

Comparison of Confirmatory Factor Analysis Estimation Methods on Binary Data

Abdullah Faruk Kilic ¹, Ibrahim Uysal ^{2,*}, Burcu Atar ³

¹Department of Educational Sciences, Faculty of Education, Adiyaman University, Adiyaman, Turkey

²Department of Educational Sciences, Faculty of Education, Bolu Abant İzzet Baysal University, Bolu, Turkey

³Department of Educational Sciences, Faculty of Education, Hacettepe University, Ankara, Turkey

ARTICLE HISTORY

Received: Dec. 17, 2019

Revised: July 01, 2020

Accepted: Aug. 09, 2020

KEYWORDS

Confirmatory factor analysis,
Estimation methods,
Binary data,
Simulation

Abstract: This Monte Carlo simulation study aimed to investigate confirmatory factor analysis (CFA) estimation methods under different conditions, such as sample size, distribution of indicators, test length, average factor loading, and factor structure. Binary data were generated to compare the performance of maximum likelihood (ML), mean and variance adjusted unweighted least squares (ULSMV), mean and variance adjusted weighted least squares (WLSMV), and Bayesian estimators. As a result of the study, it was revealed that increased average factor loading and sample size had a positive effect on the performance of the estimation methods. According to the research findings, it can be said that the methods are sufficient to estimate average factor loading and interfactor correlations, regardless of the estimation methods, in most of the conditions where the average factor loading is 0.7. In small sample sizes particularly, the interfactor correlation was underestimated for skewed indicator conditions. According to the findings of the study, although there is not the most accurate method in all conditions, it can be recommended to use ULSMV method because it performs adequately in more conditions.

1. INTRODUCTION

Most researchers conducting research in social, behavioral, and educational sciences usually work on psychological attributes. Psychological attributes, also named as constructs, are theoretical concepts. Psychological constructs cannot be directly observed: the degree to which a construct characterizes an individual can only be predicted by observing the behaviors of the individual (Crocker & Algina, 2008). To analyze the relationships among observed variables and latent constructs, researchers widely use structural equation modeling techniques (Byrne, 2016; Raykov & Marcoulides, 2006). The use of confirmatory factor analysis (CFA) is also widely accepted as one of the structural equation models to examine the construct validity of the hypothesis (AERA et al., 2014).

When the scale development and adaptation studies in the literature are examined, it is observed that CFA is frequently used for collecting evidence for construct validity. Acar-Güvendir and

CONTACT: İbrahim UYSAL ✉ ibrahimuysal06@gmail.com 📍 Department of Educational Sciences, Faculty of Education, Bolu Abant İzzet Baysal University, Bolu, Turkey

ISSN-e: 2148-7456 /© IJATE 2020

Özer-Özkan (2015) and Şahin and Boztunç Öztürk (2018) examined scale development studies and they reported that CFA was used in 61% and 52% of these studies, respectively. Deciding the estimation method used in CFA is all-important to obtain unbiased parameter estimations. For this reason, it is also important to examine which estimation method is unbiased.

When the literature is examined, there are many studies comparing CFA estimation methods. One of the most comprehensive of these studies is one conducted by Forero et al. (2009). In this study, the researchers studied 324 simulation conditions and the performance of diagonally weighted least squares (DWLS) and unweighted least squares (ULS) estimation methods was compared in terms of sample size, measurement model, test lengths, factor loadings, and categories of indicators. As a result of the research, it was reported that both methods had similar results but ULS had more accurate and less variable results for parameter estimations.

Another comprehensive study in the literature was conducted by Flora and Curran (2004). In this study, they manipulated latent response (y^*) distributions, model specifications, sample sizes, and number of categories (160 simulation conditions). As a result of the study, it was reported that, while WLS requires a large sample size, robust WLS performs better for all conditions. Also, they reported that polychoric correlation is strong against moderate violations of normality.

The study conducted by Rhemtulla et al. (2012) aimed to compare the performance of robust ML and robust categorical least squares estimation (cat-LS) method. CFA model size, underlying distribution, number of indicator categories, threshold symmetry, and sample size were manipulated. As a result of the study, it was reported that ML was more sensitive to asymmetric thresholds. The cat-LS method was suggested for indicators which have fewer than five categories.

When the other studies in the literature were researched, it was observed that there are many studies which examine datasets consists of five categories indicators (Babakus et al., 1987; Lei, 2009; Morata-Ramirez & Holgado-Tello, 2013; B. O. Muthén & Kaplan, 1985; Potthast, 1993). There are also studies examining data consisting of other than five categories. Dolan (1994) used 2, 3, 5 and 7 categories indicators, for example; Green et al. (1997) used datasets with 2, 4, and 6 categories and continuous indicators. Flora and Curran (2004) used 2 and 5 categories indicators; Beauducel and Herzberg (2006) used 2, 3, 4, 5, and 6 categories indicators; Forero et al. (2009) used 2 and 5 categories indicators; Yang-Wallentin et al. (2010) used 2, 5 and 7 categories indicators; Rhemtulla et al. (2012) used 2, 3, 4, 5, 6, and 7 categories indicators; Liang and Yang (2014) used 2 and 4 categories indicators, and Moshagen and Musch (2014) used 2 and 5 categories indicators. However, in most of these studies, frequentist estimation methods have been compared. In addition, in most of these studies, datasets were generated such that the factor loadings of the indicators were equal (Beauducel & Herzberg, 2006; Flora & Curran, 2004; Forero et al., 2009; Liang & Yang, 2014; B. O. Muthén & Kaplan, 1985; Nestler, 2013; Shi et al., 2018). However, in real life applications, it is difficult to have equal factor loadings of all indicators.

The motivation of the present study is to compare Bayesian estimation method and frequentist estimation methods under simulation conditions. Studies comparing the performance of Bayesian method with frequentist methods are also found in the literature. For example, Liang and Yang (2014) compared the performance of WLSMV and Bayesian (informative priors and non-informative priors) methods in 96 simulation conditions. However, in this study, the factor loadings of the indicators were manipulated to be 0.50 and 0.80. There was also no comparison with the performance of the ML method used by default in most software. Xu (2019) compared robust ML (MLR) and Bayesian estimation method's performance in 27 simulation conditions. However, in this study, factor loadings of all indicators were fixed as 0.7. In the study conducted by Önen (2019), ML and Bayesian estimation methods were compared in terms of detecting

model misspecification, using 0.30 and 0.80 as factor loadings in the simulation study.

When the researches in the literature were studied, although many datasets consisting of different categories of indicators were used, we could not find any paper which compared ML, ULSMV, WLSMV, and Bayesian estimation methods and, at the same time, not have fixed factor loadings of indicators. The differentiation of factor loadings of the indicators would be more appropriate for real situations. Therefore, in this study, the average factor loading, which is more suitable for real conditions, was determined as the simulation condition and the factor loadings of the indicators were generated to be different from each other (details are given in the method section). At the same time, the aim is to compare the performance of Bayesian and frequentist estimation methods (ML, ULSMV and WLSMV) for binary indicators. The present study differs from other studies in the literature in terms of compared estimation methods (frequentist vs. Bayesian) and not fixing the factor loadings of all indicators. Therefore, it is considered that the current research will contribute to the literature in order to examine which estimation methods (ML, ULSMV, WLSMV, or Bayes) perform better in binary data under different simulation conditions and will help researchers in practice.

2. METHOD

This research was designed as a Monte Carlo simulation study. Monte Carlo simulations use random sampling for a statistical model across varying conditions (Harrison, 2010). Thus, suggestions can be made by investigating the effects of different factors for the statistical model (Gilbert, 1999). The main purpose of this study was to investigate estimation methods performance under different simulation conditions. For this purpose, unlike other studies, average factor loading is considered as a simulation condition. In addition, the performance of Bayesian and frequentist estimation methods.

2.1. Estimation Methods

Estimation methods differ from each other in terms of the analysis processes they use and assumptions. In general, there are four types of estimation methods: maximum likelihood; unweighted least squares; generalized least squares, and asymptotically distribution-free (generally called as weighted least squares). Each estimation approach tries to minimize the corresponding fit function (Raykov & Marcoulides, 2006). The fit function expresses the fit between the covariances obtained from the sample and the covariances obtained from the model established by the researcher (Kline, 2016).

The maximum likelihood (ML) method is the most commonly used estimation method by researchers (Bollen, 1989). Being the default estimation method in most software, ML may be used more frequently in research. ML estimation method assumes that indicators are measured on continuous scales and requires a large sample size. Although ML requires a continuous data set, it was seen that ML is used for binary data sets in the literature. For example, ML method was used in binary data sets in Koğar and Yılmaz Koğar's (2015) study. The ULS method makes estimations under the assumption of continuous variables holding multivariate normal distribution. In addition, all variables in this process should take place on the same scale (Kline, 2016). The weighted least squares (WLS) method has no distributional assumption. Estimations can be made for both continuous and categorical indicators. However, this estimation method needs a large sample size (Kline, 2016). Bayesian methods differs from other frequency-based methods in terms of fixed and free parameters. When the ML method calculates the values which will make the obtained likelihood function maximum, Bayesian methods make estimations by combining the prior distribution of the data with the posterior distribution of parameter estimation (B. O. Muthén & Asparouhov, 2012).

The estimation methods explained above are considered as essential methods; modified estimation methods have been obtained with the help of some corrections via essential methods.

In weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean adjusted chi-square test statistic (WLSM) method, which is developed based on the WLS estimation method, the average corrected chi-square test statistics are produced by using full weight matrix. When the variances in the WLSM method are corrected, a modification of the WLS method is obtained with the WLSMV method. Full weighted matrix is used in the WLSMV method (L. K. Muthén & Muthén, 2015). In ULS parameter estimates with standard errors and a mean and variance adjusted (ULSMV) method, which is developed based on the ULS method, both the averages and the variances are corrected and the chi-square test statistics is calculated over the full weighted matrix (L. K. Muthén & Muthén, 2015).

2.2. Simulation Design

Five factors were manipulated in this simulation study: (i) sample size (200, 500 and 1,000); (ii) distribution of indicators (left skewed, normal and right skewed); (iii) test length (10 and 20 indicators); (iv) average factor loading (0.4 and 0.7), and (v) factor structure (unidimensional, two factors [$\varphi = 0$], two factors [$\varphi = 0.3$], two factors [$\varphi = 0.6$]). Full crossed design was adopted for simulation conditions.

The sample sizes were 200, 500 and 1,000. Boomsma (1985) suggests a sample size of at least 200 to avoid non-convergence and improper solutions. In addition, Mulaik (2009) states that a sample size of less than 200 is inadequate for statistical inference purposes with chi-square statistics in CFA. Liang and Yang (2014) also emphasize that there are very few studies with sample sizes of less than 200. Therefore, 200 was specified as a minimum sample size. Other sample sizes were specified to examine the effects of sample sizes on the performance of the estimation methods. In addition, the 1,000 sample size was included in the study as recommended as a minimum sample size by some researchers (Comrey & Lee, 1992; Floyd & Widaman, 1995; Gorsuch, 1974; Guadagnoli & Velicer, 1988; Streiner, 1994).

The distribution of indicators was manipulated to be left-skewed, normal and right-skewed. The ML estimation method estimates parameters under the assumption that the variables meets multivariate normal distribution (Tabachnik & Fidell, 2012). WLS and ULS are asymptotic distribution free methods (Brown, 2015). The aim was to examine the estimation methods performance when indicators were skewed. Therefore, the skewness of the indicators was specified as a simulation condition (details are in the data generation section).

The test length conditions were manipulated as 10 and 20 indicators. In order to examine the performance of the estimation methods in short tests, a test length of 10 indicators was specified as a simulation condition. The 20 item condition was determined to examine how the results change when the test length increases.

The average factor loading was specified as 0.4 (low) and 0.7 (high). Unlike other studies, the factor loadings of the indicators were generated to be different from each other. Since the lowest factor loading was suggested to be 0.4 (Stevens, 2009) or .32 (Tabachnik & Fidell, 2012), the average factor loading was specified as 0.4 for low factor loading. Since the average factor loading was used in current study, the condition of 0.40 was added as the lowest average factor loading. Because the factor loadings of the items can be smaller than 0.40 (see Table 2). As we aimed to investigate the performance of estimation methods at high factor loading, the average factor loading was specified as 0.7.

Factor structure is considered as unidimensional and two-factors ($\varphi = 0, 0.3$ and 0.6). When the studies in the literature are examined, it is observed that achievement tests are usually unidimensional (Anıl et al., 2010; Kılıç & Kelecioğlu, 2016) but, in some cases, two-factors structures may also occur (Lissitz et al., 2012; Thissen et al., 1994). Therefore, both unidimensional and two-factors structures were specified as simulation conditions. Interfactor correlations were set to $\varphi = 0, 0.3$ and 0.6 and were manipulated to examine how the magnitude

of the relationship between factors in two-factors structures affected the performance of estimation methods. While $\varphi = 0$ was specified because of the performance of the estimation methods in unrelated structures, $\varphi = 0.3$ was specified because of its frequent use in studies (Curran et al., 1996; Flora & Curran, 2004; Li, 2016). Thus, the results of the study can be compared to other studies in the literature. $\varphi = 0.6$ was specified because it offered the chance of examining how the increase of interfactor correlation affected the performance of the estimation methods. Thus, the aim is to examine the performance of estimation methods under these changing conditions. Table 1 contains a summary of the factors held constant and manipulated factors with their levels.

Table 1. Simulation conditions

Fixed Factor		Manipulated Factors			
Number of Categories of Indicators	Sample Size	Distribution of Indicators	Test Length	Average Factor Loading	Factor Structure (Model)
1-0	200	Left-Skewed			Unidimensional
	500	Normal	10	0.40	Two Factors ($\varphi = 0$)
	1,000	Right-Skewed	20	0.70	Two Factors ($\varphi = 0.3$)

Full crossed factorial design was used in the study. By crossing each condition, $3 \times 3 \times 2 \times 2 \times 4 = 72$ simulation conditions have been studied. The number of indicators is equally divided between the factors in two-factorial models. For example, in two-factors models with 10 indicators, five indicators were included in each factor. For each condition, 1,000 replications were obtained. The models examined in the study are presented in Figure 1.

The factor loadings were specified as Table 2.

Table 2. Factor loadings used in study

Item Number	Figure 1.a		Figure 1.c		Figure 1.b		Figure 1.d	
	Factor Loadings		Factor Loadings		Factor Loadings		Factor Loadings	
	Average Factor Loading							
	0.4	0.7	0.4	0.7	0.4	0.7	0.4	0.7
1	0.39	0.68	0.39	0.72	0.36	0.68	0.37	0.68
2	0.37	0.72	0.37	0.73	0.40	0.73	0.37	0.73
3	0.38	0.68	0.38	0.69	0.40	0.71	0.38	0.71
4	0.39	0.68	0.39	0.68	0.39	0.69	0.44	0.69
5	0.45	0.7	0.45	0.70	0.43	0.72	0.34	0.72
6	0.39	0.72	0.39	0.73	0.39	0.69	0.35	0.69
7	0.42	0.7	0.42	0.72	0.40	0.70	0.45	0.70
8	0.43	0.73	0.43	0.69	0.39	0.71	0.44	0.71
9	0.42	0.72	0.42	0.71	0.40	0.73	0.36	0.73
10	0.34	0.70	0.34	0.67	0.44	0.69	0.46	0.69
11					0.38	0.72	0.36	0.70
12					0.42	0.72	0.40	0.68
13					0.41	0.71	0.40	0.71
14					0.39	0.7	0.39	0.72
15					0.45	0.71	0.43	0.70
16					0.38	0.72	0.39	0.70
17					0.42	0.71	0.40	0.67
18					0.41	0.71	0.39	0.74
19					0.41	0.71	0.40	0.68
20					0.40	0.72	0.44	0.71

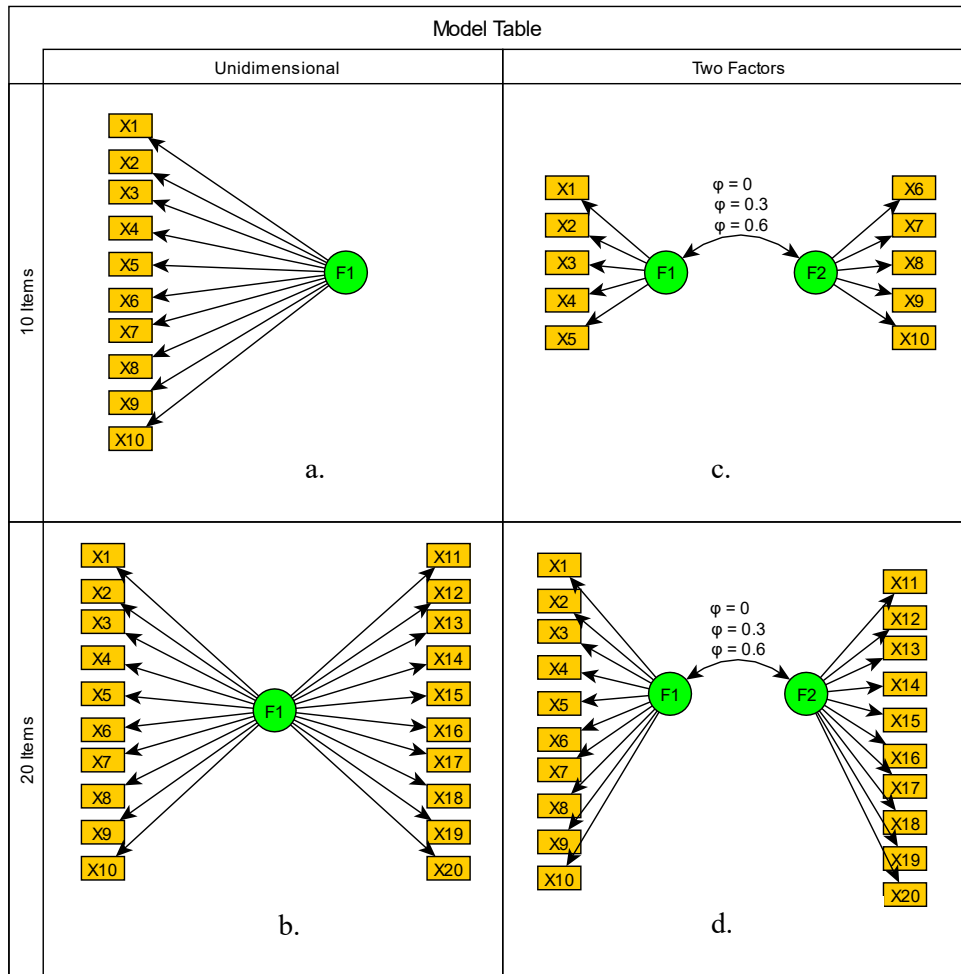


Figure 1. Models examined in the research

2.3. Data Generation

A latent response variable framework was used in data generation (Brown, 2015; B. O. Muthén & Asparouhov, 2002). Accordingly, the datasets were firstly generated as continuous which holds multivariate normal distribution. Then, the datasets were categorized according to the skewness of the indicators using threshold values. The threshold values are specified as $\{0\}$ for normal distribution, $\{1.05\}$ for right-skewed and $\{-1.05\}$ for left-skewed. In this case, the mean skewness values of the indicators are 0 for normal distribution, 2.00 for right-skewed distribution and -2.00 for left-skewed distribution. Kurtosis values are 2, 5 and 5 respectively. The lavaan package (Rosseel, 2012) in the R software (R Core Team, 2018) was used for data generation.

2.4. Data Analysis

The Mplus software (L. K. Muthén & Muthén, 2012) was used to analyze the generated data. The MplusAutomation package (Hallquist & Wiley, 2017) was used to analyze the simulated data and to obtain the outputs of the analyses. The performance of ML, mean and variance adjusted weighted least squares (WLSMV), mean and variance adjusted unweighted least squares (ULSMV) and Bayesian estimation methods were compared in terms of outcome variables. The number of iterations in ML, ULSMV and WLSMV methods is limited to 1,000, which is the default value of Mplus.

When the Bayesian estimation method was used, informative and non-informative priors could be used. Informative priors can be used if the researcher has information about the distribution of parameters. However, non-informative priors can be used if the researcher does not have information about the distribution of parameters (B. O. Muthén & Asparouhov, 2012). Non-informative priors, which is the default in Mplus used in this study, were determined as follows: for indicators, $(\tau) \sim N(0, \infty)$; for factor loadings $(\lambda) \sim N(0,5)$; for regression coefficients $(\beta) \sim N(0,5)$; for latent response variable's mean / intersection $(\alpha) \sim N(0, \infty)$, and, for latent response variable's variance, \sim inverse Gamma $(-1,0)$ was used (L. K. Muthén & Muthén, 2015). In this study, tetrachoric correlation matrix was used to conduct CFA because of the binary data.

2.5. Outcome Variables

Firstly, non-convergence solutions were investigated. Following this, improper solutions were examined. If the factor loadings of the indicators are -1.00 and smaller or $+1.00$ and greater, then this solution was treated as an improper solution and excluded from further analysis.

In order to compare data obtained from the simulation study, relative percentage bias (RPB) values were used (DiStefano & Morgan, 2014; Flora & Curran, 2004; Jin et al., 2016; Lei, 2009; Liang & Yang, 2014). The equation for RPB can be formulated as below:

$$RPB = \frac{\hat{\theta} - \theta_{True}}{\theta_{True}} \cdot 100\% \dots\dots\dots 1$$

Here, $\hat{\theta}$ is the mean of sample estimates over 1,000 replications, whereas θ_{True} presents the true value. When the formula of RPB value is examined, it is seen that the value calculated is a percentage. In studies where RPB values are used, absolute values of RPB greater than 10 are taken as the evaluation criteria (Curran et al., 1996; Flora & Curran, 2004; Rhemtulla et al., 2012). Similarly, in this research, RPB values greater than 10 were labeled biased. In this study, RPB for interfactor correlation was calculated only for $\varphi = 0.3$ and $\varphi = 0.6$. Because of the zero divided problem ($\theta_{True}=0$ in Equation 1), RPB was calculated only for $\varphi = 0.3$ and $\varphi = 0.6$ conditions.

When reporting RPB, averages of the indicators were calculated. The mean RPB value of the indicators is demonstrated in graphs. In addition, RPB values for each item are given in a table in [Appendix 3](#).

Coverage rate also was used to compare estimation methods' performances. Coverage rate examines the inclusion of the real parameter value of the confidence interval to be established around the parameter estimation. For this purpose, a 95% confidence interval was created for each estimation using the standard error of the estimation, and whether the real parameter value was in this interval was examined. Collins et al. (2001) suggest a coverage rate less than 90% is problematic. Therefore, in the present study, the cut-off point of the coverage rate was 90%.

For the relative bias of standard errors, the relative standard error bias (r-seb) was also calculated. For this:

$$r - seb = \frac{\frac{1}{nrep} \sum_{t=1}^{nrep} \widehat{se}(\hat{\theta}_{pt})}{sd(\hat{\theta}_{pt})} \dots\dots\dots 2$$

equality was used. Where $\widehat{se}(\hat{\theta}_{pt})$, standard error of parameter p for t. replication, $sd(\hat{\theta}_{pt})$ is standard deviation of parameter p for t. replication. R-seb value was classified by Holtmann et al. (2016). Holtmann et al. (2016) as $5/6 < r-seb < 6/5$ negligible, $2/3 < r-seb < 5/6$ ve $6/5 < r-seb < 3/2$ medium, and $r-seb < 2/3$ or $r-seb > 3/2$, large. In the present study, r-seb values which were negligible, and medium were considered as acceptable.

3. RESULT

In this section, results of the simulation study are provided according to the outcome variables.

3.1. Non-convergence and Improper Solutions

Non-convergence was encountered in 15 datasets (0.01%) of 144,000 datasets mostly in 200 sample sizes for the ML method. One of these datasets is in the 500 sample size and the other 14 are in the 200 sample size. The ML method was not converged under conditions where average factor loading is 0.4, indicators follow skewed distribution (right or left) and the number of indicators is 10.

Non-convergence was encountered in 19 datasets (0.01%) of 144,000 datasets, mostly in 200 sample sizes for the Bayesian method. The Bayesian method was not converged under conditions where the number of factors was two, average factor loading was 0.4 and the number of indicators was 10.

The ULSMV and WLSMV methods have more non-convergent datasets than ML and Bayesian methods. Non-convergence was encountered in 2,755 datasets (1.91%) of 144,000 datasets for the ULSMV method. When the properties of the non-converged datasets were examined, it was observed that it occurs mostly where the sample size is 200 and indicators are skewed in the ULSMV method. All non-converged datasets occurred under conditions where the average factor loading was 0.4.

The WLSMV method has non-convergent solutions in 2,856 datasets (1.98%). It was observed that non-convergence occurred mostly in conditions where the sample size is 200 and average factor loading is 0.4. Non-convergent solutions are detailed in [Appendix 1](#).

The Bayesian method has no improper solution. ML has improper solutions in 162 datasets under conditions where sample size is 200, average factor loading is 0.4, the number of indicators is 10, which were skewed. ML has improper solutions for only two factors conditions.

In the ULSMV method, there were improper solutions in 1,534 datasets (1.06%). It was observed that these datasets generally emerged under conditions where the number of indicators is 10, sample size is 200, average factor loading is 0.4, and indicators were skewed. In the WLSMV method, there were improper solutions in 1,877 datasets (1.30%). It was observed that these datasets generally emerged under conditions where the number of indicators is 10, sample size is 200, average factor loading is 0.4, and indicators were skewed. The number of datasets with the improper solution is detailed in [Appendix 2](#).

3.2. Relative Percentage Bias

3.2.1. Relative percentage bias of factor loadings

RPB values obtained from simulation conditions are presented in [Figure 2](#). In addition, the maximum, average and minimum values of the RPB obtained from the items are given in [Appendix 3](#) for researchers who want to examine them.

When the RPB values obtained from the estimation methods for factor loadings ([Figure 2](#)) are examined, it can be said that all methods have an acceptable bias for sample sizes of 500 and 1,000. It was observed that RPB values of all estimation methods are less than 10 under conditions with average factor loading of 0.7 in a sample size of 200. However, RPB can be smaller than -10 where average factor loading is 0.4 in the sample size of 200. Considering the conditions with an average factor loading of 0.4 and the number of indicators of 10, the RPB of the Bayesian method is less than -10 where the indicators are skewed and factorial structure consist of two factors (for $\phi = 0$ and 0.3). For the same conditions, except for the number of items, the RPB of the WLSMV is less than -10 where the number of indicators is 20 and two

factorial structure (for $\varphi = 0$ and 0.3). Increasing the number of indicators in skewed distributions reduced bias. In addition, increasing the interfactor correlation reduced bias. ML and ULSMV methods have acceptable bias in all conditions. The WLSMV method's RPB is less than -10 in just one condition (mean factor loading is 0.4, right-skewed indicators, the number of indicators is 20, and sample size is 200). All of the methods have negative bias.

3.2.2. Relative percentage bias of interfactor correlations

The results obtained from simulation conditions for interfactor correlations are presented in [Figure 3](#). In addition, for researchers who want to examine further detail, values are given in table in [Appendix 4](#). When the RPB values of the methods for interfactor correlation ([Figure 3](#)) are examined, RPB, obtained from all methods is within acceptable limits under conditions with an average factor loading of 0.7. However, as the sample size decreases under the conditions with an average factor loading of 0.4, the RPB of the φ parameter obtained from the methods may go beyond the limits. The RPB obtained from the Bayesian method was estimated to be less than the required value for both models under conditions where the sample size was 200 and 500, and the average factor loading was 0.4. RPB values of ML, ULSMV and WLSMV methods are within acceptable limits under conditions where sample size is 500 and average factor loading is 0.4. Under the conditions where average factor load was 0.4, sample size was 200, and skewed distribution, the number of items was increased, the RPB values of ML, ULSMV and WLSMV methods increased to acceptable range. Under the conditions where average factor loading was 0.4, sample size was 200, and normal distribution, the RPB values of ML, ULSMV and WLSMV methods were within acceptable limits.

3.3. Coverage Rate

3.3.1. Coverage rate of factor loadings

The coverage rates obtained from the simulation conditions are presented in [Figure 4](#). In addition, the maximum, minimum and average values of the coverage rates obtained from the items are given in [Appendix 5](#) for researchers who want to study the detail. When the coverage rates of the methods are examined according to the simulation conditions, it was observed that the coverage rates of the estimation methods decreased under conditions where the sample size is 200 and the items were skewed. When average factor loading increases to 0.7, the coverage rates of estimation methods increase for a sample size of 200. The coverage rate of ULSMV and WLSMV is below 90% under the conditions where sample size is 200, average factor loading is 0.4 and the items are skewed. It can be said that the Bayesian method performs better than the others under these conditions. Increasing the number of items increased the coverage rate of ML. The coverage rate of the Bayesian method is less than 90% for the conditions where the factor structure is unidimensional, the number of items is 20 and average factor loading is 0.4.

Under the conditions where the sample size is 500, the coverage rate of all methods, except for Bayesian, is over 90% for all models. However, the Bayesian method can fall below 90% in unidimensional structures. With the increase in the average factor loading, the performance of the Bayesian method in unidimensional structures increases. When the simulation conditions where the sample size is 1,000 are examined, the Bayesian method has a lower coverage rate in unidimensional structures than the other structures under conditions where average factor loading is 0.4. The coverage rates of the ML, ULSMV and WLSMV methods are adequate for all models for the conditions where the indicators are normal and distribution skewed. Under conditions with an average factor loading of 0.7, the coverage rates of all other methods, except the ML method, are sufficient. The coverage rate of ML is below 90% for some models under conditions where the number of indicators is 10 or 20. Interestingly, the increase in sample size reduced the coverage rate of ML for these conditions.

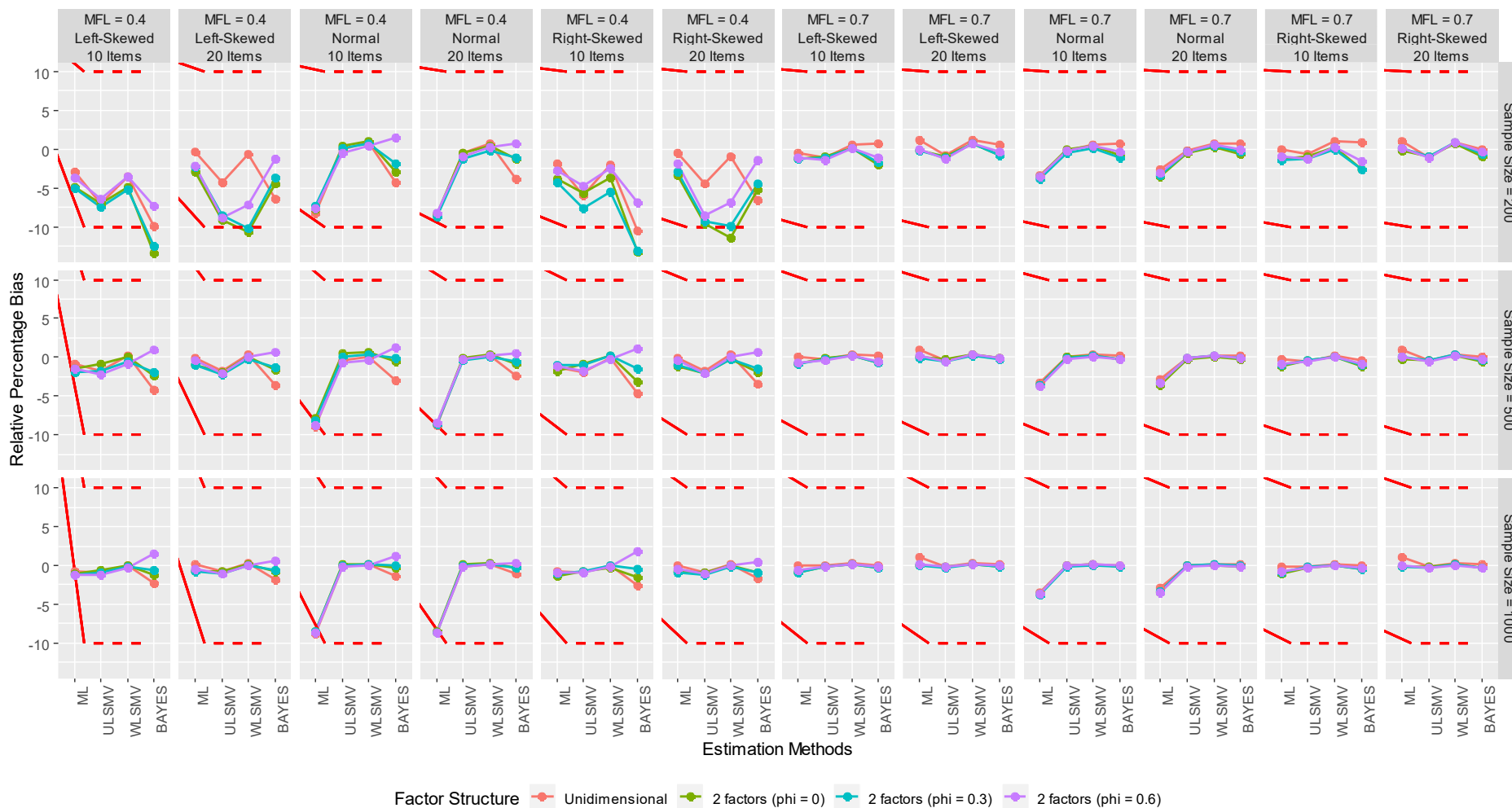


Figure 2. Relative percentage bias (RPB) of factor loadings

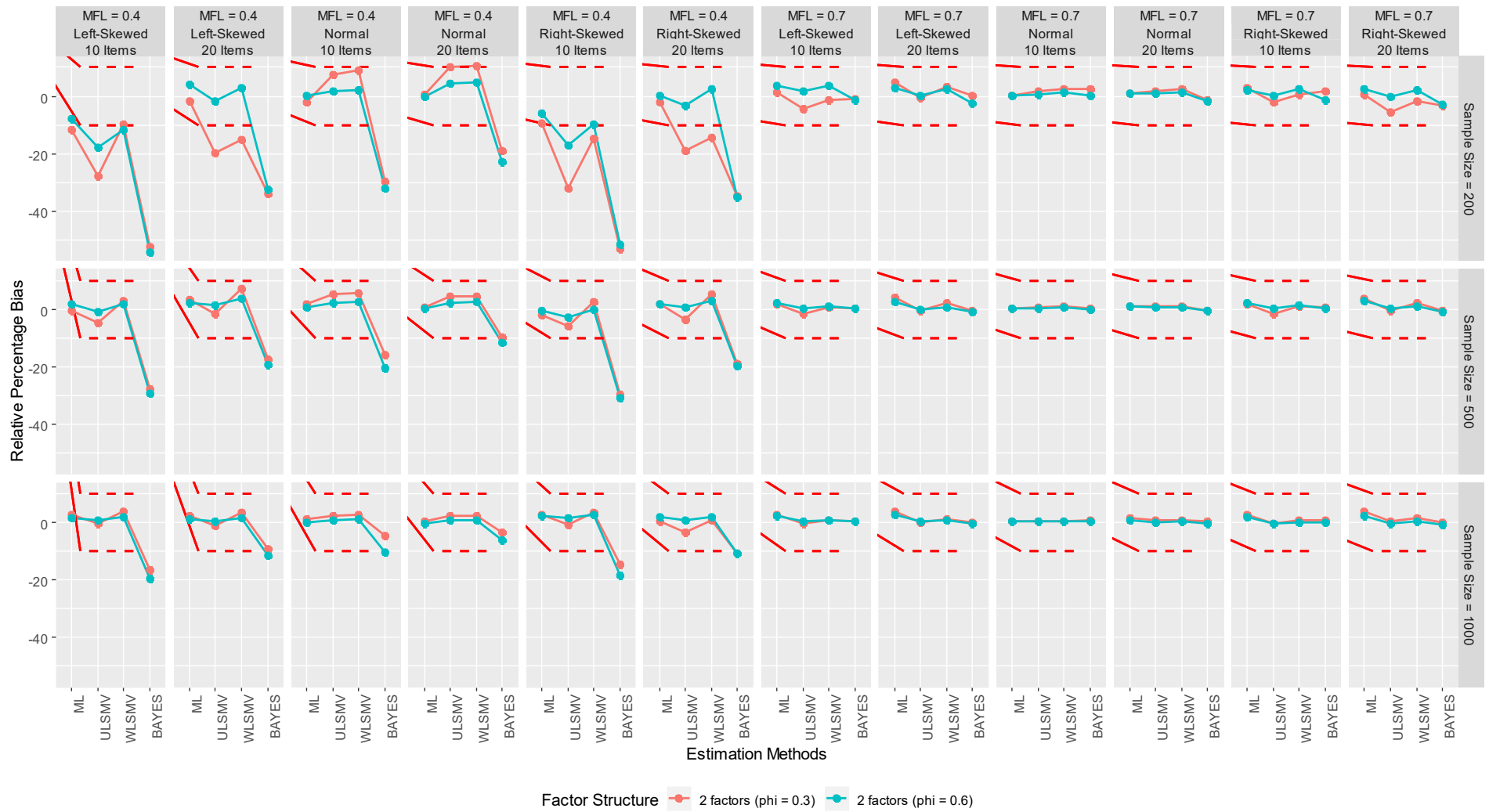


Figure 3. Relative percentage bias (RPB) of interfactor correlations

3.3.2. Coverage rate of interfactor correlations

The coverage rate of the methods for interfactor correlation are presented in [Figure 5](#), and the numerical values are presented in a table in [Appendix 6](#). Coverage rates obtained from all estimation methods are 90% and above, under conditions where the sample size is 1,000 and average factor loading is 0.7. Under the conditions where the sample size is 200 and average factor loading is 0.7, the coverage rates of the methods are above 90% under the conditions where indicators follow normal distribution. However, the performance of the WLSMV method decreased as the ϕ parameter decreases under conditions where the indicators distributions are skewed. ML and Bayesian methods have a coverage rate of over 90% under conditions where the sample size is 200, indicators are skewed and average factor loading is 0.7. The ULSMV method had a coverage rate of over 90% with the increase of the ϕ parameter under conditions where the sample size is 200, indicators were skewed and average factor loading is 0.7, and remained below 90% under conditions where ϕ was 0.

The conditions where indicators follow normal distribution and average factor loading is 0.4, increasing the sample size increased coverage rate of ML, ULSMV and WLSMV. In addition, the coverage rate of the Bayesian method is less than 90% where the ϕ parameter is 0.6. Decreasing the ϕ parameter increased the coverage rate of the Bayesian method. The conditions where indicators are normally distributed, the average factor loading is 0.4 and sample size is 200, increasing the interfactor correlation (ϕ) increased the coverage rate of ULSMV and WLSMV. With the increase in the number of indicators in these conditions, the coverage rate of ML increased but it was not affected by the magnitude of the interfactor correlation.

Under conditions where the average factor loading is 0.4, indicators follow skewed distribution, sample size is 200 and the number of items is 10, the coverage rate of the Bayesian method alone (for $\phi = 0$ and 0.3) is higher than 90%, while when the number of indicators increased to 20, the coverage rate of ML is about 90%. Under all conditions where the average factor loading is 0.4, items have skewed distribution and sample size is 500, the coverage rate of the ULSMV and WLSMV methods is higher than 90% only if the ϕ parameter is 0.6. For these conditions, the coverage rate of ML is higher than 90%. With the increase in the number of items, the coverage rate of ML also increased. While the Bayesian method has a coverage rate of less than 90% under conditions where the ϕ parameter is 0.6, under conditions where the ϕ parameter is 0 or 0.3, the coverage rate of the Bayesian method is higher than 90%.

3.4. Relative Standard Error Bias

3.4.1. Relative standard error of factor loadings

The r-seb values obtained from the simulation conditions are presented in [Figure 6](#). In addition, maximum, minimum values and averages of the r-seb values obtained from the indicators are given in [Appendix 7](#). In all conditions where the sample size is 500 and 1,000, all estimation methods have an acceptable r-seb value for all models. However, the WLSMV method has a large r-seb value in all two-dimensional models, except for unidimensional structures under 20-indicators conditions with a skewed distribution with an average factor loading of 0.4 in 200 sample size.

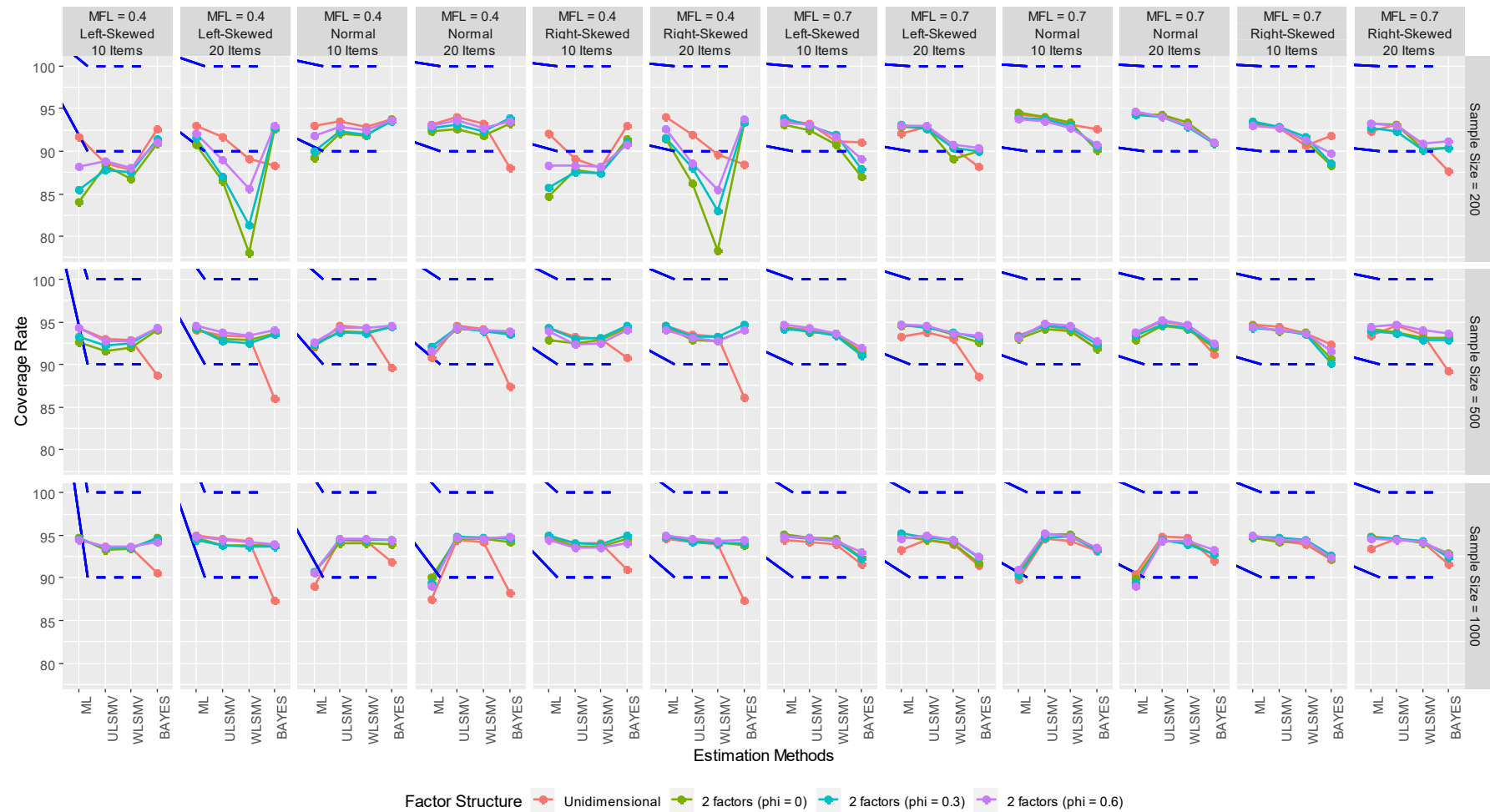


Figure 4. Coverage rate of factor loadings

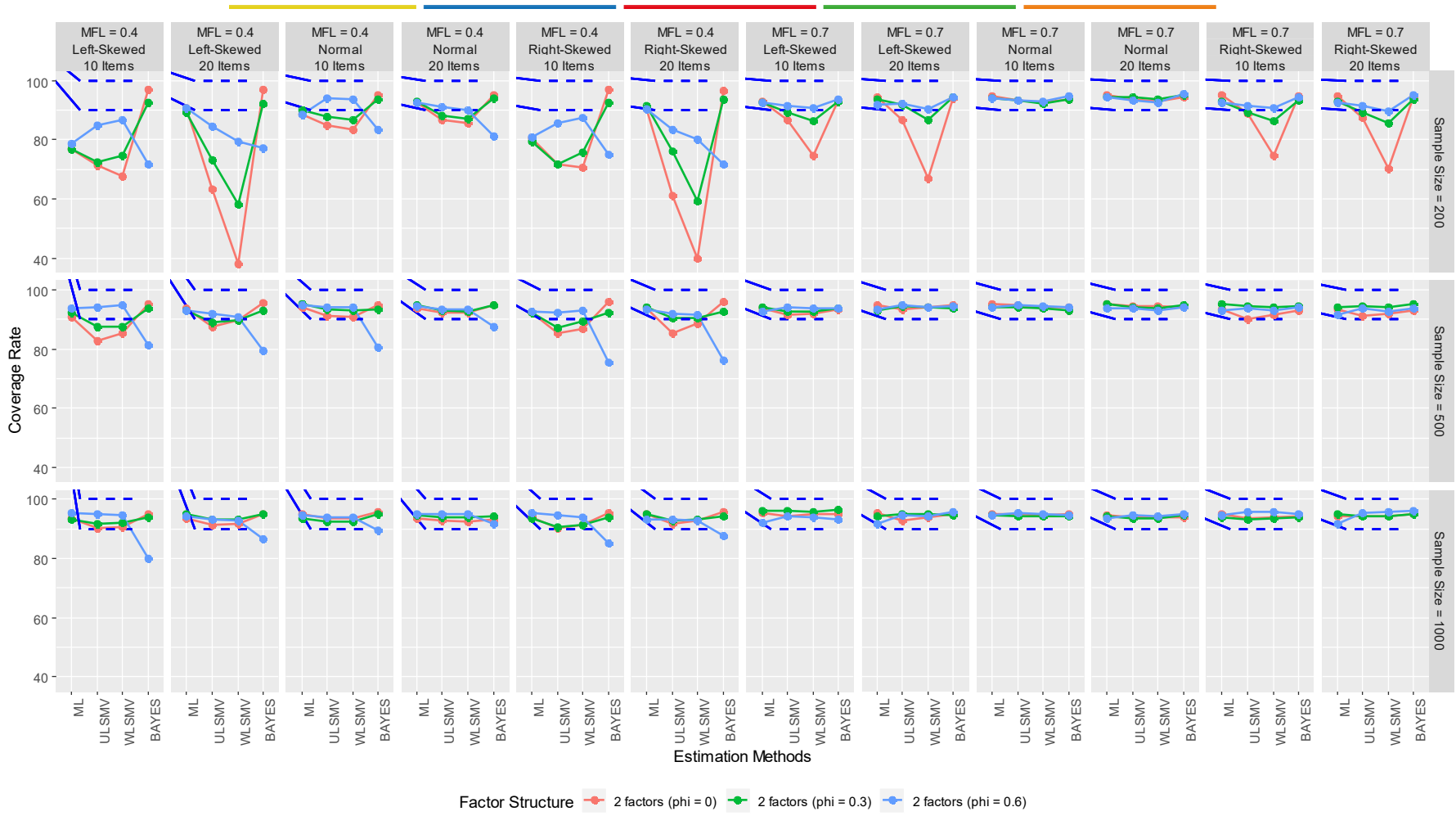


Figure 5. Coverage rate of interfactor correlations

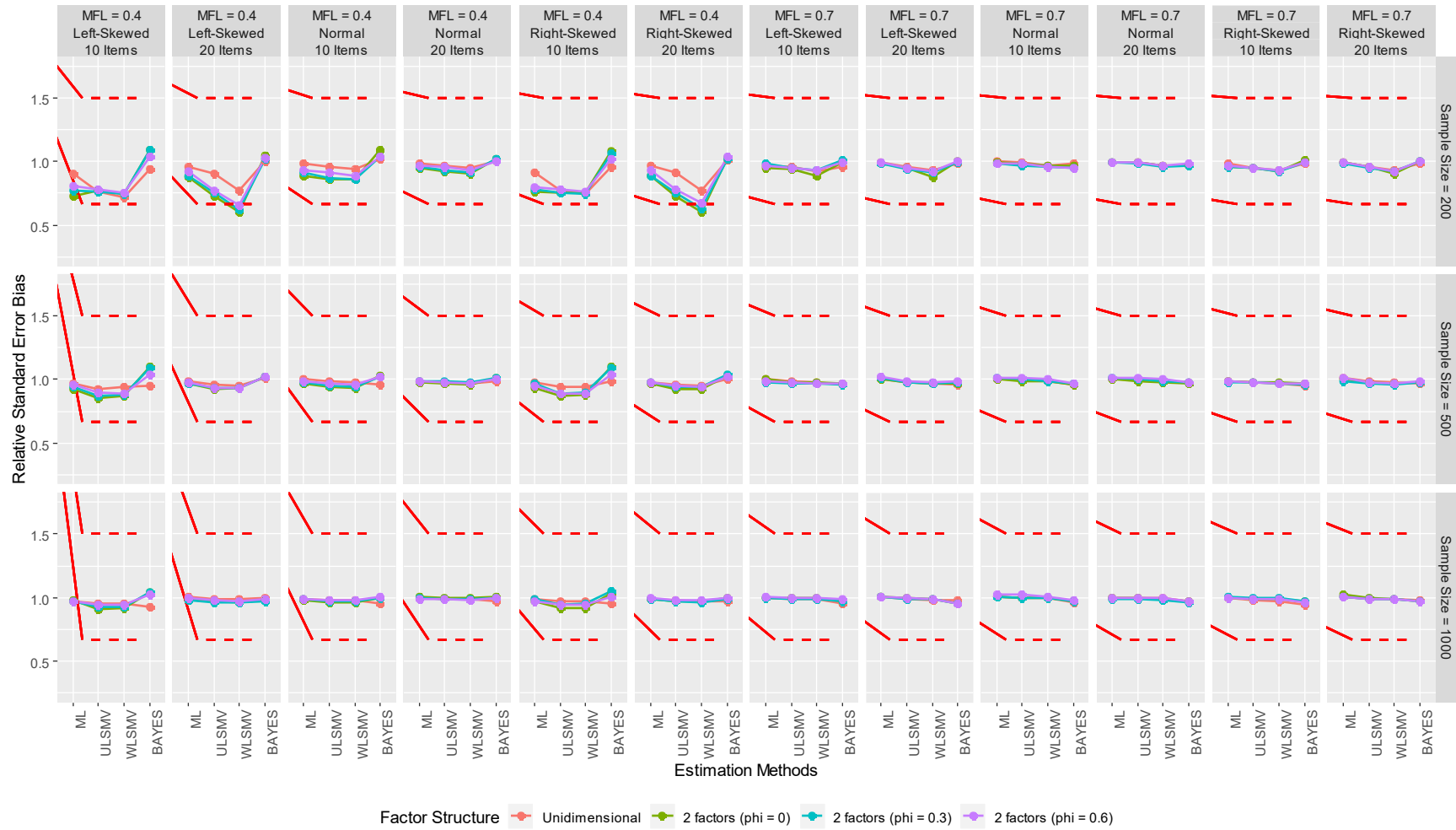


Figure 6. Relative standard error bias (r-seb) of factor loadings

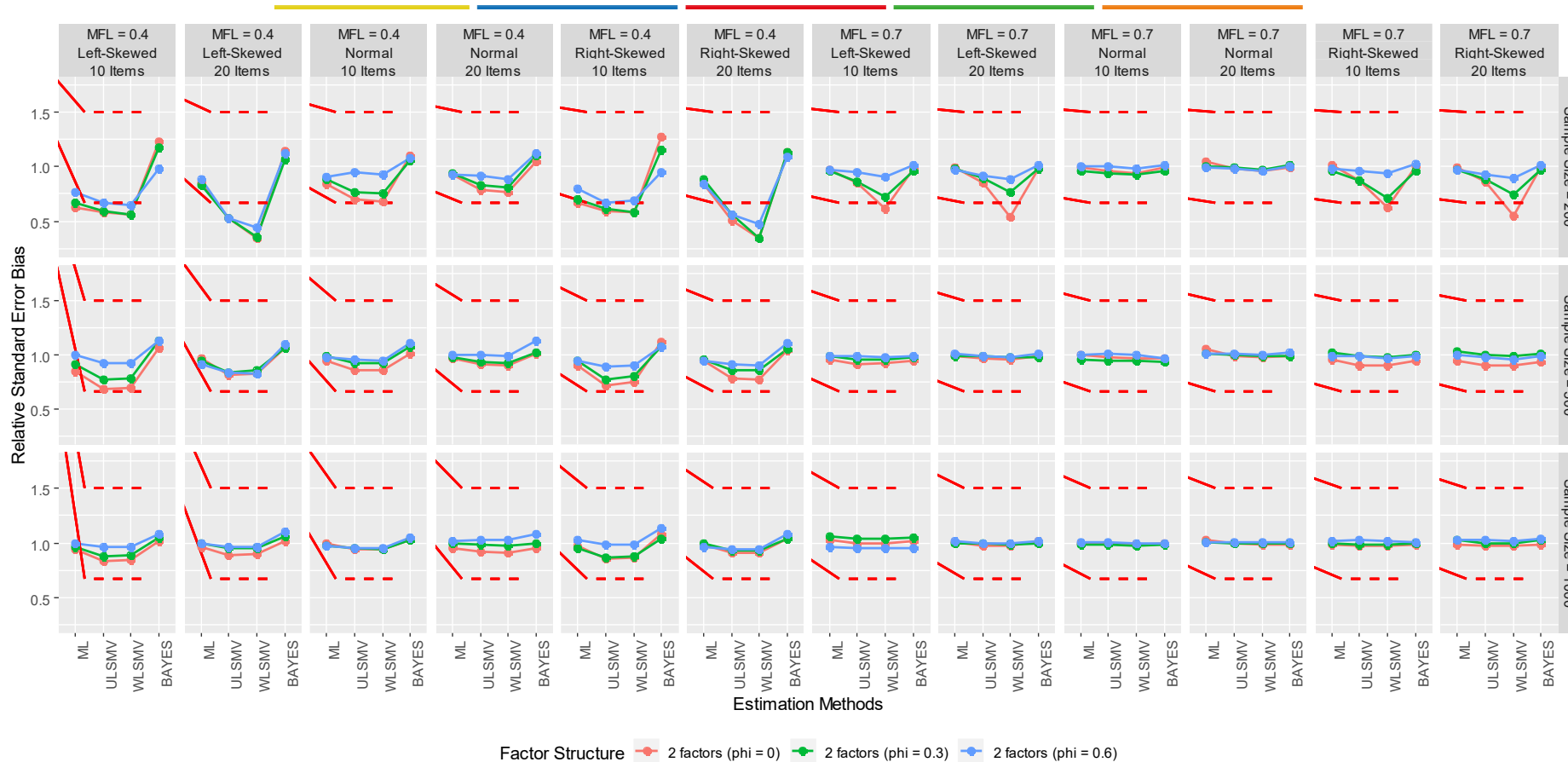


Figure 7. Relative standard error bias (r-seb) of interfactor correlations

3.4.2. Relative standard error of interfactor correlations

The r-seb values of the methods for interfactor correlation are presented in [Figure 7](#) and the numerical values are given in the table in [Appendix 8](#). When [Figure 7](#) is examined, it can be said that the r-seb values of all estimation methods are acceptable under all conditions where sample sizes are 500 and 1,000. However, the r-seb values of the estimation methods are out of range for decreasing the sample size and the indicators became skewed.

The r-seb values of the estimation methods are acceptable under conditions where the sample size is 200 and the average factor loading is 0.7. In cases where the items are skewed, the r-seb values of the WLSMV method are out of the acceptable range when ϕ parameter is 0. The r-seb values of the other methods are acceptable under these conditions.

The r-seb values of the estimation methods are acceptable under conditions where the sample size is 200, the average factor loading is 0.4 and indicators follow normal distribution. In these conditions, except for distribution of indicators, the r-seb values of the Bayesian method are acceptable under conditions where indicators follow skewed distribution. Increasing the number of items under these conditions, the r-seb values of the ML method increased to an acceptable range. The r-seb values of the ULSMV and WLSMV methods are unacceptable for these conditions. In these conditions, ML and Bayesian methods perform better in terms of r-seb values.

4. DISCUSSION

In this study, CFA estimation methods were compared by manipulating sample size, distribution of data, test length, average factor loading, and factor structure for binary data.

4.1. Non-convergence and Improper Solutions

Non-convergence frequently encounters datasets which have a two-factor structure and consist of skewed indicators for the ULSMV and WLSMV estimation methods. These methods have a less converged problem in unidimensional structures than in two-factor structures (even if items are skewed). The increase in the number of items for conditions where items are skewed decreases the non-convergence datasets for ULSMV and WLSMV. It can be said that all estimation methods converge when the average factor loading is 0.7. In other words, the ULSMV and WLSMV estimation methods are mostly non-convergent for small sample size, low average factor loading, short test length, two-dimensional models, and the magnitude of interfactor correlation is small. This result is consistent with the study by Moshagen and Musch (2014). In a study comparing MLR and WLSMV estimation methods conducted by Li (2016a), consisting of 4, 6, 8, and 10 categories indicators with skewness coefficients ranging from 1.01-1.31, it was reported that WLSMV converged under all conditions and there were no improper solutions. It is thought that the differentiation of the number of categories of indicators and the skewness of the indicators may have caused differentiation between the results. Nestler (2013) states that the DWLS (WLSMV) method had 3.7% non-convergence in sample size of 250. The present study is similar to the study conducted by Flora and Curran (2004) and Nestler (2013) in terms of non-convergence.

There are no improper solutions in the Bayesian method, whereas there are a few in the ML method. Liang and Yang (2014) also reported that there were no improper solutions in binary data in a simulation study using non-informative priors. The result obtained in this respect is consistent with the Liang and Yang's (2014) study. The ULSMV and WLSMV methods have more improper solutions under conditions where average factor loading is 0.4 and sample size is 200. There was a decrease in the number of improper solutions under conditions where the average factor load was 0.4, sample size was 200 and the indicators followed a normal

distribution. The number of improper solutions is very close to 0 even if the sample size is small under conditions where the average factor load was 0.7.

When non-convergence and improper solution results are evaluated, it can be said that ML and Bayesian methods perform better than ULSMV and WLSMV methods. It was observed that the number of non-convergence and improper solutions of the methods increased where the average factor loading was low, test length was short and the distribution of indicators were skewed, the factor structure was not unidimensional, and the sample size was small.

4.2. Relative Percentage Bias of Factor Loadings

When the RPB values calculated via factor loadings of the estimation methods were examined, the Bayesian and WLSMV methods may give biased results under conditions where the sample size is 200, average factor loading is 0.4, and items are skewed. The conditions where the sample size is 200, average factor loading is 0.4, and items are skewed show that increasing the number of indicators and number of factors and decreasing the interfactor correlation decreased the WLSMV estimation method's RPB, and decreasing the number of indicators and the interfactor correlation while increasing the number of factors decreased the RPB value of the Bayesian method. This result is similar to the findings obtained by Nalbantoğlu Yılmaz (2019) in continuous data. She stated that the WLS method has larger RPB values for small samples. In addition, this result is consistent with the research conducted by Moshagen and Musch (2014) and Lei (2009). Moshagen and Musch (2014) report that RPB was less than 10% for unidimensional structures, while Lei (2009) states that RPB values of the ML and WLSMV methods were less than 10%. Flora and Curran (2004) report that the RPB values of robust WLS estimation methods did not exceed 10%. But in the present study, the RPB value of WLSMV is more than 10% in conditions where the sample size is 200, the average factor loading is 0.4, the distribution of 20 indicators are left-skewed, and the interfactor correlation is 0 and 0.3. This difference may be due to the fact that the factor loadings are not equal for all indicators in the present study. In addition, Flora and Curran (2004) make y^* (latent continuous variable) skewed. However, in the present study, the latent variable (y^*) was generated to follow a normal distribution. Indicators were skewed and analyses were performed. It is thought that the differentiation of the results may have been due to this differentiation.

4.3. Relative Percentage Bias of Interafactor Correlations

When the RPB values of the ϕ parameter, which is interfactor correlation, are examined, the RPB performance of all methods is sufficient in all conditions with average factor loading of 0.7. However, the RPB performance of the ML, ULSMV and WLSMV methods decreased when the sample size decreased and the skewness of the items increased under conditions where average factor loading is 0.4. Beauducél and Herzberg (2006) state that interfactor correlation size is more effective on the performance of the estimation methods. In this respect, it can be said that the present study is similar to Beauducél and Herzberg's (2006) study. The RPB value of the Bayesian method is higher than 10% in almost all conditions where the average factor loading is 0.4, sample size is 1,000, distribution of indicators normal and the number of indicators is 20. This result is consistent with the findings of Liang and Yang (2014).

4.4. Coverage Rate of Factor Loadings

When the coverage rates calculated via factor loadings of the estimation methods were examined, it was observed that the increase in sample size and average factor loading increased the performance of the estimation methods. It can be said that the coverage rate of the ULSMV and WLSMV methods are not sufficient in the conditions where the distribution of indicators is skewed, sample size is small, and average factor loadings is low. This result is consistent with the findings of the simulation study conducted by Forero et al. (2009). In addition, Koğar and Yılmaz Koğar (2015) stated that ULS and DWLS methods have less standard errors when

compared to the ML method. The difference may have originated from variables that were not examined in current study included in the real data set. In the simulation study conducted by Wolf et al. (2013), it was reported that ML had sufficient coverage rates under all conditions studied. However, in this study, the data were generated as normal and continuous. It can be said that there may be a difference in this respect with the results of the present study. The coverage rate of the Bayesian method is less than 90% for the conditions where the model is unidimensional and average factor loading is 0.4. Önen (2019) states that the coverage rate of the Bayesian method is sufficient for all simulation conditions. In the present study, the difference may have arisen since non-informative priors for Bayesian estimations.

4.5. Coverage Rate of Interfactor Correlations

Coverage rates calculated for interfactor correlations for the WLSMV method remained below 90% under conditions where the correlation between the dimensions was 0 and a small sample size. The performance of the WLSMV method decreased as the ϕ parameter decreased under conditions where the indicators followed a skewed distribution. The ML and Bayesian methods had a coverage rate of over 90% under conditions where sample size is 200, average factor load is 0.7 and items are skewed. Li (2016b) likewise reported that the coverage rate of the MLR method is adequate, and that the WLSMV method may have a coverage rate of less than 90% in skewed distributions. It can be said that the current research findings are consistent with this study. The ULSMV method's coverage rate remained below 90% as interfactor correlation decreases under conditions of sample size of 200 and skewed distribution of items. Under conditions where the average factor loading is 0.4, sample is small and the indicators follow normal distribution, the coverage of the ULSMV and WLSMV methods decreases as interfactor correlation decreases. In the case of indicators with average factor loading of 0.4, the performances of other methods remained below 90%, except for the Bayesian method. However, with the increase in the sample size, the coverage rate of the methods increased.

4.6. Relative Standard Error Bias of Factor Loadings

When r -seb values are examined for factor loadings, it can be said that the methods perform sufficiently in most of the conditions. The WLSMV method went beyond the acceptable limits for r -seb values under conditions where the average factor loading is 0.4, the sample size is 200 and the number of indicators are 20, and is within the acceptable range under other conditions. Other methods are acceptable in all conditions. In the simulation study performed by Xu (2019), the MLR method has sufficient relative bias in normal, mild non-normal and moderate non-normal data. However, the Bayesian method has a relative bias greater than 10% under moderate non-normal conditions. In the present study, the Bayesian method has sufficient r -seb value under all conditions. In the study conducted by Xu (2019), the data was produced as a correlation matrix. In addition, the factor loadings of the indicators were fixed to 0.7. It can be said that the differentiation may have originated from here.

4.7. Relative Standard Error Bias of Interfactor Correlations

When the r -seb values are examined for interfactor correlations, all methods are in the acceptable range in 500 and 1,000 sample sizes, while the WLSMV method in the sample size of 200 may be out of the acceptable range in skewed distributions.

When the results of the research are evaluated in general, it can be said that the increase in the average factor loading and the sample size have a positive effect on the performances of the estimation methods. The increase in the number of indicators did not cause much difference for the indicators which follow normal distribution, but it affected the estimations of the methods for the indicators which followed skewed distribution. According to the research findings, it can be said that the methods are sufficient to estimate the average factor loading and the interfactor correlations, regardless of the estimation method used in most of the conditions

where the average factor loading is 0.7. However, as the average factor loading was 0.4, the number of skewed indicators increased, the sample size decreased and the interfactor correlations decreased, the performance of the methods decreased. Especially in small samples, the interfactor correlation was lower in the case of skewed indicators than indicators which follow normal distribution.

According to the research findings, it can be said that any estimation method can be chosen under conditions where sample size is 500 or 1000 and average factor loading is 0.7. The performance of the estimation methods differs in conditions with a sample size of 200. Therefore, the conditions where sample size is 200, the average factor loading is 0.4, indicators follow normal distribution, and the structure is unidimensional, it is recommended to use ML, ULSMV or WLSMV.

However, if the indicators are skewed, it can be recommended that to use ML or ULSMV estimation method. As the interfactor correlation decreases, the performance of the estimation methods to estimate the interfactor correlations decreases in small samples. Therefore, expanding the sample can be considered in such a case. According to the research findings, there is no method that makes the most accurate estimation under all conditions. However, it can be suggested that to use the ULSMV estimation method because it is observed that it has sufficient performance under more conditions.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author(s).

ORCID

Abdullah Faruk KILIÇ  <https://orcid.org/0000-0003-3129-1763>

İbrahim UYSAL  <https://orcid.org/0000-0002-6767-0362>

Burcu ATAR  <https://orcid.org/0000-0003-3527-686X>

5. REFERENCES

- Acar-Güvendir, M., & Özer-Özkan, Y. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi [The examination of scale development and scale adaptation articles published in Turkish academic journals on education]. *Electronic Journal of Social Sciences*, 14(52), 23–33. <https://doi.org/10.17755/esosder.54872>
- AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anıl, D., Güzeller, C. O., Çokluk, Ö., & Şekercioglu, G. (2010). Level determination exam (SBS-2008) the determination of the validity and reliability of 7th grade mathematics sub-test. *Procedia-Social and Behavioral Sciences*, 2(2), 5292-5298. <https://doi.org/10.1016/j.sbspro.2010.03.863>
- Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, 24(2), 222-228. <https://doi.org/10.2307/3151512>
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(2), 186-203. https://doi.org/10.1207/s15328007sem1302_2
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons, Inc.

- <https://doi.org/10.1002/9781118619179>
- Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in lisrel maximum likelihood estimation. *Psychometrika*, 50(2), 229-242. <https://doi.org/10.1007/BF02294248>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford.
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge.
- Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6(4), 330–351. <https://doi.org/10.1037/1082-989X.6.4.330>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Lawrence Erlbaum Associates.
- Crocker, L., & Algina, J. (2008). *Introduction of classical and modern test theory*. Cengage Learning.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29. <https://doi.org/10.1037/1082-989X.1.1.16>
- DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(3), 425-438. <https://doi.org/10.1080/10705511.2014.915373>
- Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, 47(2), 309-326. <https://doi.org/10.1111/j.2044-8317.1994.tb01039.x>
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286–299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A monte carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 625-641. <https://doi.org/10.1080/10705510903203573>
- Gilbert, N. (1999). Simulation: A new way of doing social science. *American Behavioral Scientist*, 42(10), 1485–1487. <https://doi.org/10.1177/0002764299042010002>
- Gorsuch, R. L. (1974). *Factor analysis*. W. B. Saunders.
- Green, S. B., Akey, T. M., Fleming, K. K., Hershberger, S. L., & Marquis, J. G. (1997). Effect of the number of scale points on chi-square fit indices in confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 4(2), 108-120. <https://doi.org/10.1080/10705519709540064>
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-275. <https://doi.org/10.1037/0033-2909.103.2.265>
- Hallquist, M., & Wiley, J. (2017). *MplusAutomation: Automating Mplus Model Estimation and Interpretation* [Computer software]. <https://cran.r-project.org/web/packages/MplusAutomation/>
- Harrison, R. L. (2010). Introduction to Monte Carlo Simulation. *AIP Conference Proceedings*,

- 1204, 17–21. <https://doi.org/10.1063/1.3295638>
- Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A Comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, 51(5), 661-680. <https://doi.org/10.1080/00273171.2016.1208074>
- Jin, S., Luo, H., & Yang-Wallentin, F. (2016). A simulation study of polychoric instrumental variable estimation in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 680-694. <https://doi.org/10.1080/10705511.2016.1189334>
- Kılıç, A. F., & Kelecioğlu, H. (2016). TEOG ortak ve mazeret sınavındaki Türkçe ve matematik alt testlerinin psikometrik özelliklerinin karşılaştırılması [The comparison of psychometric properties of standardised and make up maths and Turkish subtest questions in TEOG]. *Journal of Measurement and Evaluation in Education and Psychology*, 7(1), 33–58. <https://doi.org/10.21031/epod.14532>
- Kline, R. B. (2016). *Principle and practice of structural equation modeling* (4th ed.). The Guilford.
- Koğar, H., & Yılmaz Koğar, E. (2015). Comparison of different estimation methods for categorical and ordinal data in confirmatory factor analysis. *Journal of Measurement and Evaluation in Education and Psychology*, 6(2), 351-364. <https://doi.org/10.21031/epod.94857>
- Lei, P. (2009). Evaluating estimation methods for ordinal data in structural equation modeling. *Quality and Quantity*, 43(3), 495–507. <https://doi.org/10.1007/s11135-007-9133-z>
- Li, C.-H. (2016a). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Li, C.-H. (2016b). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods*, 21(3), 369–387. <https://doi.org/10.1037/met0000093>
- Liang, X., & Yang, Y. (2014). An evaluation of WLSMV and Bayesian methods for confirmatory factor analysis with categorical indicators. *International Journal of Quantitative Research in Education*, 2(1), 17-38. <https://doi.org/10.1504/IJQRE.2014.060972>
- Lissitz, R. W., Hou, X., & Slater, S. C. (2012). The contribution of constructed response items to large scale assessment: Measuring and understanding their impact. *Journal of Applied Testing Technology*, 13(3), 1-50. <http://www.jattjournal.com/index.php/atp/article/view/48366/39234>
- Morata-Ramirez, M. de los A., & Holgado-Tello, F. P. (2013). Construct validity of likert scales through confirmatory factor analysis: A simulation study comparing different methods of estimation based on Pearson and polychoric correlations. *International Journal of Social Science Studies*, 1(1), 54-61. <https://doi.org/10.11114/ijsss.v1i1.27>
- Moshagen, M., & Musch, J. (2014). Sample size requirements of the robust weighted least squares estimator. *Methodology*, 10(2), 60-70. <https://doi.org/10.1027/1614-2241/a000068>
- Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. Chapman & Hall.
- Muthén, B. O., & Asparouhov, T. (2002). *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus*. <https://www.statmodel.com/download/webnotes/CatMGLong.pdf>
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335. <https://doi.org/10.1037/a0026802>

- Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38(2), 171–189. <https://doi.org/10.1111/j.2044-8317.1985.tb00832.x>
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus statistical modeling software: Release 7.0* [Computer software]. Muthén & Muthén.
- Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (7th ed.). Muthén & Muthén.
- Nalbantoğlu Yılmaz, F. (2019). Comparison of different estimation methods used in confirmatory factor analyses in non-normal data: A monte carlo study. *International Online Journal of Educational Sciences*, 11(4), 131-140. <https://doi.org/10.15345/iojes.2019.04.010>
- Nestler, S. (2013). A monte carlo study comparing PIV, ULS and DWLS in the estimation of dichotomous confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology*, 66(1), 127-143. <https://doi.org/10.1111/j.2044-8317.2012.02044.x>
- Önen, E. (2019). A comparison of frequentist and Bayesian approaches: The power to detect model misspecifications in confirmatory factor analytic models. *Universal Journal of Educational Research*, 7(2), 494–514. <https://doi.org/10.13189/ujer.2019.070223>
- Potthast, M. J. (1993). Confirmatory factor analysis of ordered categorical variables with large models. *British Journal of Mathematical and Statistical Psychology*, 46(2), 273–286. <https://doi.org/10.1111/j.2044-8317.1993.tb01016.x>
- R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates.
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354–373. <https://doi.org/10.1037/a0029315>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Şahin, M. G., & Boztunç Öztürk, N. (2018). Eğitim alanında ölçek geliştirme süreci: Bir içerik analizi çalışması [Scale development process in educational field: A content analysis research]. *Kastamonu Education Journal*, 26(1), 191-199. <https://doi.org/10.24106/kefedergi.375863>
- Shi, D., DiStefano, C., McDaniel, H. L., & Jiang, Z. (2018). Examining chi-square test statistics under conditions of large model size and ordinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(6), 924-945. <https://doi.org/10.1080/10705511.2018.1449653>
- Stevens, J. P. (2009). *Applied multivariate statistics for the social science* (5th ed.). Taylor & Francis.
- Streiner, D. L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry*, 39(3), 135-140. <https://doi.org/10.1177/070674379403900303>
- Tabachnik, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Pearson.
- Thissen, D., Wainer, H., & Wang, X.-B. (1994). Are tests comprising both multiple-choice and free-response items necessarily less unidimensional than multiple-choice tests? An analysis of two tests. *Journal of Educational Measurement*, 31(2), 113–123. <https://doi.org/10.1111/j.1745-3984.1994.tb00437.x>
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety.

National Institutes of Health, 76(6), 913-934. <https://doi.org/10.1177/0013164413495237>

Xu, M. (2019). *A comparison of frequentist and Bayesian approaches for confirmatory factor analysis* (Publication No. 27534819) [Doctoral dissertation, The Ohio State University]. ProQuest Dissertations and Theses Global.

Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(3), 392–423. <https://doi.org/10.1080/10705511.2010.489003>

6. APPENDIX

Appendix 1. Number of datasets having convergence failure

Sample Size	Model	Method	Mean Factor Loading = 0.4						Mean Factor Loading = 0.7					
			Left-Skewed		Normal		Right-Skewed		Left-Skewed		Normal		Right-Skewed	
			10	20	10	20	10	20	10	20	10	20	10	20
200	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-
200		ULSMV	41	32	2	3	50	26	-	-	-	-	-	-
200		WLSMV	29	41	2	3	49	30	1	-	-	-	-	-
200		BAYES	1	-	-	-	2	-	-	-	-	-	-	-
200		2 factors (φ = 0)	ML	2	-	-	-	2	-	-	-	-	-	-
200			ULSMV	299	101	81	11	290	120	-	-	-	-	-
200			WLSMV	282	157	85	11	306	158	-	-	-	1	-
200		2 factors (φ = 0.3)	BAYES	3	2	1	-	-	3	-	-	-	-	-
200			ML	2	-	-	-	4	-	-	-	-	-	-
200	ULSMV		256	101	48	6	282	110	-	-	-	-	-	
200	2 factors (φ = 0.6)	WLSMV	240	157	49	6	270	160	1	-	-	-	1	
200		BAYES	-	-	-	-	1	-	-	-	-	-	-	
200		ML	4	-	-	-	-	-	-	-	-	-	-	
200	2 factors (φ = 0.6)	ULSMV	211	1-3	27	5	186	87	-	-	-	-	-	
200		WLSMV	193	1-2	27	5	165	109	-	-	-	-	1	
200		BAYES	1	2	-	-	-	2	-	-	-	-	-	
500	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	
500		ULSMV	4	-	-	-	-	-	-	-	-	-	-	
500		WLSMV	5	-	-	-	-	-	-	-	-	-	-	
500		BAYES	-	-	-	-	-	-	-	-	-	-	-	
500		2 factors (φ = 0)	ML	1	-	-	-	-	-	-	-	-	-	
500			ULSMV	65	5	1	-	70	1	-	-	-	-	
500			WLSMV	54	4	1	-	57	1	-	-	-	-	
500		2 factors (φ = 0.3)	BAYES	-	-	-	-	-	-	-	-	-	-	
500			ML	-	-	-	-	-	-	-	-	-	-	
500	ULSMV		43	6	-	-	34	2	-	-	-	-		
500	2 factors (φ = 0.6)	WLSMV	30	5	-	-	18	1	-	-	-	-		
500		BAYES	-	-	-	-	-	-	-	-	-	-		
500		ML	-	-	-	-	-	-	-	-	-	-		
500	2 factors (φ = 0.6)	ULSMV	17	1	-	-	16	5	-	-	-	-		
500		WLSMV	13	1	-	-	14	4	-	-	-	-		
500		BAYES	-	-	-	-	1	-	-	-	-	-		
1000	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-		
1000		ULSMV	-	-	-	-	-	-	-	-	-	-		
1000		WLSMV	-	-	-	-	-	-	-	-	-	-		
1000		BAYES	-	-	-	-	-	-	-	-	-	-		
1000		2 factors (φ = 0)	ML	-	-	-	-	-	-	-	-	-		
1000			ULSMV	3	-	-	-	2	-	-	-	-		
1000			WLSMV	3	-	-	-	2	-	-	-	-		
1000		2 factors (φ = 0.3)	BAYES	-	-	-	-	-	-	-	-	-		
1000			ML	-	-	-	-	-	-	-	-	-		
1000	ULSMV		2	-	-	-	-	-	-	-	-			
1000	2 factors (φ = 0.6)	WLSMV	2	-	-	-	-	-	-	-	-			
1000		BAYES	-	-	-	-	-	-	-	-	-			
1000		ML	-	-	-	-	-	-	-	-	-			
1000	2 factors (φ = 0.6)	ULSMV	-	-	-	-	-	-	-	-	-			
1000		WLSMV	-	-	-	-	-	-	-	-	-			
1000		BAYES	0	0	0	0	0	0	0	0	0	0		

Appendix 2. Number of datasets having inadmissible solution

Sample Size	Model	Method	Mean Factor Loading = 0.4						Mean Factor Loading = 0.7						
			Left-Skewed		Normal		Right-Skewed		Left-Skewed		Normal		Right-Skewed		
			10	20	10	20	10	20	10	20	10	20	10	20	
200	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-	
200		ULSMV	9	-	-	-	11	-	-	-	-	-	-	-	
200		WLSMV	8	-	-	-	11	-	-	-	-	-	-	-	
200	2 factors ($\varphi=0$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
200		ML	37	1	-	-	28	-	2	-	-	-	-	-	
200		ULSMV	219	24	57	-	205	30	21	-	-	-	26	1	
200	2 factors ($\varphi=0.3$)	WLSMV	239	96	55	-	199	111	53	8	-	-	80	15	
200		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
200		ML	25	-	1	-	27	1	1	-	-	-	1	-	
200	2 factors ($\varphi=0.6$)	ULSMV	179	20	35	1	185	34	19	1	-	-	25	1	
200		WLSMV	169	83	34	1	176	83	17	-	-	-	27	-	
200		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
200	Unidimensional	ML	19	1	-	-	10	-	-	-	-	-	1	-	
200		ULSMV	118	5	4	-	129	4	6	3	-	-	12	2	
200		WLSMV	88	29	4	-	105	23	3	3	-	-	10	2	
200	2 factors ($\varphi=0.3$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
500		Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-
500			ULSMV	-	-	-	-	-	-	-	-	-	-	-	-
500	WLSMV		-	-	-	-	-	-	-	-	-	-	-	-	
500	2 factors ($\varphi=0$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
500		ML	3	-	-	-	1	-	-	-	-	-	-	-	
500		ULSMV	52	-	-	-	35	-	-	-	-	-	-	-	
500	2 factors ($\varphi=0.3$)	WLSMV	50	1	-	-	38	4	-	-	-	-	1	-	
500		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
500		ML	1	-	-	-	1	-	-	-	-	-	-	-	
500	2 factors ($\varphi=0.6$)	ULSMV	26	-	1	-	22	-	-	-	-	-	1	-	
500		WLSMV	22	-	1	-	21	-	-	-	-	-	-	-	
500		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
500	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-	
500		ULSMV	3	-	-	-	5	-	-	-	-	-	-	-	
500		WLSMV	-	-	-	-	4	-	-	-	-	-	-	-	
500	2 factors ($\varphi=0.3$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
1000		Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-
1000			ULSMV	-	-	-	-	-	-	-	-	-	-	-	-
1000	WLSMV		-	-	-	-	-	-	-	-	-	-	-	-	
1000	2 factors ($\varphi=0$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
1000		ML	-	-	-	-	1	-	-	-	-	-	-	-	
1000		ULSMV	-	-	-	-	2	-	-	-	-	-	-	-	
1000	2 factors ($\varphi=0.3$)	WLSMV	-	-	-	-	2	-	-	-	-	-	-	-	
1000		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
1000		ML	-	-	-	-	-	-	-	-	-	-	-	-	
1000	2 factors ($\varphi=0.6$)	ULSMV	1	-	-	-	-	-	-	-	-	-	-	-	
1000		WLSMV	1	-	-	-	-	-	-	-	-	-	-	-	
1000		BAYES	-	-	-	-	-	-	-	-	-	-	-	-	
1000	Unidimensional	ML	-	-	-	-	-	-	-	-	-	-	-	-	
1000		ULSMV	-	-	-	-	-	-	-	-	-	-	-	-	
1000		WLSMV	-	-	-	-	-	-	-	-	-	-	-	-	
1000	2 factors ($\varphi=0.6$)	BAYES	-	-	-	-	-	-	-	-	-	-	-	-	

Appendix 3. Mean, maximum and minimum values of relative percentage bias

Sample Size	Model	MFL	Method	Number of Items = 10									Number of Items = 20								
				Left-Skewed			Normal			Right-Skewed			Left-Skewed			Normal			Right-Skewed		
				Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
200	Unidimensional	0.4	ML	-3.0	0.7	-5.3	-8.2	-5.3	-10.5	-1.8	0.8	-3.9	-0.3	1.4	-2.7	-8.5	-6.7	-10.2	-0.5	1.7	-1.8
200		0.4	ULSMV	-7.0	2.8	-10.8	-0.1	2.6	-2.0	-5.9	1.9	-8.8	-4.3	-2.4	-6.3	-0.5	1.3	-2.3	-4.4	-1.7	-6.8
200		0.4	WLSMV	-3.6	7.8	-6.3	0.9	3.6	-1.1	-2.0	6.7	-5.1	-0.6	3.8	-3.4	0.7	2.6	-0.9	-0.9	5.2	-2.8
200	Unidimensional	0.4	BAYES	-10.0	-5.4	-17.1	-4.3	0.3	-14.1	-10.5	-5.3	-14.6	-6.4	-0.7	-28.9	-3.8	0.5	-28.1	-6.6	-1.9	-28.7
200		0.7	ML	-0.4	0.1	-0.8	-3.3	-2.4	-4.0	0.0	0.5	-0.3	1.2	1.9	0.8	-2.6	-1.8	-3.3	1.0	1.5	0.4
200		0.7	ULSMV	-1.1	-0.6	-1.6	-0.2	0.4	-0.6	-0.7	-0.0	-1.1	-0.8	-0.0	-1.4	-0.1	0.3	-0.7	-1.1	-0.7	-1.6
200	Unidimensional	0.7	WLSMV	0.5	1.2	-0.1	0.6	1.1	0.1	1.0	1.5	0.7	1.1	1.7	0.6	0.7	1.2	0.2	0.9	1.4	0.5
200		0.7	BAYES	0.7	2.4	-1.8	0.7	1.6	-1.4	0.8	2.4	-3.5	0.6	2.5	-16.6	0.7	2.1	-9.3	-0.0	2.1	-17.5
200		0.4	ML	-4.9	1.9	-9.1	-7.7	-4.7	-10.7	-3.9	-0.9	-7.2	-3.0	-0.8	-5.8	-8.3	-6.2	-9.9	-3.4	-0.1	-7.1
200	2 factors (p = 0)	0.4	ULSMV	-6.9	5.1	-17.1	0.4	5.0	-2.3	-5.7	5.8	-14.4	-9.1	0.3	-13.4	-0.5	1.7	-2.1	-9.7	0.8	-14.0
200		0.4	WLSMV	-4.9	6.1	-13.4	1.0	5.9	-2.2	-3.6	6.6	-11.7	-10.7	1.5	-15.1	0.5	3.0	-1.1	-11.4	4.0	-16.9
200		0.4	BAYES	-13.4	6.6	-23.9	-2.9	8.4	-9.2	-13.3	8.0	-22.0	-4.4	1.1	-8.3	-1.2	2.5	-7.6	-5.2	0.5	-9.4
200	2 factors (p = 0)	0.7	ML	-1.2	-0.3	-2.2	-3.5	-3.2	-3.9	-1.2	-0.8	-1.7	-0.1	0.6	-1.0	-3.5	-2.4	-4.3	-0.1	0.7	-0.7
200		0.7	ULSMV	-0.9	0.1	-1.8	-0.0	0.6	-0.4	-0.9	-0.2	-1.6	-1.0	-0.1	-2.1	-0.4	0.1	-1.0	-0.9	-0.1	-1.6
200		0.7	WLSMV	0.1	1.2	-0.7	0.4	1.1	0.0	0.3	1.1	-0.3	0.7	1.4	-0.2	0.3	0.8	-0.3	0.7	1.9	0.1
200	2 factors (p = 0.3)	0.7	BAYES	-2.0	2.9	-16.9	-0.8	2.5	-9.7	-2.6	1.8	-16.1	-0.8	2.6	-18.5	-0.6	1.1	-12.1	-0.9	3.4	-17.2
200		0.4	ML	-5.1	0.5	-10.1	-7.2	-4.2	-9.9	-4.3	1.2	-7.5	-2.3	-0.4	-5.1	-8.7	-7.0	-10.1	-3.0	-1.4	-6.5
200		0.4	ULSMV	-7.5	5.7	-15.8	0.1	2.3	-3.4	-7.6	6.0	-16.1	-8.5	-0.3	-11.7	-1.2	0.1	-3.1	-9.3	1.5	-13.4
200	2 factors (p = 0.3)	0.4	WLSMV	-5.1	11.6	-11.4	0.7	3.0	-2.9	-5.5	8.1	-16.5	-10.1	4.8	-15.8	-0.1	1.3	-2.1	-9.9	5.8	-15.5
200		0.4	BAYES	-12.5	7.5	-23.0	-1.8	8.2	-8.5	-13.2	8.7	-21.2	-3.7	1.0	-8.2	-1.2	2.1	-6.2	-4.5	2.0	-8.6
200		0.7	ML	-1.2	-0.5	-1.8	-3.8	-3.0	-4.5	-1.4	-0.8	-1.9	-0.2	0.4	-0.9	-3.3	-2.8	-3.7	0.1	0.6	-0.9
200	2 factors (p = 0.6)	0.7	ULSMV	-1.1	-0.5	-1.7	-0.5	0.3	-1.0	-1.3	-0.6	-1.8	-1.1	-0.3	-1.8	-0.3	0.2	-0.9	-0.9	-0.3	-1.9
200		0.7	WLSMV	0.1	0.7	-0.4	0.0	0.7	-0.3	-0.0	0.5	-0.5	0.7	1.5	0.0	0.5	1.1	-0.1	0.9	1.5	-0.1
200		0.7	BAYES	-1.8	2.9	-15.6	-1.0	1.9	-8.5	-2.6	1.3	-14.9	-0.8	2.1	-17.6	-0.4	1.5	-11.8	-0.6	2.6	-16.6
200	2 factors (p = 0.6)	0.4	ML	-3.7	0.7	-7.2	-7.7	-5.9	-9.2	-2.7	-0.8	-5.2	-2.1	0.6	-4.3	-8.3	-6.7	-11.5	-1.8	-0.2	-5.1
200		0.4	ULSMV	-6.4	2.5	-11.6	-0.4	0.7	-2.1	-4.7	4.7	-9.7	-8.9	-1.9	-12.4	-0.9	0.7	-3.4	-8.6	-1.1	-12.2
200		0.4	WLSMV	-3.6	4.4	-7.3	0.3	1.4	-1.2	-2.5	7.6	-9.4	-7.1	3.1	-11.7	0.3	2.0	-2.1	-6.9	3.9	-10.6
200	2 factors (p = 0.6)	0.4	BAYES	-7.3	9.9	-16.4	1.5	10.2	-6.3	-6.8	11.3	-14.2	-1.3	4.0	-5.4	0.8	4.8	-2.6	-1.4	4.3	-6.9
200		0.7	ML	-1.0	-0.5	-1.5	-3.5	-3.0	-4.4	-0.9	-0.1	-1.9	0.0	0.5	-0.4	-3.1	-2.4	-4.2	0.1	0.9	-0.6
200		0.7	ULSMV	-1.4	-0.7	-2.1	-0.2	0.1	-0.6	-1.3	-0.1	-2.4	-1.3	-0.7	-2.0	-0.3	0.3	-1.0	-1.1	-0.3	-1.9
200	2 factors (p = 0.6)	0.7	WLSMV	0.1	0.7	-0.5	0.4	0.8	0.0	0.3	1.4	-0.7	0.7	1.2	0.0	0.6	1.2	-0.1	0.8	1.6	0.1
200		0.7	BAYES	-1.0	3.1	-12.4	-0.4	2.1	-7.4	-1.6	2.0	-12.4	-0.4	2.4	-15.5	-0.1	1.5	-10.3	-0.3	2.9	-14.7
500		Unidimensional	0.4	ML	-0.9	1.5	-2.2	-8.9	-7.5	-9.8	-1.2	-0.3	-2.6	-0.1	1.3	-1.2	-8.7	-7.6	-9.7	-0.1	1.0
500	0.4		ULSMV	-1.7	-0.2	-2.9	-0.4	0.4	-1.4	-1.9	-0.5	-3.7	-1.7	-0.2	-3.0	-0.3	0.2	-1.2	-1.8	-0.6	-2.8
500	0.4		WLSMV	0.3	2.4	-1.1	0.0	0.8	-1.0	-0.1	1.4	-1.5	0.4	1.7	-0.8	0.2	0.8	-0.7	0.3	1.4	-0.8
500	Unidimensional	0.4	BAYES	-4.3	0.4	-24.6	-2.9	0.9	-21.6	-4.6	-0.6	-22.0	-3.6	1.4	-33.7	-2.4	1.1	-27.9	-3.4	2.1	-33.5
500		0.7	ML	0.0	0.5	-0.4	-3.3	-2.7	-3.8	-0.3	-0.0	-0.5	1.0	1.5	0.6	-2.8	-2.4	-3.3	1.0	1.4	0.6
500		0.7	ULSMV	-0.3	0.1	-0.7	0.0	0.4	-0.2	-0.5	-0.3	-0.7	-0.4	-0.0	-0.8	-0.1	0.1	-0.4	-0.4	-0.0	-0.8
500	Unidimensional	0.7	WLSMV	0.4	0.8	-0.0	0.3	0.7	0.1	0.1	0.4	-0.1	0.4	0.7	-0.1	0.2	0.5	-0.1	0.4	0.8	-0.0
500		0.7	BAYES	0.3	1.7	-5.0	0.2	1.1	-4.0	-0.4	1.4	-5.8	-0.1	1.9	-12.4	0.1	1.1	-7.4	0.1	1.6	-11.6
500		0.4	ML	-1.5	1.1	-3.9	-7.8	-6.3	-9.3	-1.8	-0.2	-4.5	-1.0	0.3	-2.4	-8.6	-7.1	-9.7	-1.2	0.1	-2.6
500	2 factors (p = 0)	0.4	ULSMV	-0.9	1.0	-3.4	0.4	1.4	-0.5	-0.8	1.6	-2.4	-1.9	-0.8	-3.6	0.1	1.1	-1.1	-2.1	-0.8	-3.3
500		0.4	WLSMV	0.1	1.9	-2.6	0.7	1.6	-0.2	0.2	2.4	-1.7	0.0	1.4	-1.4	0.3	1.5	-0.6	-0.2	1.1	-1.7
500		0.4	BAYES	-2.4	3.4	-5.8	-0.5	2.0	-2.8	-3.1	0.7	-5.1	-1.6	0.8	-5.6	-0.8	1.5	-8.1	-2.0	1.7	-5.7
500	2 factors (p = 0)	0.7	ML	-0.8	-0.2	-1.4	-3.6	-3.3	-4.0	-1.1	-0.3	-1.6	0.0	0.4	-0.4	-3.5	-2.7	-4.1	-0.2	0.2	-0.5
500		0.7	ULSMV	-0.2	0.5	-0.7	0.0	0.4	-0.2	-0.4	0.3	-0.8	-0.3	0.2	-0.7	-0.2	0.1	-0.5	-0.5	-0.1	-0.9
500		0.7	WLSMV	0.3	0.9	-0.2	0.2	0.7	-0.0	0.0	0.8	-0.5	0.4	0.9	0.0	0.1	0.4	-0.2	0.2	0.6	-0.2
500	2 factors (p = 0.3)	0.7	BAYES	-0.7	1.5	-8.3	-0.3	1.8	-5.8	-1.2	1.8	-9.1	-0.2	1.6	-8.7	-0.3	1.0	-6.1	-0.6	1.3	-9.9
500		0.4	ML	-1.9	0.2	-5.4	-8.1	-6.5	-10.0	-1.0	1.0	-2.6	-1.0	1.6	-3.3	-8.6	-7.7	-11.1	-1.1	0.2	-2.9
500		0.4	ULSMV	-1.8	-0.2	-5.2	0.1	1.5	-1.4	-1.0	0.6	-2.3	-2.1	0.2	-4.8	-0.3	0.6	-2.4	-2.1	-0.8	-5.0
500	2 factors (p = 0.3)	0.4	WLSMV	-0.5	1.2	-4.0	0.4	1.8	-1.1	0.1	2.0	-1.8	-0.2	2.3	-2.7	0.1	1.0	-1.9	-0.2	1.1	-2.5
500		0.4	BAYES	-1.9	4.8	-7.7	-0.0	2.1	-2.2	-1.4	2.8	-5.6	-1.2	1.8	-4.1	-0.6	2.1	-6.5	-1.4	1.7	-4.6
500		0.7	ML	-0.9	-0.4	-1.4	-3.6	-3.0	-4.3	-1.0	-0.4	-1.6	-0.1	0.4	-0.5	-3.3	-2.5	-3.9	0.0	0.4	-0.3
500	2 factors (p = 0.6)	0.7	ULSMV	-0.3	0.2	-0.9	-0.0	0.2	-0.4	-0.4	-0.1	-1.1	-0.5	-0.1	-0.9	-0.1	0.2	-0.6	-0.3	0.1	-0.8
500		0.7	WLSMV	0.2	0.7	-0.3	0.2	0.4	-0.2	0.1	0.4	-0.5	0.2	0.6	-0.1	0.2	0.5	-0.3	0.4	0.8	-0.0
500		0.7	BAYES	-0.8	1.9	-7.7	-0.2	1.3	-5.5	-1.0	2.1	-9.0	-0.3	1.4	-8.6	-0.1	0.9	-6.0	-0.4	2.0	-9.6
500	2 factors (p = 0.6)	0.4	ML	-1.5	0.4	-3.6	-8.8	-7.5	-10.5	-1.2	-0.2	-2.4	-0.3	1.9	-1.5	-8.5	-6.8	-9.8	-0.5	1.4	-2.1
500		0.4	ULSMV	-2.3	-0.0	-5.1	-0.7	0.1	-1.6	-1.8	-0.4	-2.8	-2.0	0.6	-3.3	-0.3	0.5	-1.2	-2.0	-1.0	-4.3
500		0.4	WLSMV	-0.8	0.6	-3.3	-0.3	0.4	-1.2	-0.2	0.6	-1.5	0.1	2.4	-1.1	0.2	1.0	-0.6	0.1	1.6	-1.8

Sample Size	Model	MFL	Method	Number of Items = 10												Number of Items = 20							
				Left-Skewed			Normal			Right-Skewed			Left-Skewed			Normal			Right-Skewed				
				Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min		
1000	0.7	ML	-0.9	-0.8	-1.0	-3.7	-3.1	-4.1	-1.0	-0.7	-1.2	-0.1	0.3	-0.3	-3.4	-2.8	-3.9	-0.1	0.2	-0.6			
1000	0.7	ULSMV	-0.1	0.0	-0.2	-0.0	0.2	-0.2	-0.2	0.1	-0.5	-0.1	0.2	-0.4	-0.1	0.1	-0.3	-0.2	0.1	-0.6			
1000	0.7	WLSMV	0.1	0.3	-0.1	0.1	0.3	-0.1	-0.0	0.3	-0.3	0.2	0.5	-0.0	0.1	0.2	-0.1	0.1	0.4	-0.3			
1000	0.7	BAYES	-0.3	1.8	-5.3	-0.1	0.9	-2.5	-0.5	1.1	-5.1	-0.1	1.0	-3.1	-0.1	0.4	-2.6	-0.3	1.2	-3.0			
1000	0.4	ML	-1.3	-0.1	-2.4	-8.6	-7.3	-9.6	-1.1	-0.4	-2.8	-0.7	1.2	-1.6	-8.6	-7.3	-9.9	-0.9	0.0	-1.8			
1000	0.4	ULSMV	-0.9	0.2	-1.8	-0.0	1.2	-0.9	-0.7	0.4	-2.0	-1.0	0.8	-2.4	-0.1	0.5	-0.9	-1.2	0.2	-2.4			
1000	0.4	WLSMV	-0.2	0.6	-0.9	0.1	1.4	-0.8	-0.0	1.0	-1.3	-0.0	1.9	-1.1	0.2	0.8	-0.7	-0.2	0.8	-1.3			
1000	0.4	BAYES	-0.6	5.0	-6.6	0.0	2.7	-2.9	-0.4	2.8	-2.3	-0.6	1.8	-3.9	-0.3	1.4	-3.4	-0.9	2.0	-3.2			
1000	0.7	ML	-0.9	-0.6	-1.2	-3.7	-3.2	-4.3	-0.9	-0.4	-1.1	-0.0	0.4	-0.3	-3.4	-2.6	-4.0	-0.1	0.2	-0.3			
1000	0.7	ULSMV	-0.2	0.0	-0.5	-0.1	0.1	-0.4	-0.2	0.3	-0.6	-0.2	0.3	-0.6	-0.0	0.2	-0.2	-0.2	0.0	-0.6			
1000	0.7	WLSMV	0.1	0.4	-0.2	-0.0	0.2	-0.3	0.1	0.6	-0.3	0.1	0.7	-0.1	0.1	0.3	-0.0	0.1	0.4	-0.1			
1000	0.7	BAYES	-0.3	1.7	-5.1	-0.2	0.9	-2.8	-0.4	1.2	-5.2	-0.1	1.2	-3.3	-0.0	0.4	-2.5	-0.3	1.1	-3.4			
1000	0.4	ML	-1.3	0.3	-1.9	-8.7	-8.1	-9.7	-1.0	0.0	-3.3	-0.4	0.5	-1.4	-8.6	-7.8	-9.5	-0.5	0.6	-1.2			
1000	0.4	ULSMV	-1.1	0.2	-1.8	-0.2	0.4	-0.9	-0.9	0.2	-2.9	-1.1	-0.0	-1.8	-0.1	1.2	-1.1	-1.1	-0.1	-1.6			
1000	0.4	WLSMV	-0.4	1.0	-1.0	0.0	0.6	-0.7	-0.1	0.9	-2.2	-0.0	0.8	-0.9	0.1	1.5	-0.9	-0.0	1.0	-0.6			
1000	0.4	BAYES	1.5	10.2	-5.0	1.2	6.5	-1.9	1.9	8.0	-2.4	0.6	4.8	-2.5	0.3	1.9	-1.2	0.4	4.3	-2.0			
1000	0.7	ML	-0.6	-0.4	-1.0	-3.6	-3.2	-4.0	-0.8	-0.4	-1.0	0.2	0.3	-0.1	-3.4	-2.8	-3.9	-0.0	0.4	-0.4			
1000	0.7	ULSMV	-0.1	0.1	-0.6	-0.0	0.1	-0.1	-0.2	0.3	-0.7	-0.2	0.1	-0.4	-0.1	0.0	-0.4	-0.3	-0.0	-0.7			
1000	0.7	WLSMV	0.1	0.4	-0.2	0.1	0.3	0.0	0.0	0.5	-0.3	0.2	0.5	-0.0	0.0	0.2	-0.2	0.0	0.4	-0.3			
1000	0.7	BAYES	-0.1	1.9	-4.2	-0.0	1.0	-2.1	-0.3	1.2	-4.5	0.0	1.1	-2.7	-0.1	0.4	-2.6	-0.3	1.0	-3.2			

Appendix 4. Relative percentage bias of interfactor correlations

Model	Sample Size	Number of Items	Estimation Method	Mean Factor Loading = 0.4			Mean Factor Loading = 0.7		
				Left-Skewed	Normal	Right-Skewed	Left-Skewed	Normal	Right-Skewed
2 factors ($\rho = 0.3$)	200	10 Items	ML	-11.77	-2.13	-9.51	1.21	0.37	3.01
		10 Items	ULSMV	-27.72	7.51	-32.22	-4.45	1.88	-2.22
		10 Items	WLSMV	-9.93	8.88	-14.92	-1.24	2.55	0.59
		10 Items	BAYES	-52.33	-29.60	-53.21	-0.84	2.69	1.93
		20 Items	ML	-1.83	0.77	-2.25	4.63	1.06	0.47
		20 Items	ULSMV	-19.62	10.02	-19.00	-0.52	1.87	-5.39
		20 Items	WLSMV	-15.13	10.70	-14.29	3.41	2.49	-1.75
		20 Items	BAYES	-34.01	-19.16	-34.66	0.31	-1.17	-3.25
	500	10 Items	ML	-0.35	1.87	-1.81	1.78	0.31	1.91
		10 Items	ULSMV	-4.83	5.29	-5.98	-1.78	0.77	-1.64
		10 Items	WLSMV	2.99	5.56	2.75	0.86	1.01	0.97
		10 Items	BAYES	-27.78	-15.68	-29.73	0.39	0.25	0.76
		20 Items	ML	3.31	0.55	1.75	4.04	1.21	3.93
		20 Items	ULSMV	-1.72	4.45	-3.41	-0.46	0.89	-0.64
		20 Items	WLSMV	7.10	4.65	5.40	2.32	1.13	2.09
		20 Items	BAYES	-17.17	-9.67	-18.84	-0.36	-0.39	-0.65
	1000	10 Items	ML	2.75	0.99	2.87	2.60	0.44	2.78
		10 Items	ULSMV	-0.50	2.39	-0.68	-0.50	0.38	-0.34
		10 Items	WLSMV	3.68	2.52	3.59	0.78	0.49	0.96
		10 Items	BAYES	-16.36	-4.71	-14.68	0.31	0.80	0.80
		20 Items	ML	2.39	0.35	0.22	3.95	1.50	3.97
		20 Items	ULSMV	-0.97	2.36	-3.38	-0.13	0.81	0.26
		20 Items	WLSMV	3.37	2.46	0.94	1.21	0.92	1.58
		20 Items	BAYES	-9.30	-3.44	-10.70	0.03	0.44	-0.11
2 factors ($\rho = 0.6$)	200	10 Items	ML	-7.98	0.24	-6.08	3.56	0.30	2.05
		10 Items	ULSMV	-17.69	1.78	-17.11	1.70	0.75	0.04
		10 Items	WLSMV	-11.77	2.12	-9.89	3.79	1.33	2.34
		10 Items	BAYES	-54.18	-31.97	-51.69	-1.14	0.28	-1.18
		20 Items	ML	4.01	-0.30	0.20	2.94	1.00	2.46
		20 Items	ULSMV	-1.78	4.49	-3.28	0.15	1.06	-0.33
		20 Items	WLSMV	2.84	4.83	2.50	2.54	1.54	2.06
		20 Items	BAYES	-32.52	-22.84	-35.04	-2.30	-1.53	-2.79
	500	10 Items	ML	1.79	0.63	-0.45	2.18	0.41	2.21
		10 Items	ULSMV	-0.76	2.35	-2.84	0.15	0.43	0.35
		10 Items	WLSMV	1.72	2.53	-0.08	1.08	0.63	1.29
		10 Items	BAYES	-29.27	-20.40	-30.88	0.35	-0.05	0.44
		20 Items	ML	2.12	0.39	1.66	2.77	1.00	3.04
		20 Items	ULSMV	1.37	2.31	0.79	-0.21	0.62	0.29
		20 Items	WLSMV	3.63	2.44	3.14	0.76	0.81	1.24
		20 Items	BAYES	-19.28	-11.67	-19.67	-0.99	-0.30	-0.78
1000	10 Items	ML	1.60	0.17	2.35	2.34	0.48	1.80	
	10 Items	ULSMV	0.85	0.97	1.60	0.22	0.23	-0.40	
	10 Items	WLSMV	2.07	1.07	2.77	0.73	0.33	0.07	
	10 Items	BAYES	-19.65	-10.51	-18.50	0.38	0.38	0.01	
	20 Items	ML	1.24	-0.22	1.95	2.87	0.66	2.38	
	20 Items	ULSMV	0.26	0.74	0.91	0.26	0.14	-0.19	
	20 Items	WLSMV	1.46	0.81	2.10	0.72	0.23	0.29	
	20 Items	BAYES	-11.40	-6.10	-10.66	-0.21	-0.23	-0.70	

Sample Size	Model	MFL	Method	Number of Items = 10									Number of Items = 20								
				Left-Skewed			Normal			Right-Skewed			Left-Skewed			Normal			Right-Skewed		
				Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
1000	0.7	BAYES	92.4	95.0	84.2	93.4	95.9	90.1	92.3	94.3	87.6	92.4	96.1	82.3	92.8	94.8	84.9	93.3	96.6	85.0	
1000	0.7	ML	95.2	95.9	94.4	90.7	92.6	89.3	94.7	95.7	93.1	94.9	95.8	93.7	89.5	91.0	87.6	95.0	96.1	94.0	
1000	0.7	ULSMV	94.7	95.7	93.7	95.1	96.3	94.4	94.2	95.2	93.0	94.5	95.5	93.2	94.4	95.2	93.1	94.8	96.0	93.7	
1000	0.7	WLSMV	94.6	95.5	94.0	95.1	96.2	94.5	94.4	95.3	93.7	94.2	95.3	93.0	94.2	95.1	93.2	94.4	96.2	92.8	
1000	0.4	BAYES	94.6	97.5	93.0	94.5	96.5	92.5	94.9	98.6	92.5	93.7	96.8	91.7	94.4	96.1	90.3	94.1	98.6	91.1	
1000	0.4	ML	94.6	95.8	93.6	90.8	92.8	88.7	95.0	95.8	94.5	94.6	95.8	93.5	89.3	91.0	87.7	94.7	95.9	93.4	
1000	0.4	ULSMV	93.5	94.9	92.2	94.5	95.4	93.3	94.1	95.2	93.2	94.0	94.7	93.3	94.7	96.0	93.7	94.2	95.8	92.3	
1000	0.4	WLSMV	93.5	94.9	91.9	94.4	95.3	93.2	94.0	95.1	93.3	93.8	95.0	92.9	94.6	96.0	93.5	94.0	95.4	92.2	
1000	0.7	BAYES	92.3	96.6	86.8	93.3	95.2	87.2	92.6	95.0	86.1	92.7	95.4	82.1	92.9	95.8	85.5	92.9	96.5	84.8	
1000	0.7	ML	94.9	96.1	94.1	90.3	92.9	88.7	94.9	96.0	93.6	95.0	96.2	93.4	89.3	91.2	87.7	94.7	96.5	92.8	
1000	0.7	ULSMV	94.6	96.2	94.0	94.7	95.9	93.1	94.8	95.6	94.0	94.7	95.8	93.4	94.6	95.8	93.3	94.5	96.4	92.6	
1000	0.7	WLSMV	94.4	95.6	93.8	94.8	96.0	93.3	94.5	95.2	93.6	94.3	95.8	93.0	94.2	95.5	93.4	94.2	96.5	92.7	
1000	0.4	BAYES	94.2	98.4	90.8	94.5	95.7	93.6	94.0	99.3	91.3	94.0	96.8	91.9	94.8	96.4	93.2	94.5	97.6	91.7	
1000	0.4	ML	94.5	95.3	93.4	90.5	91.3	89.5	94.5	95.5	92.8	94.8	96.0	93.2	88.9	90.5	86.8	95.0	96.0	93.4	
1000	0.4	ULSMV	93.7	94.3	92.7	94.6	95.9	93.5	93.5	95.4	91.8	94.2	95.2	92.8	94.7	95.5	93.6	94.5	95.5	93.1	
1000	0.4	WLSMV	93.6	94.5	92.8	94.6	95.7	93.5	93.5	95.5	92.0	94.1	95.5	92.7	94.5	95.2	93.6	94.4	95.3	93.0	
1000	0.7	BAYES	93.0	95.4	87.5	93.6	96.6	90.3	92.4	95.1	87.1	92.5	95.5	85.1	93.2	95.8	85.3	93.3	95.9	84.9	
1000	0.7	ML	94.9	95.9	93.5	90.9	92.2	89.4	95.0	96.5	92.6	94.6	95.4	93.0	89.3	91.0	87.7	94.8	95.7	93.4	
1000	0.7	ULSMV	94.8	96.5	93.0	95.3	96.3	94.3	94.4	95.6	92.6	94.8	96.2	93.6	94.7	96.2	93.6	94.5	95.7	92.9	
1000	0.7	WLSMV	94.3	95.3	92.8	94.9	96.0	93.9	94.3	95.3	92.1	94.3	95.2	92.8	94.5	96.2	93.5	94.3	95.6	92.8	

Appendix 6. Coverage rate of interfactor correlations

Sample Size	Model	Mean Factor Loading	Estimation Method	Number of Items = 10			Number of Items = 20		
				Left-Skewed	Normal	Right-Skewed	Left-Skewed	Normal	Right-Skewed
200	2 factors ($\varphi = 0$)	0.4	ML	76.98	88.48	80.63	90.60	92.70	90.20
		0.4	ULSMV	71.31	84.82	71.86	63.44	86.86	61.01
		0.4	WLSMV	67.74	83.54	70.81	38.22	85.74	40.11
		0.4	BAYES	96.79	95.10	96.90	97.09	95.10	96.49
		0.7	ML	93.00	94.60	95.10	94.50	95.30	94.60
		0.7	ULSMV	86.80	93.30	89.00	86.90	93.60	87.30
	2 factors ($\varphi = 0.3$)	0.4	WLSMV	74.60	92.60	74.77	67.00	92.90	70.10
		0.7	BAYES	93.30	93.80	94.90	94.20	94.40	94.50
		0.4	ML	76.86	90.09	79.57	89.10	93.10	91.30
		0.4	ULSMV	72.52	87.66	71.77	73.36	88.03	75.97
		0.4	WLSMV	74.69	86.84	75.63	58.20	87.22	59.37
		0.4	BAYES	92.70	93.60	92.49	92.30	94.20	93.70
	2 factors ($\varphi = 0.6$)	0.7	ML	92.40	94.00	93.10	93.50	94.50	93.10
		0.7	ULSMV	89.10	93.20	89.20	92.00	94.50	89.40
		0.7	WLSMV	86.29	92.20	86.29	86.90	93.60	85.50
		0.7	BAYES	92.80	93.50	93.40	94.40	95.10	93.70
		0.4	ML	78.68	88.64	80.73	90.78	92.59	90.29
		0.4	ULSMV	84.82	94.21	85.51	84.63	91.04	83.37
	2 factors ($\varphi = 0$)	0.4	WLSMV	86.86	93.76	87.37	79.42	90.13	80.30
		0.4	BAYES	71.57	83.30	75.20	77.15	81.10	71.64
		0.7	ML	92.60	94.10	92.50	91.80	94.40	92.50
		0.7	ULSMV	91.60	93.40	91.50	92.20	93.40	91.40
		0.7	WLSMV	90.70	93.00	90.80	90.50	92.60	89.59
		0.7	BAYES	93.70	94.60	94.40	94.30	95.60	95.30
2 factors ($\varphi = 0.3$)	0.4	ML	90.69	94.00	92.19	93.90	93.80	93.80	
	0.4	ULSMV	82.74	91.19	85.22	87.44	92.40	85.19	
	0.4	WLSMV	85.23	90.89	86.79	89.76	92.30	88.79	
	0.4	BAYES	95.40	94.80	96.00	95.60	94.90	96.10	
	0.7	ML	93.60	95.30	93.40	94.90	95.40	93.30	
	0.7	ULSMV	91.50	94.70	90.20	93.30	94.40	91.20	
2 factors ($\varphi = 0.6$)	0.7	WLSMV	92.10	94.30	91.40	94.10	94.40	91.90	
	0.7	BAYES	93.40	94.20	93.20	94.80	94.00	93.00	
	0.4	ML	92.40	95.20	92.30	93.20	95.00	94.30	
	0.4	ULSMV	87.49	93.50	87.33	88.93	93.00	90.58	
	0.4	WLSMV	87.60	93.20	89.30	89.85	92.60	90.39	
	0.4	BAYES	93.70	93.30	92.20	92.90	94.90	92.80	
2 factors ($\varphi = 0$)	0.7	ML	94.30	94.00	95.20	92.90	95.10	94.00	
	0.7	ULSMV	92.70	94.30	94.40	94.60	94.00	94.60	
	0.7	WLSMV	92.50	93.70	94.30	94.00	93.80	94.10	
	0.7	BAYES	93.60	93.00	94.50	93.60	94.70	95.10	
	0.4	ML	93.67	95.00	92.56	93.00	94.40	93.50	
	0.4	ULSMV	94.07	94.29	92.41	92.08	93.40	91.86	
2 factors ($\varphi = 0.3$)	0.4	WLSMV	94.95	93.99	93.08	90.88	93.30	91.47	
	0.4	BAYES	81.20	80.70	75.48	79.50	87.50	76.30	
	0.7	ML	92.60	94.20	93.00	93.40	93.60	91.70	
	0.7	ULSMV	94.10	94.70	93.80	94.80	93.80	93.70	
	0.7	WLSMV	93.80	94.40	93.20	94.00	93.20	92.80	
	0.7	BAYES	93.60	94.20	94.00	94.60	94.20	93.90	
2 factors ($\varphi = 0$)	0.4	ML	93.40	94.80	93.70	93.60	93.40	94.80	
	0.4	ULSMV	90.17	93.60	90.08	91.30	92.80	91.70	
	0.4	WLSMV	90.57	93.60	91.18	91.80	92.40	92.60	
	0.4	BAYES	94.80	95.70	95.40	95.00	93.10	95.70	
	0.7	ML	95.50	95.10	95.10	95.20	94.70	94.40	
	0.7	ULSMV	94.40	94.80	93.50	92.90	93.80	94.10	
2 factors ($\varphi = 0.3$)	0.7	WLSMV	94.90	94.70	93.90	93.80	93.70	94.20	
	0.7	BAYES	95.10	95.10	94.40	95.00	93.80	95.00	
	0.4	ML	93.10	93.60	93.50	95.10	94.70	94.80	
	0.4	ULSMV	91.78	92.50	90.70	93.30	93.80	92.90	
	0.4	WLSMV	92.08	92.50	91.30	93.10	93.80	93.30	
	0.4	BAYES	93.70	94.80	93.90	94.90	94.30	94.30	
1000	2 factors ($\varphi = 0.3$)	0.7	ML	96.20	94.60	93.80	94.10	94.30	95.00
		0.7	ULSMV	96.00	94.10	93.30	94.90	93.60	94.40

Sample Size	Model	Mean Factor Loading	Estimation Method	Number of Items = 10			Number of Items = 20		
				Left-Skewed	Normal	Right-Skewed	Left-Skewed	Normal	Right-Skewed
2 factors ($\rho = 0.6$)		0.7	WLSMV	95.70	94.10	93.60	94.80	93.60	94.20
		0.7	BAYES	96.30	94.30	93.80	94.70	94.50	95.10
		0.4	ML	95.50	94.50	95.30	94.20	95.00	93.30
		0.4	ULSMV	94.78	94.00	94.58	93.30	95.00	93.10
		0.4	WLSMV	94.58	94.00	93.88	92.90	94.80	92.90
		0.4	BAYES	80.00	89.30	85.20	86.40	91.80	87.60
		0.7	ML	92.10	94.70	94.50	91.70	93.50	91.80
		0.7	ULSMV	94.30	95.40	95.80	94.50	94.50	95.50
		0.7	WLSMV	93.90	95.10	95.70	94.10	94.40	95.70
		0.7	BAYES	93.00	94.70	95.00	95.70	94.90	96.10

Sample Size	Model	MFL	Method	Number of Items = 10									Number of Items = 20								
				Left-Skewed			Normal			Right-Skewed			Left-Skewed			Normal			Right-Skewed		
				Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
1000	0.7	ML	1.00	1.03	0.96	1.01	1.03	0.97	1.00	1.04	0.96	1.01	1.05	0.95	1.00	1.04	0.96	1.02	1.08	0.98	
1000	0.7	ULSMV	1.00	1.03	0.96	1.00	1.02	0.97	0.99	1.04	0.95	1.00	1.04	0.93	1.00	1.04	0.95	1.00	1.04	0.96	
1000	0.7	WLSMV	1.00	1.03	0.96	1.00	1.02	0.97	0.99	1.04	0.95	0.99	1.03	0.93	0.99	1.03	0.95	0.99	1.04	0.96	
1000	0.7	BAYES	0.97	1.02	0.87	0.97	1.06	0.93	0.97	1.01	0.93	0.95	1.06	0.79	0.97	1.04	0.84	0.97	1.13	0.84	
1000	0.4	ML	0.97	1.01	0.95	0.99	1.01	0.96	0.99	1.02	0.97	0.98	1.02	0.95	1.00	1.03	0.95	0.99	1.03	0.95	
1000	0.4	ULSMV	0.93	0.96	0.90	0.97	1.00	0.95	0.94	0.96	0.92	0.96	0.99	0.93	0.99	1.02	0.96	0.97	1.01	0.93	
1000	0.4	WLSMV	0.93	0.96	0.90	0.97	0.99	0.94	0.95	0.97	0.92	0.96	0.99	0.93	0.99	1.02	0.95	0.96	1.01	0.92	
1000	0.4	BAYES	1.04	1.27	0.95	1.00	1.13	0.92	1.05	1.42	0.93	0.97	1.18	0.90	1.00	1.11	0.96	0.99	1.38	0.90	
1000	0.7	ML	1.00	1.02	0.96	1.01	1.07	0.96	1.01	1.04	0.97	1.01	1.07	0.96	0.99	1.04	0.96	1.01	1.07	0.95	
1000	0.7	ULSMV	0.99	1.02	0.96	1.00	1.06	0.95	1.00	1.03	0.97	0.99	1.04	0.94	0.99	1.03	0.95	0.99	1.05	0.94	
1000	0.7	WLSMV	0.99	1.01	0.96	1.00	1.06	0.95	1.00	1.03	0.97	0.99	1.05	0.93	0.98	1.03	0.95	0.99	1.05	0.93	
1000	0.7	BAYES	0.97	1.05	0.91	0.97	1.02	0.90	0.97	1.05	0.93	0.95	1.04	0.79	0.96	1.04	0.83	0.97	1.13	0.85	
1000	0.4	ML	0.97	1.00	0.94	0.99	1.03	0.97	0.97	1.02	0.93	1.00	1.03	0.97	0.99	1.04	0.97	1.00	1.04	0.94	
1000	0.4	ULSMV	0.94	0.97	0.91	0.98	1.02	0.96	0.94	0.98	0.90	0.98	1.02	0.95	0.99	1.03	0.96	0.98	1.02	0.94	
1000	0.4	WLSMV	0.94	0.97	0.91	0.98	1.01	0.95	0.94	0.98	0.90	0.97	1.01	0.95	0.98	1.02	0.96	0.98	1.02	0.94	
1000	0.4	BAYES	1.02	1.23	0.91	1.01	1.13	0.96	1.01	1.32	0.89	0.99	1.20	0.92	1.00	1.10	0.95	1.00	1.23	0.92	
1000	0.7	ML	1.01	1.04	0.95	1.02	1.05	0.99	1.00	1.04	0.95	1.01	1.03	0.96	1.00	1.02	0.97	1.01	1.03	0.99	
1000	0.7	ULSMV	1.00	1.03	0.93	1.02	1.05	0.98	0.99	1.03	0.95	1.00	1.03	0.95	1.00	1.03	0.97	0.99	1.03	0.95	
1000	0.7	WLSMV	1.00	1.03	0.93	1.01	1.05	0.98	0.99	1.02	0.94	0.99	1.03	0.94	1.00	1.02	0.96	0.99	1.02	0.95	
1000	0.7	BAYES	0.99	1.08	0.90	0.98	1.08	0.90	0.96	1.03	0.85	0.95	1.02	0.81	0.97	1.03	0.82	0.97	1.08	0.85	

Appendix 8. r-seb values of interfactor correlations

Sample Size	Model	Mean Factor Loading	Estimation Method	Number of Items = 10			Number of Items = 20		
				Left-Skewed	Normal	Right-Skewed	Left-Skewed	Normal	Right-Skewed
200	2 factors ($\varphi = 0$)	0.4	ML	0.62	0.84	0.67	0.85	0.93	0.84
		0.4	ULSMV	0.58	0.70	0.59	0.53	0.79	0.51
		0.4	WLSMV	0.56	0.68	0.58	0.34	0.77	0.34
		0.4	BAYES	1.23	1.10	1.28	1.15	1.05	1.13
		0.7	ML	0.97	0.99	1.01	0.99	1.05	0.99
		0.7	ULSMV	0.85	0.96	0.87	0.85	0.98	0.86
	2 factors ($\varphi = 0.3$)	0.7	WLSMV	0.61	0.94	0.62	0.54	0.96	0.55
		0.7	BAYES	0.98	0.99	1.00	0.98	0.99	0.99
		0.4	ML	0.67	0.88	0.70	0.83	0.94	0.89
		0.4	ULSMV	0.59	0.77	0.61	0.53	0.83	0.56
		0.4	WLSMV	0.56	0.76	0.58	0.35	0.81	0.34
		0.4	BAYES	1.18	1.06	1.16	1.07	1.10	1.13
	2 factors ($\varphi = 0.6$)	0.7	ML	0.96	0.96	0.96	0.98	1.00	0.97
		0.7	ULSMV	0.86	0.94	0.87	0.90	0.99	0.88
		0.7	WLSMV	0.72	0.93	0.71	0.77	0.97	0.74
		0.7	BAYES	0.96	0.96	0.96	0.98	1.01	0.97
		0.4	ML	0.77	0.91	0.80	0.89	0.93	0.84
		0.4	ULSMV	0.67	0.95	0.67	0.53	0.92	0.56
	2 factors ($\varphi = 0$)	0.4	WLSMV	0.65	0.93	0.69	0.44	0.89	0.47
		0.4	BAYES	0.98	1.08	0.95	1.12	1.12	1.09
		0.7	ML	0.97	1.00	0.98	0.97	0.99	0.97
		0.7	ULSMV	0.95	1.00	0.96	0.92	0.98	0.93
		0.7	WLSMV	0.91	0.98	0.94	0.88	0.96	0.90
		0.7	BAYES	1.01	1.01	1.03	1.01	1.00	1.01
2 factors ($\varphi = 0.3$)	0.4	ML	0.85	0.95	0.90	0.97	0.97	0.94	
	0.4	ULSMV	0.68	0.86	0.72	0.81	0.91	0.78	
	0.4	WLSMV	0.70	0.86	0.75	0.83	0.90	0.77	
	0.4	BAYES	1.06	1.01	1.12	1.06	1.01	1.04	
	0.7	ML	0.96	1.00	0.96	0.99	1.05	0.94	
	0.7	ULSMV	0.91	0.98	0.90	0.97	0.99	0.90	
2 factors ($\varphi = 0.6$)	0.7	WLSMV	0.92	0.97	0.90	0.96	0.98	0.90	
	0.7	BAYES	0.95	0.97	0.94	0.99	0.99	0.93	
	0.4	ML	0.91	0.99	0.94	0.95	0.98	0.96	
	0.4	ULSMV	0.77	0.92	0.77	0.84	0.93	0.86	
	0.4	WLSMV	0.78	0.92	0.80	0.86	0.92	0.86	
	0.4	BAYES	1.13	1.07	1.07	1.06	1.02	1.05	
2 factors ($\varphi = 0.3$)	0.7	ML	0.99	0.96	1.02	0.99	1.01	1.03	
	0.7	ULSMV	0.96	0.95	0.99	0.99	1.00	1.00	
	0.7	WLSMV	0.96	0.95	0.98	0.98	0.99	0.99	
	0.7	BAYES	0.98	0.93	1.00	0.98	0.99	1.01	
	0.4	ML	1.00	0.98	0.95	0.91	1.00	0.95	
	0.4	ULSMV	0.92	0.96	0.89	0.84	1.00	0.91	
2 factors ($\varphi = 0.6$)	0.4	WLSMV	0.92	0.95	0.90	0.83	0.99	0.90	
	0.4	BAYES	1.13	1.11	1.08	1.10	1.13	1.11	
	0.7	ML	0.99	1.00	0.99	1.01	1.01	1.00	
	0.7	ULSMV	0.99	1.01	0.99	0.99	1.01	0.98	
	0.7	WLSMV	0.98	1.00	0.97	0.98	1.00	0.96	
	0.7	BAYES	0.99	0.97	0.99	1.01	1.02	0.99	
2 factors ($\varphi = 0$)	0.4	ML	0.94	0.99	0.97	0.96	0.95	0.98	
	0.4	ULSMV	0.83	0.94	0.85	0.89	0.92	0.91	
	0.4	WLSMV	0.84	0.94	0.87	0.90	0.91	0.91	
	0.4	BAYES	1.02	1.03	1.08	1.02	0.95	1.04	
	0.7	ML	1.03	0.99	0.98	1.01	1.03	0.98	
	0.7	ULSMV	0.99	0.99	0.97	0.97	0.99	0.97	
2 factors ($\varphi = 0.3$)	0.7	WLSMV	0.99	0.98	0.97	0.97	0.98	0.97	
	0.7	BAYES	1.02	0.99	0.98	1.02	0.98	0.98	
	0.4	ML	0.96	0.97	0.95	1.00	1.00	0.99	
	0.4	ULSMV	0.88	0.95	0.87	0.95	0.98	0.93	
	0.4	WLSMV	0.89	0.94	0.88	0.95	0.97	0.93	
	0.4	BAYES	1.05	1.03	1.04	1.06	1.00	1.04	
2 factors ($\varphi = 0.3$)	0.7	ML	1.06	0.98	1.00	1.00	1.01	1.03	
	0.7	ULSMV	1.04	0.98	0.98	0.99	1.00	1.00	

Sample Size	Model	Mean Factor Estimation		Number of Items = 10			Number of Items = 20		
		Loading	Method	Left-Skewed	Normal	Right-Skewed	Left-Skewed	Normal	Right-Skewed
2 factors ($\varphi = 0.6$)		0.7	WLSMV	1.04	0.97	0.98	0.98	1.00	0.99
		0.7	BAYES	1.05	0.98	1.00	1.00	1.00	1.03
		0.4	ML	1.00	0.97	1.03	0.99	1.02	0.96
		0.4	ULSMV	0.96	0.95	0.98	0.96	1.03	0.94
		0.4	WLSMV	0.96	0.95	0.98	0.96	1.03	0.94
		0.4	BAYES	1.08	1.05	1.14	1.10	1.08	1.08
		0.7	ML	0.96	1.01	1.02	1.02	1.01	1.03
		0.7	ULSMV	0.95	1.01	1.03	1.00	1.01	1.03
		0.7	WLSMV	0.95	1.00	1.02	0.99	1.01	1.02
		0.7	BAYES	0.95	1.00	1.01	1.02	1.01	1.04