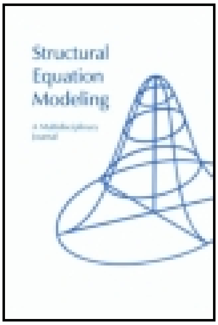

Structural Equation Modeling: A Multidisciplinary Journal
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hsem20

Why might relative fit indices differ between estimators?

Li-Jen Weng [a] & Chung-Ping Cheng [b]
[a] Department of Psychology, National Taiwan University, Taipei, 106, Taiwan, R.O.C. E-mail:
[b] Department of Psychology, National Taiwan University,
Version of record first published: 03 Nov 2009.

STRUCTURAL EQUATION MODELING, *4*(2), 121–128

Why Might Relative Fit Indices Differ Between Estimators?

Li-Jen Weng and Chung-Ping Cheng
Department of Psychology
National Taiwan University

Previous research indicates that relative fit indices in structural equation modeling may vary across estimation methods. Sugawara and MacCallum (1993) explained that the discrepancy arises from difference in the function values for the null model with no further derivation given. In this study, we derive explicit solutions for parameters of the null model. The null model specifies the variances of the observed variables as model parameters and fixes all the covariances to be zero. Three methods of estimation are considered: the maximum likelihood (ML) method, the ordinary least squares (OLS) method, and the generalized least squares (GLS) method. Results indicate that ML and LS yield an identical estimator, which is different from GLS. Function values and associated chi-square statistics of the null model vary across estimation methods. Consequently, relative fit indices using the null model as the reference point in computation may yield different results depending on the estimation method chosen. An illustration example is given and implications of this study are discussed.

Assessing the model-data fit is an essential step in structural equation modeling (SEM). The chi-square goodness-of-fit statistic is a conventional test statistic. Because this test statistic is related to sample size, many fit indices have been developed to aid in model evaluation (e.g., Bentler, 1990; Bentler & Bonett, 1980; Bollen, 1986). Fit indices can be classified as absolute or relative (La Du & Tanaka, 1995; Tanaka, 1993). Relative fit indices use a specific model as the baseline for model comparisons; absolute fit indices do not employ such a baseline model. This study concentrates primarily on cases where the null model (Bentler & Bonett, 1980, p. 596) is relevant in model evaluation.

Requests for reprints should be sent to Li-Jen Weng, Department of Psychology, National Taiwan University, Taipei, Taiwan 106, R.O.C. E-mail: ljweng@ccms.ntu.edu.tw

A null model consists of the variances of the measurement variables as free parameters and fixes all the covariances among measurement variables as zero. The null model is of particular importance to model evaluation in SEM for two reasons. First, many relative fit indices use it as the baseline model when computing the indices (e.g., Bentler, 1990; Bentler & Bonett, 1980; Bollen, 1986; McDonald & Marsh, 1990). Second, many of these indices are frequently used in SEM (Sternberg, 1992; Sugawara & MacCallum, 1993). This study demonstrates that the function values and the chi-square statistics of the null model vary across estimation methods. Estimation methods used in SEM should be considered when model evaluation involves the null model.

Tanaka (1987) observed that the Bentler–Bonett normed fit index (NFI) and nonnormed fit index (NNFI) for an identical model differed under maximum likelihood (ML) and generalized least squares (GLS) methods of estimation. Both NFI and NNFI use the null model as a reference point for calculating the indices. La Du and Tanaka (1989) further investigated the influence of sample size and estimation method on NFI with empirical data set and simulations. Their results indicated that NFI was estimation-specific and the difference between the NFI obtained from ML and GLS estimators decreased as the sample size increased. In a later study, they found similar discrepancies between ML and GLS estimators for other relative fit indices using the null model as the baseline model (La Du & Tanaka, 1995). Sugawara and MacCallum (1993) analyzed a large empirical data set, demonstrating that four incremental fit indices that employed the null model as the reference point in computation varied across different methods of estimation. They offered an explanation for the phenomenon that "it arises from the fact that different estimation methods yield very different discrepancy function values for a null model, due to the differences in the definition of the weight matrix used in the various discrepancy functions" (p. 375). A recent note by Cota, Dion, and Evans (1995) also reported different NFI values for an identical model estimated by ML and GLS methods.

The previously mentioned literature clearly suggests that relative fit indices using the null model as a reference point behave differently across estimation methods. Although Sugawara and MacCallum (1993) provided an explanation for this inconsistency, no mathematical development has been given. In this study, we explicitly solve the parameters of the null model under multivariate normal distribution theory by three estimation methods. The corresponding function values and the chi-square statistics of the model are also derived.

The rest of this article is organized as follows. The fitting functions under normal theory for these three estimation methods are given first, followed by the derivation of the parameter estimators for the null model. Next, the corresponding function values and chi-square statistics for the null model by different estimators are presented. An illustration example is provided next. Finally, implications of this study are discussed.

THREE NORMAL THEORY ESTIMATORS FOR THE NULL MODEL

Three estimation methods are considered in this study: the ML method, the ordinary least squares (OLS) method, and the GLS method. Let Σ and S be positive definite $p \times p$ population and sample covariance matrix, respectively. Let p be the number of observed variables. Tr(A) and |A| represent the trace and the determinant of matrix A, respectively. The normal theory fitting functions by these three estimation methods in structural equation models are as follows (Bentler, 1993).

$$F = \ln|\Sigma| + \mathrm{tr}(S\Sigma^{-1}) - \ln|S| - p \tag{1}$$

$$Q_{LS} = 2^{-1}\mathrm{tr}(S - \Sigma)^2 \tag{2}$$

$$Q_{GLS} = 2^{-1}\mathrm{tr}[(S - \Sigma)S^{-1}]^2 \tag{3}$$

The matrix calculus of Bentler and Lee (1975) are followed in deriving the three estimators. Let X_{diag} represent a $p \times 1$ vector consisting of the diagonal elements of $(p \times p)$ matrix X. Let Diag(X) represent a diagonal matrix containing either the diagonal elements of $(p \times p)$ matrix X or the elements in a $(p \times 1)$ vector. Also, let A*B represent the Hadamard product of matrices A and B of the same order, that is, A*B = $[a_{ij}b_{ij}]$. Taking the first-order partial derivatives with respect to the parameters in the null model (Σ_{diag}) yields the following:

$$\frac{\partial F}{\partial \Sigma_{diag}} = (\Sigma^{-1} - \Sigma^{-1}S\Sigma^{-1})_{diag} \tag{4}$$

$$\frac{\partial Q_{LS}}{\partial \Sigma_{diag}} = (\Sigma - S)_{diag} \tag{5}$$

$$\frac{\partial Q_{GLS}}{\partial \Sigma_{diag}} = (S^{-1} - S^{-1}\Sigma S^{-1})_{diag} \tag{6}$$

The three normal theory estimators of the parameters in the null model (i.e., the variances of the measurement variables) were obtained by setting Equations 4, 5, and 6 equal to zero and simplifying the results. Deriving the ML and LS estimators were straightforward and the derivation of the GLS estimator is given in the Appendix. The explicit solutions by the three methods of estimation are given later. Those results indicate that $\hat{\Sigma}_{ML}$ and $\hat{\Sigma}_{LS}$ are simply sample variances, and $\hat{\Sigma}_{GLS}$ is different from the other two estimators. Also, ML and LS yield more accurate estimators for the null model parameters than GLS.

$$\hat{\Sigma}_{ML} = \mathrm{Diag}(S) \tag{7}$$

$$\hat{\Sigma}_{LS} = \mathrm{Diag}(S) \tag{8}$$

$$\hat{\Sigma}_{GLS} = \mathrm{Diag}[(S^{-1} * S^{-1})^{-1} S^{-1}_{\mathrm{diag}}] \tag{9}$$

FUNCTION VALUES AND CHI-SQUARE STATISTICS OF THE NULL MODEL

Let R be the positive definite $p \times p$ sample correlation matrix, and let s_{ij} be the element of S. As evaluated at the explicitly derived estimators, the ML function value equals $\ln|R|$, the LS function value equals $\sum_{i>j=1}^{p} s_{ij}^2$, and the GLS function value equals $[p - \mathrm{tr}(S^{-1}\hat{\Sigma}_{GLS})]/2$, which is bounded between 0 and $p/2$ (Maiti & Mukherjee, 1991). Detailed derivation for GLS is given in the Appendix. The three estimators yield different function values.

The ML and GLS chi-square statistics equal sample size minus one times the function value evaluated at the obtained parameter estimates. Because the function values at minimum vary across estimators, the resulting chi-square statistics are different accordingly. Consequently, any relative fit index using the null model as the reference point may yield different values depending on the method of estimation employed.

AN EXAMPLE

A sample covariance matrix in EQS (Bentler, 1993, p. 19) was used to demonstrate the calculation of the parameter estimates by Equations 7, 8, and 9, as well as the corresponding function values and chi-square statistics of the null model. The covariance matrix consisted of six observed variables from 932 respondents, taken originally from Wheaton, Muthén, Alwin, and Summers (1977). The SAS IML (SAS, 1993) procedure was employed to compute the results. Similar results were also obtained by EQS Version 4.02 (Bentler, 1993) for comparison.[1]

Table 1 presents the parameter estimates of the null model by different estimators. The parameter estimates calculated from Equations 7, 8, and 9 are identical to those obtained by EQS.

Table 2 summarizes the corresponding function values and associated chi-square statistics.

[1]This comparison serves another purpose. Cota, Dion, and Evans (1995) reported an error in the null model chi-square value produced by the GLS method of EQS Version 3.00. Cota et al. (1995) were informed that the error has been corrected in EQS Version 4.02. The comparison of our results and those by EQS 4.02 would help to verify the correction made in EQS.

TABLE 1
Parameter Estimates of the Null Model by Different Estimators

Estimator / *Parameter*	*ML*	*LS*	*GLS*
σ_{11}	11.834	11.834	4.007
σ_{22}	9.364	9.364	3.306
σ_{33}	12.532	12.532	4.026
σ_{44}	9.986	9.986	3.534
σ_{55}	9.610	9.610	4.789
σ_{66}	4.503	4.503	2.579

Note. ML = maximum likelihood; LS = least squares; GLS = generalized least squares.

TABLE 2
Function Values and Chi-Square Statistics of the Null Model

	ML		*LS*		*GLS*	
Estimator	*IML*	*EQS*	*IML*	*EQS*	*IML*	*EQS*
FV	2.289	2.289	311.979	311.979	0.751	0.751
CS	2131.364	2131.364	—	2131.364	699.636	699.636

Note. ML = maximum likelihood; LS = least squares; GLS = generalized least squares. FV = function value; CS = chi-square statistic.

Table 2 reveals that the function values evaluated at the obtained estimates vary substantially across estimators. EQS 4.02 and Equations 1, 2, and 3 evaluated at the obtain estimated values produced by IML yield identical results. The chi-square statistics again vary dramatically across estimators.[2] The chi-square statistic from ML is three times that from GLS. Sample size minus one times the LS function value does not have a chi-square asymptotic distribution. The resulting value of 290452.449 is different from the statistic given by EQS. Because the least squares method does not use an optimal weight matrix, EQS corrects the associated statistic for LS solution (Bentler, 1993, p. 217). The LS chi-square statistic computed in EQS equals the chi-square statistic for ML minus a term involving the gradient, $\frac{\partial F}{\partial \Sigma_{\text{diag}}}$ evaluated at the ML estimates, which happens to be zero for the null model. Consequently, EQS produces identical chi-square statistics for LS and ML methods.

[2]The identical GLS null model chi-square values by IML and EQS 4.02 support the claim that the GLS error in EQS 3.00 has been corrected.

DISCUSSION

Three normal theory parameter estimators for the null model in SEM were explicitly derived. The ML and LS methods yield an identical estimator, which is different from GLS. The corresponding function values and chi-square statistics are different, as pointed out by Sugawara and MacCallum (1993). Those results clearly demonstrate the reason why relative fit indices using the null model as the baseline model are estimation-specific as shown by previous studies. In view of the derived estimators of the null model by different methods, the relative fit indices by ML may be more appropriate than those by GLS because ML yields better parameter estimates than GLS in terms of its ability to recover the sample variances of the measurement variables. This conjecture deserves further empirical verification.

Although the chi-square values given by these estimators differ, the effects of this discrepancy on relative fit indices may diminish as sample size increases. Let's take NFI as an example. If a model fits the data and the sample size is very large, ML and GLS give very close chi-square statistics (Browne, 1974). Moreover, the null model chi-square value is quite likely to be much greater than that of the proposed model—regardless of the method of estimation used. Under such circumstances, the NFI will be close to 1 regardless of estimation method used. This phenomenon may account for the finding by La Du and Tanaka (1989) that the difference between NFI by ML and GLS decreases as sample size increases. In our example, although the chi-square statistics of the null model for ML and GLS differed substantially, the NFI for the proposed stability of alienation model (Bentler, 1993, p. 34) is .994 and .981 for ML and GLS, respectively. The chi-square statistics for the hypothesized model of 9 degrees of freedom yielded a chi-square statistic of 13.476 for ML and 13.508 for GLS. Further research may be necessary to understand under what conditions the discrepancy between relative fit indices obtained from different estimation methods can be regarded as trivial and the indices can be treated as estimation-free.

The commonly employed rule of thumb for model evaluation with relative fit indices may be required to be used with caution. For instance, a model with NFI less than .90 is normally regarded as inadequate (Bentler & Bonett, 1980; Tanaka, 1987). However, because of the differences in the chi-square values of the null model across estimation methods, experiences may suggest different criteria for evaluating the meaningfulness of the hypothesized model by various estimation methods. More studies are needed to facilitate appropriate use of various goodness-of-fit indices so that researchers in the scientific community can better assess the degree of model-data fit in SEM.

ACKNOWLEDGMENT

We thank three anonymous reviewers for their comments on an earlier version of this article.

REFERENCES

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P. M. (1993). *EQS structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Bentler, P. M., & Lee, S. Y. (1975). Some extensions of matrix calculus. *General Systems, 20,* 145–150.

Bollen, K. A. (1986). Sample size and Bentler and Bonett's nonnormed fit index. *Psychometrika, 51,* 375–377.

Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal, 8,* 1–24.

Cota, A. A., Dion, K. L., & Evans, C. R. (1995). An error in Cota, Dion, and Evans (1993) produced by a GLS error in EQS Version 3.00. *Educational and Psychological Measurement, 55,* 684–686.

La Du, T. J., & Tanaka, J. S. (1989). Influence of sample size, estimation method, and model specification on goodness-of-fit assessments in structural equation models. *Journal of Applied Psychology, 74,* 625–635.

La Du, T. J., & Tanaka, J. S. (1995). Incremental fit index changes for nested structural equation models. *Multivariate Behavioral Research, 30,* 289–316.

Maiti, S. S., & Mukherjee, B. N. (1991). Two new goodness-of-fit indices for covariance matrices with linear structures. British *Journal of Mathematical and Statistical Psychology, 44,* 153–180.

McDonald, R. P., & Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin, 107,* 247–255.

SAS Institute Inc. (1993). *SAS/IML user's guide (Version 6).* Cary, NC: Author.

Sternberg, R. J. (1992). Psychological Bulletin's Top 10 "Hit Parade." *Psychological Bulletin, 112,* 387–388.

Sugawara, H. M., & MacCallum, R. C. (1993). Effect of estimation method on incremental fit indexes for covariance structural models. *Applied Psychological Measurement, 17,* 365–377.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness-of-fit in structural equation models with latent variables. *Child Development, 58,* 134–146.

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–39). Newbury Park, CA: Sage.

Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing reliability and stability in panel models. In D. R. Heise (Ed.), *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey-Bass.

APPENDIX

GLS estimator of the null model and its associated function value are derived as follows.

$$\text{Given } \frac{\partial Q_{\text{GLS}}}{\partial \Sigma_{\text{diag}}} = (S^{-1} - S^{-1}\Sigma S^{-1})_{\text{diag}} \text{ from Equation 6, and let } \frac{\partial Q_{\text{GLS}}}{\partial \Sigma_{\text{diag}}} = 0$$

We have

$$S_{\text{diag}}^{-1} = (S^{-1} * S^{-1})\Sigma_{\text{diag}}$$

And accordingly,

$$\hat{\Sigma}_{\text{diag}} = (S^{-1} * S^{-1})^{-1} S_{\text{diag}}^{-1}$$

or equivalently,

$$\hat{\Sigma}_{\text{GLS}} = \text{Diag}[(S^{-1} * S^{-1})^{-1} S_{\text{diag}}^{-1}]$$

Note that

$$\begin{aligned} Q_{\text{GLS}} &= 2^{-1}\text{tr}(S^{-1} - S^{-1}\Sigma S^{-1})^2 \\ &= 2^{-1}[p - 2\text{tr}(S^{-1}\Sigma) + \text{tr}(S^{-1}\Sigma S^{-1}\Sigma)] \end{aligned}$$

Because under the null model $\text{Diag}(S^{-1}) = \text{Diag}(S^{-1}\hat{\Sigma}_{\text{GLS}}S^{-1})$ and $\hat{\Sigma}_{\text{GLS}}$ is diagonal, we then have

$$\text{tr}(S^{-1}\hat{\Sigma}_{\text{GLS}}) = \text{tr}(S^{-1}\hat{\Sigma}_{\text{GLS}}S^{-1}\hat{\Sigma}_{\text{GLS}})$$

and consequently,

$$Q_{\text{GLS}} = 2^{-1}[p - \text{tr}(S^{-1}\hat{\Sigma}_{\text{GLS}})]$$

Because $\text{tr}(S^{-1}\hat{\Sigma}_{\text{GLS}}) \in (0, p)$, according to Proposition 2.4 of Maiti and Mukherjee (1991), $Q_{\text{GLS}} \in (0, p/2)$ for the null model.