

**Structural Equation Models with Missing Data:  
Exploring Situations in Which EM Behaves Better Than FIML**

**Xiaoxu Li** The Chinese University of Hong Kong  
**Kit-Tai Hau** The Chinese University of Hong Kong  
**Herbert W. Marsh** Oxford University

Paper presented at the American Educational Research Association Annual Conference, Chicago, IL, April 9-13, 2007. Address all correspondence to [Lixiaoxu@cuhk.edu.hk](mailto:Lixiaoxu@cuhk.edu.hk).

### Abstract

The expectation maximization (EM) and full information maximum likelihood (FIML) procedures have been generally recommended to handle missing data. In this simulation study, we test their robustness with increasing missingness. Results show that FIML behaves unsatisfactorily at high missingness even with  $N$  as large as 4000. With mild missingness, indefinite solutions are found with both FIML and EM solutions at some levels of population discrepancy and only in certain model structures. The effects of population discrepancy will also be discussed. Although inefficient in estimating standard error size, EM is advantageous over FIML in estimating population discrepancy with large sample size, or with reasonably large sample size and a not-close-fit population discrepancy.

## Structural Equation Models with Missing Data: Exploring Situations in Which EM Behaves Better Than FIML

### *Two ML Approaches for Missing Data in SEM*

It is common to have missing data in empirical research involving structural equation models (SEM) and two of the widely used treatment procedures are examined here. Expectation maximization (EM) refers to the powerful algorithm introduced by Dempster, Laird, and Rubin (1977). Although it plays a pivotal role in the development of many other important statistical methods (e.g., the popular HLM, see Raudenbush & Bryk, 2002), it was originated from Rubin's (1976) initial research on missing data (1976). In our SEM context, EM is used to estimate summary statistics that replaces raw data as input for model analyses (Jamshidian & Bentler, 1997).

While full information maximum likelihood (FIML, see Allison, 1987) is EM's competitor and directly use original missing data for SEM analyses, both are maximizing likelihood (ML) procedures. However, they differ in that EM presumes a full population space, while FIML presumes a zero population discrepancy from its specific model subspace. Within the full population space, EM could provide an unbiased estimation of the population discrepancy, but could not provide the unbiased  $p$  value for each assumed population discrepancy bound. FIML, in contrast, could fix the population discrepancy to just-fit, and report the respective  $p$  value.

In contrast to situations with missing data, when data are complete, the RMSEA bound assumption in the population could be used to provide its respective  $p$  value (Li, Hau & Marsh, 2007). Moreover, an unbiased point estimate of the population RMSEA could also be provided. With a sample size  $n$  and  $df$  in a complete data set, any pair of (RMSEA,  $p$ ) could be used to deduce the whole set of (RMSEA  $p$ ) with *any arbitrary* RMSEA criterion, or with any arbitrary  $p$  criterion (Li, Hau & Marsh, 2007).

An obvious problem with missing data is that the sample size  $n$  is inflated considering the empirical data with missing information is of comparatively smaller size. Assuming there are  $n_c + 1$  complete observations, the appropriate rescaled sample size should be between  $n_c$  and the original  $n$ . If the missing observations are few in number relative to the total sample size, then EM, or FIML, or even any traditional method like listwise-deletion or pairwise-deletion, could provide good approximation to the complete data.

The typical problem is that relative to the rather big set of variables, only few variables are missing in each incomplete observation, resulting in a sizable number of cases with few missing variables. The total missing observations percentage remains high. In such a case, neither  $n$  nor  $n_c$  can be used contentedly as a good approximation in calculating the RMSEA- $p$  mapping.

### ***Availability in Software Packages***

Sometimes both procedures are simply called ML methods and are analogous to the multiple imputation (MI) approach as introduced by Rubin (1987, see Rubin, 1996), which is too flexible to be fully realized in any popular SEM-specific software package. For example, LISREL8.72 (Jöreskog & Sörbom, 2003) provides an incomplete MI routine that, without the reliance on another MI-specific software, could only give a single averaged summary statistics as a simple input to the second stage of model analysis, similar to EM. In contrast, the more general-purpose software SAS9.1.3 (SAS Institute Inc., 2003) provided a full package (Proc MI, Proc MIANALYZE) for both MI and EM.

While most SEM-specific packages implement a convenient FIML user interface, the general-purpose SAS9.1.3 (SAS Inc., 2003) does not despite its high programming power in covariance analysis (Proc CALIS). EM is widely available in both general-purpose and SEM-specific software. In fact, most MI and FIML routines by default utilize the efficient EM estimates as starting values (e.g., LISREL8.72).

### ***Comparison of Performance***

Recent literature usually recommend ML and MI approaches over other traditional listwise-deletion, pairwise-deletion, mean imputation, regression imputation procedures in ignorable or non-ignorable missing conditions (Collins, Schafer & Kam, 2001; Enders, 2001, 2006; Enders & Bandalos, 2001; Graham, 2003; Schafer & Gradham, 2002). Usually, deletion of variables could also be [seen-viewed](#) as a viable method when only a small unimportant portion of variables have a high percentage of missingness while all the other variables are complete. Geometrically, listwise-deletion and variable-wise-deletion differ only in their dimension of deletion.

Another more general approach is to delete some rows and some columns to maximize the remaining complete data matrix with important variables. When the condition is Missing Completely at Random (MCAR, see Rubin, 1976), such a deletion method will lose its efficiency or power only without introducing any bias. However, with a typical missing data set, this approach would lose too many partitions of data and subsequently lowering its efficiency substantially. While in contrast, in situations when this method can attain its efficiency, the ML and MI methods would also be [efficient-as-well](#), thus leaving little advantage of this deletion by rows and columns method.

It is also generally concluded that the distinctions among EM, FIML and MI in terms of their bias and efficiency are trivial and ignorable (Enders, 2006; Schafer & Gradham, 2002). These descriptions are typically limited to the point estimates of model parameters, and may not be applicable to the estimates of standard errors (Schafer & Gradham, 2002). Furthermore, in addition to the possible inconvenience in their implementation in some software, there are potential drawbacks and differences

among the three approaches. The present paper concentrates on comparing EM and FIML in the SEM context.

Enders (2006) may be one of the few who explicitly has an overall preference of FIML to EM. However, he also noted the minor drawback of FIML, e.g., Graham's (2003) approach for FIML including auxiliary variables is only applicable for most but not all types of possible models. Besides there are two earlier notes on FIML (Allison, 1987) that were almost ignored recently. Firstly, when the number of missing patterns is very large, enormous resource will be needed for FIML. Secondly, FIML fails for missing patterns with small sample sizes, which are hard to be modeled.

We would examine whether large sample sizes help FIML to compete with EM under high missingness that contains lots of missing patterns and we would analyze the mechanism of advantages of EM over FIML in terms of model population discrepancy.

### Method

There are five experimental variables in our simulation research in a factorial design, 2 Model Type  $\times$  3 Discrepancy Levels  $\times$  8 Levels of Missing Ratio  $\times$  3 Sample Sizes  $\times$  2 Types of Treatments (FIML vs. EM). We have 2 model types both with 8 degree of freedom (df), one being a two-factor CFA, adapted from LISREL8.72 FIML examples ("MissingEx2A.ls8"), the other being a single-factor CFA with correlated errors between Items 1 and 6. Model 2 could be seen as evolved from Model 1, with the two latent factors highly correlated and a method effect simultaneously affecting Items 1 and 6. To construct populations with identical discrepancies on these two different models, we solved the equation of PH(1 2) as the only unknown variable,

$$F_{ML}(\Sigma_0(\text{PH}_{(1\ 2)}, \text{TD}_{(6\ 1)}), \text{Subspace}_{\text{Model 1}}) = F_{ML}(\Sigma_0(\text{PH}_{(1\ 2)}, \text{TD}_{(6\ 1)}), \text{Subspace}_{\text{Model 2}})$$

while  $\text{TD}_{(6\ 1)}$  is given, the 6 loadings are fixed respectively at .626, .529, .732, .872, .825 and .821, and factor variances are fixed to 1.

If both  $\text{PH}_{(1\ 2)}$  and  $\text{TD}_{(1\ 2)}$  are known, the population  $\Sigma_0$  could be constructed with matrix computation. The fitting of either model would result in the respective RMSEA estimates. We explore the range of  $\text{PH}_{(1\ 2)}$  by introducing small incremental steps to minimize the difference between the two RMSEA estimates. The smaller the difference, the closer the approximation the equation is. With three settings of increasing  $\text{TD}_{(6\ 1)}$  (e.g., .077, .115, and .145), those respective  $\text{PH}_{(1\ 2)}$  (e.g., 1 - .07756, 1 - .12132, and 1 - .15916) were selected by the stepping simulation. Then, we construct three population covariances with a 7-df model whose  $\text{PH}_{(1\ 2)}$  and  $\text{TD}_{(6\ 1)}$  are freed. The 3 levels of discrepancy are operationally rescaled to a population

value of RMSEA (.049, .074 and .094). With population covariance limited in the subspace of the 7-df model, our RMSEA is obviously smaller than those generally observed without respective constraints (.139, .209 and .264, computed with  $df_{\text{relative}}=1$ ). The subtle reason is that a discrepancy caused by only one single dimension would have a smaller RMSEA after being averaged out by the other dimensions. The model space with six fixed loadings saves six dimensions and leaves only one, so the  $df_{\text{relative}}$  in the latter should be one rather than seven.

We limit the study to the MCAR condition with missing ratio levels mimicking a general depreciating mechanism with data availability reduces by year independently and multiplicatively. The resulting ratios are in a power series with a constant adjacent ratio  $(.45)^{1/8} = .905$ . We generate  $4^8$  complete data records for each of discrepancy level. Each observation of the original complete data has eight missing versions with missingness being cumulated and embedded across years, while data within any year still satisfying the MCAR condition.

The data is reused and divided into replicates of sample size 256, 1024 and 4096. Thus the respective number of replications is 256, 64 and 16. With SAS9.1.3 (SAS Institute Inc., 2003), Proc MI is used for the EM estimation. The EM covariance estimations and those records with various levels of missingness for FIML approach are saved as input for LISREL8.72 (Jöreskog & Sörbom, 2003).

For cases with both definite proper EM and FIML solutions, we compute two discrepancy estimations as  $\hat{F}_{0,FIML} = \frac{\chi_{FIML}^2}{n} - \frac{df}{n}$  and  $\hat{F}_{0,EM} = F_{EM} - \frac{df}{n}$ , where

$\chi_{FIML}^2$  is the chi-square statistic reported by FIML and  $F_{EM}$  is the discrepancy between EM estimated covariance and the model subspace, labeled as the Minimum Fit Function Value in LISREL8.72. This value would not change even another arbitrary incorrect sample size is provided. To compare the two  $F_0$  estimations, we calculate the

ratio of their roots of mean squared errors denoting as  $r \equiv \sqrt{\frac{\sum (\hat{F}_{0,EM} - F_0)^2}{\sum (\hat{F}_{0,FIML} - F_0)^2}}$ . The

sum is applicable to each qualified replication in each experimental sub-group with different respective sample size, discrepancy level, and missing level. The results of the two types of matching models are aggregated for simplification purpose.

## Result

All Maximal Likelihood Estimations for EM under Proc MI converge within the criteria .005, .001 and .0001 respectively for the three sample sizes. LISREL reports only one non-converging solution for all 16128 replications of EM covariance.

Indefinite solutions are found only in the two-factor model, with frequency proportionate to the level of missingness (2.5%-19.9% respectively in the worst sample size subgroup  $N=256$ ), and inversely to the level of population discrepancy (15.4%,10.6% and 4.5% respectively in the worst sample size  $N=256$ ) and sample size (0% in the best sample size  $N=4096$ ). It suggests that our population for two-factor model are too close to the positive definite boundary where  $PH(1\ 2)=1$ .

While FIML performance dramatically drops when missingness is high (over 39%), strikingly large sample sizes do not help (for  $N=1024$  and  $4096$ , neither model reports any proper solution at the 5<sup>th</sup> or higher level missingness; for  $N=256$ , the best proper solution rate is 12.5% at the highest 4 levels of missingness, while the proper solution rate is 100% at the first 3 levels of missingness). The two model structures and the three discrepancy levels do not behave differently under the above condition. Here operationally, a proper solution is defined as one having a chi-square of likelihood ratio bigger than 0 and less than  $T_{MAX}=CInv(1-(10E-13),df,n*F_0)$  (SAS Institute Inc., 2003), which means the error probability of losing any proper case is less than  $1-(1-(10E-13))^{8*3*2*(256+64+16)}=1.6133E-9$  and is ignorable (see Figure 1, chi-square limited from 0 to  $T_{MAX}$ ). At the better performing first 3 missingness levels, all the FIML indefinite solutions occur in places similarly to those in EM and no indefinite solutions are reported with the single factor model. In fact, among the FIML proper cases, most (>86%) of their indefinite solutions are also respectively indefinite in EM and vice versa. Indefinite cases do not affect the failure of FIML.

The quantities of  $r$  (Table 1) are listed only for missingness levels 1 to 3 where all FIML solutions are proper. For  $n=1023$  at discrepancy levels 2 or 3, or  $n=4095$ ,  $r$  ranges from .3079 to .9989. In other subgroups,  $r$  ranges from 1.1621 to 2.9023. Generally, the larger the sample size or discrepancy is, the smaller  $r$  will become. The higher the missingness level, the larger the absolute value of  $\log(r)$  will become.

### Discussion

Inherent in most simulation studies is the concern whether the results are software-specific or algorithm-specific. Enders (2006) provides detailed differences between FIML or FIML-like approaches under LISREL, Amos, EQS, and Mplus, highlighting the possible different implementations. However, considering the popularity of LISREL, it is still useful to note the inadequacy of FIML at high missingness. Under this particular setting, when FIML fails, LISREL either reports the failure or give obviously improper chi-square without detailed explanation, which susceptibly may be related to Allison's (1987) earlier concern. It also remains unknown whether a more powerful software platform with a larger sample size can help FIML with high missingness.

Applied users usually compare alternative or competitive missing treatment methods and choose the one that performs relatively better. Particularly worth

alarming, when we want to evaluate a certain method, it is conducted under specific conditions when one method outperforms the other. For example, if an estimation of some goodness-of-fit is unbiased unconditionally but biased conditionally on the practical choice of the method among its rivals, researchers would either give up the estimation if the method fails in competition, or obtain a biased estimation when the method outperforms others in this special case. So, if further quantitative attributes of FIML and EM are to be investigated, we should be cautious to exclude the examination of FIML under failure conditions.

FIML seems not as inadequate as documented here because extreme missingness is rare for most practitioners. Moreover, there are further differences between the FIML and EM approaches. In the MI literature, there is a thorough discussion on the imputer's model versus analyst's model (Schafer, 1997, pp.139-143). A parallel comparison can be made between ML procedures on saturated versus analyst's models. As for the "full information" of FIML, it is in fact based on the "model-specific information" from the posited model subspace only, without considering the population discrepancy, as seen in the inverse relation between  $r$  and discrepancy level.

It should also be noted that FIML never impute missing data nor produce the sample covariance matrix, although an estimated population matrix could be deducted after model estimation. So, neither the population discrepancy nor the non-zero RMSEA population could be modeled theoretically in FIML. The chi-square statistic reported by FIML is derived from the likelihood ratio test rather than from the discrepancy between the sample covariance and the model subspace (multiplied with  $n$ ; Allison, 1987; Enders, 2006). Without an ideal rescaling of  $n$ , no statistical hypothesis could be tested on the non-zero population RMSEA, despite that the FIML chi-square has a quantity comparable to that with complete data with a comparable sample size. Consequently researchers are forced either to test a zero RMSEA (just-fit) null hypothesis or to compromise on an unknown bias from the direct usage of  $n$  without appropriate rescaling (Davey, Savla & Luo, 2005).

One fatal shortcoming with the zero discrepancy (true model) approach is that we could only obtain statistical conclusion on rejecting the model at a certain Type I error. Once researchers accept the hypothesis, the power, or the Type II error probability (1 - power) could only be correctly calculated with difficulty for the discrepancy RMSEA population criterion (MacCallum, Browne & Sugawara, 1996). Another consequence is that the zero-RMSEA hypothesis will always be rejected for any sufficiently large sample size. An implication is that FIML faces dilemma similar to that in EM, e.g., there is still no proper estimation of the rescaled sample size. In the Appendix, we analyze the mathematical condition for an estimation of the rescaled sample-size in the EM context. It is unfortunate that in general the rescaling estimation is model-specific. So it is impossible to develop a black box so as to give EM estimation both an estimation of correlation matrix and sample size with EM. To estimate the

rescaled sample size through  $\hat{n} = \frac{\chi_{FIML}^2 - df}{F_{EM}}$ , however, remains an active problem.

Direct usage of the estimation outputs has demonstrated to provide significantly poorer results in comparisons to  $\sum (\hat{F} - F_0)^2$  in our simulation data.

Besides, results suggest that EM is more robust than FIML under extremely high missingness conditions, and if the sample size is sufficiently large relative to the reciprocal of squared RMSEA population, the EM approach can even provide  $\sqrt{\frac{F_{MLEM}}{df}}$

which asymptotically converges to the population RMSEA. Importantly the larger the sample size is, the poorer the  $F_0$  or RMSEA estimation provided by FIML relative to those by EM will become.

The ML theory suggests both EM and FIML will perform successfully with MAR cases as they do with MCAR cases. However, the cases of MNAR are not covered by the ML theory. Neither simulation study nor theoretical investigation can ~~not~~ assure the generality-generalization of MAR to MNAR cases. Generally, it is difficult to test whether a certain data set is MAR in nature. However, some promising methods utilizing the planned missing data designs have been successfully put forward (Graham, Cumsille, Olchowski, 2006; Graham, Hofer, MacKinnon, 1996), making the MAR condition easier to be operationalized.

## References

- Allison, P. D. (1987). Estimation of linear models with incomplete data. *Sociological Methodology*, 17, 71-103.
- 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, 330-351.
- Davey, A., J. Savla, & Luo, Z. (2005) Issues in evaluating model fit with missing data, *Structural Equation Modeling*, 12, 578-597.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, 39, 1-38.
- Enders, C. K. (2001). A primer on maximum likelihood algorithms for use with missing data. *Structural Equation Modeling*, 8, 128-141.
- Enders, C. K. (2006). Analyzing structural equation models with missing data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: a second course*, (pp. 313-342). Greenwich, CT: Information Age Publishing.
- Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, 8, 430-457.
- Graham, J. W. (2003). Adding missing-data relevant variables to FIML-based structural equation models. *Structural Equation Modeling*, 10, 80-100.
- Graham, J. W., Cumsille, P. E., & Olchowski, A. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, 11(4), 323-343.
- Graham, J. W., Hofer, S. M., & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, 31, 197-218.
- Jamshidian, M., & Bentler, P. M. (1999). Using complete data routines for ML estimation of mean and covariance structures with missing data. *Journal of Educational and Behavioral Statistics*, 23, 21-41.
- Jöreskog, K. G., & Sörbom, D. (2003). *LISREL 8.72*. Chicago, IL: Scientific Software International.
- Li, X., Hau, K. T., & Marsh, H. W. (2007). Relation between population RMSEA and p-value: 2-dimensional approach in assessing and reporting goodness of fit in Structural Equation Modeling. Paper to be presented at the annual meeting of the American Educational Research Association, Chicago, IL, 9-13 April.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1, 130-149.

Raudenbush, S. W, & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks, CA: Sage.

Rubin, D. B. (1976). Inference and missing data, *Biometrika*, 63, 581-592.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley. ~~on: Chapman & Hall.~~

Schafer, J. L. & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7, 147–177.



Table 1 EM/FIML Ratio of roots of mean squared errors by different conditions

$r \equiv \sqrt{\frac{\sum (\hat{F}_{0,EM} - F_0)^2}{\sum (\hat{F}_{0,FIML} - F_0)^2}}$		Missing Level		
		1	2	3
n	Discrepancy Level			
255	1	1.4615	2.1329	2.9023
	2	1.2648	1.5957	1.9477
	3	1.1621	1.3467	1.6347
1023	1	1.1845	1.2463	1.3252
	2	0.9989	0.8937	0.8111
	3	0.8729	0.7510	0.6985
4095	1	0.8892	0.7019	0.5835
	2	0.6828	0.4621	0.3348
	3	0.4985	0.3757	0.3079

Appendix: Geometrical Interpretation of RMSEA with a Minimum Fit Function Value from the EM approach

Model subspace is conceptualized as the  $\vec{X}$  axis and the distance between the population covariance  $\Sigma_0$  and the subspace  $\vec{X}$  is  $z_0 \geq 0$ . Set the new orthogonal

coordinates with reference to  $\Sigma_0$  as  $\begin{pmatrix} 0 \\ 0 \\ z_0 \end{pmatrix}$ . In fact,  $\vec{X}$  contains  $(p^* - df)$

dimensions;  $\vec{Y}$  contains  $(df - 1)$  dimensions; only  $\vec{Z}$  contains the common single-dimensional axis.

Denote the sample covariance as  $S$ , which is derived from  $(n + 1)$  observations.

There is a relation between  $S$  and  $\Sigma_0$  such that  $S = \begin{pmatrix} x \\ y \\ z + z_0 \end{pmatrix}$  and that the asymptotic

distribution of  $\sqrt{n} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  is  $N(0, I_{p^*})$ . The distance between  $S$  and  $\vec{X}$  is denoted by

$F = \left\| \begin{pmatrix} 0 \\ y \\ z + z_0 \end{pmatrix} \right\|^2$ . According to the definition of non-central  $\chi^2$  distribution, the

asymptotic distribution of  $n \times F$  is the non-central  $\chi^2(df, n \times z_0^2)$ . While the population value of RMSEA is defined as  $\frac{z_0}{\sqrt{df}}$ .

Denote the EM estimation of the covariance by the missing data of  $(n + 1)$

observations as  $S_{EM}$ . If the relation between  $S$  and  $S_{EM}$  is such that  $S_{EM} = S + \begin{pmatrix} x_d \\ y_d \\ z_d \end{pmatrix}$

and that the asymptotic distribution of  $\sqrt{n_d} \times \begin{pmatrix} x_d \\ y_d \\ z_d \end{pmatrix}$  is  $N(0, I_{p^*})$ , where  $n_d$  is some

unknown value decided by the ability of EM procedure, then with the independence

assumption, the asymptotic distribution of  $\sqrt{n_e} \times \begin{pmatrix} x_d + x \\ y_d + y \\ z_d + z \end{pmatrix}$  is  $N(0, I_{p^*})$ , wherein

$$n_e \equiv \frac{1}{\frac{1}{n_d} + \frac{1}{n}}.$$

—Because  $\|S_{\text{list-deletion}} - S\|^2 \geq \|S_{\text{list-deletionEM}} - S\|^2$ , we have the inequality constrains for  $n_e, n > n_e \geq n_{\text{list-deletion}}$ .

The final conclusion is on the distance between  $S_{EM}$  and  $\vec{X}$ , denoted as

$F_{EM} \equiv \left\| \begin{pmatrix} \mathbf{0} \\ y_d + y \\ z_d + z + z_0 \end{pmatrix} \right\|^2$ . The asymptotic distribution of  $n_e \times F_{EM}$  is the non-central  $\chi^2(df, n_e \times z_0^2)$ .

The most questionable foregoing link is the asymptotic distribution of  $\sqrt{n_d} \times \begin{pmatrix} x_d \\ y_d \\ z_d \end{pmatrix}$ . In most cases,  $y_d$  and  $z_d$  are non-symmetrical along different directions

because different positions of  $S_{EM}$  will often deviate from  $S$  by different scales. Without this asymptotic distribution assumption, we could only calculate

$E(F_{EM}) = \frac{df}{n} + z_0^2 + E \left\| \begin{pmatrix} \mathbf{0} \\ y_d \\ z_d \end{pmatrix} \right\|^2$ . If the EM software package could provide the

estimation of  $E \left\| \begin{pmatrix} x_d \\ y_d \\ z_d \end{pmatrix} \right\|^2$ , we could make use of it in the following inequalities to

estimate the upper and lower bounds of the point estimates of RMSEA.

$$E(F_{EM}) - \frac{df}{n} - E \left\| \begin{pmatrix} x_d \\ y_d \\ z_d \end{pmatrix} \right\|^2 \leq z_0^2 \leq E(F_{EM}) - \frac{df}{n}$$