Supplementary materials for:

Exploratory Structural Equation Modeling

To appear in
Hancock, G. R., & Mueller, R. O. (Eds.). (2013). Structural equation modeling: A second course (2nd ed.). Charlotte, NC: Information Age Publishing, Inc.

Alexandre J. S. Morin

University of Western Sydney, Australia & University of Sherbrooke, Canada

Herbert W. Marsh

University of Western Sydney, Australia, University of Oxford, UK, & King Saud University, Saudi Arabia

and

Benjamin Nagengast

University of Tübingen, Germany

Acknowledgements

The authors would like to thank Tihomir Asparouhov and Bengt Muthén for helpful comments at earlier stages of this research, and each of the co-authors of ESEM studies cited in this chapter (Tihomir Asparouhov, Rhonda Craven, Linda Hamilton, Gregory Arief Liem, Oliver Lüdtke, Christophe Maïano, Andrew Martin, Bengt Muthén, Roberto Parada, Alexander Robitzsch, and Ulrich Trautwein). This research was supported in part by a grant to the second author from the UK Economic and Social Research Council.

An extended technical presentation of the ESEM model

In the ESEM model (Asparouhov & Muthén, 2009; Marsh et al., 2009), there are p dependent variables $\mathbf{Y} = (Y_1, ..., Y_p)$, q independent variables $\mathbf{X} = (X_1, ..., X_q)$, and m latent variables $\boldsymbol{\eta} = (\eta_1, ..., \eta_m)$, forming the following general ESEM model:

$$Y = \nu + \Lambda\eta + \mathrm{K}\mathrm{X} + \varepsilon \quad (1)$$

$$\eta = \alpha + \mathrm{B}\,\eta + \Gamma\mathrm{X} + \zeta \quad (2)$$

Standard assumptions of this model are that the $\boldsymbol{\varepsilon}$ and $\boldsymbol{\zeta}$ residuals are normally distributed with mean 0 and variance covariance matrix $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ respectively. The first equation represents the measurement model where $\boldsymbol{\nu}$ is a vector of intercepts, $\boldsymbol{\Lambda}$ is a factor loading matrix, $\boldsymbol{\eta}$ is a vector of continuous latent variables, $\mathbf{K}$ is a matrix of $\mathbf{Y}$ on $\mathbf{X}$ regression coefficients, and $\boldsymbol{\varepsilon}$ is a vector of residuals for $\mathbf{Y}$. The second equation represents the latent variable model where $\boldsymbol{\alpha}$ is a vector of latent intercepts, $\mathbf{B}$ is a matrix of $\boldsymbol{\eta}$ on $\boldsymbol{\eta}$ regression coefficients, $\boldsymbol{\Gamma}$ is a matrix of $\boldsymbol{\eta}$ on $\mathbf{X}$ regression coefficients, and $\boldsymbol{\zeta}$ is a vector of latent variables residuals.

In ESEM, $\boldsymbol{\eta}$ can include multiple sets of EFA factors and CFA factors. More precisely, the CFA factors are identified as in traditional SEM where each factor is associated with a different set of indicators. EFA factors can be divided into blocks of factors so that a series of indicators is used to estimate all EFA factors within a single block, and a different set of indicators is used to estimate another block of EFA factors. However, specific items may be assigned to more than one set of EFA or CFA factors. Assignments of items to CFA and/or EFA factors is usually determined based of a priori theoretical expectations, practical considerations, or preliminary tests conducted on the data.

In a basic version of the ESEM model including only CFA factors (and thus equivalent to the classical SEM model), all parameters can be estimated with the maximum likelihood (ML) estimator or robust alternatives. However, when EFA factors are posited, further constraints are required to achieve an identified solution (Asparouhov & Muthén, 2009, Marsh et al., 2009). In the first step, an unconstrained factor structure is estimated. Given the need to estimate all loadings, a total of m^2 constraints are required to achieve identification for the EFA factors (Jöreskog, 1969). These constraints are generally implemented by specifying the factor variance-covariance matrix as an identity matrix and constraining factor loadings in the right upper corner of the factor loading matrix to be 0 (for the i^{th} factor, i-1 factor loadings are restricted to 0). Consider any m x m square matrix (m = number of factors), a square matrix that we refer to as H. In this (mxm) square matrix H one can replace the η vector by H η in the ESEM model (1-2) which will also alter the parameters in the model as well; Λ to $\Lambda\ H^{-1}$, the α vector H α, the Γ matrix to H Γ, the B matrix to HBH^{-1} and the Ψ matrix to $H\Psi H^{\mathbf{T}}$. Since H has m^2 elements, the ESEM model has a total of m^2 indeterminacies that must be resolved. Two variations of this model are considered; one where factors are orthogonal so that the factor variance-covariance matrix (Ψ) is an identity matrix, and an oblique model where Ψ is an unrestricted correlation matrix (i.e., all correlations and residual correlations between the latent variables are estimated as free parameters). This model can also be extended to include a structured variance-covariance matrix (Ψ).

For an orthogonal matrix H (i.e., a square mxm matrix H such that $HH^T = I$), one can replace the η vector by H η and obtain an equivalent model in which the parameters are changed. EFA can resolve this non-identification problem by minimizing f(Λ*) = f(Λ H−1), where f is a function called the rotation criteria or simplicity function (Asparouhov & Muthén, 2009; Jennrich & Sampson, 1966), typically such that among all equivalent Λ parameters the simplest solution is obtained. There are a total of m(m−1)/2 constraints in addition to m(m + 1)/2 constraints that are directly imposed on the Ψ matrix for a total of m^2 constraints needed to identify the model. The identification for the oblique model is developed similarly such that a total of m^2 constraints needed to identify the model are imposed. Although the requirement for m^2 constraints is only a necessary condition and in some cases it may be insufficient, in most cases the model is identified if and only if the Fisher information matrix is not singular (Silvey, 1970). This method can be used in the ESEM framework as well (Asparouhov & Muthén, 2009; also see Hayashi & Marcoulides, 2006).

The estimation of the ESEM model consists of several steps (Asparouhov & Muthén, 2009). Initially a SEM model is estimated using the ML estimator. The factor variance covariance matrix is specified as an identity matrix (ψ = I), giving $m(m + 1)/2$ restrictions. The EFA loading matrix (Λ), has all entries above the main diagonal (i.e., for the first m rows and column in the upper right hand corner of factor loading matrix, Λ), fixed to 0, providing remaining $m(m - 1)/2$ identifying restrictions. This initial, unrotated model provides starting values that can be subsequently rotated into an EFA model with m factors. The asymptotic distribution of all parameter estimates in this starting value model is also obtained. Then the ESEM variance covariance matrix is computed (based only on $\Lambda\ \Lambda^T + \theta$ and ignoring the remaining part of the model).

The correlation matrix is also computed and, using the delta method (Asparouhov & Muthén, 2009), the asymptotic distribution of the correlation matrix and the standardization factors are obtained. In addition, again using the delta method, the joint asymptotic distribution of the correlation matrix, standardization factors and all remaining parameters in the model are computed and used to obtain the standardized rotated solution based on the correlation matrix and its asymptotic distribution (Asparouhov & Muthén, 2009). This method is also extended to provide the asymptotic covariance of the standardized rotated solution, standardized unrotated solution, standardization factors, and all other parameters in the model. This asymptotic covariance is then used to compute the asymptotic distribution of the optimal rotation matrix H and all unrotated parameters which is then used to compute the rotated solution for the model and its asymptotic variance covariance. In Mplus multiple random starting values are used in the estimation process to protect against non-convergence and local minimums in the rotation algorithms.

With ESEM models it is possible to constrain the loadings to be equal across two or more sets of EFA blocks in which the different blocks represent multiple discrete groups or multiple occasions for the same group. This is accomplished by first estimating an unrotated solution with all loadings constrained to be equal across the groups or over time. If the starting solutions in the rotation algorithm are the same, and no loading standardizing is used, the optimal rotation matrix will be the same as well as the subsequent rotated solutions. Thus obtaining a model with invariant rotated Λ^* amounts to simply estimating a model with invariant unrotated Λ, a standard task in maximum likelihood estimation.

For an oblique rotation it is also possible to test the invariance of the factor variance-covariance matrix (Ψ) matrix across the groups. To obtain non-invariant Ψs an unrotated solution with Ψ = I is specified in the first group and an unrestricted Ψ is specified in all other groups. Note that this unrestricted specification means that Ψ is not a correlation matrix as factor variances are freely estimated. It is not possible in the ESEM framework to estimate a model where in the subsequent groups the Ψ matrix is an unrestricted correlation matrix, because even if the factor variances are constrained to be 1 in the unrotated solution, they will not be 1 in the rotated solution. However, it is possible to estimate an unrestricted Ψ in all but the first group and after the rotation the rotated Ψ can be constrained to be invariant or varying across groups. Similarly, when the rotated and unrotated loadings are invariant across groups, it is possible to test the invariance of the factor intercept and the structural regression coefficients. These coefficients can also be invariant or varying across groups simply by estimating the invariant or group-varying unrotated model. However, in this framework only full invariance can be tested in relation to parameters in Ψ and Λ in that it is not possible to have measurement invariance for one EFA factor but not for the other EFA factors. Similar restrictions apply to the factor variance covariance, intercepts and regression coefficients, although it is possible to have partial invariance in the ε matrix of residuals. (It is however, possible to have different blocks of ESEM factors such that invariance constraints are imposed in one block, but not the other). Furthermore, if the ESEM model contains both EFA factors and CFA factors, then all of the typical strategies for the SEM factors can be pursued with the CFA factors.

Selecting the optimal number of factors in exploratory ESEM.

An important issue when an EFA or ESEM model is used for purely exploratory purposes is to determine the optimal number of factors required to best represent the data. Many criteria were proposed over the years to help in this decision such as (i) the Kaiser (1960; Guttman, 1954) criterion of retaining all factors with eigenvalues greater or equal to one (the default in many statistical packages such as SPSS), (ii) Cattell's (1966) scree test which consist of plotting the eigenvalues and retaining as many factors as there are before the first break point in the lines; (iii) Velicer's (1976) minimum average partial (MAP) method in which the minimum average of the squared partial correlation indicates the optimal number of factors; and (iv) Horn's (1965; Glorfeld, 1995) parallel analysis, which consists of complementing the scree test with eigenvalues calculated from a set of random variables – the crossing point of the two lines indicates the number of components to retain. Research evidence clearly shows that, although most studies still tend to rely on either Kaiser criterion or the scree test, both of these tests tended to produce biased results and are outperformed by the less accessible MAP tests and Parallel analysis (Fabrigar et al., 1999; Hayton, Allen, & Scarpello, 2004; Henson, & Roberts, 2006; Kahn, 2006; Zwick, & Velicer, 1982, 1986). O'Connor (2000) developed SPSS, SAS, and MATLAB macros that allows for the easy calculation of these tests (see https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html). Unfortunately, these macros still rely on the Listwise deletion of cases with missing data on any of the variables. When there are missing data, we thus recommend conducting parallel analysis in the following manner. First, the data's eigenvalues should be calculated from any statistical package allowing for full information maximum likelihood handling of missing data (Enders, 2010; Graham, 2009) and provide EFA capabilities. In Mplus (Muthén & Muthén, 2010), the focal package of this chapter as it is the only one including ESEM capabilities, the data's eigenvalues are obtained by default with the EFA command (see chapter 4 of the users' manual freely available online at statmodel.com). Second, O'Connor (2000) macros can be used to generate the random variables eigenvalues, either based on normal theory assumptions or on random permutations of the real data, which as the advantages of preserving the properties of the real data in the calculation of the random eigenvalues.

Exploring different rotational procedures with the simulated data set.

Following Asparouhov and Muthén (2009) suggestions, different forms of rotations were compared, as shown in Table S1 (it should be noted that no matter which specific rotation is selected, the fit indices and items' uniquenesses are unchanged). In fact, we compared Target rotation, Geomin rotation based on an ε value of .5 as recommended by Marsh et al. (2009, 2010), Geomin rotation based on Mplus defaults. The results show that the rotation that is apparently the most successful at deflating the factor correlations is Geomin with an ε value of .5. As in the current example we have the advantage of knowing the population parameter values, we also know that this specific rotation is also the method that is most accurate in relation to the real population value of .30. Deviations from this value could be due to sampling variation in the data simulation and to the fact that the real simulated data set was a longitudinal multiple group model, whereas these models are estimated on the full sample and on a single measurement point at a time. Another difference that is apparent from the examination of the results is that the relative size of cross loadings differs according to the rotational algorithm. However, once again it is the Geomin rotation with an ε value of .5 that most accurately represented the cross loadings of zero in the population generating model for the 1st and 6th items. Thus, we retain a Geomin rotation with an ε value of .5 for the remainder of this chapter. However, we emphasize that this conclusion is specific to this chapter and that any ESEM study should start with similar comparison before a final form of rotation is selected. In particular, we note that target rotation provides a useful link between CFA and EFA methods, as well as with Bayesian estimation methods relying on priors. In target rotation, a priori defined cross loadings are "targeted" to be close to 0 or some other pre-specified value, but zero is the default target rotation method that was used in the current test. However, nothing precludes the use of other "target" values when working with a well replicated factor structure. For instance, in the present study, if we "target" cross loadings to be .10 (rather than 0), a value indicating that small cross loadings are expected, the results of this "informed" target rotation would be as effective as geomin rotation with an ε value of .5.

Table 3. Standardized parameters from the CFA and ESEM models.

	CFA			ESEM geomin, ε = .5			ESEM geomin default			ESEM target			Population values		
Item	**F1**	**F2**	**Uniq.**	**F1**	**F2**	**Uniq.**	**F1**	**F2**	**Uniq.**	**F1**	**F2**	**Uniq.**	**F1**	**F2**	**Uniq.**
Time 1	(CFA-Time1.inp)			(ESEM-Geo.5-Time1.inp)			(ESEM-Geo-def-T1.inp)			(ESEM-target-Time1.inp)					
X1	.722		.478	.833	-.038	.333	.820	-.006	.333	.932	-.207	.333	.860	.000	.400
X2	.843		.289	.690	.232	.330	.632	.282	.330	.745	.110	.330	.650	.220	.300
X3	.742		.450	.536	.318	.461	.466	.369	.461	.564	.232	.461	.560	.300	.500
X4		.835	.303	.139	.766	.300	-.002	.838	.300	.075	.787	.300	.250	.840	.400
X5		.774	.401	.166	.674	.420	.041	.739	.420	.114	.684	.420	.190	.560	.300
X6		.510	.739	-.046	.566	.700	-.146	.612	.700	-.111	.611	.700	.000	.450	.500
Correlations	.728			.439			.540			.637			.300		
Time 2	(CFA-Time2.inp)			(ESEM-Geo.5-Time2.inp)			(ESEM-Geo-def-T2.inp)			(ESEM-target-Time2.inp)					
Y1	.737		.457	.837	-.032	.322	.827	-.007	.322	.932	-.196	.322	.860	.000	.400
Y2	.841		.292	.698	.222	.330	.643	.265	.330	.752	.100	.330	.650	.220	.300
Y3	.751		.436	.560	.303	.449	.492	.349	.449	.590	.213	.449	.560	.300	.500
Y4		.846	.284	.149	.762	.300	.002	.836	.300	.087	.780	.300	.250	.840	.400
Y5		.776	.398	.140	.704	.400	.005	.772	.400	.084	.720	.400	.190	.560	.300
Y6		.492	.758	-.054	.551	.719	-.157	.599	.719	-.117	.595	.719	.000	.450	.500
Correlations	.709			.429			.543			.625			.300		

Note. Names of the input file in the supplementary materials are reported in parentheses; All coefficients significant at the .05 level; CFA: Confirmatory factor analysis; ESEM: Exploratory Structural Equation Modeling; F1: standardized loadings on the first factor; F2: standardized loadings on the second factor; Uniq.: standardized uniquenesses.

References used in the supplementary materials

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16,* 397-438.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Enders, C.K. (2010). *Applied missing data analysis*. New York, NY: Guilford.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Glorfeld, L. W. (1995). An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain. *Educational and Psychological Measurement, 55,* 377-393.

Graham, J.W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika, 19,* 149-162.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7,* 191-205.

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66,* 393-416.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179-185.

Hayashi, K., & Marcoulides, G. A. (2006). Examining identification issues in factor analysis. *Structural Equation Modeling, 13*, 631–645.

Jennrich, R.I., & Sampson, P. F. (1966). Rotation to simple loadings. *Psychometrika, 31*, 313-323.

Kaiser, H.F., (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141-151.

Marsh, H.W., Lüdtke, O., Muthén, B.O., Asparouhov, T., Morin, A.J.S., & Trautwein, U. (in press, 2010). A new look at the big-five factor structure through Exploratory Structural Equation Modeling. *Psychological Assessment, 22,* 471-491.

Marsh, H.W., Muthén, B., Asparouhov. T., Lüdtke, O., Robitzsch, A., Morin, A.J.S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16,* 439-476 .

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. B*ehavior Research Methods, Instruments, & Computers, 32*, 396-402.

Silvey, S. D. (1970). *Statistical inference*. Harmondsworth, UK: Penguin.

Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321-327.

Zwick, W.R., & Velicer, W.F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17,* 253-269.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432-442.