



Taylor & Francis
Taylor & Francis Group



A Review of Eight Software Packages for Structural Equation Modeling

Author(s): A. Narayanan

Source: *The American Statistician*, Vol. 66, No. 2 (MAY 2012), pp. 129-138

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: <https://www.jstor.org/stable/23339473>

Accessed: 07-11-2019 21:09 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Taylor & Francis, Ltd., American Statistical Association are collaborating with JSTOR to digitize, preserve and extend access to *The American Statistician*

A Review of Eight Software Packages for Structural Equation Modeling

A. NARAYANAN

This article reviews eight different software packages for linear structural equation modeling. The eight packages—Amos, SAS PROC CALIS, R packages *sem*, *lavaan*, *OpenMx*, *LISREL*, *EQS*, and *Mplus*—can help users estimate parameters for a model where the structure is well specified. Capabilities for handling single group, multiple group, nonnormal variables, and missing data are considered and the eight packages are compared across a variety of criteria from documentation to parameter estimation. The main difference between the packages is the presence of a graphical interface for model specification and presentation of results. Each package differs in terms of strengths, areas of improvement, and unique features that may dictate the choice of selection. Some suggestions on areas of improvement for all software packages are made. This article has supplementary materials online.

KEY WORDS: Analysis of moment structures; Confirmatory factor analysis; Covariance structure analysis; Latent variable modeling; Linear structural relations; Path analysis.

Software Reviewed:

Amos 18: Available from IBM; phone: +1-312-651-3000; fax: +1-312-675-2140; e-mail: salesbox@us.ibm.com; website: <http://www-01.ibm.com/software/analytics/spss/products/statistics/>

SAS PROC CALIS (version 9.22): Available from SAS Institute Inc., 100 SAS Campus Drive, Cary, NC, 27513-2414, USA; phone: +1-919-677-8000; fax: +1-919-677-4444; e-mail: mccenter@sas.com; website: <http://www.sas.com>

R package *sem* (2.1-1): Available from CRAN R-project website: <http://cran.r-project.org/web/packages/sem/index.html>

R package *lavaan* (0.4-11): Available from CRAN R-project website: <http://cran.r-project.org/web/packages/lavaan/index.html>

A. Narayanan is Principal Scientist, Quantitative Sciences, Procter & Gamble Company, Cincinnati, OH (E-mail: narayanan.a@pg.com) and Adjunct Professor, Department of Operations, Business Analytics and Information Systems, University of Cincinnati, Cincinnati, OH 45221-0130. All program scripts are available from the author via E-mail. The views expressed are those of the author and do not represent those of Procter & Gamble Company. The author thanks the reviewers, the editor, and the associate editor of software reviews for providing comments and suggestions that greatly improved the contents of the article.

R package *OpenMx*: Available from <http://openmx.psyc.virginia.edu/installing-openmx>

LISREL 8.80: Available from Scientific Software International Inc., 7383 N. Lincoln Avenue, Lincolnwood, IL, 60712-1747, USA; phone: +1-847-675-0720; fax: +1-847-675-2140; e-mail: sales@ssicentral.com; website: <http://www.ssicentral.com>

Mplus 6.11: Available from Muthen & Muthen, 3463 Stoner Avenue, Los Angeles, CA, 90066, USA; phone: +1-310-391-9971; fax: +1-310-391-8971; website: <http://www.statmodel.com>

EQS 6.1: Available from Multivariate Software Inc., 15720 Ventura Blvd., Suite 306, Encino, CA, 91436-2989, USA; phone: +1-818-906-0740; fax: +1-818-906-8205; website: <http://www.mvsoft.com>

1. INTRODUCTION

Structural equation modeling (SEM) is a flexible class of models that allows for modeling complex relationships between variables that are either observed (manifest variables) or unobserved (latent variables). It can be thought of as a combination of factor analysis models (called measurement models) and regression models (called structural models). The motivation behind this structure is that manifest variables are usually measured with error and factor analysis models are able to account for this measurement error; then, the true structural relationship between the variables devoid of measurement error is modeled through regression models among the latent variables. In the social and behavioral sciences, such models are sometimes called “causal” models; the terminology is unfortunate since these models do not establish causality but only test proposed relationships between the variables under study (see Freedman 1987). Although much of the development of SEM has taken place in the social and behavioral sciences, SEM research and reviews have been published in mainstream statistics journals and book chapters (see Yuan and Bentler 1997, 2007; Steiger 2001; Sanchez et al. 2005, and references therein).

The software packages reviewed here use what are known as the *LISREL*, reticular action model (RAM), and Bentler–Weeks representations. In the popular *LISREL* notation, the structural model is represented as

$$\eta = \mathbf{B}\eta + \mathbf{\Gamma}\xi + \zeta, \quad (1)$$

where η and ξ are vectors of latent dependent and independent variables, \mathbf{B} and $\mathbf{\Gamma}$ are regression coefficient matrices, and ζ is a

vector of latent error variables. The corresponding measurement equations are given by

$$\mathbf{y} = \mathbf{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon}, \quad (2)$$

$$\mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \quad (3)$$

where \mathbf{y} and \mathbf{x} are vectors of observed variables of latent vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$; $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ are regression coefficient matrices in the respective equations; and $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ are errors in the measurement equations. The model hypothesis can be written as

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}), \quad (4)$$

where $\boldsymbol{\Sigma}$ represents the population covariance matrix, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ represents the model-implied covariance matrix, and $\boldsymbol{\theta}$ represents the vector of all parameters to be estimated that includes the coefficient matrices and the covariance matrices of error terms. To analyze a model, one usually begins with a path diagram relating the observed variables and the latent variables using Equations (2) and (3) and the latent variables to each other using Equation (1). Following model specification and identification, the next step is estimation. To estimate the parameters, a fit function is defined and optimized with respect to $\boldsymbol{\theta}$:

$$Q = (\mathbf{S} - \boldsymbol{\sigma}(\boldsymbol{\theta}))^T \mathbf{W}(\mathbf{S} - \boldsymbol{\sigma}(\boldsymbol{\theta})), \quad (5)$$

where \mathbf{S} is the sample covariance matrix written in vector form, $\boldsymbol{\sigma}(\boldsymbol{\theta})$ is the corresponding vector of the model-implied covariance matrix, and \mathbf{W} is an appropriate weight matrix. In SEM estimation, techniques vary by choice of \mathbf{W} (see Jöreskog and Sörbom 1996a). Another method of model specification that is used in the R packages *sem* and *OpenMx* is the RAM proposed by McArdle (1980) and McArdle and McDonald (1984). In this specification, the manifest variables and the latent variables are organized into a single vector \mathbf{v} and the model is represented as

$$\mathbf{v} = \mathbf{A}\mathbf{v} + \mathbf{u},$$

where \mathbf{A} is the parameter matrix, $\mathbf{I} - \mathbf{A}$ is nonsingular, $\mathbf{u} \sim \mathbf{N}(\mathbf{0}, \mathbf{P})$, and \mathbf{u} and \mathbf{v} are independent. The model-implied covariance matrix is

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{J}(\mathbf{I} - \mathbf{A})^{-1} \mathbf{P}(\mathbf{I} - \mathbf{A})^{-1T} \mathbf{J}^T,$$

where \mathbf{J} is the selection matrix that chooses the observed variables. It is easy to show that this representation is same as the LISREL representation (see Timm 2000).

In the Bentler–Weeks representation used by EQS, the model is represented as

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi},$$

where $\boldsymbol{\eta}$ is a vector of dependent variables (DVs), $\boldsymbol{\xi}$ is a vector of independent variables (IVs) that includes error terms, \mathbf{B} is a matrix of regression coefficients between DVs, and $\boldsymbol{\Gamma}$ is a matrix of regression coefficients among IVs. Variances and covariances among all IVs are estimated. Note that the \mathbf{B} and $\boldsymbol{\Gamma}$ in the Bentler–Weeks model include the measurement component and are not of the same size as the LISREL model. It is interesting to note that if $\boldsymbol{\Gamma} = \mathbf{I}$, the Bentler–Weeks representation reduces to the RAM representation.

1.1 Scope of the Review

SEM has lately been expanded to include several new methods. Generalizing the concept of latent variable by including continuous, categorical, and a combination of the two opens up a range of different models. This includes conventional SEM, growth curve modeling, multilevel modeling, latent class analysis with and without covariates, latent transition analysis, finite mixture modeling, latent profile analysis, and growth mixture modeling. Using a general latent variable framework, Muthén (2002) unified these methods and implementation using a single software system (Mplus), which is unique to these developments. Due to space limitations, this review will focus attention on the conventional SEM using continuous latent variables.

2. THE EIGHT SEM SOFTWARE PACKAGES

The eight software packages considered are

- Amos (version 18), available from <http://www-01.ibm.com/software/analytics/spss/products/statistics/>.
- SAS PROC CALIS (version 9.22), available from <http://www.sas.com>.
- R package *sem* (version 0.9-20), available from <http://cran.r-project.org/web/packages/sem/index.html>.
- R package *lavaan* (version 0.4-11), available from <http://cran.r-project.org/web/packages/lavaan/index.html>.
- R package *OpenMx* (version 1.1.1-1785), available from <http://openmx.psyc.virginia.edu/installing-openmx>.
- LISREL (version 8.80), available from <http://www.ssicentral.com>.
- Mplus (version 6.11), available from <http://www.statmodel.com>.
- EQS (version 6.1), available from <http://www.mvsoft.com>.

While previous reviews in this journal have considered Amos or EQS, this is the first review where all the eight software packages are considered together. This review excludes nontraditional alternatives to SEM such as partial least squares. These eight packages were chosen since they include a mixture of commercial and freeware packages. Only the windows version of all these programs is considered here for review. Further description of each software package can be found in Section 2 of the online supplementary materials.

3. DOCUMENTATION

Documentation for all the software packages reviewed can be found in the following sources: Amos—*Amos User's Guide* and *Amos Programming Reference Guide* (Arbuckle 2009a, 2009b); SAS PROC CALIS—*SAS STAT User's Guide: The CALIS Procedure* (<http://support.sas.com/documentation/cdl/en/statugcalis/63660/PDF/default/statugcalis.pdf>); R package *sem*—Fox (2006) and Fox et al. (2010); R package *lavaan* (Rosseel 2011; Rosseel et al. 2011); R package *OpenMx*—user's guide and reference manual (OpenMx Development Team 2011); LISREL—*LISREL 8: User's Reference Guide* (Jöreskog

and Sorböm 1996a), *PRELIS 2: User's Reference Guide* (Jöreskog and Sorböm 1996c), *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language* (Jöreskog and Sorböm 1996b), *LISREL 8: New Statistical Features* (Jöreskog et al. 2001), and *Interactive LISREL: User's Guide* (du Toit and du Toit 2001); *Mplus—Mplus User's Guide* (Muthen and Muthen 1998–2010); and *EQS—user's guide* (Bentler and Wu 2002) and program manual (Bentler 2006). The documentation for each of these packages was comprehensive and user-friendly. Details of the documentation for each software package can be found in Section 3 of the online supplementary materials.

4. DATA ENTRY AND DATA MANAGEMENT

Data entry and data management capabilities are extensive and varied among the software packages reviewed. The data management capabilities of Amos are very specific and are the most limited, while those of EQS are the most extensive. Details of the capabilities for each software package can be found in Section 4 of the online supplementary materials.

5. CAPABILITIES FOR SINGLE-GROUP ANALYSIS

The focus of this section is to explore the capabilities of the different software packages for conducting a single-group analysis of a general SEM using the observed and latent variables.

5.1 Description of the Data

The data for the single-group analysis illustrate the relationship between political democracy and industrialization. These are panel data with industrialization measured in 1960 and political democracy measured in 1960 and 1965. For industrialization, the indicators in 1960 are gross national product per capita (GNPPC60), energy consumption per capita (ENPC60), and percentage of labor force in industry (INDLF60). For political democracy, the same indicators are used in 1960 and 1965: freedom of press (PRESS60, PRESS65), freedom of political opposition (FREOP60, FREOP65), fairness of elections (FREEL60, FREEL65), and effectiveness of elected legislature (LEGIS60, LEGIS65). The model proposed by Bollen (1989, p. 324) in equating the loading coefficients of political democracy indicators and correlating their error terms is used here for illustration. More details about the capabilities of each software package for single-group analysis can be found in Section 5 of the online supplementary materials.

The results of analyses using all the eight packages for the political democracy data are identical within two decimal places. The only difference seems to be in the error variance estimates where some packages use the default as maximum likelihood (ML) estimates and other packages use unbiased estimates. This has the effect of using N or $N - 1$ in the denominator for variance estimates. Some packages allow the user one to change the default while others do not. For example, in Amos, this can be changed using View → Analysis Properties → Bias → Covariances To Be Analyzed; in lavaan, the likelihood = <argument> can be used when calling the fitting function.

6. CAPABILITIES FOR MODEL EVALUATION

Once the parameters are estimated, the specified model can be tested using the minimum fit function value. Using the unbiased estimator for the sample covariance matrix, the following statistic can be computed for model evaluation for a single-group analysis:

$$T = Q_{\min}(N - 1).$$

Under the appropriate choice of \mathbf{W} in (5) and the correctness of model in (4), T has an asymptotic chi-squared distribution, which can be used to test the proposed model. While the overall test statistic T has been widely used for model evaluation, it has at least two limitations: one is the sensitiveness to sample size where even trivial differences can be amplified by a large sample size leading to the rejection of the proposed model; and the other is the rigid hypothesis testing framework of representing a model that usually has some misspecification. For these and other reasons, there are many other goodness-of-fit indices that have been proposed in the literature. These supplemental fit indices can be broadly divided into absolute fit indices (with no reference model) and incremental fit indices (in comparison with a target model). Hu and Bentler (1999) investigated the performance of several of these indices and recommended a multiindex strategy with appropriate cutoff values for indices in each class. While many users had long waited for such guidelines for model testing, Marsh, Hau, and Wen (2004) pointed out that such guidelines cannot be taken as “golden rules” under all situations. Under such conditions, a safe rule might be to look at several types of fit indices and evaluate the model in totality.

Currently, different software packages report different number of fit indices. While LISREL and SAS PROC CALIS report a broad range of fit indices, Mplus and R packages sem and OpenMx report relatively fewer indices. However, for R users, it is possible to compute these missing indices from the model object, which contain the relevant information. Among the incremental fit indices, naming convention is not standard. While some packages report them as Bollen normed index (Rho1) and Bollen nonnormed fit index (Delta2), other packages report the same indices as relative fit index (RFI) and incremental fit index (IFI). The Tucker Lewis index (TLI) is also sometimes referred to as nonnormed fit index (NNFI). For ML estimation, the model chi-squared value is reported in two different ways. In Mplus and lavaan (default setting), they are reported based on sample size N using the normal likelihood approach, whereas in other packages, they are reported based on $N - 1$ using the Wishart likelihood approach. In LISREL, the independence model chi-squared and the related fit index (such as normed fit index) depend on estimation method and normality assumption. Details on how the selection is made are given in a technical report (<http://www.ssicentral.com/lisrel/techdocs/ftb.pdf>). There is also a discrepancy in the reporting of Hoelter's N , which is the largest sample size for which the model is not rejected. LISREL reports the value at a significance level of 0.01, SAS PROC CALIS reports at a significance level of 0.05, and Amos reports values at significance levels of 0.05 and 0.01. Besides these minor variations, all fit indices (other than information

Table 1. Selected fit indices of several SEM packages for single-group analysis

Fit indices	Amos	SAS PROC CALIS	R package sem	R package lavaan	R package OpenMx	LISREL	Mplus	EQS
Overall indices								
Model chi-square	Y	Y	Y	Y	Y	Y	Y	Y
Baseline chi-square	Y	Y	Y	Y	N	Y	Y	Y
Fit function value	Y	Y	N	N	N	Y	N	Y
Incremental fit indices								
NFI (Delta1)	Y	Y	Y	N	N	Y	N	Y
NNFI (or TLI)	Y	Y	Y	Y	Y	Y	Y	Y
IFI (or Delta2)	Y	Y	N	N	N	Y	N	Y
CFI	Y	Y	Y	Y	Y	Y	Y	Y
Information theory based								
AIC	Y	Y	Y	Y	Y	Y	Y	Y
CAIC	Y	Y	Y	N	N	Y	N	Y
BIC (or SBC)	Y	Y	Y	Y	Y	N	Y	N
General								
GFI	Y	Y	Y	N	N	Y	N	Y
RMSR	Y	Y	N	N	N	Y	N	Y
SRMR	Y	Y	Y	Y	N	Y	Y	Y
Hoelter's <i>N</i> (0.05)	Y	Y	N	N	N	Y	N	N
RMSEA	Y	Y	Y	Y	Y	Y	Y	Y
Parsimony fit indices								
AGFI	Y	Y	Y	N	N	Y	N	Y
PGFI	Y	Y	N	N	N	Y	N	N

CFI, comparative fit index; AIC, Akaike information criterion; CAIC, consistent AIC; BIC, Bayesian information criterion; SBC, Schwarz's Bayesian criterion; GFI, goodness-of-fit index; RMSR, root mean square residual; SRMR, standardized root mean square residual; RMSEA, root mean square error of approximation.

theory-based measures that are relative) match from the different software packages. Table 1 gives a summary of the selected model fit indices and their availability across the eight software packages.

7. CAPABILITIES FOR MULTIPLE-GROUP ANALYSIS

Multiple-group analysis in SEM is used to test for model equivalence across groups. In the measurement context, multiple-group analysis can be used to test for measurement invariance across groups (e.g., do men and women assign the same meaning to measurement of a construct such as intolerance). While all the packages reviewed here (including SAS PROC CALIS) can handle true multiple-group situations by maximizing a composite fit function, the sem package in R can only handle single-group analysis. However, in the special case of equal sample size, the sem package in R can be made to handle multiple-group analysis by creating a block diagonal covariance matrix where the blocks are the individual covariance matrices. The fit function is the sum of the fit in each group. Dividing this function by the number of groups yields a value needed for multiple groups (Evermann 2010). The degrees of freedom and certain fit indices, which are functions of degrees of freedom, need to be recomputed.

7.1 Description of Data for Multiple-Group Analysis

This example is taken from the Amos 18.0 user's guide (example 12). In this example, it is hypothesized that visual perception (visperc), spatial visualization (cubes), and spatial orientation

(lozenges) measure spatial ability (spatial); paragraph completion score (paragrap), sentence completion score (sentence), and word meaning test score (wordmean) measure verbal ability (verbal). Data from 72 girls and 72 boys were taken to test measurement invariance across these two groups (the last observation from the original data for girls was dropped to equalize the sample size across the two groups to accommodate the sem package in R).

A nested hierarchy of models was tested for measurement invariance across the two groups. This ranged from a completely unconstrained model to a completely constrained model of equal loadings, equal factor covariance matrix, and equal error variances. All software packages gave very similar estimates to at least two decimal places accuracy. The same capabilities of single-group analysis (Section 5) are also available in multiple-group analysis in terms of graphical and programming options for all the software packages. Figure S4 (see the online supplementary materials) shows how multiple-group modeling can be completed in LISREL using the path diagram by creating a SIMPLIS project. The top half of the figure shows a part of the SIMPLIS syntax and the bottom half shows a dialog box for setting different constraints for testing different levels of measurement invariance. To facilitate model fit across groups, SAS PROC CALIS provides a comparison table, which is shown in Table S1 (see the online supplementary materials). Such a table is useful to compare fit across groups with overall fit. In addition, LISREL, Mplus, and R package lavaan also report the chi-squared contribution from each group to assess the fit (or misfit) in each group. All comments regarding fit indices made in the previous section on single-group analysis also apply to multiple groups. In LISREL models, the parsimony fit indices [adjusted goodness-of-fit index (AGFI), parsimony

goodness-of-fit index (PGFI)] are not defined in the multiple-group situation.

8. CAPABILITIES FOR HANDLING NONNORMAL VARIABLES

While the ML method of estimation assumes multivariate normality, there are at least three options available in many SEM software packages to handle nonnormal observed variables: (1) asymptotically distribution-free (ADF) estimators, (2) scaling corrections to fit statistics and standard errors based on ML estimation, and (3) bootstrapping.

8.1 ADF Estimator

Given the prevalence of nonnormal data in practice, Browne (1984) developed the ADF estimator that does not make assumptions about normality. While theoretically appealing, there are practical limitations with the ADF estimator. Empirical research has shown that ADF estimator does not perform well unless the sample size is very large ($n \geq 5000$). Also, the method is somewhat computationally demanding since it requires computing the inverse of a weight matrix of size $(p + q)(p + q + 1)/2$, where $p + q$ = number of the observed variables. Due to the practical limitations, ADF estimator has not become a viable alternative.

The ADF estimator is sometimes called the weighted least squares (WLS) approach in many software packages. In LISREL and Mplus, the WLS method of estimation has to be specified, along with continuous observed variables, to obtain ADF estimators. In LISREL, a two-step process is used to obtain the ADF estimator; a preprocessor called PRELIS should be used to obtain the weight matrix using the raw data. In a second step, the weight matrix and the covariance matrix (or raw data) are used to compute the ADF estimator. Table 2 shows the capability of various software packages in terms of ADF estimator and other nonnormal options discussed below. In EQS, there are other classes of estimators to deal with nonnormality: elliptical theory estimators (for distributions that are homogeneously kurtotic), heterogeneous kurtosis theory estimators, and case-robust estimators (differentially weights individual observations). For more details on these estimators, refer to EQS program manual (Bentler 2006).

8.2 Scaling Correction to Fit Statistic and Standard Errors

Another method of dealing with nonnormal data is to scale the fit indices and standard errors of parameter estimates by a scaling factor that is determined by the amount of nonnormality in the data, particularly the kurtosis. This scaling procedure (sometimes called the Satorra–Bentler scaling method) is typically used with ML estimation, thus alleviating the computational problems associated with ADF estimation. In the nonnormal situation, the parameter estimates are usually not biased and are not scaled (Finney and DiStefano 2006). The usual chi-squared difference test for nested model comparison is not applicable here; Satorra and Bentler (2001, 2010) gave formulas to make this comparison using standard results from SEM software. Similar to ADF estimation, the LISREL software requires the two-step approach with a preprocessing step using PRELIS.

8.3 Bootstrapping

A third method of dealing with nonnormality is bootstrapping, where the original data are used as the parent data for repeated sampling with replacement. In SEM, there are two types of bootstrapping: the naïve bootstrapping and the Bollen–Stine bootstrapping. The naïve bootstrapping is the typical bootstrap method used for computing standard errors of parameter estimates. Bollen and Stine (1992) noted that the naïve bootstrapping is not appropriate for the model fit statistic or the associated p -value since the covariance structure in the parent data is not consistent with the null hypothesis. Transformation of the original data is required so that the covariance structure is consistent with the null hypothesis and the generated values reflect only the sampling variability and nonnormality and not model misfit. The empirical distribution of model fit statistic is then used as a reference distribution (instead of the theoretical chi-squared distribution) to obtain the corrected proportion of samples that exceed the computed statistic of the parent sample.

Bootstrapping capability in SEM software is generally not well known and is rarely used. This capability in SEM software packages should be exploited since many linear multivariate models such as regression, simultaneous equation, path analysis, and canonical correlation can be considered as special cases of SEM. Even simple correlation analysis can be conducted as a saturated confirmatory factor analysis (Fan 2003). SEM

Table 2. Table of options for handling nonnormal variables and other types of estimators in SEM packages

Nonnormal options and other estimators	Amos	SAS PROC CALIS	R package sem ^a	R package lavaan	R package OpenMx ^b	LISREL	Mplus	EQS
ADF estimator	Y	Y	N	Y	N	Y	Y	Y
Robust standard error	N	N	N	Y	N	Y	Y	Y
Naïve bootstrap	Y	N	Y	Y	N	Y	Y	Y
Bollen–Stine bootstrap	Y	N	N	Y	N	N	Y	Y
Other estimators								
ML	Y	Y	Y	Y	Y	Y	Y	Y
Unweighted least squares	Y	Y	N	N	N	Y	Y	Y
Generalized least squares	Y	Y	Y	Y	N	Y	Y	Y
Diagonally weighted least squares	N	Y	N	N	N	Y	Y	N
Two-stage least squares	N	N	Y	N	N	Y	N	N

^aIn sem, the user can supply an objective function to be minimized.
^bIn OpenMx, package users can write their own objective functions.

software packages vary in the degree of automation with which bootstrap capability is implemented. Amos and EQS seem to offer a number of choices regarding bootstrap capability. To perform bootstrap using Amos graphics mode, after the model is specified, choose View → Analysis Properties and select the bootstrap tab. The resulting window with different options is shown in Figure S5 (see the online supplementary materials). The Perform Bootstrap option should be checked and when the estimates are computed, the bootstrap computations take place at the same time. Nonconvergent solutions are automatically rejected until valid bootstrap samples for the requested number are obtained. The results can be viewed under View → Text Output.

In EQS, bootstrap can be done using the simulation option of the Build.EQS menu. The options under Type of Simulation include regular bootstrap (naïve), model-based bootstrap (Bollen–Stine), and jackknife. Once the options are selected, the Build.EQS menu automatically populates the syntax window. In LISREL, bootstrapping can be done using Statistics → Bootstrapping and choosing Output Options to save the results. In lavaan, bootstrapping can be done using the bootstrapLavaan function. In Mplus, bootstrapping can be done using the BOOTSTRAP option in the ANALYSIS command. In SAS PROC CALIS, bootstrapping is not implemented, although the naïve bootstrapping can be implemented via user-written macros.

9. CAPABILITIES FOR HANDLING MISSING DATA

A practical challenge when using multivariate data is handling missing data issues. Similar to bootstrapping capabilities, SEM software packages offer a number of options for handling missing data. These include not only the traditional methods such as listwise and pairwise deletion, but also more modern methods such as full information maximum likelihood (FIML) and multiple imputation (MI). Since many of the standard linear multivariate models can be cast in a latent variable framework, SEM software packages offer an ideal platform to handle missing data. The latent variable modeling framework also allows for easy incorporation of auxiliary variables into the model, which helps mitigate bias and improve efficiency in the presence of missing data (Enders 2006).

9.1 Full Information Maximum Likelihood (FIML)

Under the assumption of normality, this method computes ML estimates in the presence of missing data. Amos was one of the first software packages to implement FIML in SEM (Arbuckle 1996) and since then, many other SEM packages have implemented this option. Missing data introduce some nuances into FIML estimation. To begin with, raw data are needed for computing individual log-likelihoods as opposed to inputting a covariance matrix, which is typical in complete data analysis. Second, standard error computations require special attention; observed information and expected information can produce diverging results and studies have shown observed information to be a better choice for FIML estimation under the missing at random (MAR) mechanism (Enders 2010).

An advantage of using SEM software packages for missing data is the ability to incorporate auxiliary variables into the model. An auxiliary variable is a variable that is not of substantive interest in the analysis but is a potential correlate of missingness in the model. Inclusion of this variable helps to satisfy the assumption of MAR and also helps in mitigating the bias and improving efficiency (Enders 2010). Among the packages reviewed here, Mplus and EQS are the only packages that have the ability to include auxiliary variables in the analysis for handling missing data.

Many of the other packages also make the FIML missing data analysis fairly straightforward with minimal changes compared with a complete data analysis. For example, using the graphics mode in Amos18, once the model is specified, FIML can be implemented by selecting View → Analysis Properties → Estimation Tab (choose Means and Intercepts) → Output Tab (choose Observed Information Matrix). In LISREL, FIML is implemented by default whenever the raw data matrix is analyzed and a missing value code is specified. LISREL also supports a variant of hot-deck imputation by matching the response patterns of incomplete and complete data. The PRELIS manual (Jöreskog and Sorböm 1996c) gives complete details of the algorithm. In EQS, FIML can be chosen with the option MISSING = ML in the /SPECIFICATIONS section in the syntax (or via Build.EQS → Title/Specification → Missing data handling). EQS also offers a range of missing value options available, including listwise, pairwise, regression, stochastic regression imputation, and diagnostic plots of missing data patterns. SAS 9.22 and later offer FIML as an option within PROC CALIS, which is easy to implement. With SAS 9.3, PROC CALIS also offers additional missing data diagnostics such as coverage proportions and dominant missing patterns and their summary statistics, which could help identify problematic variables. In lavaan, FIML can be invoked using the missing = <FIML> option. In Mplus, type = missing in the ANALYSIS section invokes the FIML method.

9.2 Multiple Imputation (MI)

Unlike FIML, MI creates multiple copies of the data by filling in the missing values in each copy with different estimates. These multiple copies of complete datasets are analyzed separately and the resulting estimates are pooled to a single estimate using rules outlined by Rubin (1987). SEM software packages differ in the number of options available in the three stages of imputation, analysis, and pooling.

The analysis and pooling phases are currently offered only in Mplus, which seems to be the only software to offer all three phases of MI. Currently, LISREL and Amos offer the option of generating multiply imputed datasets that can be read into Mplus or other programs. In Amos, MI can be done using Analyze → Data Imputation. Amos 18.0 offers regression imputation, stochastic regression imputation, and Bayesian imputation. Amos offers imputation based on any constrained model such as a confirmatory factor model. According to Amos 18.0 user's guide, the estimates will be more accurate than the saturated model and estimates of latent variable scores are also obtained. The analysis and pooling phases are not automated in Amos. In

Table 3. FIML, MI, and complete data estimates for various SEM packages using girls' data

Parameters	Amos	SAS PROC CALIS	R package lavaan	LISREL	Mplus	EQS	Complete data results	MI using SAS
spatial → visperc	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
spatial → cubes	0.511	0.512	0.511	0.512	0.511	0.499	0.610	0.511
spatial → lozenges	1.047	1.047	1.047	1.047	1.047	0.956	1.198	1.017
verbal → paragraf	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
verbal → sentence	1.259	1.259	1.259	1.259	1.259	1.257	1.334	1.245
verbal → wordmean	2.14	2.14	2.14	2.14	2.14	2.188	2.234	2.176
Variance parameters								
visperc	18.776	18.784	18.775	18.778	18.776	15.305	23.872	16.095
cubes	8.034	8.035	8.034	8.034	8.034	7.961	11.601	9.059
lozenges	36.625	36.620	36.625	36.625	36.625	36.60	28.275	37.008
paragraf	2.825	2.824	2.825	2.825	2.825	2.813	2.834	2.929
sentence	7.875	7.876	7.875	7.875	7.875	8.019	7.967	8.503
wordmean	22.677	22.676	22.677	22.678	22.677	22.467	19.925	23.854
spatial	29.563	29.549	29.563	29.559	29.563	34.259	23.303	34.999
verbal	10.814	10.814	10.814	10.814	10.814	10.757	9.682	11.210
Covariance parameter								
spatial with verbal	7.993	7.992	7.993	7.993	7.993	7.956	7.315	8.005
Intercepts								
visperc	28.885	28.884	28.885	28.885	28.885	28.883	29.315	28.809
cubes	24.998	24.998	24.998	24.998	24.998	25.154	24.699	25.125
lozenges	15.153	15.153	15.153	15.153	15.153	14.962	14.836	15.003
paragraf	10.987	10.987	10.987	10.987	10.987	10.976	10.589	10.971
sentence	18.864	18.864	18.864	18.864	18.864	18.802	19.301	18.759
wordmean	18.097	18.097	18.097	18.097	18.097	18.262	18.014	18.304

LISREL, MI can be done once the raw data are converted into a PRELIS System File by choosing Statistics → Multiple Imputation. The imputed values can be stored in an external file and multiple copies of imputed data are stacked into the same file. Similar to Amos 18.0, the analysis and pooling phases are not automated; however, the estimates based on any constrained model can be output to a separate file and simple summary statistics of these estimates can be computed. Details are given in technical documents available at the vendor's website (<http://www.ssicentral.com/lisrel/resources.html>).

Another option is to use SAS PROC CALIS in combination with SAS PROC MI (to generate imputed datasets) and SAS PROC MIANALYZE to analyze and pool the results (Zhang and Yung 2011). While there is some extra work involved in programming these steps, this option allows all three phases, including graphical convergence diagnostics and multiple parameter significance testing, which makes SAS a complete package for MI using SEM.

Finally, the assessment of fit when pooling results from multiple datasets is an area that has not been well explored. Consequently, there are no rules for assessing fit other than combining the likelihood ratio tests as outlined by Enders (2010). In terms of implementation, MI is more involved than FIML, which is usually invoked as an option or a method of estimation. However, MI is more general since the imputed datasets are not tied to SEM or any specific model. For a detailed comparison of the two methods in the context of SEM, see Enders (2010).

Table 3 gives a comparison of MI and FIML estimates for the girls' data of the multiple-group analysis in Section 7. The MI estimates were computed using SAS PROC MI (for imputation), PROC CALIS (for analysis), and PROC MIANALYZE (for pooling) based on 50 multiply imputed datasets. The data are

taken from Amos user's guide (example 17) where missing values were generated at random. The estimates from complete data are also shown for comparison. Given that approximately 27% of the data are missing and only 7 out of 73 cases had complete data, the MI and FIML estimates are not too far away from the complete data analysis for most of the estimates. With missing data, some of the fit indices that depend on the saturated model and the independence model are not reported in some packages.

10. CONCLUSIONS

This article has presented a review of eight software packages for linear SEM, including capabilities for handling single-group analysis, multiple-group analysis, nonnormal variables, and missing data handling. All packages produce estimates that are fairly close to each other in accuracy (up to two decimals). The main difference between the packages comes in the user interface and the availability of different options; while Amos, LISREL, and EQS can be used in programming mode and graphical interface, the rest of the packages can only be used in a programming environment. SAS PROC CALIS and lavaan currently do not render path diagram output. With R packages sem and OpenMx, it is possible to produce a path diagram through graphviz or other third-party applications.

Each package has its own strengths and special features as shown in Table 4. Users of SEM may choose one package over the other based on the special features needed for a particular application or the computing environment they are already familiar with based on other applications. While the R packages (sem, lavaan, and OpenMx) are attractive for users who prefer open source environments, rapid developments are also taking

Table 4. Comparison of strengths and unique features of eight structural equation modeling software

Software	Strengths	Special features
Amos	Excellent graphical interface Well-organized and quickly accessible output format	Bayesian estimation in a limited way Several options for bootstrapping Specification search in the absence of theory
SAS PROC CALIS	Integrating other analysis capabilities available in SAS (e.g., MI for missing data analysis)	Flexibility of model specification in many different formats such as COSAN, FACTOR, LINEQS, MSTRUCT, and RAM
R package sem	Open source environment with complete features of R including graphics capabilities	Incorporation of two-stage least squares for econometric modeling; ability to create path diagram via graphviz
R package lavaan	Open source environment with complete features of R including graphics capabilities Compact scripting even for complicated models	Ability to compute two different likelihoods for ML estimation (normal and Wishart) Ability to “mimic” Mplus and EQS results
R package OpenMx	Open source environment with complete features of R including graphics capabilities Integrates with other R packages to enable multicore processing and large-scale distributed processing	Ability to optimize user-specified objective functions Hierarchical model definition to link a series of models in a tree structure Ability to use the program as an optimizer and matrix algebra calculator
LISREL	General linear modeling software for SEM, multilevel modeling, generalized linear modeling, and recursive modeling	Estimation of observational residuals in factor analysis and structural equation models
Mplus	Integrated modeling framework to handle continuous, categorical, observed, and latent variables; incorporation of latest missing data handling methods	Incorporation of auxiliary variables for missing data handling. Easy implementation of interaction effects among latent variables. Unique capability including Bayesian SEM, exploratory SEM, handling of count and censored values, and missing data under not missing at random (NMAR) mechanism
EQS	Structural modeling made simpler by offering different ways to create models; more capability for exploratory analysis	Unique capability for handling nonnormal variables in complete and missing data situations Ability to compute multivariate Lagrange multiplier and Wald tests Estimates of reliability for the measurement model

place in commercial software packages (particularly Mplus) as observed in Table 4. A unified approach could take advantage of the strength of each environment.

Software packages for SEM have some room for improvement. Three areas of opportunity involve help with model identification, model misspecification, and comparison of models. A first area is when a model is not identified, many of these programs offer less help to the user to fix the problem. While identification can be thought of a fundamental prerequisite to estimation in fitting structural equation models, many users on a limited budget with only a handful of measured variables are likely to face this situation. Pointing to the right source of identification issue would be helpful to users. Also, different programs have different ways of handling inadmissible values (von Eye and Fuller 2006). A second area of opportunity is model misspecification. In the absence of rigorous theory, this is likely to happen more often than nonconvergence of well-specified models. SEM programs could implement some of the algorithms of graphical models as outlined in Haughton, Kamis, and Scholten (2006) by possibly identifying equivalent models that are consistent with the covariance structure of the data. A third area of opportunity is in comparison of models. In many commercial packages, there is no easy way to store the results of several models along with their fit statistics for comparison. This is useful for exploratory analysis. In many commercial software packages, for nested model comparison, the user has to do programming to compute the difference in chi-squared values and degrees of freedom. The Manage Models feature available in Amos for multiple-group analysis is along this direction. The same idea would also be helpful to compare models in

single-group analysis. The `anova()` function in R is also useful in this respect.

The MplusAutomation package in R (Hallquist 2011) is useful for postprocessing results from Mplus by extracting and tabulating parameters and fit indices. The package can read output from several directories, extract summary information, and display them in a concise tabular form. An example of such a display for a series of nested models testing measurement invariance from the multiple-group example in Section 7 is shown in Table S2 (see the online supplementary materials). The `reqs` package is a similar interface between R and EQS to call, execute, and read results from EQS into R for further processing. Mair, Wu, and Bentler (2010) showed how to combine the best of both “worlds” by studying the effect of normality violations on the distributions of test statistics and power analysis for a structural parameter. These examples show how to combine the powerful computational and graphical capabilities of R with the unique implementation in some of the SEM software packages. A trend in the future may be more integration of SEM software packages with R statistical computing environment. Such integration would allow for methodological developments to take place rapidly.

SUPPLEMENTARY MATERIALS

Section 2: Eight structural equation modeling (SEM) packages (contd.)

Section 3: Documentation

Section 4: Data management and data entry

Section 5: Capabilities for single-group analysis (contd.)

Figure S4. SIMPLIS project from path diagram for multiple-group analysis using

LISREL 8.80.

Figure S5. Available bootstrap options in Amos 18.0.

Table S1. Fit comparison among groups for boys' and girls' data from SAS PROC CALIS

Table S2. Mplus automation summary table for boys' and girls' data

[Received November 2010. Revised March 2012.]

REFERENCES

- Arbuckle, J. L. (1996). "Full Information Estimation in the Presence of Incomplete Data," in *Advanced Structural Equation Modeling: Issues and Techniques*, eds. G. A. Marcoulides and R. E. Schumaker, Mahwah, NJ: Erlbaum, pp. 243–277. [134]
- (2009a), *Amos 18.0 User's Reference Guide*, Chicago: SPSS, Inc. [130]
- (2009b), *Amos 18.0 Programming Reference Guide*, Chicago: SPSS, Inc. [130]
- Bentler, P. M. (2006), *EQS 6 Structural Equations Program Manual*, Encino, CA: Multivariate Software, Inc. [130,133]
- Bentler, P. M., and Wu, E. J. C. (2002), *EQS 6 for Windows User's Guide*, Encino, CA: Multivariate Software, Inc. [130]
- Blunch, N. J. (2008), *Introduction to Structural Equation Modeling Using SPSS and Amos*, Boca Raton, FL: Chapman & Hall/CRC. [Supplement]
- Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., Spies, J., Estabrook, R., Kenny, S., Bates, T., Mehta, P., and Fox, J. (2011), "OpenMx: An Open Source Extended Structural Equation Modeling Framework," *Psychometrika*, 76 (2), 306–317. [Supplement]
- Bollen, K. A. (1989), *Structural Equations With Latent Variables*, New York: Wiley. [131]
- Bollen, K. A., and Stine, R. A. (1992), "Bootstrapping Goodness of Fit Measures in Structural Equation Modeling," *Sociological Methods and Research*, 21, 205–229. [133]
- Browne, M. W. (1984), "Asymptotic Distribution Free Methods in the Analysis of Covariance Structures," *British Journal of Mathematical and Statistical Psychology*, 37, 62–83. [133]
- du Toit, M., and du Toit, S. H. C. (2001), *Interactive LISREL: User's Guide*, Lincolnwood, IL: Scientific Software International, Inc. (not in print) [130]
- Enders, C. (2006), "Analyzing Structural Equation Models With Missing Data," in *Structural Equation Modeling: A Second Course*, eds. G. R. Hancock and R. O. Mueller, Greenwich, CT: Information Age, pp. 315–344. [134]
- (2010), *Applied Missing Data Analysis*, New York: Guilford Press. [134,135]
- Evermann, J. (2010), "Multiple Group Analysis Using the SEM Package in the R System," *Structural Equation Modeling*, 17 (4), 677–702. [132]
- Fan, X. (2003), "Using Commonly Available Software for Bootstrapping in Both Substantive and Measurement Analysis," *Educational and Psychological Measurement*, 63 (1), 24–50. [133]
- Finney, S., and DiStefano, C. (2006), "Non-Normal and Categorical Data in Structural Equation Modeling," in *Structural Equation Modeling: A Second Course*, eds. G. R. Hancock and R. O. Mueller, Greenwich, CT: Information Age, pp. 269–314. [133]
- Fox, J. (2006), "Structural Equation Modeling With the SEM Package in R," *Structural Equation Modeling*, 13 (3), 465–486. [130]
- Fox, J., Nie, Z., Byrnes, J., Culbertson, M., Friendly, M., Kramer, A., and Monette, G. (2010), "SEM: Structural Equation Models," R package version 0.9-20 [online]. Available at <http://CRAN.R-project.org/package=sem> (accessed July 24, 2010). [130]
- Freedman, D. A. (1987), "As Others See Us: A Case in Path Analysis" (with discussion), *Journal of Educational Statistics*, 12, 101–129. [129]
- Hallquist, M. (2011), "MplusAutomation: Automating Mplus Model Estimation and Interpretation," R package version 0.5 [online]. Available at <http://cran.r-project.org/web/packages/MplusAutomation/index.html> (accessed March 2, 2012). [136]
- Houghton, D., Kamis, A., and Scholten, P. (2006), "A Review of Three Directed Acyclic Graphs Software Packages: MIM, Tetrad, and Win Mine," *The American Statistician*, 60 (3), 272–286. [136]
- Hu, L., and Bentler, P. M. (1999), "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, 6 (1), 1–55. [131]
- Jöreskog, K. G., and Sorböm, D. (1996a), *LISREL 8: User's Reference Guide*, Lincolnwood, IL: Scientific Software International, Inc. [130]
- (1996b), *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*, Lincolnwood, IL: Scientific Software International, Inc. [130]
- (1996c), *PRELIS 2: User's Reference Guide*, Lincolnwood, IL: Scientific Software International, Inc. [130,134]
- Jöreskog, K. G., Sorböm, D., du Toit, S. H. C., and du Toit, M. (2001), *LISREL 8: New Statistical Features* (3rd Printing with revisions), Lincolnwood, IL: Scientific Software International, Inc. [130]
- Mair, P., Wu, E., and Bentler, P. M. (2010), "EQS Goes R: Simulations for SEM Using Package REQS," *Structural Equation Modeling*, 17, 333–349. [136]
- Marsh, H. W., Hau, K.-T., and Wen, Z. (2004), "In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings," *Structural Equation Modeling*, 11 (3), 320–341. [131]
- McArdle, J. J. (1980), "Causal Modeling Applied to Psychonomic Systems Simulation," *Behavior Research Methods and Instrumentation*, 12, 193–209. [130]
- McArdle, J. J., and McDonald, R. P. (1984), "Some Algebraic Properties of the Reticular Action Model," *British Journal of Mathematical and Statistical Psychology*, 37, 234–251. [130]
- Muthen, B. O. (2002), "Beyond SEM: General Latent Variable Modeling," *Behaviormetrika*, 29, 81–117. [130]
- Muthen, L. K., and Muthen, B. O. (1998–2010), *Mplus User's Guide* (6th ed.), Los Angeles, CA: Muthen & Muthen. [130]
- OpenMx Development Team (2011), "OpenMx Documentation," Release 1.1.1-1785 [online]. Available at <http://openmx.psyc.virginia.edu/docs/OpenMx/latest/>. [130]
- Rossee, Y. (2011), "lavaan: An R Package for Structural Equation Modeling and More," version 0.4-11 (BETA) [online]. Available at <http://lavaan.ugent.be> (accessed December 24, 2011). [130]
- Rossee, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Arminger, G., and Merkle, E. (2011), "lavaan : Latent Variable Analysis," R package version 0.4-9 [online]. Available at <http://CRAN.R-project.org/package=lavaan> (accessed July 1, 2011). [130]
- Rubin, D. B. (1987), *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley. [134]
- Sanchez, B. N., Budtz-Jorgensen, E., Ryan, L. M., and Hu, H. (2005), "Structural Equation Models: A Review With Applications to Environmental Epidemiology," *Journal of American Statistical Association*, 100, 1443–1455. [129]
- Satorra, A., and Bentler, P. M. (2001), "A Scaled Difference Chi-Square Test Statistic for Moment Structure Analysis," *Psychometrika*, 66, 507–514. [133]
- (2010), "Ensuring Positiveness of the Scaled Difference Chi-Square Test Statistic," *Psychometrika*, 75, 243–248. [133]

- Steiger, J. H. (2001), "Driving Fast in Reverse: The Relationship Between Software Development, Theory and Education in Structural Equation Modeling," *Journal of American Statistical Association*, 96, 331–338. [129]
- Timm, N. (2000), *Applied Multivariate Analysis*, New York: Springer. [130]
- von Eye, A., and Fuller, B. E. (2006), "A Comparison of the SEM Software Packages Amos, EQS, and LISREL," in *Structural Equation Modeling: Applications in Ecological and Evolutionary Biology*, ed. B. Pugsek, New York: Cambridge University Press, pp. 355–391. [136]
- Wothke, W. (1993), "Nonpositive Definite Matrices in Structural Modeling," in *Testing Structural Equation Model*, eds. K. Bollen and S. Long, Newbury Park, CA: Sage, pp. 256–293. [Supplement]
- Yuan, K. H., and Bentler, P. M. (1997), "Mean and Covariance Structure Analysis: Theoretical and Practical Improvements," *Journal of American Statistical Association*, 92, 767–774. [129]
- (2007), "Structural Equation Modeling", in *Handbook of Statistics, Volume 26: Psychometrics*, eds. C. R. Rao and S. Sinharay, Amsterdam: North-Holland, pp. 297–358. [129]
- Zhang, W., and Yung, Y. F. (2011), "A Tutorial on Structural Equation Modeling With Incomplete Observations: Multiple Imputation and FIML Methods Using SAS" [online], Poster presented at the International Meeting of Psychometric Society, Hong Kong. Available at http://support.sas.com/rnd/app/stat/papers/imps2011_FIML.pdf. [135]