b>0,700</b>	0,219
I.6	0,745	0,186	<b>0,843</b>
I.7	0,687	0,212	<b>0,828</b>

I.8	0,576	0,248	0,791
I.9	0,689	0,351	0,752
I.10	0,730	0,366	0,665

**Table 1: EFA (Rotated Principal Components Analysis) Results**

Upon examining the eigenvalues of the factor structure of the ICT014 and ICT015 scales, it was calculated with IBM SPSS25 that the scales were a two-factor structure, and the eigenvalues of each factor were greater than one. This two-factor structure explains 68.270% of the total variance. The factor loads that provided the item load values of 0.30 were not considered, and items that loaded onto the same factors were determined. Additionally, it was observed that the difference between the factor loadings of items was greater than 0.29. Therefore, no items were removed from the scales. The factors with the highest factor loading were examined based on the condition that there was no overlap; the factor loadings of the items based on rotated principal components analysis results are presented in Table 1. The two-factor structure was determined after the graph of the factor loads was examined. In this study, two-factor CFA was conducted using PLS-SEM and PLSc-SEM method with SmartPLS software. Additionally, the two-factor CFA was conducted using the CB-SEM method and ML estimation in Mplus software. The ML method is based on the principle of maximizing the probability of obtaining the covariances of the observed data. At the same time, in the ML method, the sum of the squares of the differences between the sample covariance and the elements of the researcher's predicted covariance matrix is minimized. The

ML method has some limitations, such as being inefficient in small samples, having variables with normal distribution, no missing data, and independence of exogenous variables [5].

### 3. Findings

A two-factor CFA was conducted using the PLS-SEM and PLSc-SEM methods with the SmartPLS software, utilizing data from 5963 students who completed the ICT014 and ICT015 scales in Turkey. The latent variable for the ICT014 scale was named "Comfort," while the latent variable for the ICT015 scale was named "Self-Efficacy." The relationship between the implicit structure and the variables was found to be reflective. In the CFA analysis conducted using CB-SEM, PLS-SEM, and PLSc-SEM, the factor loads of the items were examined; the lower values in the 0.5% confidence interval are given in Table 2. The Mplus software took ML estimation as 1000 iterations and convergence criteria as 0.500D-04. When the factor loads of the items were examined, the factor loads of the PLS-SEM are higher than the factor loads of the CB-SEM and that the factor loads are reduced when the PLSc-SEM is used. Further, the path coefficients (the coefficient between the latent variables) are larger in CB-SEM when compared to the other two methods.

n = 5963							
Item	Comfort			Item	Self-Efficacy		
	Factor Loads				Factor Loads		
	CB	PLS	PLSc		CB	PLS	PLSc
I.1	0,616	0,719	0,674	I.6	0,762	0,812	0,711
I.2	0,769	0,839	0,787	I.7	0,775	0,826	0,751
I.3	0,764	0,822	0,746	I.8	0,735	0,802	0,816
I.4	0,864	0,872	0,839	I.9	0,820	0,858	0,858
I.5	0,854	0,867	0,833	I.10	0,787	0,828	0,730

**Table 2: Factor Loads of CB-SEM, PLS-SEM and PLSc-SEM**

The fit indices, NFI, TLI, SRMR, validity, and reliability values of the three methods compared were examined and are given in Table

3. It was found that the method with the largest NFI value and the smallest SRMR value was the CB-SEM method.

n = 5963	NFI /TLI	SRMR	R <sup>2</sup>	$\rho_A$			AVE
				Comfort	Self-Efficacy	Comfort	Self-Efficacy
CB	0,934	0,034	0,415	0,913	0,920	0,666	0,679
PLS	0,909	0,060	0,365	0,914	0,914	0,682	0,681
PLSc	0,934	0,040	0,464	0,884	0,882	0,605	0,601

**Table 3: Fit, Reliability, and Validity Indices for CB-SEM, PLS-SEM and PLSc-SEM**

When the NFI and TLI values were examined, it was seen that the values of CB-SEM and PLSc-SEM were equal and greater than the value of PLS-SEM. It is thus evident that CB-SEM provides more stable results in establishing the relationship between the estimated model and the initially given model. Therefore, this indicates that the SEM structure obtained with CB-SEM and PLSc-SEM is more compatible.

With regard to the SRMR values, PLS-SEM values were greater than the CB-SEM and PLSc-SEM values. This is because, in the Lohmöller algorithm used in PLS-SEM, the errors are ignored when examining the weights of variables.

When examining the studies conducted, it was noted using the ML estimation used in CB-SEM in the normal distribution posed a limitation. To compare CB-SEM, PLS-SEM, and PLSc-SEM in samples with different distributions based on sample size, skewness, and kurtosis coefficients, the R program was used. When the CFA results conducted using CB-SEM, PLS-SEM, and PLSc-SEM were examined, it was observed that the linear relationship

between the latent variables was positive. The data of 3000, 2500, 2000, 1500, 1000, 750, 500, 250, and 100 students among the data of 5963 students in the PISA 2018 ICT Scale Turkey sample were selected in the appropriate kurtosis range using the R program as normal, right, and left skewed distribution. EFA was conducted using CB-SEM, PLS-SEM, and PLSc-SEM with the obtained data. The findings of NFI/TLI, SRMR values obtained because of the analysis are given in Table 4. When PLS-SEM and PLSc-SEM were used, for data groups that meet the assumptions of normal and non-normal distributions, no suitable solution could be found. When examining other parameters of the conducted analyses, it was found that AVE and  $\rho_A$  values were below the acceptable limit ( $\rho_A < 0.70$  and  $AVE < 0.50$ ). At the same time, there was a negative correlation between latent variables. The biggest disadvantage of PLS-SEM and PLSc-SEM is that the relationship established between latent variables is unidirectional. Since the algorithm used in PLSc-SEM starts with the value of  $\rho_A$  as 1 at the beginning,  $\rho_A$  and AVE values decrease as the NFI value decreases in the data group. In Table 4, if there is no suitable solution, it was determined as n/a.

Distribution		TLI/NFI	SRMR				
Normal	N	CB	PLS	PLSc	CB	PLS	PLSc
	5963	0,934	0,909	0,934	0,034	0,060	0,040
	3000	0,848	0,755	0,722	0,048	0,090	0,088
	2500	0,862	0,802	0,794	0,049	0,085	0,078
	2000	0,820	0,747	0,757	0,051	0,090	0,087
	1500	0,849	0,763	0,710	0,049	0,086	0,082
	1000	0,865	0,772	0,689	0,051	0,087	0,102
	750	0,877	0,823	0,826	0,052	0,083	0,070
	500	0,865	0,810	0,837	0,048	0,082	0,067
	250	0,896	0,832	0,777	0,055	0,081	0,088
	100	0,853	0,787	0,777	0,058	0,083	0,071
Right skewed	N	CB	PLS	PLSc	CB	PLS	PLSc
	3000	0,889	0,854	0,860	0,047	0,077	0,061
	2500	0,718	0,144	n/a	0,071	0,193	n/a
	2000	0,720	0,564	0,437	0,071	0,106	0,131
	1500	0,879	0,839	0,821	0,053	0,081	0,074
	1000	0,806	0,668	0,326	0,060	0,085	0,140
	750	0,879	0,839	0,830	0,052	0,081	0,065
	500	0,848	0,347	n/a	0,058	0,218	n/a
	250	0,802	0,669	0,267	0,065	0,100	0,197
	100	0,815	0,118	n/a	0,072	0,285	n/a
Left-skewed	N	CB	PLS	PLSc	CB	PLS	PLSc
	3000	0,930	0,886	0,927	0,035	0,070	0,041
	2500	0,890	0,738	0,655	0,049	0,085	0,114
	2000	0,936	0,887	0,934	0,034	0,070	0,038
	1500	0,940	0,896	0,937	0,031	0,067	0,037
	1000	0,851	0,357	n/a	0,058	0,171	n/a
	750	0,917	0,877	0,906	0,042	0,068	0,043

	500	0,810	0,347	n/a	0,065	0,183	n/a
	250	0,772	0,105	0,102	0,080	0,656	0,642
	100	0,933	0,840	0,716	0,059	0,085	0,082

**Table 4: Fit Indices for CB-SEM, PLS-SEM and PLSc-SEM**

#### 4. Results and Discussion

In this study, data from 5963 students who answered the ICT014 and ICT015 scales in Turkey were used. Using the same sample data, CFA and EFA were conducted on data with sample sizes of 3000, 2500, 2000, 1500, 1000, 750, 500, 250, and 100, both with normal and non-normal distributions. The results obtained using CB-SEM, PLS-SEM, and PLSc-SEM can be summarized as follows:

1) When CB-SEM, PLS-SEM, and PLSc-SEM were used, for data groups that met the assumption of normal distribution, the method with the highest NFI/TLI value and the smallest SRMR value was found to be CB-SEM. The main reason for this is that the ML estimation used in CB-SEM minimizes the sum of the squares of the differences between the sample covariance and the researcher's predicted covariance matrix elements. Therefore, the resulting TLI value is higher. When trying to make the PLSc-SEM factor loads consistent, there is no definitive evidence regarding whether the distribution assumption of NFI values is effective. However, when other parameters of the analysis are examined, it is seen that the correlation between latent variables is negative in data groups with very low NFI values. When examining the correlation between latent variables in PLSc-SEM, it can be said that it is insufficient in the case of PLS-SEM.

2) When CB-SEM, PLS-SEM, and PLSc-SEM were used, for data groups that met the assumptions of normal and non-normal distributions, CB-SEM was the method with the smallest SRMR value. This is because the Lohmöller and Dijkstra algorithms used in PLS-SEM and PLSc-SEM work in the first step of the linear equation between the latent variables and the observed variables without considering the error.

3) When CB-SEM, PLS-SEM, and PLSc-SEM were used, for data groups that met the assumption of normal distribution, the factor loads of all items were greater in PLS-SEM compared to the factor loads obtained in CB-SEM. It can thus be stated that, in PLS-SEM, the convergence is fast, and the errors are not considered important. Since PLSc-SEM attempts to reduce factor loadings to be consistent with the correlation matrix, the factor loadings obtained in PLSc-SEM are smaller than those obtained using PLS-SEM. However, when the results obtained were examined, for some items, the factor loadings in PLSc-SEM were greater than those in PLS-SEM. When the other results of the analyses were examined, there were items with high weights in the equation established linearly between the latent variables and the observed variables.

4) When PLS-SEM and PLSc-SEM were used for data groups that met the assumptions of normal and non-normal distributions, different methods were followed when calculating variances. When the other parameters of the analysis obtained were examined, it was seen that while CB-SEM examined the variance shared with other indicators, the item's unique variance and error variances for

each item, PLS-SEM, and PLSc-SEM work on total variance.

5) When CB-SEM is used, model fit indices are a better method for explaining how well a hypothetical model fits the experimental data. Based on the analysis results, CB-SEM calculates confidence intervals, estimations, item variances, and fit indices. However, CB-SEM is not sufficient for calculating reliability and validity parameters. PLS-SEM and PLSc-SEM, on the other hand, have sufficient reliability and validity parameters for the weight of the items, while their confidence intervals, estimations, and variances of the items are insufficient. Therefore, PLS-SEM and PLSc-SEM can be recommended for EFA, while CB-SEM is better suited for CFA.

6) PLS-SEM and PLSc-SEM estimate the structural model and evaluate the model's explanatory and predictive power. Therefore, both methods are more sensitive than CB-SEM regarding examining the reliability and validity parameters. CB-SEM focuses on estimations and fit indices.

7) In the existing field literature, there are different opinions regarding data distribution and sample size. In this study, it was found that PLS-SEM and PLSc-SEM do not work efficiently in small sample sizes and non-normal distributions. Additionally, PLS-SEM and PLSc-SEM offer alternatives for modifying the structural model. Supporting the studies of Hair et al., it can be concluded that as the sample size decreases, the statistical consistency of PLS-SEM and PLSc-SEM decreases [4]. Therefore, it would be misleading to state that PLS-SEM and PLSc-SEM work better as the sample size decreases.

#### 5. Suggestions

In future studies, when the sample size is small and in non-normal distributions using experimental datas (such as PISA), the choice between these two methods should be made by examining the assumptions and objectives of both approaches. The observed dataset should be in accordance with the formulated hypothesis and theory. Otherwise, the results obtained from the analysis will contain errors and misconceptions. It has been observed that PLS-SEM and PLSc-SEM are better suited for determining complex models in SEM (reflective and formative, internal and external latent variables) and offer alternatives to researchers. However, it should not be forgotten that, as the characteristics of the data distribution and sample size change, values outside limits can be found in PLS-SEM and PLSc-SEM. If theoretical or conceptual assumptions support large models and there is sufficient data, PLS-SEM and PLSc-SEM should be used. In the field of Educational Sciences and the international exams such as the PISA exam, for SEM studies with experimental data, it is recommended to use PLS-SEM and PLSc-SEM in EFA, while CB-SEM is recommended for CFA. Despite the limitations of CB-SEM in terms of data distribution and sample size, it offers various estimation methods, such as GLS and ULS. In future studies, comparisons can be made

with other estimations of CB-SEM [23-62].

## References

1. Konishi, S. (2014). *Introduction to multivariate analysis: linear and nonlinear modeling*. CRC Press.
2. Hair Jr, J. F., Babin, B. J., & Krey, N. (2017). Covariance-based structural equation modeling in the Journal of Advertising: Review and recommendations. *Journal of Advertising*, 46(1), 163-177.
3. Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding statistics*, 3(4), 283-297.
4. Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial least squares structural equation modeling (PLS-SEM) using R: A workbook.
5. Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
6. Wold, H. (1975). Soft modelling by latent variables: the non-linear iterative partial least squares (NIPALS) approach. *Journal of Applied Probability*, 12(S1), 117-142.
7. Joreskog, K. G. (1982). The ML and PLS techniques for modeling with latent variables: Historical and comparative aspects. *Systems under indirect observation, part I*, 263-270.
8. 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.
9. Lohmoller, J. (1989). Predictive vs. structural modeling: Pls vs. ml. Latent Variable Path Modeling with Partial Least Squares.
10. Dijkstra, T. K., & Schermelleh-Engel, K. (2014). Consistent partial least squares for nonlinear structural equation models. *Psychometrika*, 79(4), 585-604.
11. Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational statistics & data analysis*, 81, 10-23.
12. Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, 3(4), 424.
13. Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, 8(2), 23-74.
14. Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *sociological methods & research*, 21 (2), 230-258.
15. Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6 (1), 53-60.
16. Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (Vol. 2). Ankara: Pegem akademi.
17. Hair Jr, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of business research*, 109, 101-110.
18. Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, 173, 121092.
19. Civelek, M. E. (2018). Comparison of covariance-based and partial least square structural equation modeling methods under non-normal distribution and small sample size limitations. *Eurasian Academy of Sciences Eurasian Econometrics, Statistics & Empirical Economics Journal*, 10, 39-50.
20. Polat, M. (2022). Validity and reliability of the Turkish-adapted school participant empowerment scale (SPES) for teachers with the PLS-SEM approach. *Journal of STEM Teacher Institutes*, 2(1), 10-23.
21. Youssef, A. B., & Dahmani, M. (2008). The impact of ICT on student performance in higher education: Direct effects, indirect effects and organisational change. *Rev. U. Soc. Conocimiento*, 5, 45.
22. Yıldırım, A., & Simsek, H. (1999). Sosyal bilimlerde nitel araştırma yöntemleri (11 baski: 1999-2018).
23. Afthanorhan, W. M. A. B. W. (2013). A comparison of partial least square structural equation modeling (PLS-SEM) and covariance based structural equation modeling (CB-SEM) for confirmatory factor analysis. *International Journal of Engineering Science and Innovative Technology*, 2(5), 198-205
24. Albright, J. J., & Park, H. M. (2006). Confirmatory factor analysis using AMOS, LISREL, and MPLUS. *The Trustees of Indiana University*, 2008.
25. AlNuaimi, B. K., Khan, M., & Ajmal, M. M. (2021). The role of big data analytics capabilities in greening e-procurement: A higher order PLS-SEM analysis. *Technological Forecasting and Social Change*, 169, 120808.
26. Andy, F. (2000). Discovering statistics using spss for windows: Advanced techniques for the beginner.
27. Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the academy of marketing science*, 40(1), 8-34.
28. 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.
29. Bartlett, M. S. (1950). Tests of significance in factor analysis. *British journal of psychology*.
30. Cepeda-Carrión, G., Hair, J. F., Ringle, C. M., Roldán, J. L., & García-Fernández, J. (2022). Guest editorial: Sports management research using partial least squares structural equation modeling (PLS-SEM). *International Journal of Sports Marketing and Sponsorship*, 23(2), 229-240.
31. Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, 295(2), 295-336.
32. Cohen, J. (2013). *Statistical power analysis for the behavioral*

sciences. routledge.

33. Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: potentially serious and misleading consequences due to uncorrected measurement error. *Psychological methods, 19*(2), 300.
34. Cramer III, R. D. (1993). Partial least squares (PLS): its strengths and limitations. *Perspectives in drug discovery and design, 1*(2), 269-278.
35. Jöreskog, K. (2001). *Structural equation modeling: present and future; a Festschrift in honor of Karl Jöreskog*. SSI, Scientific Software Int.
36. Çakır, F. S. (2019). Kısmi en küçük kareler yapısal eşitlik modellemesi (PLS-SEM) ve bir uygulama. *Sosyal Araştırmalar ve Davranış Bilimleri, 5*(9), 111-128.
37. Garson, G. D. (2016). Partial least squares. Regression and structural equation models. *Statistical Publishing Associates*.
38. Geladi, P. (1988). Notes on the history and nature of partial least squares (PLS) modelling. *Journal of Chemometrics, 2*(4), 231-246.
39. Götz, O., Liehr-Gobbers, K., & Krafft, M. (2009). Evaluation of structural equation models using the partial least squares (PLS) approach. In *Handbook of partial least squares: Concepts, methods and applications* (pp. 691-711). Berlin, Heidelberg: Springer Berlin Heidelberg.
40. Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice, 19*(2), 139-152.
41. Hair, J., & Alamer, A. (2022). Partial Least Squares Structural Equation Modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics, 1*(3), 100027.
42. Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In *New challenges to international marketing* (pp. 277-319). Emerald Group Publishing Limited.
43. Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the academy of marketing science, 43*(1), 115-135.
44. James, L. R., Mulaik, S. A., & Brett, J. M. (1983). *Causal analysis: Assumptions, models, and data*. Beverly Hills (Calif.): Sage, 1983..
45. John, O. P., & Benet-Martínez, V. (2000). Measurement: Reliability, construct validation, and scale construction. *Handbook of research methods in social and personality psychology*.
46. Jöreskog, K. G. (1973). Analysis of covariance structures. In *Multivariate analysis-III* (pp. 263-285). Academic Press.
47. Kock, N. (2014). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in PLS-based SEM. *International Journal of e-Collaboration (ijec), 10*(1), 1-13.
48. Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information systems journal, 28*(1), 227-261.
49. Lowry, P. B., & Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE transactions on professional communication, 57*(2), 123-146.
50. MEB. (2019, Aralık). PISA 2018 Türkiye ön raporu. [http://www.meb.gov.tr/meb\\_iys\\_dosyalar/2019\\_12/03105347\\_PISA\\_2018\\_Turkiye\\_On\\_Raporu.pdf](http://www.meb.gov.tr/meb_iys_dosyalar/2019_12/03105347_PISA_2018_Turkiye_On_Raporu.pdf). adresinden alındı.
51. Muthen, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British journal of mathematical and statistical psychology, 45*(1), 19-30.
52. OECD. (2019). <https://www.oecd.org/pisa/data/2018database> adresinden alındı.
53. Rasoolimanesh, S. M. (2022). Discriminant validity assessment in PLS-SEM: A comprehensive composite-based approach. *Data Analysis Perspectives Journal, 3*(2), 1-8.
54. Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial least squares structural equation modeling. In *Handbook of market research* (pp. 587-632). Cham: Springer International Publishing.
55. Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and psychological measurement, 80*(3), 421-445.
56. Smartpls. (2022). <https://www.smartpls.com/documentation/choosing-pls-sem/pls-sem-compared-with-cbsem> adresinden alındı.
57. Tenenhaus, M. (2008). Component-based structural equation modelling. *Total quality management, 19*(7-8), 871-886.
58. Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education*. Macmillan Publishing Co, Inc.
59. Esposito Vinzi, V., Chin, W. W., Henseler, J., & Wang, H. (2010). Handbook of partial least squares: Concepts, methods and applications.
60. Wold, H. (1982). Models for knowledge. *The making of statisticians, 189-212*.
61. Wold, S., Ruhe, A., Wold, H., & Dunn, III, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing, 5*(3), 735-743.
62. Wong, K. K. K. (2019). Mastering partial least squares structural equation modeling (PLS-Sem) with Smartpls in 38 Hours. IUiverse.

**Copyright:** ©2025 Duygu Vargör, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.