Estimating Reliability for

Multidimensional Composite Scale Scores

Akihito Kamata

Ahmet Turhan

Eqbal Darandari

Florida State University

307 Stone Building
Tallahassee, FL 32306-4453
kamata@coe.fsu.edu

Abstract

When a test is consisted of multiple unidimensional scales, such as subscales, the composite of the entire test items is not likely meet the essentially tau-equivalent condition. Therefore, Coefficient-alpha, a widely used reliability estimator, likely underestimates the true reliability at large degree. This paper examined three alternative methods (Stratified-alpha, Maximal Reliability, and Multidimensional omega) to estimate the reliability under five different multidimensional factor-structure conditions. The result revealed that the three alternative methods estimated the true reliability much better than Coefficient-alpha did in all conditions. It also appeared that Stratified-alpha generally performed the best. Comparisons between the three alternative methods are discussed in detail.

Estimating Reliability for

Multidimensional Composite Scale Scores

Scale reliability is of interest to many researchers who develop or utilize psychological and educational measurement. To date a number of methods have been proposed to estimate the reliability of composite scores based on a single administration of a test. Among them, Coefficient-alpha (Guttman, 1945; Cronbach, 1951) is probably the most frequently reported reliability estimate in educational and psychological literature. It is a well-known fact that the essentially tau-equivalent condition is required for Coefficient-alpha to be an unbiased estimator of the reliability (Lord & Novic, 1968). By definition, the essentially tau-equivalent condition requires that the true scores for any two components in the test have to be different only by a constant (Lord & Novic, 1968). Operationally, the essentially tau-equivalent condition requires equal discriminating power for all components in the test, as well as the unidimensionality of the test, which is represented by equal factor loadings for all components under the one-factor factor analytic model (McDonald, 1999). If the essentially tau-equivalent condition is not satisfied, Coefficient-alpha is a lower bound of reliability; that is, it underestimates the true reliability to some extent.

In practice it is almost impossible to achieve the tau-equivalent assumption perfectly, in terms of both equal discrimination power for all test components and unidimensionality of the test. When the essentially tau-equivalent condition is not satisfied only because of unequal discrimination power of test components (i.e., the unidimensionality of the test is satisfied), the test is identified as a congeneric test. Underestimation of the true reliability because of unequal discriminations is rather optimistic, as far as unidimensionality of the test is secured, because there are some attempts to estimate reliability in such a case and have been shown to be quite

effective (e.g., Jöreskog, 1971; McDonald, 1970, 1999). However, underestimation of the true reliability may become serious when the test is not unidimensional (e.g., Cortina, 1993; Schmitt, 1996; Osburn, 1999). This would be a concern when one is interested in estimating the reliability of multidimensional scale scores. For example, one might have a multidimensional scale to measure the self-efficacy for collegiate athletes (Czerniack, 2002), where the scale is composed of three dimensions of the self-efficacy (athletic, academic, and social life). In this example, one might be interested in measuring the overall self-efficacy as a summation of the scores from all three dimensions, as well as in measuring the self-efficacy for three dimensions separately. It is clear that Coefficient α for the total scale scores will underestimate the true reliability, because of the multidimensionality nature of the scale. Another example is the total GRE scores as an admission criterion for a graduate school. The problem extended to approaches where one corrects a validity coefficient for attenuation due to unreliability, or intends to estimate the standard error of measurement (Schmitt, 1996; Rogers et al., 2002).

Recent studies points out that with pre-specified set of items that do not assess the same underlying dimensions, popular methods, such as Coefficient-alpha, may lead to biased reliability estimates (Zimmerman et al., 1993; Komaroff, 1997; Murphy & DeShon, 2000; Osbourn, 2000; Raykov, 1998; 2001). Particularly, Osbourn (2000) demonstrates that when the scale was multidimensional, most of the reliability coefficients, including Coefficient-alpha, Standardized alpha, and Feldt's coefficient, underestimated the true reliability of the scale. Exceptions were Stratified-alpha and Maximal Reliability that provided consistent and accurate reliability estimates. This paper, therefore, investigates the effect of multidimensionality of a scale on Coefficient-alpha, as well as some procedures that may be suitable to estimate reliability under multidimensional conditions, such as Stratified-alpha (Cronbach, Shonenman, & McKie,

1965), and Maximal Reliability (Li, Rosenthal, & Rubin, 1996). This paper also examine the

multidimensional extension of McDonald's omega (McDonald, 1970, 1999).

Stratified-alpha

Stratified-alpha was proposed by Cronbach, Shonenman, and McKie (1965). It is

intended for cases where components of a test can be grouped into subtests on the basis of

content. It assumes k components, where ith component ($i = 1, \ldots, k$) consisted of n_i

components. Stratified a is obtained by

$$Stratified\ \alpha = 1 - \frac{\sum_{i=1}^{k} \sigma_i^2 (1 - \alpha_i)}{\sigma_x^2},$$

where σ_i^2 is the variance of items in the ith component, α_i is the reliability of ith component, and

σ_x^2 is the variance of the test. Stratified α may be suitable for estimating reliability for

multidimensional composite scores.

Maximal Reliability

Maximal Reliability was derived by Li, Resenthal, and Rubin (1996). It extends the

Spearman-Brown formula (Lord & Novick, 1968) to k components, where ith component ($i = 1,$

$\ldots, k$) consisted of n_i parallel components. Operationally, it assumes that a test can be divided

into k subtests, where all items within a subtest are parallel. In other words, it assumes that all

items within a subtest have the same reliability and the same variances. Maximal Reliability R_k^*

is obtained by

$$R_k^* = \frac{\dfrac{n_1 r_1}{1 - r_1} + \dfrac{n_2 r_2}{1 - r_2} + \cdots + \dfrac{n_k r_k}{1 - r_k}}{\dfrac{k}{1 + (k-1)\rho} + \dfrac{n_1 r_1}{1 - r_1} + \dfrac{n_2 r_2}{1 - r_2} + \cdots + \dfrac{n_k r_k}{1 - r_k}}$$

where r_i is the item reliability of the ith subtest, ρ is the common correlation between traits.

Maximal Reliability should be suitable for estimating reliability for multidimensional composite scores when items within subtests are parallel. However, if items within subtests are not parallel, underestimation of reliability is expected.

Omega

McDonald (1970, 1999) proposed the estimate of the reliability for a composite score from a congeneric a test. McDonald defines omega (ω) as the ratio of modeled variance to the total variance in the one-factor factor analytic model. In the one-factor factor model

$$X_j = \lambda_j F + E_j ,\tag{1}$$

where X_j is a random score on the jth item ($j = 1, \ldots , k$), F is the factor score (i.e., examinee's ability) E_j is the uniqueness (i.e., error), and λ_j is the factor loading for item j. Then, assuming the variance of F is fixed to be 1, the modeled variance for the composite score of k items (σ^2_C) is

$$\sigma^2_C = \left(\sum_{j=1}^{n} \lambda_j \right)^2 = \mathbf{1}'_n \boldsymbol{\lambda}\boldsymbol{\lambda}' \mathbf{1}_n ,\tag{2}$$

where $\boldsymbol{\lambda}$ is a column vector of n factor loadings, and $\mathbf{1}_n$ is an $n \times 1$ column vector with all elements are 1. The total variance (σ^2_Y) of composite scores ($\sum_{j=1}^{n} X_j$) is then obtained from the sample variance-covariance matrix $\mathbf{S}$ by

$$\sigma^2_X = \mathbf{1}'_n \mathbf{S} \mathbf{1}_n .\tag{3}$$

Then, the reliability ω is

$$\omega = \frac{\sigma^2_C}{\sigma^2_X} = \frac{\mathbf{1}'_n \boldsymbol{\lambda}\boldsymbol{\lambda}' \mathbf{1}_n}{\mathbf{1}'_n \mathbf{S} \mathbf{1}_n} .\tag{4}$$

Even though ω does not assume the essentially tau-equivalent condition, it still assumes a

congeneric assumption, in which components are measuring one latent trait. Therefore, ω itself is

not suitable for estimating reliability when the test is multidimensional.

Multidimensional Extension of ω

McDonald's ω is extended to cases with a multi-factor factor analytic model. For *n*

variables and *p* dimensions, the factor model can be written as

$$\mathbf{X} = \mathbf{F}\boldsymbol{\Lambda}' + \mathbf{E}, \tag{5}$$

where **X** is a $k \times n$ matrix of observed scores on *n* variables and *k* examinees, **F** is a $k \times p$ matrix

of factor scores on *p* factors, $\boldsymbol{\Lambda}$ is an $n \times p$ matrix of factor loadings, $\boldsymbol{\Lambda}'$ is a transpose of $\boldsymbol{\Lambda}$, and **E**

is an $n \times k$ matrix of error scores. Here, the reliability is still conceptualized as the ratio of the

modeled variance to the total variance, which is consistent with the definition of reliability for

the unidimensional ω. Now, the covariance matrix of **X**, *cov*(**X**), is

$$\begin{aligned} cov(\mathbf{X}) &= \boldsymbol{\Lambda} cov(\mathbf{F})\boldsymbol{\Lambda}' + cov(\mathbf{E}) \\ &= \boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \end{aligned}, \tag{6}$$

where $\boldsymbol{\Phi} \equiv cov(\mathbf{F})$, and $\boldsymbol{\Psi} \equiv cov(\mathbf{E})$. Assuming that the variances in **F** are constrained to be 1, $\boldsymbol{\Phi}$

is the correlation matrix for *p* factors. Therefore, the modeled variance for the *p*-dimensional

composite score ($\sum_{j=1}^{n} X_j$) with *n* variables is the sum of all elements in $\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}'$, and it is expressed

as

$$\sigma_C^2 = \mathbf{1}_n'\boldsymbol{\Lambda}\boldsymbol{\Phi}\boldsymbol{\Lambda}'\mathbf{1}_n, \tag{7}$$

where $\mathbf{1}_n$ is an $n \times 1$ column vector with all elements are 1. On the other hand, the total variance

is still expresses in the same way as for the unidimensional case in (3), which is the sum of all

elements in the data matrix **S**. As a result, Multidimensional-omega (MD-omega) is obtained by

$$\omega_{MD} = \frac{1'_n \Lambda \Phi \Lambda' 1_n}{1'_n S 1_n}.$$ (8)

Since the numerator in (8) can be reduced to the numerator of (4) with $\Lambda = \lambda$ and $\Phi = 1$, we can

confirm that ω is a special case of ω_{MD} when $p = 1$.

Simulation Methods

Reliability estimation methods, including Coefficient-alpha, Stratified-alpha, Maximal

reliability, and MD-omega, were evaluated through simulations. In order to generate data, the

reliability values for the 10 components and covariance matrix of the true scores for the 10

components were predetermined. Based on the reliabilities and the true-score covariance matrix,

the observed-score covariance matrix $Var(\mathbf{X})$ was obtained by $Var(\mathbf{X}) = \rho^{-1} Var(\mathbf{T})$, where ρ^{-1} is

the inverse of 10×10 matrix with the reliability of the ith component is the ith diagonal element

in the matrix ($i = 1, \ldots , 10$), and $Var(\mathbf{T})$ is the covariance matrix of the true scores. Then, the

observed scores for the 10 components were randomly generated from the multivariate normal

distribution with 0 means and with the derived observed-score covariance matrix.

A two-factor structure with 5 components per factor was assumed to generate data, and 5

different factor-structure conditions were investigated. The reliabilities for 10 components were

the same for all 5 factor-structure conditions, except in 2 conditions where one of the component

reliability was altered (see descriptions of the 5 conditions below). The reliabilities for the 10

components were .42, .45, .47, .50, .53, .43, .46, .49, 52, and .55 for components 1 through 10.

On the other hand, the structure of the covariance matrix for the true scores was altered for the 5

factor-structure conditions. Also, the diagonals in the true-score covariance matrix were fixed to

1. In other words, the variances of the true scores for 10 components were all 1. True-score

covariances between 10 components that were in the same dimension were all 1 (perfect

correlations). On the other hand, true-score covariances between components that were in different dimensions were all one of .8, .5, or .3, depending on the correlation between the two factors.

The first condition was the perfect simple-structure, where the first 5 components were loaded onto one factor, and the other 5 components were loaded onto another factor. The second condition was the same as the first condition, except that one of the components had considerably lower reliability (i.e., large measurement error) than the other components. Specifically, the reliability of the ninth component was .25, instead of .52, while reliabilities for the other components were the same as in the first condition. In the third condition, the factor structure to generate the data was exactly the same as the first condition. However, one component was miss-placed in the wrong dimension when the reliability was estimated by the four estimation methods. Specifically, the fourth component in the first factor was misplaced in the second factor. The fourth condition was the same as the third condition, except that the misplaced component had considerably lower reliability than the other components. The reliability of the fourth component was .25 instead of .50. The fifth condition had a complex structure, where the first component loaded on both factors. When reliability was estimated, this complex structure was ignored.

In addition to the 5 factor-structure conditions, two additional simulation factors were considered. The true correlation between the two factors was varied among .30, .50, and .80, and the sample size was varied among 50, 200, 500, 1000, and 2000. As a result, 5 factor-structures × 5 sample sizes × 3 correlations = 75 simulation conditions were examined, and each simulation condition was replicated 200 times. For each of the 75 simulation conditions, we evaluated the true reliability, Coefficient-alpha, Stratified-alpha, Maximal reliability, and MD-omega for the

composite of the 10 components. Coefficient-alpha, Stratified-alpha, and Maximal reliability were computed based on the formulas from original literatures, while equation 8 was used for MD-omega.

<div align="center">Simulation Results</div>

Overall, we found that Coefficient-alpha largely underestimated the true reliability. This is consistent with previous literature (e.g., Cortina, 1993; Schmitt, 1996; Osburn, 1999). This underestimation was more prominent when the correlation between dimensions was lower (see Figures). Also, Coefficient-alpha underestimated the true reliability more when the sample size was the smallest at $n = 50$; however, the difference from other sample size conditions was very small. This tendency was consistent across all 5 factor-structure conditions. Therefore, it is not appropriate to use Coefficient-alpha as an estimate of the reliability of a multidimensional composite scale score, unless the correlation between dimensions is high.

On the other hand, the three alternative methods (MD-omega, Stratified-alpha, and Maximal Reliability) estimated the true reliability much better than Coefficient-alpha in all simulation conditions. Therefore, these three methods were compared in details for each of the 5 factor-structure conditions. The root mean square error (RMSE), bias, and standard error were evaluated for each simulation condition.

In the first factor-structure condition, where a perfect simple-structure was assumed, the difference between the three methods was almost negligible. The bias was almost identical for the three methods, except in the $n = 2000$ condition, where MD-omega had larger bias than the other two. However, the difference was only at the third decimal place. The standard error was almost identical for the three methods in all simulation conditions. When RMSE were examined, the effect of the sample size was noticeable for a small sample size ($n = 50$), but the difference

from a larger sample size ($n = 200$) was between .01 and .02. On the other hand, the difference

between $n = 200$ and the largest sample size ($n = 2000$) was less than .01 in all 3 correlation

conditions. The effect of the correlation between factors was very small, and the effect became

smaller as sample size became larger.

In the second factor-structure condition, where one component had significantly lower

reliability than the other components in the test, MD-omega tended to overestimate the true

reliability. The magnitude of the overestimation was between approximately.004 and .010.

Maximal Reliability also had a tendency of overestimation of the true reliability, but the

magnitude of the overestimation was relatively small (.003 maximum). On the other hand,

Stratified-alpha had very small bias in all conditions. The standard errors were smallest for the

MD-omega in all conditions, although the differences were small (.0003 to .0035, mainly

depending on the sample size). Assessment of RMSE revealed the existence of a method ×

sample size interaction. RMSE tended to be smaller when sample size increased. However, this

tendency was weaker for MD-omega, while the tendency and the magnitudes of RMSE were

almost identical for Stratified-alpha and Maximal Reliability, in all 3 correlation conditions.

In the third factor-structure condition, where one component was miss-specified, all three

methods underestimated the true reliability. The magnitude of underestimation was larger when

the correlation between factors was smaller. Also, the difference between the three methods was

larger with smaller correlation, although the difference was negligible with $r = 0.5$ and $r = 0.8$

conditions. The standard error was always largest for Maximal Reliability, but the difference

from the smallest standard error was very small (0.009 to 0.0001) When RMSE were examined,

the interaction effect of method × correlation was observed. The three methods estimated the true

reliability equally well when the correlation between factors was the highest ($r = 0.8$). However,

when $r = 0.5$, Maximal Reliability's performance was worse than the performance of the other two methods, although the difference was very small. When $r = 0.3$, the performance of Stratified-alpha was clearly better than the performance of the other two. However, the effect of method was consistent across difference sample sizes.

In the fourth factor-structure condition, where one component was miss-specified and had much lower reliability than the other components in the test, Stratified-alpha and Maximal Reliability tended to approach the true reliability as the correlation between traits became larger. On the other hand, MD-omega tended to overestimate the true reliability when the correlation was the highest ($r = 0.8$). However, it underestimated the true reliability, and the magnitude of underestimation was smaller than the other two methods when $r = 0.3$ and $r = 0.5$. The standard error was always smallest for MD-omega, but the difference from the largest standard error was small (0.0004 to 0.007). When RMSE was examined an interaction effect of method $\times$ correlation was observed. The three methods estimated the true reliability equally well when the correlation between factors was the highest ($r = 0.8$). However, when $r = 0.5$, MD-omega performed better than the other two methods, while the performance of Stratified-alpha and Maximal Reliability were very similar. When $r = 0.3$, MD-omega still performed better than the other two methods, while Stratified-alpha's performance was clearly better than the performance of Maximal Reliability. The effect of methods was consistent across different sample size, except in the condition of $r = 0.8$ where the effect of sample size was slightly smaller for MD-omega than the other two methods. This was due to MD-omega's tendency to overestimate the true reliability with larger sample sizes ($n \geq 200$) in the $r = 0.8$ conditions. Overall, MD-omega performs the best among the three methods. However, one has to realize the fact that MD-omega may overestimate the true reliability.

In the last factor-structure condition, where one component loaded on two traits, MD-omega tended to overestimate the true reliability, while maximal reliability tended to underestimate the true reliability. This tendency was more prominent when the correlation between traits was smaller. On the other hand, the bias for Stratified-alpha was almost zero in all correlation conditions, except when the sample size was small ($n = 50$). The standard error was always smallest for MD-omega, although the difference between the three methods was very small (.0002 to 0.0075). When RMSEs are examined, there is a method × correlation interaction, as well as a method × sample size interaction. The performance of the three methods were similar when $r = 0.8$, while the difference was larger at $r = 0.5$ and $r = 0.3$. Also, the effect of sample size was stronger for Stratified-alpha than MD-omega, especially in $r = 0.3$ conditions.

Conclusions

Generally, any of the three alternative methods can be recommended when the scale is not unidimensional. In fact, all three alternative methods estimated the true reliability very well under the perfect simple structure condition. Also, the magnitudes of the standard errors were very small, although MD-omega always had the smallest standard errors.

Stratified-alpha estimated the true reliability with almost no bias also in the factor-structure condition 2, where one of the components had significantly lower reliability than the other components, and the factor-structure condition 5, where a complex structure was ignored. On the other hand, Stratified-alpha tended to underestimate the true reliability in the factor-structure condition 3, where one component was miss-specified in the wrong dimension, and in the factor-structure condition 4, where one component had a lower reliability than the other component and was miss-specified. However, the magnitude of the underestimation was smaller than the other two methods and became smaller as the correlation between the dimensions

became larger. On the other hand, the magnitude of the underestimation was not as small as MD-omega but became smaller as the correlation between the dimensions became larger. In conclusion, Stratified-alpha estimated the true reliability with almost no bias as far as the test components were specified to appropriate dimensions.

Maximal Reliability estimated the true reliability with almost no bias in the factor-structure condition 2, where one component had a low reliability. However, it underestimated the true reliability in the factor-structure condition 3, where one component was miss-specified, factor-structure condition 4, where one component had a lower reliability and was miss-specified, and factor-structure condition 5, where complex structure was ignored. However, the magnitude of the underestimation was very small when the correlation between the dimensions was the highest. In conclusion, Maximal Reliability estimated the true reliability with almost no bias as far as the test components were specified in simple structure.

MD-omega estimated the true reliability with almost no bias in the factor-structure condition 4, where one component had a lower reliability and was miss-specified. In fact, this is the condition that the other two alternative methods did not do as well as MD-omega. Therefore, this is the condition that MD-omega has its strength. However, in practice, it is almost impossible to detect a miss-specification. (If it is known, we will certainly try to correct the miss-specification!) On the other hand, MD-omega had tendency of overestimation and underestimation depending on factor-structure conditions. In factor-structure condition 2, where one component had low reliability, MD-omega had a tendency to overestimate the true reliability, while it had a tendency to underestimate the true reliability in factor-structure conditions 3 and 5. However, like the other two alternative methods, the magnitude of the bias was very small when

the correlation between the dimensions was the highest. Overall, MD-omega provides a good

estimate of true reliability, but one has to be aware that it may overestimate the true reliability.

In summary Stratified-alpha appeared to be the most "reliable" procedure among the

three alternative methods we investigated, although the other two methods may outperform

stratified alpha depending on the condition. However, the difference between the three methods

was very small. Also, this study revealed that the reliability could be overestimated by alternative

methods, especially by MD-omega, which can result in a larger error. Therefore, it is not

recommended to take an approach to estimate reliability by multiple methods and use the highest

value as the best estimate.

Although we investigated five different factor-structure conditions, the performance

difference between the three methods was rather subtle. We will further investigate the

characteristics of these three reliability estimates to make more specific recommendations,

including the number of items and the number of dimensions.

References

Cortina, J. M (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology*, *78*, 98-104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Cronbach, L. J., Schönemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement, 25*, 291-312.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*, 255-282.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*, 109-133.

Komaroff, E. (1997). Effect of simultaneous violations of essential tau-equivalence and uncorrelated error on coefficient alpha. *Applied Psychological Measurement*, *21*, 337-348.

Li, H., Rosenthal, R., & Rubin, D. B. (1996). Reliability of measurement in psychology: From Spearman-Brown to maximal reliability. *Psychological Methods*, *1*, 98-107.

Lord, F. M. & Novick, R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.

McDonald, R. P. (1999). *Test Theory: Unified treatment*. Lawrence Erlbaum Associates.

McDonald, R. P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1-21

Murphy, K. R. & DeShon, R. P. (2000). Inter-rater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology*, *53*, 873-900.

Osbourn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, *5*, 343-355.

Raykov, T. (1998). Cronbach's alpha and reliability of composite with interrelated

nonhomogenous items. *Applied Psychological Measurement*, *22*, 375-385.

Raykov, T. (2001). Bias of coefficient alpha for congeneric measures with correlated errors.

Applied Psychological Measurement, *25*, 69-76.

Rogers, W. M., Schmitt, N., & Mullins, M. (2002). Correction for unreliability of multifactor

measures: Comparison of alpha and parallel forms approaches. *Organizational Research

Methods*, *5*, 184-199.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *4*, 350-353.

Zimmerman, D. W., Zumbo, B. D., & Lalonde, C. (1993). Coefficient alpha as an estimate of

test reliability under violation of two assumptions. *Educational and Psychological

Measurement*, *53*, 33-49.

Table 1.

Simulation Results for Perfect Simple Structure – Factor Structure 1

Conditions			Coefficient α	Stratified α	Maximal Reliability	MD-ω
$\rho = 0.3$						
True Reliability = 0.86	$n = 50$	BIAS	-0.0554	-0.0026	-0.0027	0.0003
		SE	0.0482	0.0336	0.0346	0.0334
	$n = 200$	BIAS	-0.0556	-0.0033	-0.0030	-0.0013
		SE	0.0211	0.0155	0.0156	0.0153
	$n = 500$	BIAS	-0.0531	-0.0014	-0.0012	0.0003
		SE	0.0126	0.0093	0.0093	0.0091
	$n = 1000$	BIAS	-0.0525	-0.0012	-0.0010	0.0005
		SE	0.0095	0.0069	0.0070	0.0069
	$n = 2000$	BIAS	-0.0516	-0.0001	0.0000	0.0014
		SE	0.0072	0.0052	0.0052	0.0051
$\rho = 0.5$						
True Reliability = 0.87	$n = 50$	BIAS	-0.0418	-0.0059	-0.0045	-0.0028
		SE	0.0397	0.0297	0.0295	0.0290
	$n = 200$	BIAS	-0.0339	-0.0010	-0.0006	0.0008
		SE	0.0177	0.0140	0.0140	0.0136
	$n = 500$	BIAS	-0.0334	-0.0005	-0.0004	0.0010
		SE	0.0117	0.0092	0.0092	0.0090
	$n = 1000$	BIAS	-0.0324	-0.0002	-0.0001	0.0013
		SE	0.0077	0.0062	0.0063	0.0062
	$n = 2000$	BIAS	-0.0320	0.0005	0.0006	0.0019
		SE	0.0055	0.0042	0.0042	0.0041
$\rho = 0.8$						
True Reliability = 0.89	$n = 50$	BIAS	-0.0134	-0.0019	-0.0010	0.0008
		SE	0.0264	0.0233	0.0233	0.0228
	$n = 200$	BIAS	-0.0122	-0.0011	-0.0009	0.0005
		SE	0.0146	0.0133	0.0133	0.0130
	$n = 500$	BIAS	-0.0104	0.0006	0.0008	0.0020
		SE	0.0081	0.0075	0.0075	0.0073
	$n = 1000$	BIAS	-0.0112	-0.0001	0.0001	0.0012
		SE	0.0059	0.0054	0.0054	0.0053
	$n = 2000$	BIAS	-0.0110	0.0001	0.0002	0.0013
		SE	0.0043	0.0040	0.0040	0.0039

Table 2.

Simulation Results for Perfect Simple Structure with Low Reliability for One Component –

Factor Structure 2

Conditions			Coefficient α	Stratified α	Maximal Reliability	MD-ω
ρ = 0.3						
True Reliability = 0.83	n = 50	BIAS	-0.0615	-0.0078	-0.0048	0.0048
		SE	0.0479	0.0355	0.0341	0.0320
	n = 200	BIAS	-0.0511	-0.0006	0.0023	0.0096
		SE	0.0259	0.0197	0.0193	0.0177
	n = 500	BIAS	-0.0514	-0.0013	0.0013	0.0086
		SE	0.0150	0.0115	0.0114	0.0107
	n = 1000	BIAS	-0.0490	0.0006	0.0032	0.0104
		SE	0.0101	0.0077	0.0076	0.0070
	n = 2000	BIAS	-0.0496	0.0001	0.0027	0.0098
		SE	0.0073	0.0055	0.0053	0.0050
ρ = 0.5						
True Reliability = 0.85	n = 50	BIAS	-0.0397	-0.0072	-0.0044	0.0039
		SE	0.0419	0.0347	0.0356	0.0329
	n = 200	BIAS	-0.0333	-0.0018	0.0006	0.0077
		SE	0.0205	0.0172	0.0172	0.0156
	n = 500	BIAS	-0.0332	-0.0014	0.0012	0.0078
		SE	0.0130	0.0107	0.0106	0.0099
	n = 1000	BIAS	-0.0316	0.0000	0.0022	0.0088
		SE	0.0095	0.0077	0.0076	0.0072
	n = 2000	BIAS	-0.0312	0.0003	0.0025	0.0089
		SE	0.0059	0.0049	0.0048	0.0045
ρ = 0.8						
True Reliability = 0.87	n = 50	BIAS	-0.0176	-0.0063	-0.0040	0.0034
		SE	0.0314	0.0290	0.0287	0.0264
	n = 200	BIAS	-0.0131	-0.0023	0.0000	0.0060
		SE	0.0148	0.0135	0.0132	0.0122
	n = 500	BIAS	-0.0114	-0.0003	0.0018	0.0076
		SE	0.0091	0.0083	0.0083	0.0077
	n = 1000	BIAS	-0.0115	-0.0007	0.0013	0.0071
		SE	0.0065	0.0060	0.0059	0.0055
	n = 2000	BIAS	-0.0111	-0.0002	0.0019	0.0075
		SE	0.0043	0.0040	0.0040	0.0037

Table 3.

Simulation Results for Perfect Simple Structure with Miss-Specified Component – Factor

Structure 3

Conditions			Coefficient α	Stratified α	Maximal Reliability	MD-ω
ρ = 0.3						
True Reliability = 0.86	n = 50	BIAS	-0.0555	-0.0243	-0.0340	-0.0289
		SE	0.0460	0.0393	0.0482	0.0439
	n = 200	BIAS	-0.0486	-0.0185	-0.0273	-0.0239
		SE	0.0215	0.0176	0.0211	0.0200
	n = 500	BIAS	-0.0495	-0.0194	-0.0290	-0.0249
		SE	0.0145	0.0119	0.0145	0.0134
	n = 1000	BIAS	-0.0492	-0.0190	-0.0283	-0.0247
		SE	0.0094	0.0078	0.0094	0.0087
	n = 2000	BIAS	-0.0483	-0.0181	-0.0272	-0.0238
		SE	0.0064	0.0055	0.0065	0.0061
ρ = 0.5						
True Reliability = 0.87	n = 50	BIAS	-0.0339	-0.0149	-0.0150	-0.0144
		SE	0.0383	0.0337	0.0365	0.0349
	n = 200	BIAS	-0.0296	-0.0104	-0.0131	-0.0104
		SE	0.0177	0.0152	0.0170	0.0161
	n = 500	BIAS	-0.0315	-0.0122	-0.0146	-0.0124
		SE	0.0103	0.0091	0.0100	0.0095
	n = 1000	BIAS	-0.0307	-0.0115	-0.0143	-0.0118
		SE	0.0075	0.0065	0.0073	0.0068
	n = 2000	BIAS	-0.0311	-0.0119	-0.0148	-0.0123
		SE	0.0053	0.0047	0.0053	0.0049
ρ = 0.8						
True Reliability = 0.89	n = 50	BIAS	-0.0154	-0.0081	-0.0068	-0.0055
		SE	0.0275	0.0256	0.0253	0.0252
	n = 200	BIAS	-0.0107	-0.0043	-0.0031	-0.0028
		SE	0.0121	0.0114	0.0115	0.0112
	n = 500	BIAS	-0.0108	-0.0041	-0.0031	-0.0028
		SE	0.0077	0.0073	0.0073	0.0072
	n = 1000	BIAS	-0.0102	-0.0037	-0.0027	-0.0025
		SE	0.0061	0.0057	0.0057	0.0056
	n = 2000	BIAS	-0.0107	-0.0041	-0.0032	-0.0029
		SE	0.0041	0.0038	0.0038	0.0037

Table 4.

Simulation Results for Perfect Simple Structure with Miss-Specified and Low Reliability

Component – Factor Structure 4

Conditions			Coefficient α	Stratified α	Maximal Reliability	MD-ω
$\rho = 0.3$						
True Reliability = 0.84	$n = 50$	BIAS	-0.0596	-0.0277	-0.0300	-0.0151
		SE	0.0508	0.0429	0.0483	0.0413
	$n = 200$	BIAS	-0.0497	-0.0201	-0.0234	-0.0102
		SE	0.0215	0.0185	0.0211	0.0182
	$n = 500$	BIAS	-0.0503	-0.0203	-0.0233	-0.0110
		SE	0.0144	0.0127	0.0143	0.0122
	$n = 1000$	BIAS	-0.0480	-0.0184	-0.0215	-0.0094
		SE	0.0100	0.0087	0.0098	0.0084
	$n = 2000$	BIAS	-0.0471	-0.0177	-0.0206	-0.0087
		SE	0.0068	0.0058	0.0069	0.0058
$\rho = 0.5$						
True Reliability = 0.85	$n = 50$	BIAS	-0.0333	-0.0151	-0.0138	-0.0039
		SE	0.0405	0.0369	0.0417	0.0347
	$n = 200$	BIAS	-0.0312	-0.0124	-0.0118	-0.0020
		SE	0.0196	0.0175	0.0184	0.0163
	$n = 500$	BIAS	-0.0300	-0.0113	-0.0109	-0.0013
		SE	0.0113	0.0102	0.0112	0.0095
	$n = 1000$	BIAS	-0.0308	-0.0118	-0.0111	-0.0017
		SE	0.0084	0.0076	0.0083	0.0072
	$n = 2000$	BIAS	-0.0301	-0.0112	-0.0106	-0.0014
		SE	0.0062	0.0055	0.0060	0.0052
$\rho = 0.8$						
True Reliability = 0.88	$n = 50$	BIAS	-0.0201	-0.0139	-0.0104	-0.0026
		SE	0.0334	0.0327	0.0342	0.0296
	$n = 200$	BIAS	-0.0122	-0.0058	-0.0021	0.0041
		SE	0.0148	0.0142	0.0142	0.0128
	$n = 500$	BIAS	-0.0112	-0.0047	-0.0014	0.0046
		SE	0.0089	0.0085	0.0086	0.0079
	$n = 1000$	BIAS	-0.0105	-0.0040	-0.0007	0.0051
		SE	0.0062	0.0059	0.0059	0.0054
	$n = 2000$	BIAS	-0.0107	-0.0043	-0.0012	0.0049
		SE	0.0047	0.0045	0.0045	0.0041

Table 5.

Simulation Results for Complex Structure – Factor Structure 5

| Conditions			Coefficient α	Stratified α	Maximal Reliability	MD-ω
ρ = 0.3						
True Reliability = 0.87	n = 50	BIAS	-0.0424	-0.0046	-0.0167	0.0043
		SE	0.0374	0.0274	0.0330	0.0255
	n = 200	BIAS	-0.0391	-0.0009	-0.0128	0.0064
		SE	0.0192	0.0141	0.0165	0.0132
	n = 500	BIAS	-0.0369	0.0002	-0.0116	0.0069
		SE	0.0125	0.0095	0.0110	0.0090
	n = 1000	BIAS	-0.0371	0.0003	-0.0117	0.0068
		SE	0.0077	0.0060	0.0071	0.0056
	n = 2000	BIAS	-0.0372	0.0004	-0.0115	0.0069
		SE	0.0056	0.0043	0.0050	0.0041
ρ = 0.5						
True Reliability = 0.88	n = 50	BIAS	-0.0306	-0.0053	-0.0122	0.0012
		SE	0.0322	0.0258	0.0296	0.0243
	n = 200	BIAS	-0.0257	-0.0010	-0.0078	0.0036
		SE	0.0154	0.0124	0.0137	0.0120
	n = 500	BIAS	-0.0251	-0.0007	-0.0076	0.0038
		SE	0.0096	0.0078	0.0087	0.0074
	n = 1000	BIAS	-0.0251	-0.0006	-0.0075	0.0037
		SE	0.0068	0.0057	0.0063	0.0054
	n = 2000	BIAS	-0.0246	-0.0003	-0.0072	0.0039
		SE	0.0048	0.0040	0.0045	0.0039
ρ = 0.8						
True Reliability = 0.89	n = 50	BIAS	-0.0148	-0.0055	-0.0069	-0.0013
		SE	0.0267	0.0243	0.0253	0.0231
	n = 200	BIAS	-0.0102	-0.0015	-0.0033	0.0010
		SE	0.0125	0.0115	0.0119	0.0112
	n = 500	BIAS	-0.0090	-0.0004	-0.0024	0.0019
		SE	0.0078	0.0072	0.0076	0.0070
	n = 1000	BIAS	-0.0084	0.0002	-0.0018	0.0023
		SE	0.0054	0.0050	0.0053	0.0049
	n = 2000	BIAS	-0.0089	-0.0003	-0.0023	0.0018
		SE	0.0040	0.0038	0.0039	0.0037

(a) r = 0.3

(b) r = 0.5

(c) r = 0.8
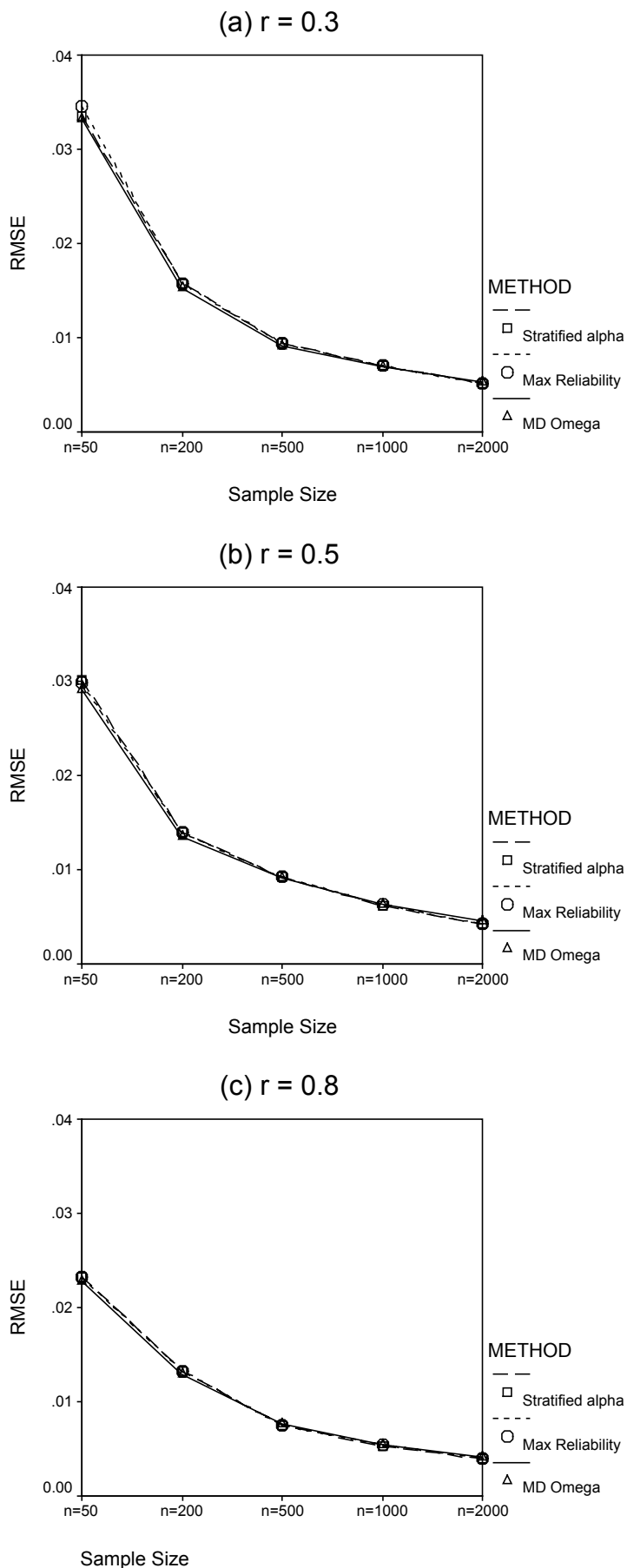
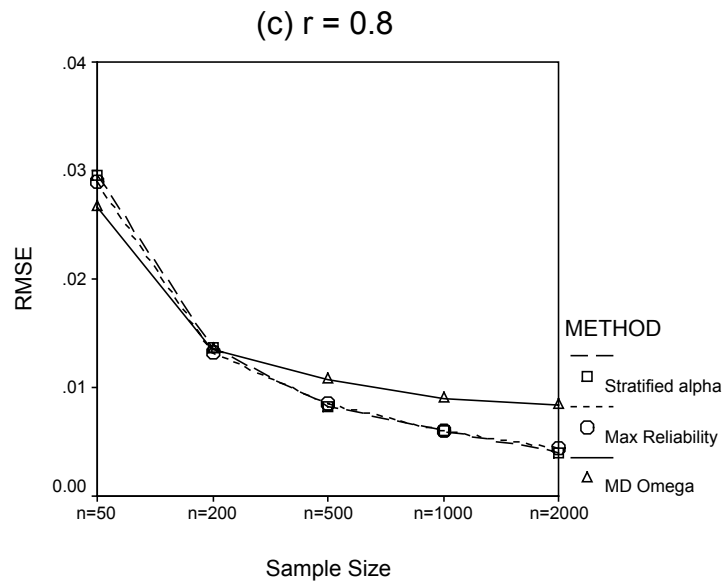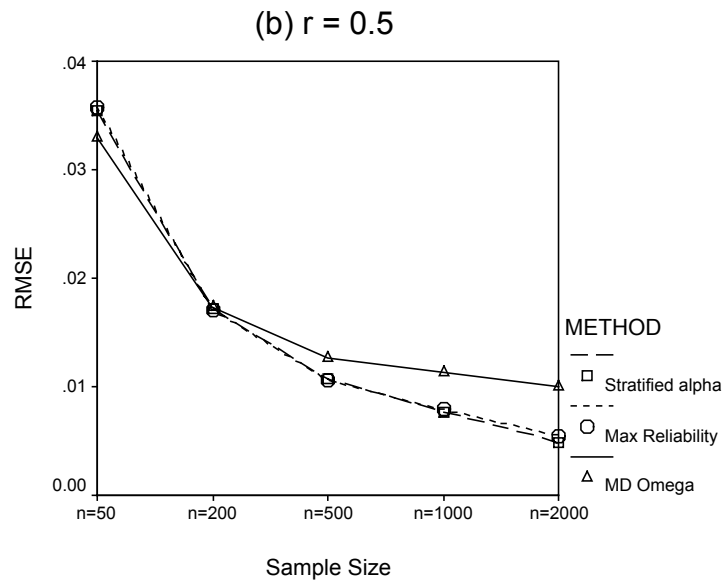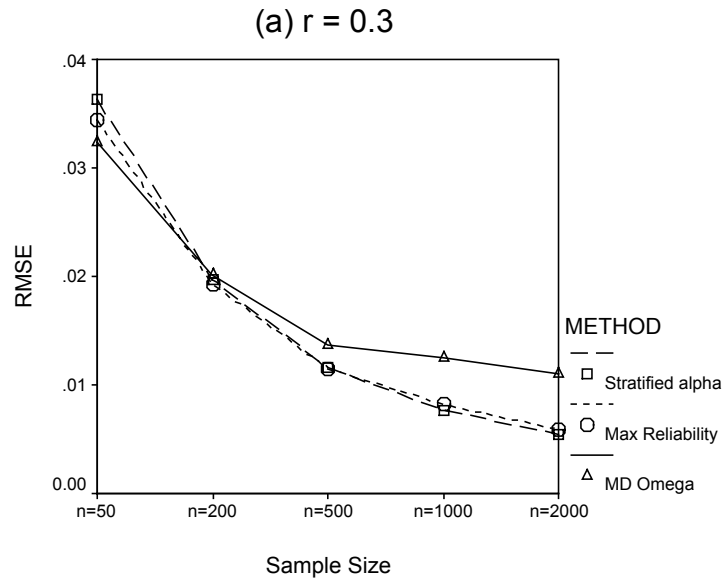Figure 1. RMSE for Perfect Simple Structure – Factor Structure 1

Figure

Figure 2. RMSE for Perfect Simple Structure with Low Reliability for One Component – Factor Structure 2
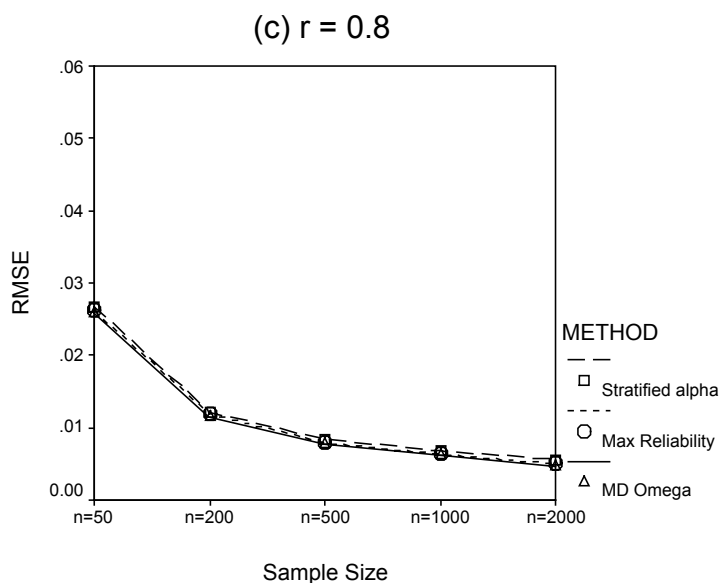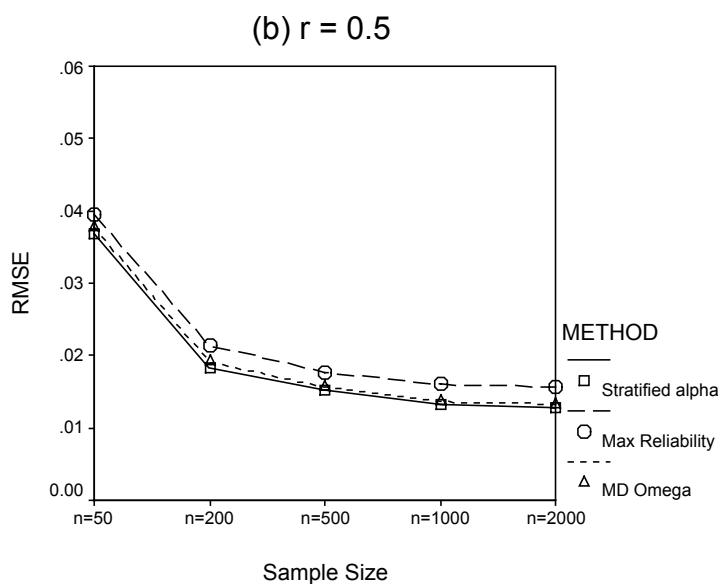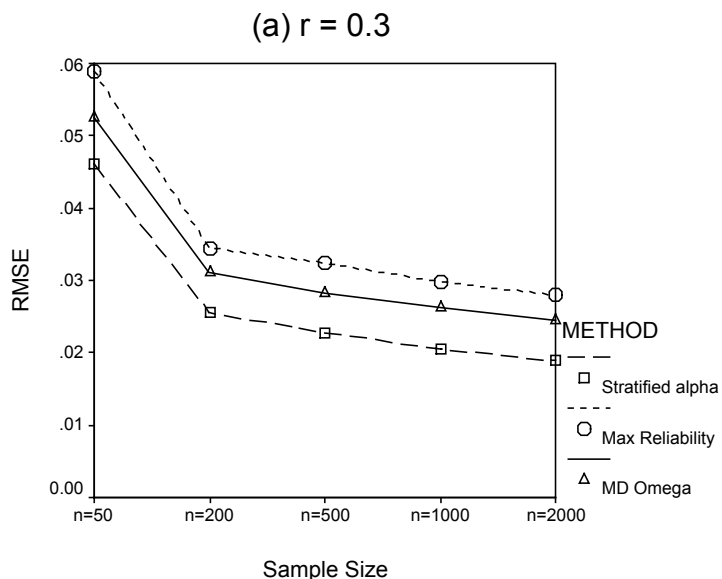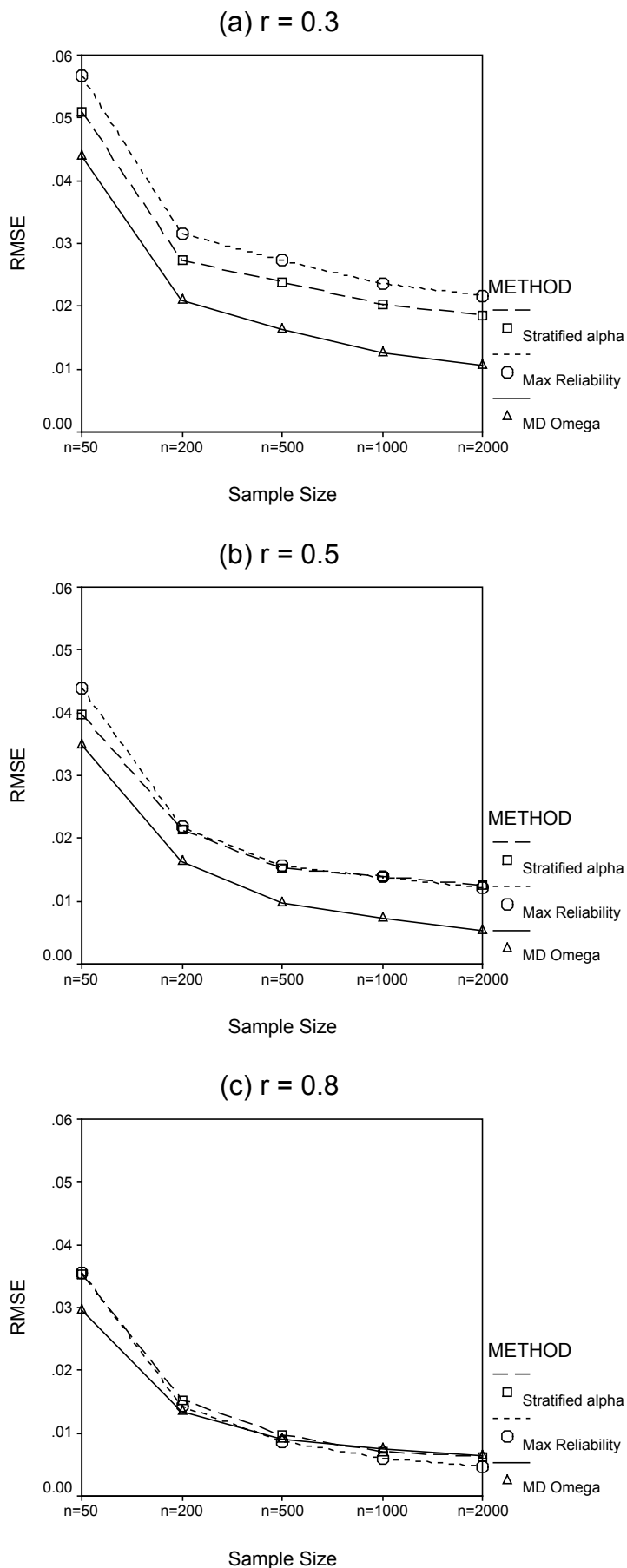
Figure

Figure 3. RMSE for Perfect Simple Structure with Miss-Specified Component – Factor Structure 3

Figure

(a) r = 0.3

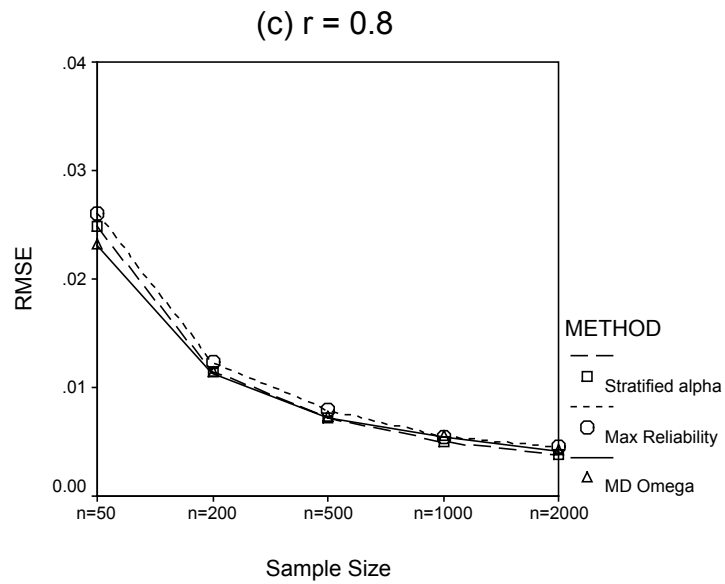

(b) r = 0.5

(c) r = 0.8

Figure 4. RMSE for Perfect Simple Structure with Miss-Specified and Low Reliability Component – Factor Structure 4
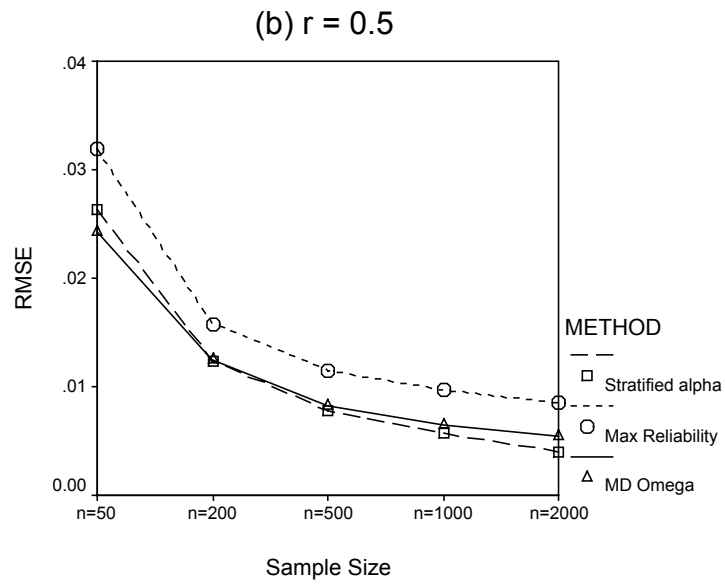
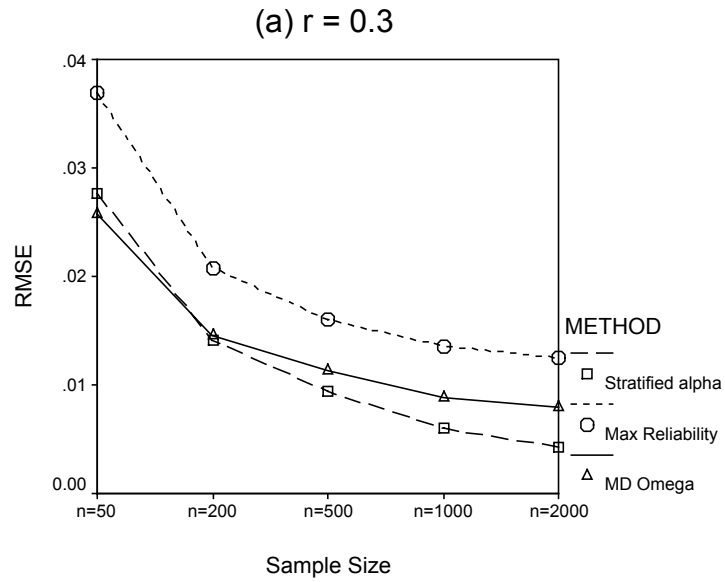Figure

(a) r = 0.3

(b) r = 0.5

(c) r = 0.8

Figure 5. RMSE for Complex Structure – Factor Structure 5

Figure