Exploratory Structural Equation Modeling

Tihomir Asparouhov and Bengt Muthén

Version 3[*]

May 23, 2008

Abstract

Exploratory factor analysis (EFA) has been said to be the most
frequently used multivariate analysis technique in statistics. Jennrich
and Sampson (1966) solved a significant EFA factor loading matrix
rotation problem by deriving the direct Quartimin rotation. Jennrich
was also the first to develop standard errors for rotated solutions al-
though these have still not made their way into most statistical soft-
ware programs. This is perhaps because Jennrichs achievements were
partly overshadowed by the subsequent development of confirmatory
factor analysis (CFA) by Joreskog (1969). The strict requirement
of zero cross-loadings in CFA, however, often does not fit the data
well and has led to a tendency to rely on extensive model modifi-
cation to find a well-fitting model. In such cases, searching for a
well-fitting measurement model may be better carried out by EFA
(Browne, 2001). Furthermore, misspecification of zero loadings tends
to give distorted factors with over-estimated factor correlations and
subsequent distorted structural relations. This paper describes an
EFA-SEM (ESEM) approach, where in addition to or instead of a
CFA measurement model, an EFA measurement model with rotations
can be used in a structural equation model. The ESEM approach has
recently been implemented in the Mplus program. ESEM gives ac-
cess to all the usual SEM parameter and the loading rotation gives a
transformation of structural coefficients as well. Standard errors and
overall tests of model fit are obtained. Geomin and Target rotations
are discussed. Examples of ESEM models include multiple-group EFA
with measurement and structural invariance testing, test-retest (lon-
gitudinal) EFA, EFA with covariates and direct effects, and EFA with
correlated residuals. Testing strategies with sequences of EFA and
CFA models are discussed. Simulated data are used to illustrate the
points.

1 Introduction

The latent variable measurement specification in structural equation modeling (SEM; Joreskog and Sorbom, 1979; Muthén, 1984; Bollen, 1989; Browne and Arminger, 1995) uses the Joreskog (1969) confirmatory factor analysis (CFA) model. Based on theory and prior analyses, the CFA measurement model specifies a number of factor loadings fixed at zero to reflect a hypothesis that only certain factors influence certain factor indicators. Often a simple structure is specified where each indicator is influenced by a single factor, i.e. there are no cross-loadings, sometimes referred to as variable complexity of one. The number of such zero loading restrictions is typically much larger than the number of restrictions needed to identify the factor analysis measurement model, which as in exploratory factor analysis with m factors is m^2 restrictions on the factor loadings, factor variances, and factor covariances. The use of CFA measurement modeling in SEM has the advantage that researchers are encouraged to formalize their measurement hypotheses and develop measurement instruments that have a simple measurement structure. Incorporating a priori substantive knowledge in the form of restrictions on the measurement model makes the definition of the latent variables better grounded in subject-matter theory and leads to parsimonious models.

The use of CFA measurement modeling in SEM also has disadvantages and these are likely to have contributed to poor applications of SEM where the believability and replicability of the final model is in doubt. While technically appealing, CFA requires strong measurement science which is often not available in practice. A measurement instrument often has many small cross-loadings that are well motivated by either substantive theory or by the formulation of the measurements. The CFA approach of fixing many or all cross-loadings at zero may therefore force a researcher to specify a more parsimonious model than is suitable for the data. Because of this, models often do not fit the data well and there is a tendency to rely on extensive model modification to find a well-fitting model. Here, searching for a well-fitting measurement model is often aided by the use of model modification indices. A critique of the use of model searches using modification indices is given for example in MacCallum, Roznowski, and Necowitz (1992). In such situations of model uncertainty, Browne (2001) advocates exploratory rather than confirmatory approaches:

"Confirmatory factor analysis procedures are often used for ex-

ploratory purposes. Frequently a confirmatory factor analysis, with pre-specified loadings, is rejected and a sequence of modifications of the model is carried out in an attempt to improve fit. The procedure then becomes exploratory rather than confirmatory — In this situation the use of exploratory factor analysis, with rotation of the factor matrix, appears preferable. — The discovery of misspecified loadings ... is more direct through rotation of the factor matrix than through the examination of model modification indices."

Furthermore, misspecification of zero loadings in CFA tends to give distorted factors. When non-zero cross-loadings are specified as zero, the correlation between factor indicators representing different factors is forced to go through their main factors only, leading to over-estimated factor correlations and subsequent distorted structural relations.

For the reasons given above, it is important to extend structural equation modeling to allow less restrictive measurement models to be used in tandem with the traditional CFA models. This offers a richer set of a priori model alternatives that can be subjected to a testing sequence. This paper describes an exploratory structural equation modeling (ESEM) approach, where in addition to or instead of CFA measurement model parts, EFA measurement model parts with factor loading matrix rotations can be used. For each EFA measurement model part with m factors, only m^2 restrictions are imposed on the factor loading matrix and the factor covariance matrix. ESEM gives access to all the usual SEM parameters, for example residual correlations, regressions of factors on covariates, and regressions among factors. Multiple-group analysis with intercept and mean structures are also handled. The ESEM approach has recently been implemented in the Mplus program.

Exploratory factor analysis (EFA) has been said to be the most frequently used multivariate analysis technique in statistics. Jennrich and Sampson (1966) solved a significant EFA factor loading matrix rotation problem by deriving the direct Quartimin rotation. Jennrich was also the first to develop standard errors for rotated solutions. Cudeck and O'Dell (1994) provide a useful discussion on the benefits of considering standard errors for the rotated factor loadings and factor correlation matrix in EFA. However, EFA standard errors have still not made their way into most statistical software programs (Jennrich, 2007), perhaps because Jennrichs achievements were partly overshadowed by the subsequent development of CFA by Joreskog (1969). The

4

work to be presented can therefore also be seen as a further development and modernization of EFA, continuing the classic psychometric work that was largely abandoned. Three examples can be mentioned. Correlated residuals among factor indicators sharing similar wording can confound the detection of more meaningful factors using conventional EFA; allowing such parameters in an extended EFA can now give new measurement insights. Comparing EFA factor loadings across time in longitudinal studies or across groups of individuals can now be done using rigorous likelihood-ratio testing without the researcher being forced to switch from EFA to CFA.

It should be made clear that the development in this paper is not intended to encourage a complete replacement of CFA with EFA measurement modeling in SEM. Instead, the intention is to add further modeling flexibility by providing an option that in some cases is more closely aligned with reality, reflecting more limited measurement knowledge of the researcher or a more complex measurement structure. There will still be many situations where a CFA approach is preferred. Apart from situations where the measurement instruments are well understood, this includes applications where a CFA specification lends necessary stability to the modeling. As one example, multi-trait, multi-method (MTMM) modeling relies on CFA specification of both the trait and the methods part of the model. Although it is in principle possible with the methods presented here to let the trait part be specified via EFA, leaving the methods part specified as CFA, this may not provide easy recovery of the data-generating parameter values.

In ESEM, the loading matrix rotation gives a transformation of both measurement and structural coefficients. Extending the work summarized in Jennrich (2007), ESEM provides standard errors for all rotated parameters. Overall tests of model fit are also obtained. With EFA measurement modeling, the reliance on a good rotation method becomes important. The paper discusses the Geomin rotation (Yates, 1987) which is advantageous with variable complexity greater than one (Browne, 2001; McDonald, 2005). Target rotation (Browne, 2001) is a less-known rotation technique that conceptually is situated in between of EFA and CFA, which is also implemented in the general ESEM framework. Examples of ESEM models are presented including multiple-group EFA with measurement invariance [1]. Testing strategies with sequences of EFA and CFA models are discussed. Simulated

---

[1]Examples of ESEM models illustrating structural invariance testing, EFA

data are used to illustrate the points.

The outline of this paper is as follows. A simple ESEM model is presented in Section 2. The general ESEM model is presented in Section 3. Section 4 outlines the estimation method. In Section 5 the ESEM modeling is expanded to include constrained rotation methods that are used to estimate for example measurement invariant ESEM models and multiple group EFA models. Various rotation criteria and their properties are described in Section 6. Several simulation studies are presented in Section 7. Discussion on the choice of the rotation criterion is given in Section 8. Section 9 concludes.

2 Simple Exploratory Structural Equation Model

Suppose that there are p dependent variables $Y = (Y_1, ..., Y_p)$ and q independent variables $X = (X_1, ..., X_p)$. Consider the general structural equation model with m latent variables $\eta = (\eta_1, ..., \eta_m)$

$$Y = \nu + \Lambda\eta + KX + \varepsilon \tag{1}$$

$$\eta = \alpha + B\eta + \Gamma X + \xi \tag{2}$$

The standard assumptions of this model are that the ε and ξ are normally distributed residuals with mean 0 and variance covariance matrix Θ and Ψ respectively. The model can be extended to multiple group analysis, where for each group model (1-2) is estimated and some of the model parameters can be the same in the different groups. The model can also be extended to include categorical variables and censored variables as in Muthen (1984) using limited-information weighted least squares estimation. For each categorical and censored variable Y^* is used instead of Y is equation (1), where Y^* is an underlying unobserved normal variable. For each categorical variables there is a set of parameters τ_k such that

$$Y = k \Leftrightarrow \tau_k < Y^* < \tau_{k+1}. \tag{3}$$

---

with covariates and direct effects, and EFA with correlated residuals are available in Bengt Muthén's multimedia presentation on this topic available at http://www.ats.ucla.edu/stat/mplus/seminars/whatsnew_in_mplus5_1/default.htm.

Thus a linear regression for Y^* is equivalent to a Probit regression for Y. Similarly, for censored variables with a censoring limit of c

$$Y = \begin{cases} Y^* & \text{if } Y \geq c \\ c & \text{if } Y \leq c \end{cases} \tag{4}$$

All of the parameters in the above model can be estimated with the maximum likelihood estimation method, however, this structural model is generally unidentified and typically many restrictions need to be imposed on the model. Otherwise the maximum likelihood estimates will be simply one set of parameter estimates among many equivalent solutions.

One unidentified component is the scale of the latent variable. Two different approaches are generally used to resolve this non-identification. The first approach is to identify the scale of the latent variable by fixing its variance to 1. The second approach is to fix one of the Λ parameters in each column to 1. The two approaches are generally equivalent and a simple reparameterization can be used to obtain the parameter estimates from one to the other scales. In what follows the first approach is taken. It is assumed that the variance of each latent variable is 1. Later on the model is expanded to include latent factors with scale identified by method 2. It is also assumed in this section that all Λ parameters are estimated.

Even when the scale of the latent variable is identified, however, there are additional identifiability issues when the number of latent factors m is greater than 1. For each square matrix H of dimension m one can replace the η vector by $H\eta$ in model (1-2). The parameters in the model will be altered as well. The Λ will be replaced by ΛH^{-1}, the α vector is replaced by $H\alpha$, the Γ matrix is replaced by $H\Gamma$, the B matrix is replaced by HBH^{-1} and the Ψ matrix is replaced by $H\Psi H^T$. Since H has m^2 elements the model has a total of m^2 indeterminacies. In the discussion that follows two specific models are considered. The first model is the orthogonal model where Ψ is restricted to be the identity matrix, i.e., the latent variables have no residual correlation. The second model which is the oblique model where Ψ is estimated as an unrestricted correlation matrix, i.e., all residual correlations between the latent variables are estimated as free parameters. Later on the model is generalized to include structured variance covariance matrices Ψ.

First consider the identification issues for the orthogonal model. For each orthogonal matrix H of dimension m, i.e., a square matrix

H such that $HH^T = I$, one can replace the η vector by $H\eta$ and obtain an equivalent model. That is because the variance $H\eta$ is again the identity matrix. Again the Λ matrix is replaced by ΛH^{-1} and similarly the rest of the parameters are changed. Exploratory factor analysis (EFA) offers a solution to this non-identification problem. The model is identified by minimizing

$$f(\Lambda^*) = f(\Lambda H^{-1}) \tag{5}$$

over all orthogonal matrices H, where f is a function called the rotation criteria or simplicity function. Several different simplicity functions have been utilized in EFA, see Jennrich and Sampson (1966) and Appendix A. For example, the Varimax simplicity function is

$$f(\Lambda) = -\sum_{i=1}^{p} \left(\frac{1}{m} \sum_{j=1}^{m} \lambda_{ij}^4 - \left(\frac{1}{m} \sum_{j=1}^{m} \lambda_{ik}^2 \right)^2 \right). \tag{6}$$

These functions are usually designed so that among all equivalent Λ parameters the simplest and most interpretable solution is obtained.

Minimizing the simplicity function is equivalent to imposing the following constraints on the parameters Λ, see Archer and Jennrich (1973),

$$R = ndg\left(\Lambda^T \frac{\partial f}{\partial \Lambda} - \frac{\partial f}{\partial \Lambda}^T \Lambda \right) = 0. \tag{7}$$

where the ndg refers to the non-diagonal entries of the matrix. Note that the above matrix is symmetric and therefore these are $m(m-1)/2$ constraints. These constraints are in addition to the $m(m+1)/2$ 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. The simplicity function

$$f(\Lambda^*) = f(\Lambda H^{-1}) \tag{8}$$

is minimized over all matrices H such that $diag(H\Psi H^T - I) = 0$, i.e., matrices H such that all diagonal entries of $H\Psi H^T$ are 1. In this case minimizing the simplicity function is equivalent to imposing the following constraints on the parameters Λ and Ψ

$$R = ndg(\Lambda^T \frac{\partial f}{\partial \Lambda} \Psi^{-1}) = 0 \tag{9}$$

The above equation specifies $m(m-1)$ constraints because the matrix is not symmetric. These constraints are in addition to the m constraints that are directly imposed on the Ψ matrix for a total of m^2 constraints needed to identify the model.

If the dependent variables are on different scales the elements in the Λ matrix will also be on different scales which in turn can lead to imbalance of the minimization of the simplicity function and consequently lead to a suboptimal Λ^* solution. In EFA this issue is resolved by performing a standardization of the parameters before the rotation. Let Σ_d be a diagonal matrix of dimension p where the $i-$th diagonal entry is the standard deviation of the Y_i variable. The standardized parameters Λ are then $\Sigma_d^{-1}\Lambda$, i.e., in EFA analysis

$$f(\Sigma_d^{-1}\Lambda H^{-1}) \tag{10}$$

is minimized over all oblique or orthogonal matrices H. An equivalent way of conducting the EFA analysis is to first standardize all dependent variables so that they have 0 mean and variance 1 and then complete the rotation analysis using the unstandardized Λ matrix. Alternative standardization techniques are described in Appendix B.

The structural equation model (1-2) is similarly standardized to avoid any undesired effects from large variation in the scales of the dependent variables. Define the diagonal matrix

$$\Sigma_d = \sqrt{diag(\Lambda\Psi\Lambda^T + \Theta)} \tag{11}$$

and the normalized loadings matrix Λ_0 as

$$\Lambda_0 = \Sigma_d^{-1}\Lambda. \tag{12}$$

The simplicity function

$$f(\Lambda_0 H^{-1}) \tag{13}$$

is then minimized over all oblique or orthogonal matrices H. Denote the optimal matrix H by H^*. Call this matrix the rotation matrix. Denote the optimal Λ_0 by Λ_0^*. Call Λ_0^* the rotated standardized solution. Note that after the rotation the optimal Λ^* matrix should be obtained in the original scale of the dependent variables

$$\Lambda^* = \Sigma_d\Lambda_0^*. \tag{14}$$

Note here that formally speaking the squares of the diagonal entries of Σ_d are not the variances of Y_i. That is because the standardization

factor as defined in (11) does not include the remaining part of the structural model such as the independent variables X as well as equation (2). Nevertheless the simpler standardization factor defined in (11) will reduce generally any discrepancies in the scales of the dependent variables. In addition, formula (11) simplifies the computation of the asymptotic distribution of the parameters because it does not include the variance covariance of the independent variables X which typically is not part of the model. The model usually includes distributional assumptions and estimation only for the dependent variables. Note also that if the model does not include any covariates or other structural equations, i.e., if the model is equivalent to the standard EFA model then the standardization factor Σ_d is the standard deviation just like in EFA analysis.

The exploratory structural equation model (ESEM) described above can be estimated simply by constrained maximum likelihood estimation. This however, is not the algorithm implemented in Mplus. The parameter constraints (9-7) are rather complicated and constrained maximization is more prone to convergence problem in such situations. The algorithm used in Mplus is based on the gradient projection algorithm (GPA) developed in Jennrich (2001) and Jennrich (2002).

In the traditional EFA analysis the rotation of the factors affects only the parameters Λ and the Ψ matrix. In the exploratory structural equation model (ESEM) described above nearly all parameters are adjusted after the optimal rotation H^* is determined. The following formulas describe how the rotated parameters are obtained

$$\nu^* = \nu \tag{15}$$

$$\Lambda^* = \Lambda (H^*)^{-1} \tag{16}$$

$$K^* = K \tag{17}$$

$$\Theta^* = \Theta \tag{18}$$

$$\alpha^* = H^* \alpha \tag{19}$$

$$B^* = H^* B (H^*)^{-1} \tag{20}$$

$$\Gamma^* = H^* \Gamma \tag{21}$$

$$\Psi^* = (H^*)^T \Psi H^* \tag{22}$$

3 General Exploratory Structural Equation Model

The general ESEM model is described again by the equations

$$Y = \nu + \Lambda\eta + KX + \varepsilon \tag{23}$$

$$\eta = \alpha + B\eta + \Gamma X + \xi \tag{24}$$

where the factors η_i can be divided in two groups, exploratory factors and confirmatory factors. Let η_1, η_2, ..., η_r be the exploratory factors and η_{r+1}, ..., η_m be the confirmatory factors. The confirmatory factors are identified the same way factors are identified in traditional SEM models, for example, by having different factor indicator variables for each of the factors. The group of exploratory factors is further divided into blocks of exploratory latent variables that are measured simultaneously. Suppose that a block of exploratory latent variables consists of η_1, η_2, ..., η_b. For each exploratory block a block of dependent factor indicator variables are assigned. Suppose the Y_1, Y_2, ... , Y_c are the indicator variables assigned to the exploratory block. Note that different exploratory blocks can use the same factor indicators. Similarly exploratory factors can use the same factor indicators as confirmatory factors. The measurement model for η_1, η_2, ..., η_b based on the indicators Y_1, Y_2, ... , Y_c is now based and identified as the model in the previous section, using an optimal rotation for the exploratory factor block. Equation (24) uses all the confirmatory and exploratory factors. If H^* represent a combined optimal rotation matrix which consists of the optimal rotations for each of the exploratory factor blocks the rotated estimates are obtained from the set of unidentified parameters again via formulas (15-22).

There are certain restrictions that are necessary to impose on the flexibility of this model. Exploratory factors have to be simultaneously appearing in a regression or correlated with. For example if a factor in an exploratory block is regressed on a covariate X_i all other factors in that block have to be regressed on that covariate. Similarly if a variable is correlated with an exploratory factor, the variable has to be correlated to all other variables in that exploratory block, i.e., these covariance parameters can either be simultaneously 0 or they have to be simultaneously free and unconstrained.

4   Estimation

This section describes the procedure used to estimate the ESEM model, including the estimates for the asymptotic distribution of the parameter estimates. The estimation consist of several steps. In the first step using the ML estimator a SEM model is estimated where for each exploratory factor block the factor variance covariance matrix is specified as $\Psi = I$, giving $m(m+1)/2$ restrictions, and the exploratory factor loading matrix for the block has all entries above the main diagonal, in the upper right hand corner, fixed to 0, giving the remaining $m(m-1)/2$ identifying restrictions. This model is referred to as the starting value model or the initial model or the unrotated model. It is well known that such a model can be subsequently rotated into any other exploratory factor model with m factors. The asymptotic distribution of all parameter estimates in this starting value model is also obtained. Then for each exploratory block / simple ESEM the variance covariance matrix implied for the dependent variable based only on

$$\Lambda\Lambda^T + \Theta \tag{25}$$

and ignoring the remaining part of the model is computed. The correlation matrix is also computed and using the delta method the asymptotic distribution of the correlation matrix and the standardization factors are obtained. In addition, using the delta method again the joint asymptotic distribution of the correlation matrix, standardization factors and all remaining parameter in the model is computed. A method developed in Asparouhov and Muthen (2007) is then used to obtain the standardized rotated solution based on the correlation matrix and its asymptotic distribution, see Appendix C for a summary of this method. 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 model parameters. The optimal rotation matrix H is computed as follows

$$H = M_0^{-1} M_0^* \tag{26}$$

where M_0 is a square matrix that consists of the first m rows of Λ_0 and similarly Λ_0^* is a square matrix that consists of the first m rows of L_0^*. Finally all rotated parameters and their asymptotic distribution is obtained using formulas (15-22) and the delta method.

This estimation method is equivalent to the constrained maximum likelihood method based on (7) or (9). The estimation of the starting value model may give non-convergence. A random starting value procedure is implemented in Mplus for this estimation. In addition a random starting value procedure is implemented in Mplus for the rotation algorithms which are prone to multiple local minima problems.

5   Constrained Rotation

Factor analysis is often concerned with invariance of measurements across different populations such as defined by gender and ethnicity (see, e.g. Meredith, 1993). Studies of measurement invariance and population differences in latent variable distributions is commonplace through multiple-group analysis (Joreskog and Sorbom, 1979). A similar situation occurs for longitudinal data where measurement invariance is postulated for a factor model at each of the time points. Analysis of measurement invariance, however, has been developed and used only with CFA measurement specifications. Although related methods have been proposed in EFA settings, see Meredith (1964) and Cliff (1966), they only attempt to rotate to similar factor patterns. The methods of this paper introduce multiple-group exploratory factor analysis, and multiple-group analysis of EFA measurement parts in structural equation modeling. This development makes it possible for a researcher to not have to move from EFA to CFA when wanting to study measurement invariance.

This section describes ESEM models constraining the loadings to be equal across two or more sets of EFA blocks. For example in multiple group analysis it is of interest to evaluate a model where the loading matrices are constrained to be equal across the different groups. This can easily be achieved in the ESEM framework by first estimating an unrotated solution with all loadings constrained to be equal across the groups. 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 and in turn the rotated solutions will also be the same. Thus obtaining a model with invariant rotated Λ^* amounts to simply estimating a model with invariant unrotated Λ and that is a standard task in maximum likelihood estimation.[2]

---

[2]Note again, however, that Mplus will automatically use RowStandardization=Covariance, so that differences across groups in the residual variances Θ do not

When an oblique rotation is used an important modeling possibility is to have the Ψ matrix also be invariant across the groups or alternatively to be varying across the groups. These models are obtained as follows.[3] To obtain varying Ψ across the groups one simply estimates an unrotated solution with $\Psi = I$ in the first group and an unrestricted Ψ matrix in all other groups. Note that unrestricted here means that Ψ is not a correlation matrix but the variances of the factors are also free to be estimated. It is not possible in this framework to estimate a model where in the subsequent groups the Ψ matrix is an unrestricted correlation matrix, because even if in the unrotated solution the variances of the factors are constrained to be 1, in the rotated solution they will not be 1. However, it is possible to estimate an unrestricted variance Ψ in all but the first group and after the rotation the rotated Ψ will also be varying across groups.

Similarly, when the rotated and unrotated loadings are invariant across groups one can estimate two different models in regard to 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. Note that in this framework only full invariance can be estimated, i.e., it is not possible to have measurement invariance for one EFA factor but not for the other, if the two EFA factors belong to the same EFA block. Similar restrictions apply to the factor variance covariance, intercepts and regression coefficients. If the model contains both EFA factors and CFA factors all of the usual possibilities for the CFA factors are available.

6   Rotation Criteria

When the EFA specification is used in ESEM instead of CFA the choice of the rotation procedure becomes important. This section considers the properties of some key rotation criteria: Quartimin, Geomin, and the Target criteria. Further rotation criteria are given in Appendix A.[4]

---

cause differences in the rotated solutions, see Appendix B.

[3]Using again RowStandardization=Covariance the estimated unrotated solution with equality of the loadings across groups and all $\Psi = I$ leads to rotated solution with equality in the rotated loadings as well as in the Ψ matrix, see Appendix B.

[4]All of these rotation criteria are implemented in Mplus.

The choice of the rotation criterion is to some extent still an open research area. Generally it is not known what loading matrix structures are preserved by each rotation criterion. The simulation studies presented in this article, however, indicate that the Geomin criterion is the most promising rotation criterion when little is known about the true loading structure. [5] Geomin appears to be working very well for simple and moderately complicated loading matrix structures. However, it fails for more complicated loading matrix structures involving 3 or more factors and variables with complexity 3 and more, i.e., variables with 3 or more non-zero loadings. Some examples are given in the simulation studies described in Section 7. For more complicated examples the Target rotation criterion will lead to better results. Additional discussion on the choice of the rotation criterion is presented in Section 8.

Following are some general facts about rotation criteria. Let f be a rotation criterion, Λ_0 be the loading matrix and Ψ be the factor covariance. The oblique rotation algorithm minimizes

$$f(\Lambda) = f(\Lambda_0 H^{-1}) \tag{27}$$

over all matrices H such that $diag(H\Psi H^T) = 1$, while the orthogonal rotation algorithm minimizes (27) over all orthogonal matrices H. The matrix Λ_0 is called an $f-$ invariant loading structure if (27) is minimized at $H = I$, i.e., (27) is minimized at the loading matrix Λ_0 itself, regardless of the value of Ψ. The invariant structures presented here are the ones that attain the global unconstrained minimum for the rotation criteria. Typically the global unconstrained rotation function minimum is 0. If Λ_0 is the true simple structure, rotations based on f will lead to Λ_0 regardless of the starting solution. There is a second requirement for this to happen, namely, Λ_0 has to be the unique minimum of f, up to a sign change in each factor and factor permutation. If it is not, the rotation algorithm will have multiple solutions and generally speaking the rotation algorithm may not be identified sufficiently.

A sufficient condition for rotation identification has been described in Howe (1955), Joreskog (1979) and Mulaik and Millsap (2000). Consider a factor analysis model with m factors. In general, m^2 restrictions have to be imposed on the parameters in Λ and Ψ for identification purposes. For oblique rotation m factor variances are fixed to

---

[5]The Geomin rotation is now the default rotation criterion in Mplus.

15

1 and therefore additional $m(m-1)$ constraints have to be imposed. It should be noted that not all sets of $m(m-1)$ constraints will lead to identification. Consider the case when the constraints are simply $m(m-1)$ loading parameters fixed at 0. The following two conditions are sufficient conditions for rotation identifiability.

(a) Each column of Λ has $m-1$ entries specified as zeroes.

(b) Each submatrix Λ_s, $s = 1, ..., m$, of Λ composed of the rows of Λ that have fixed zeros in the $s-$th column must have rank $m-1$.

These conditions are sufficient for identification purposes regardless of what the value of the correlation matrix Ψ is. Conditions (a) and (b) can also be used to establish identifiability of the rotation criteria. Rotation functions are generally designed so that the optimally rotated loading matrix has many zero loadings. If these zero loadings satisfy conditions (a) and (b) then the rotation method is sufficiently identified. This approach will be used with the Geomin and the Target rotation method.

In general one needs to know what structures are invariant under which rotation criteria so that one can make a proper rotation criterion selection for the type of structure that one is searching for. In the next three sections the Quartimin, Geomin and Target rotation criteria and their invariant loading structures are described. Let the loading matrix Λ be a matrix with dimensions p and m.

6.1 Quartimin

The rotation function for the Quartimin criterion is

$$f(\Lambda) = \sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{l \neq j}^{m} \lambda_{ij}^2 \lambda_{il}^2. \tag{28}$$

If each variable loads on only one factor, i.e., each row in Λ has only one non-zero entry, then Λ is Quartimin invariant, and this rotation criterion will work perfectly for recovering such a factor loading structure. Note that in this case the minimum of the rotation function is the absolute minimal value of 0. Note also that this fact is independent of the number of variables or the number of factors. Usually no other rotation criteria can be as effective as Quartimin for these kind of simple loading structures in terms of MSE of the parameters

estimates. However, rotation criteria such as Geomin will generally produce rotation results similar to Quartimin.

6.2 Geomin

The rotation function for the Geomin rotation criterion is

$$f(\Lambda) = \sum_{i=1}^{p} \left(\prod_{j=1}^{m} (\lambda_{ij}^2 + \epsilon) \right)^{1/m} \tag{29}$$

where ϵ is a small constant. The original purpose of this constant is to make the rotation function differentiable when there are zero loadings, but by varying the constant one can actually create different rotation criteria.

Note that if $\epsilon = 0$ and one entry in each row is zero, the Geomin rotation function is zero, i.e., the rotation function is already minimized and the minimization process does not help in the identification of the remaining entries in the row. If however $\epsilon > 0$ this problem is resolved to some extent. Note also that the Geomin rotation function is simply the sum of the rotation functions for each of the rows, but the rotation function for each row can not be minimized separately because the loading parameters are not independent across rows. They can only vary according to an oblique or orthogonal rotation. Thus even when $\epsilon = 0$ and each row contains a zero the non-zero entries in the row can be identified through the sufficient conditions (a) and (b) in Section 6.

The known Geomin invariant loading structures will now be described. Consider first the case when the parameter ϵ is 0 (or a very small number such as 10^{-5}). The Geomin function is 0 for all Λ structures that have at least one 0 in each row, i.e., structures with at least one zero in each row are Geomin invariant. This is a very large set of loading structures. However, in many cases there are more than one equivalent Λ structure with at least one zero in each row. Suppose that $p \geq m(m-1)$ for oblique rotations (and $p \geq m(m-1)/2$ for orthogonal rotations) where p is the number of dependent variables and m is the number of factors and that conditions (a) and (b) described in Section 6 are satisfied. Then the Λ structure is unique and will therefore be completely recovered by the Geomin criterion. Even in this case however, there could be multiple solutions that reach the 0 rotation function value because a different set of 0 locations can lead to a different rotated solution.

The Geomin rotation criterion is known to frequently produce multiple solutions, i.e., multiple local minima with similar rotation function values. The role of the ϵ value is to improve the shape of the rotation function, so that it is easier to minimize and to reduce the number of local solutions. Models with more factors are more likely to have more local solutions and are more difficult to minimize. Thus larger ϵ values are typically used for models with more factors[6]. Note however, that multiple solutions is not a problem but rather an opportunity for the analysis, see Rozeboom (1992) and Section 8 below.

Another reason to include a positive ϵ value in the Geomin rotation function is the fact that if $\epsilon = 0$ the rotation function is not differentiable. Differentiability is important for convergence purposes as well as standard error estimation. For example, if $\epsilon < 10^{-5}$ the convergence can be very slow and the prespecified maximum number of iteration can be exceeded.

6.3   Target

Conceptually, target rotation can be said to lie in between the mechanical approach of EFA rotation and the hypothesis-driven CFA model specification. In line with CFA, target loading values are typically zeros representing substantively motivated restrictions. Although the targets influence the final rotated solution, the targets are not fixed values as in CFA, but zero targets can end up large if they do not provide good fit. An overview with further references is given in Browne (2001), including reference to early work by Tucker (1944).

The target rotation criterion is designed to find a rotated solution Λ^* that is closest to a prespecified matrix B. Not all entries in the matrix B need to be specified. For identification purposes at least $m - 1$ entries have to be specified in each column for oblique rotation and $(m-1)/2$ entries have to be specified in each column for orthogonal rotation. The rotation function is

$$f(\Lambda) = \sum_{i=1}^{p} \sum_{j=1}^{m} a_{ij}(\lambda_{ij} - b_{ij})^2 \tag{30}$$

where a_{ij} is either 1 if b_{ij} is specified and 0 if b_{ij} is not specified.

The known Target invariant loading structures can be described as follows. If all targets in the rotation function are correct then the Λ

---

[6]The Mplus default for ϵ for 2 factors is 0.0001, for 3 factors is 0.001, and for 4 or more factors it is 0.01.

matrix minimizes the rotation criteria. In addition, if at least $m(m-1)$ zero targets are specified that satisfy conditions (a) and (b) in Section 6 then the Λ matrix is the unique minimum and therefore it is Target invariant.

For example, consider a 3-factor EFA model with 9 measurement variables. Data is generated and estimated according to this model with the following parameter values. The factor variance covariance Ψ is the identity matrix and the loading matrix Λ is as follows

$$\begin{pmatrix} 1 & (0) & (0) \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ (0) & 1 & (0) \\ (0) & 1 & 0 \\ 0 & 1 & 0 \\ 0 & (0) & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \tag{31}$$

The residual variances of the dependent variables are 1. The simulation study is based on 100 samples of size 1000. The data are analyzed using an EFA model with target rotation where the targets are the entries in the parentheses in the above matrix

$$\lambda_{41} = \lambda_{51} = \lambda_{12} = \lambda_{72} = \lambda_{13} = \lambda_{43} = 0 \tag{32}$$

Obviously condition (a) is satisfied. Consider now the submatrices Λ_s. Since the s-th column of Λ_s by definition consists of all zeroes, that column will not contribute to the rank of Λ_s and thus the s-th column can be removed for simplicity. In the above example the submatrices are

$$\Lambda_1 = \begin{pmatrix} \lambda_{42} & \lambda_{43} \\ \lambda_{52} & \lambda_{53} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

$$\Lambda_2 = \begin{pmatrix} \lambda_{11} & \lambda_{13} \\ \lambda_{71} & \lambda_{73} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Lambda_3 = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{41} & \lambda_{42} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The ranks of these matrices are as follows: $rank(\Lambda_1) = 1$, $rank(\Lambda_2) = 2$, $rank(\Lambda_3) = 2$. Thus the submatrix Λ_1 does not satisfy the identifying condition (b) and it has to be modified, i.e., the targets in column 1 have to be modified. This is confirmed in the simulation. From the 100 samples, 13 samples recognized the model as a non-identified model. For the remaining samples many of the parameters have large standard error estimates and generally all parameter estimates are biased. The average absolute bias for all loading parameters is 0.511. The average standard error for the loading parameters is 1.393. Such large standard errors indicate a poorly identified model.

The reason that the non-identification is not recognized in all samples is as follows. While for the true parameter values $rank(\Lambda_1) = 1$, for individual samples the $rank(\hat{\Lambda}_1)$ may actually be 2 because of variation in the data generation and thus 87 of the 100 samples were considered identified. However, that identification is very poor because $\hat{\Lambda}_1$ is generally quite close to Λ_1, i.e., it is nearly singular and has deficiency in the rank.

Now consider an alternative target specification. Replace the target $\lambda_{51} = 0$ with the target $\lambda_{71} = 0$. All other targets remain the same. The new submatrix Λ_1 now is

$$\Lambda_1 = \left(\begin{array}{cc} \lambda_{42} & \lambda_{43} \\ \lambda_{72} & \lambda_{73} \end{array} \right) = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right)$$

which clearly has rank 2 and the model is now well identified. The results of the simulation confirm this. The average absolute bias for the loading estimates is now 0.003, and the average standard error for the loading estimates is 0.039.

Note that conditions (a) and (b) are generally speaking only sufficient conditions for identification. These conditions are strictly speaking not necessary. A necessary condition is the fact that there should be at least $m(m-1)$ targets, because that will lead to the $m(m-1)$ constraints needed for identification purposes. The above simulation example, however, suggests that for practical purposes one could treat conditions (a) and (b) also as necessary conditions.

For orthogonal rotations the identification requirements are similar, however, now only $(m-1)/2$ targets should be specified in each column, because the Ψ matrix has $m(m-1)/2$ additional constraints, beyond the m factor variances fixed at 1. If m is even $(m-1)/2$ is not an integer, so in that case the total number of targets has to be at

least $m(m-1)/2$ while each column can contain a different number of targets. Again, however, all submatrices Λ_s have to be of full rank.

7 Simulation Studies

A series of simulation studies will now be presented to illustrate the performance of the ESEM analysis. General considerations of the use of simulation studies with EFA and ESEM are presented in Appendix D. The simulation studies are conducted with Mplus 5.1. The Mplus input for the first simulation is given in Appendix E [7].

7.1 Small Cross Loadings

One of the advantages of ESEM modeling is that small cross loadings do not need to be eliminated from the model. Given the lack of standard errors for the rotated solution in most EFA software, common EFA modeling practice is to ignore all loadings below a certain threshold value such as 0.3 on a standardized scale, see Cudeck and O'Dell (1994). In subsequent CFA analysis such loadings are typically fixed to 0, see e.g. van Prooijen and van der Kloot (2001). Small model misspecifications such as these, however, can have a relatively large impact on the rest of the model.

In the following simulation study data are generated according to a 2-factor model with 10 indicator variables Y_j and one covariate X. Denote the two factors by η_1 and η_2. The model is specified by the following two equations

$$Y = \nu + \Lambda\eta + \varepsilon \tag{33}$$

$$\eta = B\ X + \xi \tag{34}$$

where ε is a zero mean normally distributed residuals with covariance matrix Θ and ξ are zero mean normally distributed residuals with covariance matrix Ψ. The following parameter values are used to generate the data. The intercept parameter $\nu = 0$, the residual

---

[7] A tutorial on Mplus simulation studies with ESEM is available in Mplus V5.1 Examples Addendum available at www.statmodel.com/ugexcerpts.shtml. In addition, all Mplus input and outputs for the simulation studies presented in this article are available by email from the second author: bmuthen@ucla.edu.

covariance Θ is a diagonal matrix with the value 0.36 on the diagonal. The loading matrix Λ is

$$\Lambda = \begin{pmatrix} 0.8 & 0 \\ 0.8 & 0 \\ 0.8 & 0 \\ 0.8 & 0.25 \\ 0.8 & 0.25 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix} \tag{35}$$

The values $\lambda_{42} = \lambda_{52} = 0.25$ represent the small cross loadings. The true value for Ψ is

$$\Psi = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

The true values for the regression slopes are

$$B = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}.$$

The covariate X has a standard normal distribution. The simulation study uses 100 samples of size 1000. The samples are then analyzed by ESEM based on Geomin rotation with $\epsilon = 0.0001$, ESEM based on Geomin rotation with $\epsilon = 0.01$, ESEM based on Quartimin rotation, as well as by the CFA-SEM model where the two cross loadings λ_{42} and λ_{52} are held fixed to 0. All methods produced unbiased estimates for ν and Θ parameters. The results for the remaining parameters are presented in Tables 1 and 2.

It is clear from these results that the consequences of eliminating small cross loadings in the SEM analysis can result in substantial bias in the rest of the parameters estimates as well as poor confidence interval coverage. Among the three ESEM methods the best results were obtained by the Geomin method with $\epsilon = 0.0001$. The Quartimin method and Geomin with $\epsilon = 0.01$ showed some small biases which leads to poor confidence interval coverage. In contrast, ESEM based on Geomin rotation with $\epsilon = 0.0001$ produces results with little bias for all parameters and coverage near the nominal 95% level. A simulation study based on samples with only 100 observations reveals

22

very similar results to the ones presented in Tables 1 and 2, i.e., these results appear to be independent of the sample size.

The chi-square test of fit for the model is also affected by the elimination of small cross loadings. Using a simulation with 500 samples of size 1000 the SEM model is rejected 100% of the time while the ESEM model is rejected only 7% of the time. For sample size of 100 the rejection rate for the SEM model is 50% and for the ESEM model it is 10%.

The simulation study presented here is not as easy to interpret as traditional simulation studies especially when it comes to comparing different rotation methods. To provide proper interpretation of the results one has to first accept the notion that the loading matrix presented in (35) is the simplest possible loading matrix among all rotated versions of that matrix. In particular one has to accept the notion that Λ given in (35) is simpler than rotations of Λ that have no zero loading values. If this simplicity notion is accepted then the simulation study can be interpreted in the traditional sense, i.e, the matrix Λ given in (35) is the true loading matrix that has to be estimated by the rotated loading matrix $\hat{\Lambda}$. Now suppose that, for some reason, an analyst decides that another rotated version of Λ is simpler than the one given in (35). In that case, the above simulation study would be irrelevant and a different rotation criterion, that targets the alternative rotated version of Λ, would have to be explored.

To illustrate the above point, consider the rotation results on the population level. Using the rotation algorithms with the true population parameters Λ, Ψ and Θ one can obtain the optimal rotations on the population level[8]. Denote the Λ rotations obtained by Quartimin, Geomin with $\epsilon = 0.01$, and Geomin with $\epsilon = 0.0001$ by Λ_q, $\Lambda_{0.01}$, and $\Lambda_{0.0001}$, respectively. Denote the corresponding Ψ rotations by Ψ_q, $\Psi_{0.01}$, and $\Psi_{0.0001}$, respectively[9]. These matrices are presented in Table 3 and 4. Finite sample based rotated parameter estimates are essentially consistent estimates of the rotated population values presented in Table 3 and 4. All four of these rotated solutions are equivalent in terms of model fit because the matrices are rotations of each other. To decide which rotation is optimal one has to consider the notion of simplicity. Which of the four Λ matrices should

_____

[8]Note that Θ also influences the rotation through the correlation standardization.

[9]In Mplus the population level rotations are obtained by generating a large sample, such as a sample with 1,000,000 observations. In such a large sample the estimated parameters are nearly identical with the population parameters.

23

Table 1: Comparison of ESEM and CFA-SEM with small cross loadings. Average parameter estimates.

Para-meter	True Value	CFA-SEM	ESEM Quartimin	ESEM Geomin(0.01)	ESEM Geomin(0.0001)
λ_{11}	0.80	0.75	0.84	0.82	0.81
λ_{21}	0.80	0.75	0.83	0.82	0.80
λ_{31}	0.80	0.75	0.83	0.82	0.81
λ_{41}	0.80	0.99	0.84	0.82	0.81
λ_{51}	0.80	0.99	0.84	0.83	0.81
λ_{61}	0.00	0.00	0.01	0.01	0.00
λ_{71}	0.00	0.00	0.01	0.01	0.00
λ_{81}	0.00	0.00	0.01	0.01	0.00
λ_{91}	0.00	0.00	0.01	0.01	0.00
λ_{101}	0.00	0.00	0.01	0.01	0.00
λ_{12}	0.00	0.00	-0.06	-0.03	-0.01
λ_{22}	0.00	0.00	-0.06	-0.03	-0.01
λ_{32}	0.00	0.00	-0.06	-0.04	-0.01
λ_{42}	0.25	0.00	0.18	0.21	0.24
λ_{52}	0.25	0.00	0.18	0.21	0.23
λ_{62}	0.80	0.80	0.80	0.80	0.80
λ_{72}	0.80	0.80	0.80	0.79	0.80
λ_{82}	0.80	0.80	0.80	0.80	0.80
λ_{92}	0.80	0.80	0.80	0.79	0.80
λ_{102}	0.80	0.80	0.80	0.80	0.80
β_1	0.50	0.61	0.56	0.54	0.52
β_2	1.00	1.00	1.00	1.00	1.00
ψ_{12}	0.50	0.61	0.55	0.53	0.51

Table 2: Comparison of ESEM and CFA-SEM with small cross loadings. Confidence intervals coverage.

Parameter	CFA-SEM	ESEM Quartimin	ESEM Geomin(0.01)	ESEM Geomin(0.0001)
λ_{11}	0.54	0.77	0.85	0.90
λ_{21}	0.48	0.87	0.97	0.97
λ_{31}	0.48	0.82	0.93	0.95
λ_{41}	0.00	0.78	0.86	0.95
λ_{51}	0.00	0.76	0.88	0.95
λ_{61}	1.00	0.98	0.97	1.00
λ_{71}	1.00	0.95	0.94	0.97
λ_{81}	1.00	0.96	0.98	1.00
λ_{91}	1.00	0.95	0.95	1.00
λ_{101}	1.00	0.95	0.92	0.97
λ_{12}	1.00	0.05	0.50	0.95
λ_{22}	1.00	0.05	0.46	0.96
λ_{32}	1.00	0.02	0.38	0.97
λ_{42}	0.00	0.24	0.66	0.91
λ_{52}	0.00	0.09	0.67	0.89
λ_{62}	0.99	0.98	0.97	0.98
λ_{72}	0.99	0.95	0.94	0.97
λ_{82}	0.94	0.96	0.96	0.96
λ_{92}	0.95	0.97	0.97	0.99
λ_{102}	0.94	0.97	0.97	0.97
β_1	0.13	0.59	0.83	0.94
β_2	0.96	0.97	0.97	0.97
ψ_{12}	0.01	0.44	0.77	0.93

Table 3: Rotation of population loading matrix.

Λ		Λ_q		$\Lambda_{0.01}$		$\Lambda_{0.0001}$	
0.80	0.00	0.80	-0.07	0.82	-0.03	0.80	-0.01
0.80	0.00	0.80	-0.07	0.82	-0.03	0.80	-0.01
0.80	0.00	0.80	-0.07	0.82	-0.03	0.80	-0.01
0.80	0.25	0.80	0.18	0.82	0.21	0.80	0.24
0.80	0.25	0.80	0.18	0.82	0.21	0.80	0.24
0.00	0.80	0.01	0.83	0.01	0.79	0.00	0.80
0.00	0.80	0.01	0.83	0.01	0.79	0.00	0.80
0.00	0.80	0.01	0.83	0.01	0.79	0.00	0.80
0.00	0.80	0.01	0.84	0.01	0.79	0.00	0.80
0.00	0.80	0.01	0.84	0.01	0.79	0.00	0.80

Table 4: Rotation of population correlation matrix.

Ψ		Ψ_q		$\Psi_{0.01}$		$\Psi_{0.0001}$	
1.00	0.50	1.00	0.55	1.00	0.52	1.00	0.51
0.50	1.00	0.55	1.00	0.52	1.00	0.51	1.00

be considered the simplest and the most interpretable? Regardless of the arguments and notion of simplicity in this example, one inevitably reaches the conclusion that the matrix Λ is the simplest. Therefore in the estimation process this matrix should be considered the desired matrix. It is clear that $\Lambda_{0.0001}$ is the closest to Λ and that is the reason why the Geomin rotation with $\epsilon = 0.0001$ produced the best results in the simulation study. If however for some reason one decides that Λ_q is the simplest possible matrix, then obviously the Quartimin rotation would be the optimal rotation method to use. Section 7.4 below describes a realistic example where two different loading matrices are quite likely to be considered as the simplest and most interpretable.

7.2 Chi-Square test of fit and likelihood ratio testing

Testing various aspects of ESEM can be done the same way as for regular SEM models. The standard chi-square test of fit which compares a structural model against an unrestricted mean and variance model can be done for ESEM the same way, i.e., using the likelihood ratio test (LRT) for the two models. For example, consider the question of how many factors are needed in the ESEM model. One standard approach is to sequentially fit models with 1, 2, ... etc. factors and then use the smallest number of factors for which the test of fit does not reject the model. Consider the simulation example in Section 7.1. Estimating the model with one factor leads to an average chi-square test of fit of 1908 with 44 degrees of freedom and 100% rejection rate, i.e., the LRT testing correctly identifies the 1-factor ESEM model as insufficient. In contrast, for the 2-factor ESEM model the average chi-square test of fit is 35 and with 34 degrees of freedom the rejection rate dropped to 9%, i.e., the LRT correctly finds the 2-factor ESEM model well fitting. It is possible to estimate even a 3-factor ESEM model, although convergence problems occur in 30 out of the 100 replications. The average chi-square test of fit for the 3-factor ESEM model is 20 and with 25 degrees of freedom this leads to 0% rejection rate.

Alternatively, the LRT can be used to test directly an $m-1$-factor ESEM model against an m- factor ESEM model, without testing the models against the unrestricted mean and variance models. In the above example, testing the 1-factor model against the 2-factor model gives an average chi-square test statistic of 1873 and with 8 degrees of freedom this leads to 100% rejection rate. In certain cases such direct testing can be preferable as it directly tests the hypothesis of interest, namely, whether or not the additional factor is needed. The direct test will also be more powerful than the general test of fit model, i.e., it will outperform the test of fit approach in small sample size problems.

In practice however not all of the residual correlation will be picked up by the unrestricted loading structure of the ESEM model and strictly using the chi-square test of fit will often lead to an unreasonable number of factors in the model, many of which contribute little to the overall model fit. In such cases one can use approximate fit indices such as SRMR, CFI and TLI to evaluate the fit of the model. One can also use the SRMR index to evaluate the improvement in the fit due to each additional factor. For example, if an additional factor

contributes less than 0.001 decrease in the SRMR, it seems unreasonable to include such factors in the model. Instead one can use the new ESEM modeling feature, extending the standard EFA analysis, by including residual covariance parameters in the model in addition to the exploratory factors. Furthermore it is possible to point out which residual covariances should be included in the model, and thereby improve factor stability and overall fit, by using standard modeling tools such as modification indices, standardized and normalized residuals.

The LRT testing can be used also to test an EFA model against a CFA model. Consider again the simulation example in Section 7.1 and the LRT test of this model against the CFA model based on all non-zero loadings, so including the two small cross loadings. Note first that the two models are nested. This is not very easy to see because of the parameter constraints imposed on the ESEM parameters by the rotation algorithm. There are 8 loading parameters that are fixed at 0 in the CFA-SEM but not in the ESEM. However, the ESEM model has 2 parameter constraints, imposed by the rotation algorithm, that involve all loading and factor covariance parameters. To see that the CFA model is nested within the ESEM model first note that the ESEM model is equivalent to its starting unrotated solution. The rotated solution has the same log-likelihood value as the unrotated starting value solution, and any testing of a model against an ESEM model is essentially a test against the unrotated starting value model. A number of different unrotated solutions can be used at this point. Two of these are generally convenient in assessing the model nesting. The first one is the orthogonal starting value where the factor variance covariance matrix is the identity matrix and the loadings above the main diagonal in the upper right hand corner are all fixed to 0. The second unrotated starting value solution that can be used is the oblique starting value where the factor variances are fixed to 1, the factor covariances are free and each loading column contains exactly $m - 1$ zeroes in locations that satisfy condition (b) given in Section 6. For example, a square submatrix of size m, can be selected from the loading matrix and in this submatrix all values except the main diagonal entries can be fixed to 0. In the above example one can use the oblique starting value solution to assert the nesting of the CFA and ESEM models. The ESEM model is equivalent to an unrotated oblique starting value solution with any 2 loadings from different rows fixed to 0. It is now clear that the CFA model can be thought of as more constrained than the ESEM model where the

28

additional constraints simply fix the remaining 6 loadings at 0.

Conducting the LRT test between the ESEM and CFA models for the simulation example in Section 7.1, using 100 samples of size 5000, the average test statistic is 5.73 and with 6 degrees of freedom that leads to a rejection rate of only 2%, i.e., the LRT correctly concludes that the CFA model with all 8 loadings fixed to 0 is well fitting.

Now consider the situation when both nested models are approximately fitting models, i.e., the models have small misspecifications but the sample size is large enough that even small misspecifications lead to poor tests of fit. For example, if the data generation in Section 7.1 is altered by adding a residual covariance between Y_7 and Y_8 of 0.05, using a sample size of 5000, both the ESEM and CFA models are rejected by the test of fit 100% of the time with average chi-squares test of fit statistics of 88 and 97 respectively. The average SRMR measures are 0.004 and 0.005 respectively, i.e., both models are fitting approximately in all 100 replications. Conducting the LRT between the CFA and ESEM models provides relatively good results here as well. The average LRT test statistic for testing the CFA model against the ESEM model is 8.65 and with 6 degrees of freedom, this leads to a 14% rejection rate. This suggests that even when the models are fitting the data only approximately, the LRT can be used to distinguish between ESEM and CFA models. The relatively small inflation in the rejection rate is due to the fact that the more flexible ESEM model is able to accommodate more of the model misspecifications than the CFA model. The inflation however is relatively small and the LRT can clearly be recommended. Even though both the ESEM and CFA models are incorrect in this simulation, the LRT correctly concludes that the 8 loadings are indeed 0.

7.3   Multiple Group ESEM

This section describes a multiple group example and demonstrates the constrained rotation technique described in Section 5 for group invariant loading matrices. Consider a two- group two-factor model with 10 dependent variables

$$Y = \nu_g + \Lambda_g \eta + \varepsilon \tag{36}$$

$$\eta = \alpha_g + \xi \tag{37}$$

where ε and ξ are zero mean residuals with covariance matrices Θ_g and Ψ_g. One common application of multiple group analysis is to

test measurement invariance across the groups, that is to test the hypothesis $\Lambda_1 = \Lambda_2$, see Joreskog and Sorbom (1979). Estimating the measurement invariance model is of interest as well. This simulation study evaluates the performance of the ESEM modeling technique for the measurement invariance model. Data is generated using the following parameter values $\nu_1 = \nu_2 = 1$, $\alpha_1 = 0$, $\alpha_2 = (0.5, 0.8)$, Θ_1 is a diagonal matrix with all diagonal values 1, Θ_2 is a diagonal matrix with all diagonal values 2,

$$\Lambda_1 = \Lambda_2 = \begin{pmatrix} 0.8 & 0 \\ 0.8 & 0 \\ 0.8 & 0 \\ 0.8 & 0 \\ 0.8 & 0 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \\ 0 & 0.8 \end{pmatrix}$$

$$\Psi_1 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

$$\Psi_2 = \begin{pmatrix} 1.5 & 1 \\ 1 & 2 \end{pmatrix}.$$

The simulation study is conducted for samples with 100 observations in each group as well as samples with 500 observations in each group. The simulation study is based on 100 samples for each of the two sample size specifications. For each of the samples the ESEM model is estimated with the following constraints. The loadings and the intercepts are held equal across the two groups

$$\Lambda_1 = \Lambda_2 \tag{38}$$

$$\nu_1 = \nu_2. \tag{39}$$

In the first group the factor variances are fixed to 1 and the factor means are fixed to 0

$$\psi_{111} = \psi_{221} = 1 \tag{40}$$

$$\alpha_1 = 0. \tag{41}$$

In addition, Θ_1 and Θ_2 are estimated as diagonal matrices, α_2 is estimated as a free vector, Ψ_2 is estimated as unrestricted variance matrix, while Ψ_1 is estimated as unrestricted correlation matrix. This model specification is a typical measurement invariance model. Other sets of identifying restrictions can be similarly specified. The model described above has a total of 54 independent parameters, 10 ν parameters, 10 Θ_1 parameters, 10 Θ_2 parameters, 20 Λ parameters, 3 Ψ_2 parameters, 2 α_2 parameters and ψ_{121}, minus the two parameter restrictions imposed on Ψ and Λ by the rotation algorithm. The ESEM model is estimated with the Geomin rotation and $\epsilon = 0.0001$. The average estimate for some of the parameters in the model and their confidence interval coverage are reported in Table 5. For sample size 500, all parameter estimates have negligible bias and the coverage is near the nominal 95% level. For sample size 100, the coverage is near the nominal 95% level, however, some of the parameter estimates show substantial bias, namely, the factor covariance parameter in both groups.

The results in Table 5 indicate that the small sample size properties of the ESEM models may be somewhat inferior to those of the traditional SEM. To investigate the small sample size parameter biases in the above simulation study the samples with 100 observations in each group are analyzed by the following three methods: the ESEM method with Geomin rotation and $\epsilon = 0.0001$, the ESEM method with Target rotation using all 0 loadings as targets, and the SEM with all 0 loadings fixed to 0. In practice both the ESEM-Target method and the SEM method can be used as a follow up model to the ESEM-Geomin method. Based on the ESEM-Geomin method, the ESEM-Target model is constructed by setting all loadings that are not significantly different from 0 as targets. Similarly, the SEM model is constructed by setting all loadings that are not significantly different from 0 as loadings that are fixed to 0. Note that while the parameter estimates for ESEM-Geomin show some small sample size bias for some parameters, the standard errors produced correct coverage for all parameters, i.e., when evaluating the significance of small loadings for purposes of constructing ESEM-Target model and the SEM model the ESEM-Geomin model will point out correctly all zero loadings.

The results of this simulation study are presented in Table 6, which contains the average parameter estimates and the mean squared error (MSE) for the parameter estimates. Small sample size results should be interpreted very cautiously. Usually there is no theoretical justi-

31

fication for preferring one method over another for small sample size and usually simulation studies are used to draw general conclusions. However, there is no guarantee that the results in one simulation study would be similar to the results of the same simulation study with different parameters and even in the same simulation study the results can be inconsistent. For example, in this simulation the covariance in the first group is best estimated by the SEM model, while the covariance in the second group is best estimated by the ESEM-Target model. Nevertheless, Table 6 seems to give general guidance for reducing small sample size biases. It appears that the additional information that ESEM-Target and SEM facilitate, namely that some loadings are small or even 0, does result in a reduction of the small sample size biases and the MSE of the parameter estimates. In addition, the SEM model does appear to have slightly smaller biases overall than the ESEM-Target method although this does not appear to be a consistent trend and for some parameters ESEM-Target produces better results. For many of the parameters the three methods produce nearly identical results. The SEM model has fewer number of parameters overall and thus can be expected in general to produce somewhat smaller biases and smaller MSE. Conducting this simulation for a sample size of 500 does not lead to any substantial difference between the three methods. Thus the differences presented in Table 6 are likely to occur only in small samples.

In addition, the usual chi-square test of fit which compares the estimated ESEM model against the unrestricted mean and variance two-group model can be used to evaluate the fit of the model. In this simulation study the model has 76 degrees of freedom. For a sample size of 100, the average test of fit statistic is 78.25 with a rejection rate at 5%. For a sample size of 500, the average test of fit statistic is 76.05 with a rejection rate at 5%. This shows that the chi-square test of fit works well for the ESEM models.

7.4 General factor

In certain EFA applications there is one main factor on which all items load. In addition, there can be other factors that are specific to the different items. This structure is also referred to as a bi-factor solution in the classic factor analysis text of Harman (1976). For example, consider a 3-factor model with 10 items with the following

Table 5: Two-group ESEM-Geomin analysis

Para- meter	True Value	n=100 Average Estimate	n=500 Average Estimate	n=100 Coverage	n=500 Coverage
λ_{11}	0.80	0.76	0.79	0.92	0.95
λ_{12}	0.00	0.04	0.01	0.97	0.99
ψ_{121}	0.50	0.42	0.49	0.98	0.98
ν_{11}	1.00	0.99	0.99	0.94	0.98
θ_{111}	1.00	0.97	1.00	0.93	0.99
α_{12}	0.50	0.47	0.51	0.93	0.91
α_{22}	0.80	0.81	0.82	0.96	0.95
ψ_{122}	1.00	0.92	0.98	0.92	0.96
ψ_{112}	1.50	1.58	1.50	0.92	0.96
ψ_{222}	2.00	2.03	2.02	0.93	0.95
θ_{112}	2.00	1.96	1.99	0.96	0.96

loading matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0.5 & 0 \\ 1 & 0.5 & 0 \\ 1 & 0.5 & 0 \\ 1 & 0.5 & 0 \\ 1 & 0 & 0.5 \\ 1 & 0 & 0.5 \\ 1 & 0 & 0.5 \end{pmatrix}. \tag{42}$$

Table 6: Two-group ESEM analysis, small sample size comparison of ESEM-Geomin, ESEM-Target, and SEM.

Para-meter	True Value	ESEM Geomin Average Estimate	ESEM Target Average Estimate	SEM Average Estimate	ESEM Geomin MSE	ESEM Target MSE	SEM MSE
λ_{11}	0.80	0.76	0.77	0.78	0.021	0.022	0.010
λ_{12}	0.00	0.04	0.02	0.00	0.014	0.012	0.000
ψ_{121}	0.50	0.42	0.45	0.48	0.021	0.014	0.012
ν_{11}	1.00	0.99	0.99	0.99	0.016	0.016	0.017
θ_{111}	1.00	0.97	0.97	0.98	0.027	0.027	0.025
α_{12}	0.50	0.47	0.48	0.49	0.044	0.043	0.041
α_{22}	0.80	0.81	0.82	0.82	0.041	0.040	0.040
ψ_{122}	1.00	0.92	0.99	1.04	0.107	0.095	0.101
ψ_{112}	1.50	1.58	1.59	1.60	0.284	0.291	0.275
ψ_{222}	2.00	2.03	2.04	2.05	0.305	0.298	0.292
θ_{112}	2.00	1.96	1.96	1.96	0.097	0.097	0.096

If one considers oblique rotations, there is a rotation of the above matrix that will have just one non-zero entry in each row

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1.12 & 0 \\ 0 & 1.12 & 0 \\ 0 & 1.12 & 0 \\ 0 & 1.12 & 0 \\ 0 & 0 & 1.12 \\ 0 & 0 & 1.12 \\ 0 & 0 & 1.12 \end{pmatrix} \tag{43}$$

$$\Psi = \begin{pmatrix} 1 & 0.89 & 0.89 \\ 0.89 & 1 & 0.80 \\ 0.89 & 0.80 & 1 \end{pmatrix} \tag{44}$$

Thus rotation criteria such as Quartimin that converge to complexity 1 solutions will not be able to recover the general factor structure (42). Geomin with $\epsilon = 0$ has two different optimal solutions, namely (42) and (43), both leading to a rotation function value of 0. For very small positive values of ϵ one can expect this to remain so. However, as ϵ increases, the rotation function can change sufficiently so that some of these multiple solutions are no longer local solutions. As ϵ increases the rotation function value for (43) will be lower because it has 2 zeroes in each row, i.e., the loadings matrix (43) will be the global minimum and (42) will be at best a local solution. In fact it is not clear whether (42) will represent a local solution at all. Even with $\epsilon = 10^{-4}$ using 30 random starting values, the GPA algorithm converged to solution (43) in all 30 replications. In general it is not easy to force a minimization algorithm to find local solutions, because minimization algorithms are designed to find global solutions. The rotation function value for (43) is 0.027 while for (42) it is 0.214, i.e., the two solutions are of different magnitude. If ϵ is chosen to be a smaller value, such as 10^{-6}, the rotation function values are closer, however, the convergence process is substantially more difficult. Many more replications are needed for convergence and the convergence criteria have to be relaxed as well. Using $\epsilon = 10^{-6}$ again most replications converge to solution (43), but another local solution is found that is different from both (42) and (43). In addition, in a simulation study, even if the GPA algorithm

35

is able to find consistently a particular local solution in all samples it is difficult to implement constraints that will always recognize that particular local solution so that when the results of the simulation are accumulated the same local solution is used. This investigation shows that relying on local Geomin solutions may not work well and that from practical perspective the loading matrix (42) should not be considered Geomin invariant.

For orthogonal rotations, however, the loading matrix (42) is Geomin invariant. This is demonstrated in the following simulation study that compares Geomin with $\epsilon = 0.001$ with another popular rotation method, Varimax. The simulation study is based on 100 samples of size 5000. The data is generated according to the above model and using the loading matrix (42). The intercept parameters $\nu = 0$, the residual variance for the indicator variables is 1, and the factor covariance matrix Ψ is the identity matrix. The results of the simulation study are presented in Table 7 for a representative set of parameters. The Geomin method produces unbiased parameter estimates with good confidence interval coverage. In contrast, the Varimax method produces biased parameter estimates and poor confidence interval coverage.

The Geomin method, however, has two solutions. The first solution is given in (42) and has rotation function values 0.28. The second solution

$$
\Lambda = \begin{pmatrix}
0.94 & 0.33 & 0 \\
0.94 & 0.33 & 0 \\
0.94 & 0.33 & 0 \\
1.06 & 0 & 0.35 \\
1.06 & 0 & 0.35 \\
1.06 & 0 & 0.35 \\
1.06 & 0 & 0.35 \\
1.06 & 0 & -0.35 \\
1.06 & 0 & -0.35 \\
1.06 & 0 & -0.35
\end{pmatrix}
\tag{45}
$$

has rotation function value 0.30. Using random starting values and the population parameters, the GPA algorithm converged to the global minimum of 0.28 about half of the time and the other half it converged to the local minimum of 0.30. When the sample size is sufficiently large, such as the 5000 used in this simulation, there will be two solutions, but they will consistently appear in the same order, i.e., the global minimum in all finite sample size replications will correspond

36

Table 7: General Factor ESEM analysis with orthogonal rotation

Para-meter	True Value	Geomin Average	Varimax Average	Geomin Coverage	Varimax Coverage
λ_{11}	1.00	1.00	0.58	0.91	0.00
λ_{12}	0.00	-0.01	0.57	1.00	0.00
λ_{13}	0.00	0.01	0.58	0.98	0.00
λ_{41}	1.00	1.01	0.90	0.96	0.71
λ_{42}	0.50	0.49	0.37	0.94	0.00
λ_{43}	0.00	0.00	0.47	0.98	0.00
λ_{81}	1.00	1.00	0.45	0.96	0.00
λ_{82}	0.00	-0.01	0.31	0.98	0.00
λ_{83}	0.50	0.50	0.96	0.97	0.00

to the global minimum solution in the population model. Thus an algorithm that always selects the global minimum, will essentially always select the same solution. If however, the sample size is smaller, the global and the local solutions will switch orders across the replications, and thus an algorithm that always selects the global minimum will essentially average the two different solutions and thus will render useless results[10]. A more advanced algorithm that includes a method for picking the same local solution would avoid that problem. This issue is important only in simulation studies. In single replication studies such as real data analysis, one has to simply evaluate all local solutions and choose the one that is simplest and easiest to interpret.

When a general factor model is anticipated and oblique rotation is used, the Target rotation method may be a better alternative. The next section illustrates the Target rotation with a complex loading structure.

7.5 Complexity 3

In this section the advantages of the Target rotation are demonstrated with a complexity 3 example, i.e., an example with 3 non-zero loadings in a row. The three methods compared in this section are the Target rotation, the Geomin rotation with $\epsilon = 0.01$, and the Geomin rotation

---

[10]Future version of Mplus will include tools for resolving this problem.

with $\epsilon = 0.0001$. Consider a 4-factor 12-indicator factor analysis model with the intercept parameter $\nu = 0$, the covariance matrices Ψ and Θ as the identity matrices, and Λ as follows

$$\Lambda = \begin{pmatrix} 1 & (0) & (0) & (0) \\ 1 & 0 & 0 & 0 \\ 1 & 0.5 & 0 & 0 \\ (0) & 1 & (0) & (0) \\ 0 & 1 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \\ (0) & (0) & 1 & (0) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0.5 & 0.5 & 1 \\ (0) & (0) & (0) & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{46}$$

The complexity of Y_{10} is 3. The entries in the parentheses represent the targets for the Target rotation. One easy way to select targets and avoid any identification problems is to identify pure factor indicators, i.e., identify one variable for each factor that loads only on that variable just like in this example. The rank condition is then automatically satisfied. When each factor has a pure indicator one can set all zero loadings for the pure indicators as targets and the loading matrix is then Target invariant, i.e, the estimates are asymptotically unbiased under the Target rotation. Tables 8 and 9 contain the results of the simulation study based on the above model and conducted over 100 samples of size 5000. A representative set of loadings parameters is presented in the tables. Both Geomin-based estimations produced biased estimates. The bias of the estimates based on Geomin with $\epsilon = 0.01$ is smaller. The coverage of the Geomin-based estimation is also quite poor. In contrast, the Target rotation shows negligible bias and coverage near the 95% nominal level.

One can investigate the source of the Geomin bias by conducting the rotation on the population values and investigating all local solutions. Using $\epsilon = 0.0001$ Geomin has more than 5 local solutions that have similar rotation function values. One of these solutions corresponds to (46). Thus the simulation study presented here somewhat unfairly evaluates Geomin. If the algorithm included evaluation of the different local Geomin solutions and included a constraint to make the additional selection among these solutions so that the solution corre-

Table 8: Complexity 3 ESEM analyses. Average estimates.

Para- meter	True Value	Geomin $\epsilon = 0.0001$	Geomin $\epsilon = 0.01$	Target
λ_{11}	1.00	1.00	1.00	1.00
λ_{12}	0.00	0.00	-0.03	0.00
λ_{13}	0.00	0.00	0.01	0.00
λ_{14}	0.00	0.00	0.00	0.00
λ_{51}	0.00	0.00	0.00	0.00
λ_{52}	1.00	1.00	0.99	1.00
λ_{53}	0.50	0.49	0.45	0.50
λ_{54}	0.00	0.00	0.00	0.00
λ_{101}	0.00	0.00	0.00	0.00
λ_{102}	0.50	0.25	0.44	0.50
λ_{103}	0.50	0.25	0.41	0.50
λ_{104}	1.00	1.12	1.01	1.01
ψ_{12}	0.00	0.00	0.03	-0.01
ψ_{34}	0.00	0.22	0.06	0.00

sponding to (46) is always selected, there would be no bias. The bias in the simulation study is caused by the fact that the average estimates really represent the average estimates among a mixed sets of local Geomin solutions, instead of the same solution. In real data examples this is essentially a non-existent problem because one simply has to consider the various Geomin local solutions.

8   Choosing the right rotation criterion

In most ESEM applications the choice of the rotation criterion will have little or no effect on the rotated parameter estimates. In some applications, however, the choice of the rotation criterion will be critical and in such situations one has to make a choice. This section describes the underlying principles that one can follow to make that choice.

Choosing the right rotation is essentially a post estimation decision and there is no right or wrong rotation. The goal of the rotation

Table 9: Complexity 3 ESEM analyses. Coverage.

Para-meter	Geomin $\epsilon = 0.0001$	Geomin $\epsilon = 0.01$	Target
λ_{11}	0.94	0.94	0.94
λ_{12}	1.00	0.12	1.00
λ_{13}	0.97	0.87	1.00
λ_{14}	0.97	0.92	1.00
λ_{51}	0.98	0.97	0.94
λ_{52}	0.98	0.97	0.99
λ_{53}	0.99	0.26	0.99
λ_{54}	0.90	0.90	0.95
λ_{101}	0.99	0.93	0.97
λ_{102}	0.50	0.22	0.95
λ_{103}	0.45	0.05	0.97
λ_{104}	0.00	0.94	0.94
ψ_{12}	0.97	0.61	0.94
ψ_{34}	0.45	0.08	0.94

algorithms is to select the simplest and most interpretable loading structure. It is ultimately the analyst's choice and perception of what the simplest and most interpretable loading structure is. It is the analyst's choice of what the rotation criterion should be and which of the multiple rotated solutions represents the best loading structure for that particular application. Understanding the properties of the different rotation criteria will help the analyst in exploring the various rotation criteria. In particular, understanding the type of loading structures that each of the rotation criteria can reproduce, i.e., the invariant loading structures, is essential.

Estimation methods based on fit function optimizations such as the maximum-likelihood and least squares estimation methods would only accept the global optimum as the proper solution and local optima are perceived as estimation problems that have to be resolved so that the global optimum is always obtained. This is not the case, however, when it comes to local minima for the rotation criteria. Understanding and exploring the ability of rotation criteria such as Geomin to produce multiple optimal solutions can help the analyst in finding the best loading structure. It will generally be useful to consider the alternative top 2 or 3 Geomin solutions when such solutions are available[11]. Similarly changing the ϵ value in Geomin is equivalent to changing the rotation criterion. There is no correct or incorrect ϵ value. Different values for this parameter produce different rotation criteria that can enable the analyst to fine tune the loading matrix. In fact it is important that the analyst explores the sensitivity of the Geomin solution with respect to the ϵ value. In particular ϵ values such as $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$ should always be used.

To summarize, there is no statistical reason to prefer one rotation criterion over another, one ϵ value over another, or one local minimum over another. It is entirely in the hands of the analyst to make the choice and interpret the results. It is not the data that decides what a simple loading structure is, it is not the estimator, and it is not the rotation method. The analyst alone has to decide that. While for many simple loading structures, such as (31), most analysts will agree that no alternative rotation of Λ is simpler and more interpretable,

---

[11]Mplus will automatically run 30 random starting values with the Geomin rotation. More random starting values can be requested using the *rstarts=* command. In addition the different rotation values are presented in regular EFA analysis, as well as the loading structures for the different local minima. The ESEM output in Mplus 5.1 presents only the Geomin solution with lowest rotation function value.

that is not the case for other loading structures such as (42) and (43). For more complicated loading structures analysts can disagree on what the simplest loading structure is, even when the same rotation criterion is used and different local minima are selected. There is no statistical tool to resolve such disagreement and multiple equally valid solutions can be used.

9   Discussion

This paper has presented a new approach to structural equation modeling (SEM) which extends the types of measurement models that can be used. Adding the possibility of an exploratory factor analysis (EFA) measurement specification, strict loading restrictions in line with confirmatory factor analysis (CFA) are not necessary. The resulting ESEM approach has the full generality of regular SEM. From an EFA perspective, this implies that EFA can be performed while allowing correlated residuals, covariates including direct effects on the factor indicators, longitudinal EFA with across-time invariance testing, and multiple-group EFA with across-group invariance testing. Several factor loading rotation methods are available, including Geomin and Target rotation.

The main advantage of the ESEM model over existing modeling practices is that incorporates seamlessly the EFA and SEM models. In most applications with multiple factors the EFA analysis is used to discover and formulate factors. Usually the EFA analysis is followed by an ad-hoc procedure that mimics the EFA factor definitions in a SEM model with a CFA measurement specification. The ESEM model accomplishes this task in a one step approach and thus it is a simpler approach. In addition, the ESEM approach is more accurate because it avoids potential pitfalls due to the challenging EFA to CFA conversion. For example, EFA-based CFA model may lead to poor fit when covariates are added to the model. The ESEM approach avoids this problem by estimating the measurement and structural model parts simultaneously.

Typical CFA approaches draw on EFA to formulate a simple structure loading specification. The EFA is typically carried out without obtaining standard errors and instead rules of thumb such as ignoring loadings less than 0.3 are used. A CFA based on such an EFA often leads to a misspecified model using chi-square testing of model

fit. Model modification searches may not lead to the correct model and fit indices such as CFA may show sufficiently high values for the model not to be rejected. The paper illustrates the possible distortion of estimates that such a CFA-SEM approach can lead to and shows how ESEM avoids the misestimation.

ESEM makes possible better model testing sequences. Starting with an EFA measurement specification of only the number of factors, CFA restrictions can be added to the measurement model. Chi-square difference testing can be carried out to study the appropriateness of the CFA restrictions. Previously such testing sequences have been available only outside the SEM model structure, but they can now be integrated into SEM.

For many applications the ESEM model can be considered as a replacement of the more restrictive SEM model. Unlike EFA analysis, which is typically followed by a CFA analysis, the ESEM model does not need to be followed by a SEM model, because it has all of the features and flexibilities of the SEM model. Nevertheless, in certain cases it may be beneficial to follow an ESEM model by a SEM model. For example, in small sample size studies a follow-up SEM model may have more precise estimates because it has fewer parameters. Constructing a follow-up SEM model from a given ESEM model is fairly easy, amounting to fixing at 0 all insignificant loadings. In addition, because the ESEM and SEM models are typically nested a rigorous test can be conducted to evaluate the restrictions imposed by the SEM model.

The ESEM modeling framework does not limit the researcher's ability to incorporate substantive information in the model. The researcher can use different rotation criteria to reach the factor pattern that most closely represents the substantive thinking, without sacrificing the fit of the model.

The paper also discusses the performance of rotation techniques in Monte Carlo studies, showing the advantage of Geomin. Target rotation is shown to provide an approach that bridges EFA and CFA measurement specification.

Longitudinal and multiple-group analysis with EFA measurement structures greatly expands the possibilities of both EFA and SEM. The paper illustrates multiple group analysis in a simulation study.

Another advantage of the ESEM framework is that it easily accommodates EFA simulation studies. Such studies have been rarely published previously. In this new framework EFA simulation stud-

ies are as simple as SEM simulation studies. Simulation studies can greatly enhance this research field.

The ESEM approach is implemented in Mplus Version 5.1 and is developed not only for continuous outcomes with maximum-likelihood estimation but also for dichotomous, ordered categorical, censored and combinations of such outcomes with continuous outcomes with limited-information weighted least squares estimation. Other analysis features available include model modification indices, standardized co-efficients and their standard errors, estimation of indirect effects and their standard errors, factor scores, and Monte Carlo simulations.

10 Appendix A. Additional Rotation Criteria

Following is a list of additional rotation criteria implemented in Mplus.

- Varimax .or. CF-Varimax

$$f(\Lambda) = -\sum_{i=1}^{p} \left(\frac{1}{m} \sum_{j=1}^{m} \lambda_{ij}^4 - \left(\frac{1}{m} \sum_{j=1}^{m} \lambda_{ik}^2 \right)^2 \right). \qquad (47)$$

- CF-Quartimax

$$f(\Lambda) = -\frac{1}{4} \sum_{i=1}^{p} \lambda_{ij}^4 \qquad (48)$$

- CF-Equamax

$$f(\Lambda) = \frac{2p-m}{2p} \sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{l \neq j}^{m} \lambda_{ij}^2 \lambda_{il}^2 + \frac{m}{2p} \sum_{j=1}^{m} \sum_{i=1}^{p} \sum_{l \neq i}^{p} \lambda_{ij}^2 \lambda_{lj}^2 \qquad (49)$$

- CF-Parsimax

$$f(\Lambda) = \frac{p-1}{p+m-2} \sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{l \neq j}^{m} \lambda_{ij}^2 \lambda_{il}^2 + \frac{m-1}{p+m-2} \sum_{j=1}^{m} \sum_{i=1}^{p} \sum_{l \neq i}^{p} \lambda_{ij}^2 \lambda_{lj}^2 \qquad (50)$$

- CF-Facparsim, Factor Parsimony

$$f(\Lambda) = \sum_{j=1}^{m} \sum_{i=1}^{p} \sum_{l \neq i}^{p} \lambda_{ij}^2 \lambda_{lj}^2 \qquad (51)$$

- Crawfer, Crawford-Ferguson family

$$f(\Lambda) = (1-k) \sum_{i=1}^{p} \sum_{j=1}^{m} \sum_{l \neq j}^{m} \lambda_{ij}^2 \lambda_{il}^2 + k \sum_{j=1}^{m} \sum_{i=1}^{p} \sum_{l \neq i}^{p} \lambda_{ij}^2 \lambda_{lj}^2 \qquad (52)$$

where k is a parameter.

- Oblimin

$$f(\Lambda) = \sum_{j=1}^{m} \sum_{l \neq j}^{m} \left(p \sum_{i=1}^{p} \lambda_{ij}^2 \lambda_{il}^2 - k \sum_{i=1}^{p} \lambda_{ij}^2 \sum_{i=1}^{p} \lambda_{il}^2 \right) \qquad (53)$$

where k is the parameter.

11   Appendix B. Row Standardization

Typically the optimal rotation is determined by minimizing the rotation criteria using the standardized loadings, i.e., the loadings standardized to correlation scale as in equations (10) and (11). An alternative standardization frequently used in practice is the Kaiser standardization. In that case the optimal rotation is determined by minimizing the rotation criteria

$$f(D_d^{-1}\Lambda H^{-1}) \tag{54}$$

over all oblique or orthogonal matrices H where

$$D_d = \sqrt{diag(\Lambda\Lambda^T)} \tag{55}$$

Another alternative approach implemented in Mplus is to determine the optimal rotation by using the raw loadings matrix, using the original scales of the variables. In that case

$$f(\Lambda H^{-1}) \tag{56}$$

is minimized over all oblique or orthogonal matrices H.[12]

12   Appendix C. EFA Standard Errors

The asymptotic distribution of the rotated solution is based on the following general fit function method. Suppose that S_0 is a correlation matrix and Σ_0 is the estimated correlation matrix, based on an EFA model. Let $F(S_0, \Sigma_0)$ be a general fit function that is minimized to obtain the EFA parameters Λ and Ψ under the rotation constraints (7) or (9), and denote these constraint equations by R. Two examples of such functions are the likelihood fit function

$$F(S_0, \Sigma_0) = ln(|\Sigma_0|) + Tr(\Sigma_0^{-1}S_0) \tag{57}$$

and the least squares fit function

$$F(S_0, \Sigma_0) = \sum_{i<j}(\sigma_{0ij} - s_{0ij})^2. \tag{58}$$

---

[12]The standardization option is controlled in Mplus by the *RowStandardization=* command and the three options described above are *RowStandardization= Correlation, Kaiser* or *Covariance.*

It is possible to obtain the asymptotic distribution of the rotated solutions using the asymptotic distribution of S_0. Using the Lagrange multipliers method the rotated solution is also the local extremum for the augmented function

$$F_1(S_0, \Sigma_0) = F(S_0, \Sigma_0) + L^T R \qquad (59)$$

where L is a vector of new parameters. The asymptotic distribution for the parameters that minimize the new fit function is obtained, see Theorem 4.1 in Amemiya (1985), by the sandwich estimator

$$(F_1'')^{-1} Var(F_1')(F_1'')^{-1} \qquad (60)$$

where the second derivative with respect to the model parameters and the new parameters L is given by

$$F_1'' = \begin{pmatrix} F'' & R' \\ R' & 0 \end{pmatrix}. \qquad (61)$$

The middle term is the variance of the score and is computed as follows

$$Var(F_1') = \frac{\partial^2 F_1}{\partial \theta \partial S_0} Var(S_0) \left(\frac{\partial^2 F_1}{\partial \theta \partial S_0} \right)^T \qquad (62)$$

where θ is the vector of model parameters and

$$\frac{\partial^2 F_1}{\partial \theta \partial S_0} = \begin{pmatrix} \frac{\partial^2 F}{\partial \theta \partial S_0} \\ 0 \end{pmatrix}. \qquad (63)$$

The general fit function method described above is utilized in ESEM as follows. Using the asymptotic distribution of the unrotated solution, the asymptotic distribution of the estimated correlation matrix is computed via the delta method. The asymptotic distribution of the rotated solution is then obtained from the general fit function method by substituting the estimated correlation matrix for S_0 above and using either the (57) or (58) fit functions. Because the fit of the model is perfect, both fit functions lead to the same result.

13   Appendix D. Simulation Studies with ESEM and EFA

In ESEM as well as EFA analysis the order of all factors is interchangeable and each factor is interchangeable with its negative. These

indeterminacies are typically not important. However, they are important in simulation studies where accumulations across the different replications is done to evaluate mean-squared error (MSE), parameter estimates bias and, confidence interval coverage.

To avoid this problem additional parameter constraints are used. For example, to identify a factor over its negative the following restriction on the loadings is incorporated

$$\sum_i \lambda_{ij} > 0. \tag{64}$$

In addition, to make sure that the factors appear consistently in the same order across the replications the following quantities are computed

$$d_j = \frac{average \quad index \quad of \quad the \quad large \quad loadings}{\sum_i \lambda_{ij}^2} \tag{65}$$

where the large loadings are the loadings that are at least 0.8 of the largest loading. For example suppose that the loadings of a factor are (0.2, 1, 0.9, 0.9, 0, 0.1). The large loadings are loadings 2, 3 and 4, and therefore the average index of the large loadings is 3. The factors are ordered so that

$$d_1 < d_2 < ... < d_m. \tag{66}$$

This rule guarantees that factors with large loadings on the first dependent variables will tend to appear first[13]. In addition, factors that explain more of the dependent variables' covariance matrix will appear first. This is the effect of the denominator in the definition of d_j.

Simulation studies that are presented here are constructed in a way that ensures that the order of the factors is the same across the replications as well as the sign of the factors. The constraints (64) and (66), however, will not work for any simulation study and a different set of constraints may have to be used to ensure stable factor order and factor signs. Simulation studies that do not include proper constraints similar to (64) and (66) will lead to meaningless results as they will

---

[13]In simulation studies for SEM models Mplus uses user specified starting values to ensure that the order of the factors is the same across the replications. However, ESEM and EFA analysis in Mplus do not use user specified starting values.

combine factor loadings from different factors across the replications[14]. Such simulation studies will not give good results and will not provide any information for the quality of the estimation method. Parameter constraints (64) and (66) are important only for simulation studies. These constraints have no implication for a single replication analysis such as real data analysis. It is well known that the order of the factor is exchangeable and that each factor can be replaced with its negative. Because the data does not contain any information about the order of the factors or their signs, it is up to the analyst to make that choice[15].

14 Appendix E. Mplus Input

Following is the Mplus input for the small cross loadings simulation study presented in Section 7.1. Comments lines begin with (!) and are provided here only for clarity. They are not needed in general.

! this section specifies the simulation framework
montecarlo:

> *names = y1-y10 x;*
> *nobs = 1000;*
> *nreps = 100;*

! this section specifies the parameters for the data generation
model population:

> *[x@0]; x@1;*
> *f1 by y1-y5*.8 y6-y10*0;*
> *f2 by y1-y3*0 y4-y5*.25 y6-y10*.8;*

---

[14]It is not possible to specify an alternative set of constraints in Mplus. To conduct proper simulation studies in Mplus the dependent variables should be carefully ordered so that the factors do not reverse the order. For example, the variables that load on the first factor should be placed first in the *names=* command. However, for some models, even carefully arranged simulation studies may lead to factor switching, especially when the sample size is small and the number of replications is large. It is recommended that the parameter estimates obtained from each replication are inspected for possible factor switching. Future versions of Mplus will include additional tools to prevent factor switching.

[15]Mplus will use the constraints (64) and (66) even for real data analysis, so the factors and their signs are always uniquely determined by Mplus.

y1-y10*.36; [y1-y10*0];
f1-f2@1;
f1 with f2*.5;
f1 on x*.5;
f2 on x*1;

! this section specifies the rotation type
analysis: rotation = geomin(0.0001);

! this section specifies the model to be estimated and the true
! values to be used for confidence interval coverage rates
model:

f1 by y1-y5*.8 y6-y10*0 (*1);
f2 by y1-y3*0 y4-y5*.25 y6-y10*.8(*1);
y1-y10*.36; [y1-y10*0];
f1 with f2*.5;
f1 on x*.5;
f2 on x*1;

References

[1] Amemiya, T. (1985). *Advanced Econometrics.* Harvard University Press.

[2] Archer, C. & Jennrich, R. I. (1973). Standard errors for rotated factor loadings. *Psychometrika* **38**, 581-592.

[3] Asparouhov T. & Muthen B. (2007). Computing the Asymptotic Distribution of the Rotated Solution in Exploratory Factor Analysis Using Mplus.

[4] Bollen, K.A. (1989). *Structural equations with latent variables.* New York: John Wiley & Sons.

[5] Browne, M.W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research* **36**, 111-150.

[6] Browne, M.W. & Arminger, G. (1995). Specification and estimation of mean- and covariance-structure models. In G. Arminger, C.C. Clogg & M.E. Sobel (eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 311-359). New York: Plenum Press.

[7] Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, **31**, 33-42.

[8] Cudeck, R. & O'Dell, L.L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin*, **115**, 475-487.

[9] Harman, H. H. (1976). *Modern factor analysis.* Third edition. Chicago: The University of Chicago Press.

[10] Howe, W. G. (1955). Some contributions to factor analysis (Rep. No. ORNL1919). Oak Ridge, TN: Oak Ridge National Laboratory.

[11] Jennrich, R.I. (2001). A simple general procedure for orthogonal rotation. *Psychometrika*, **66**, 289306.

[12] Jennrich, R.I. (2002). A simple general method for oblique rotation. *Psychometrika*, **67**, 719.

[13] Jennrich, R.I. (2007). *Rotation methods, algorithms, and standard errors. In Factor Analysis at 100: Historical Developments and*

Future Directions. Edited by Robert C. MacCallum and Robert Cudeck.

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

[15] Joreskog, K.G. (1969). A general approach to confirmatory maximum-likelihood factor analysis. *Psychometrika* **34**, 183-202.

[16] Joreskog, K.G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika* **43**, 443-477.

[17] Joreskog, K.G. (1979). A general approach to confirmatory maximum likelihood factor analysis, with addendum. In K. G. Joreskog & D. Sorbom, *Advances in factor analysis and structural equation models* (J. Magidson, Ed., pp. 2143). Cambridge, MA: Abt Books.

[18] Joreskog, K.G. & Sorbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge, MA: Abt Books.

[19] MacCallum, R.C., Roznowski, M., & Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin* **111**, 490-504.

[20] McDonald, R.P. (2005). Semiconfirmatory factor analysis: The example of anxiety and depression. *Structural Equation Modeling*, **12**, 163-172.

[21] Meredith, W. (1964). Rotation to achieve factorial invariance. *Psychometrika*, **29**, 187-206.

[22] Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance, *Psychometrika*, **58**, 525-543.

[23] Mulaik, S. & Millsap, R. (2000) Doing the Four-Step Right. *Structural Equation Modeling Journal*, **7**, 3673.

[24] Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115-132.

[25] Rozeboom W. (1992). The Glory of Suboptimal Factor Rotation: Why Local Minima in Analytic Optimization of Simple Structure are More Blessing Than Curse. *Multivariate Behavioral Research*, **27**, 585-599.

[26] Thurstone, L. L. (1947). Multiple factor analysis. Chicago: University of Chicago Press.

[27] Tucker, L.R. (1944). A semi-analytical method of factorial rotation to simple structure. *Psychometrika*, **9**, 43-68.

[28] van Prooijen, J. & van der Kloot, W. (2001). *Educational and Psychological Measurement*, **61**, 777-792

[29] Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis.* Albany: State University of New York Press.