Multiple group alignment for exploratory and structural equation models

Tihomir Asparouhov and Bengt Muthén

Mplus

May 4, 2022

Abstract

The multiple group alignment methodology is adapted to the general structural equation model. This includes models with cross-loadings and covariates. Group specific model for the factors can be estimated even when measurement invariance doesn't hold. The methodology is also extended to the weighted least squares estimation method to accommodate models with continuous, binary and ordered variables. The alignment loss function is combined with the rotation loss function to extend alignment to multiple group EFA and ESEM models with group-specific factor means, variance/covariance, rotation and loadings. Simulations studies and an empirical example are used for illustration purposes.

1 Introduction

The alignment methodology was introduced in Asparouhov and Muthén (2014) for continuous variables and in Muthén and Asparouhov (2014) for categorical variables. This method aims to compare latent variables across groups without requiring measurement invariance. The differences in the observed variables across groups are primarily attributed to differences in the latent variables. Remaining differences that cannot be explained by latent variable differences are interpreted as evidence of partial non-invariance. The alignment method automates this process within a single stage estimation. The model fit is the same as the model fit of the configural model, i.e., the fit of the alignment model is as good as or better than any other measurement invariance model. The alignment method attempts to minimize the amount of non-invariance without altering the fit of the model.

Alignment utilizes the EFA methodology in the following sense. In EFA, an unrotated model is estimated as a first step which determines the best fitting variance covariance matrix for the observed variables given a fixed number of factors. The unrotated model can be rotated with an infinite number of rotations without altering the model fit. This provides an indeterminacy in the model, i.e., there is no information in the data that can illuminate the best possible rotation for the factors. This indeterminacy is resolved by specifying a rotation criterion. The role of the rotation criterion is to eliminate the indeterminacy in the model by quantifying our preference for simple loading structures. These are the loading structures where each observed variable loads primarily on one factor only and the number and size of cross-loadings is minimized. Alignment uses the same logic. The configural model plays the role of the unrotated solution, i.e, this is the best fitting model given the number of factors and factor structure. The configural model can be reparameterized to include arbitrary values for the factor means and variances, without altering the model fit. The factor means and variances are unidentifiable. This indeterminacy is resolved by specifying an alignment criterion. The role of the alignment criterion is to eliminate the indeterminacy in the model by quantifying our preference for measurement invariant structures. The alignment method gives preference to as many invariant parameters and as few non-invariant parameters as possible. This parallelism between EFA and alignment can be very useful in understanding the alignment methodology.

The alignment methodology was implemented in Mplus version 7.1 Several extensions have been added in Mplus since the original release. In Version 7.2 the alignment methodology was extended to binary items and in Version 7.3 to ordered polytomous items. Three estimation methods have been developed previously: maximum-likelihood, Bayesian estimation, as well as BSEM based alignment. The BSEM based alignment starts with a BSEM measurement invariance estimation, instead of the configural model, and is followed up with alignment.

Mplus 8.8 extends the alignment methodology in several important ways. First, the alignment method is implemented for the WLS estimators with the delta and theta parameterizations for categorical variables. This extension is valuable for those situations where the ML estimation is slow due to numerical integration, i.e., factor models with categorical indicators that have more than 1 factor. Furthermore, the WLS estimators can accommodate residual correlations between all factor indicators, including categorical indicators, which is not available with the ML estimation for categorical indicators. The implementation of the WLS alignment estimation for the two parameterizations also allows us to study the effect of the parameterization on the measurement invariance across groups. The WLS alignment estimation also includes test of fit and modification indices which simplifies the overall analysis in terms of the number of steps that need to be taken in analyzing multiple group measurement models.

The second important generalization of the alignment method in Mplus 8.8 is the possibility of complex loading structures. Prior to Mplus 8.8, the alignment method was available only for simple loading structures, i.e., models without cross-loadings. Most practical examples however include multiple factors where the loading structure is not pure/simple and cross-loadings are present. This generalization therefore allows us to apply the alignment method in most practical situations.

Another important generalization of the alignment method in Mplus 8.8 is the possibility to apply the methodology for a general structural equation model. Prior to Mplus 8.8, only factor analysis models could be estimated with the alignment method. It is now possible to estimate general SEM models with alignment, including adding covariates / factor predictors, adding direct effects from the covariates to the factor indicators, and correlating the factors with other dependent variables. This approach resembles the ESEM extension of EFA models to general SEM models where the measurement is simply an EFA analysis rather than a CFA analysis. This extension simplifies the overall analysis by reducing the number of steps and the number of models that must be estimated. Previously, adding a factor predictor as part of the alignment estimation required the AwC two-stage estimation described in Marsh et al. (2018). With this new automated approach, a single model estimation is conducted that includes both the factor analysis alignment and the factor covariates. Thus, the perils of multistage estimation are avoided. We will abbreviate the aligned SEM model as ASEM.

Another alignment generalization implemented in Mplus 8.8. is the possibility to align multiple group EFA and ESEM models, see Asparouhov and Muthén (2009), with continuous variables (ML estimation) and with the combination of continuous and categorical variables (WLS estimation). We will abbreviate the aligned ESEM model as AESEM.

The Mplus alignment language has also been simplified. The original Mplus implementation was based on a Mixture model setup with a known-class specification for the grouping variable. In Mplus 8.8 the simpler multiple group specification can be used in all alignment models except those that are based on numerical integration.

The article proceeds as follows. In Section 2 we describe the generalized alignment methodology. In Section 3 we illustrate the quality of the alignment estimation with several simulation studies. In Section 4 we illustrate the methodology with an empirical example. Section 5 describes the alignment R-square measure of invariance in the generalized settings. Some practical guidelines are provided in Section 6 and Section 7 concludes.

2 The alignment methodology

In this section we describe the alignment methodology. First, we review the alignment methodology for the factor analysis model with one factor. We then describe several extensions and generalizations.

2.1 The simple alignment model

In this section we review the basic alignment model for a factor analysis model with one factor. Consider the multiple-group factor analysis model with a single factor η measured by P observed variables in G groups. Let Y_{ipg} be the p-the observed variable for individual i in group g. The factor model is given by the following equation.

$$Y_{ipg} = \nu_{pg} + \lambda_{pg}\eta_{ig} + \varepsilon_{ipg}, \quad (1)$$

where ν_{pg} and λ_{pg} are the intercept and loading parameters, $\varepsilon_{ipg} \sim N(0, \theta_{pg})$ is the residual variable, and $\eta_{ig} \sim N(\alpha_g, \psi_g)$ is the factor for individual i in group g. The alignment method estimates all of the parameters ν_{pg}, λ_{pg}, α_g, ψ_g, θ_{pg} as group specific parameters.

In particular, the method estimates group specific factor mean and variance without assuming measurement invariance.

The first step in the alignment method is the estimation of the configural model. In the configural model $\alpha_g = 0$, $\psi_g = 1$ for every g, and all loading, intercept and residual variance parameters are estimated as group-specific parameters. Denote the configural model estimates by $\nu_{pg,0}$,$\lambda_{pg,0}$, and $\theta_{pg,0}$, and let the configural factor be $\eta_{ig,0}$. Because the aligned model has the same model fit as the configural model the following relationships must hold

$$\eta_{ig} = \alpha_g + \sqrt{\psi_g}\eta_{ig,0}, \tag{2}$$

$$V(y_{ipg}) = \lambda_{pg}^2\ \psi_g + \theta_{pg} = \lambda_{pg,0}^2 + \theta_{pg,0}, \tag{3}$$

$$E(y_{ipg}) = \nu_{pg} + \lambda_{pg}\ \alpha_g = \nu_{pg,0}, \tag{4}$$

where $E(Y_{ipg})$ and $V(Y_{ipg})$ are the model estimated mean and variance for Y_{ipg}. Setting $\theta_{pg,0} = \theta_{pg}$, we get that

$$\lambda_{pg} = \frac{\lambda_{pg,0}}{\sqrt{\psi_g}}, \tag{5}$$

$$\nu_{pg} = \nu_{pg,0} - \alpha_g \frac{\lambda_{pg,0}}{\sqrt{\psi_g}}. \tag{6}$$

The aligned model chooses α_g and ψ_g as to minimize the amount of measurement non-invariance, i.e., the differences in λ_{pg} and ν_{pg} across groups.

To formalize this, we minimize with respect to α_g and ψ_g the alignment function F which accumulates all measurement non-invariance

$$F = \sum_p \sum_{g_1<g_2} w_{g_1,g_2} f(\lambda_{pg_1} - \lambda_{pg_2}) + \sum_p \sum_{g_1<g_2} w_{g_1,g_2} f(\nu_{pg_1} - \nu_{pg_2}), \tag{7}$$

where f is a component loss function and w_{g_1,g_2} are weights. The weights w_{g_1,g_2} are set to reflect the group size and the amount of certainty we have in the group estimates for a particular group. We use $w_{g_1,g_2} = \sqrt{N_{g_1} N_{g_2}}$. With these weights, bigger groups will contribute more to the total loss function than smaller groups. The component loss function is set to

$$f(x) = \sqrt[4]{x^2 + \epsilon} \tag{8}$$

where ϵ is a small number such as 0.0001. This function is approximately equal to $\sqrt{|x|}$. We use a positive ϵ so that F has a continuous first derivative which makes the optimization easier and more stable. This choice of f, as compared to other choices such as $|x|$ and x^2, has the advantage that it overemphasizes the penalty for medium size losses/non-invariance and underemphasizes the penalty for larger losses/non-invariance. Thus, the optimal invariance losses are expected to be either close to zero (invariant parameters) or not zero (non-invariant parameters). The medium range losses are meant to be eliminated with this choice of f. This is a key feature of the alignment methodology that distinguishes the method from other methods. BSEM measurement invariance or multilevel models with random intercepts and slopes tend to minimize mean squared error functions which can lead to many parameters with medium sized non-invariance. The alignment method typically will result in many approximately invariant measurement parameters, a few large non-invariant measurement parameters, and no medium-sized non-invariant measurement parameters. This is similar to the fact that EFA rotation functions aim for either large or small loadings, but not mid-sized loadings. Minimizing the loss function F will generally identify the parameters α_g and ψ_g in all but the first group. In the first group these parameters remain fixed to 0 and 1 respectively.

The alignment methodology works very well when most of the measurement parameters are invariant. The method will automatically separate the invariant and non-invariant parameters and all estimates will be consistent. It is somewhat difficult to quantify, however, the amount of non-invariance for which the alignment method will perform well. A rule of thumb is that as long as the number of non-invariant parameters is less than 20%, we can expect the alignment method to work correctly. However, the exact percentage of non-invariant parameters is not really what determines the alignment performance. Rather, the alignment performance is determined by the following question. Is the true parameter set, i.e., the data generating parameter set the simplest and most invariant representation of the data? Alignment will always pick α_g and ψ_g that produce the smallest amount of non-invariance. If a data-generating parameter set has a simpler alternative we can not expect alignment to produce estimates consistent with the data-generating parameter set. Alignment will converge to the simpler alternative instead. If 100% of the data generating parameters are not invariant, we know that a simpler representation exists (at least one indicator can be made group invariant by adding factor means and variances) and so alignment will not recover the data generating parameters. On the other hand, in most situations when less than 20% of the data generating parameters are non-invariant, a simpler alternative will likely not exist and the alignment estimates will be consistent.

Consider as an example the situation where data is generated using intercept and loading parameters that are group-specific random effects, see Muthén and Asparouhov (2018). In this case, all parameters are non-invariant and the data-generating parameters will have a simpler (more invariant) alignment alternative. Alignment is not expected to recover the data-generating parameters. This phenomenon is exactly as in EFA. EFA produces simple loading structures. If the data-generating loadings include a large amount of cross-loadings, EFA will not recover the parameters and will produce a simpler loading structure instead. Alignment and EFA parameters in such situations are not biased. Both methods would produce more optimal (in terms of their optimization criterion) representation of the data than the data-generating parameters. Furthermore, because the intercepts and slopes are random, the alignment results are expected to show many non-invariant parameters, which probably will become challenging to interpret. The alternative methodology of estimating the model with random intercepts and loadings will in fact recover the data-generating parameters and will provide a simpler model interpretation. Thus, we conclude that the alignment methodology is not universally applicable for all measurement invariance studies. If the amount of non-invariance found with alignment is so large that model interpretation is challenging, alternative measurement invariance methodologies should be pursued.

Extensive simulation studies on the alignment methodology can be found in Flake and McCoach (2018) and practical illustrations can be found in Munck et al. (2018) and Lomazzi (2018). A brief tutorial on the alignment method is provided in Rudnev (2019).

2.2 Extending alignment to the WLS estimators

This extension of the alignment method simply parallels the alignment method for IRT models described in Muthén and Asparouhov (2014). The configural model is estimated with the WLS estimator. The configural estimates are then used in the minimization of (7) to determine the factor mean and variance in each group. The threshold parameters for all categorical indicators are treated as the intercept parameters for continuous variables. Categorical variables with more than 2 categories are treated as having more than one mean parameter, i.e., alignment is conducted for every level of these categorical variables.

As with the ML alignment, the model fit of the WLS alignment is the same as the model fit of the configural model. If the WLS alignment model is rejected, the correct interpretation of that rejection is that the configural model is rejected. Thus, model misfit is never due to poor alignment, it is always due to configural model misfit. As usual, such misfit can be addressed using model modification indices, by adding additional factors, or by adding residual covariances between the factor indicators.

There are two separate parameterizations available for a factor analysis model with the WLS estimator: the theta (unstandardized) and the delta (standardized) parameterizations, see Muthén and Asparouhov (2002). With the theta parameterization, the residual variance parameters for all categorical variables are fixed to 1 for identification purposes, while in the delta parameterization the total variance is fixed to 1. The configural factor analysis model can be estimated in either one of these two metrics and the models will be equivalent in terms of data fit (i.e.the models are reparameterizations of each other). The alignment methodology is implemented for both parameterizations. In multiple group situations, scalar measurement invariance in one metric does not translate into scalar measurement invariance in the other metric. Since alignment's goal is the measurement invariance, the alignment results will depend on the metric. The metric may affect which parameters are considered invariant and which are not. This leads to the question regarding which metric alignment should be done in. One way to determine the most optimal metric is to estimate the model in both metrics and the more invariant metric would be preferred. Some further practical guidelines may be necessary in this regard.

The WLS alignment with the theta parameterization is equivalent in terms of the estimated model to the alignment for the ML estimator with numerical integration method and the probit link function. We can expect these estimation settings in Mplus to behave similarly. The ML and Bayes estimators in Mplus are currently available only for the theta parameterization.

Another important issue regarding alignment with categorical data with the WLS estimators is related to estimated residual variances in the theta parameterization and estimated delta parameters in the delta parameterization. Currently, for alignment models with categorical variables, these parameters remain fixed to 1 (for the aligned models). Conceivably, however, allowing the alignment function to be minimized with respect to these parameters as well, one could find an even better measurement invariance. In fact, measurement invariance models for categorical data without the alignment methodology and the WLS estimators typically involves estimating these parameters for all but the first group and is done by default in Mplus for the scalar invariance model. Our attempts to pursue this idea in the context of alignment, however, have fallen short so far. It appears that the sample size requirements make such an approach impractical and the additional parameters are rarely significantly different from 1, i.e., categorical variables alignment with residual variances fixed to 1 is expected to be sufficient in most situations. It should be noted, however, that in the configural model, the theta and delta parameters are always fixed to 1. This means that the alignment models, which have the same data fit as the configural models, do not suffer from suboptimal data fit due to fixed theta and delta parameters. For the scalar invariance model, this is not the case. Fixing the theta and delta parameters in the scalar invariance model may result in worse model fit. The alignment models, however, will not. From this point of view, the benefit of free delta and theta parameters for the alignment models is somewhat marginal. It affects only the classification of invariance and not the model fit.

Furthermore, consider the case of binary factor indicators. The alignment procedure must align the threshold and the loading parameter for that indicator and extract some information out of this process for the estimation of the factor mean and variance. If a free theta parameter is introduced in the alignment optimization, which is indicator specific, either the threshold or the loading parameter can be made invariant by the free theta parameter. The consequence of that is as follows. If the threshold is made invariant, the information from that variable is essentially used to estimate the theta parameter and not the factor mean. Similarly if the optimization function uses the free theta parameter to align the loading parameter fully, this indicator will contribute nothing to estimating the factor variance. All the information will be used to estimate the theta parameter. This line of argument points to two conclusions. If alignment is used with free theta/delta parameters, models with only binary indicators are unlikely to be able to extract much information regarding the scale of the parameters. Second, adding free theta and delta parameters to alignment,with binary and ordered categorical variables, will greatly diminish the ability of the alignment optimization to extract information regarding the factor scales, which generally is expected to be seen in large standard errors for most model parameters. This again leads to the conclusion that the free theta and delta models are somewhat incompatible with alignment because of the large sample sizes that would be needed for such models. In a practical context, if alignment estimation is to be converted to standard CFA model as for example it is done in Marsh et al. (2018), the practical issue arises regarding the estimation of the theta/delta parameters. We recommend that such parameters are considered carefully. These parameters should be included in the model only if they are significantly different from 1. If they are not statistically different from 1, they should remain fixed to preserve the parsimony and power of the model.

2.3 Extending alignment to the general factor analysis model with complex loading structures

The alignment procedure implemented in Mplus prior to version 8.8 applies only to factor analysis models with multiple factors and simple loading structures, i.e., without cross-loadings. In Mplus 8.8, the alignment procedure is extended to factor models with complex loading structures, i.e., models with cross-loadings and bi-factor models. To accommodate such models, equation (6) is replaced by

$$\nu_{pg} = \nu_{pg,0} - \sum_{m=1}^{M} \alpha_{mg} \frac{\lambda_{pmg,0}}{\sqrt{\psi_{mg}}}. \tag{9}$$

where M is the number of factors, α_{mg} and ψ_{mg} are the $m-$th factor mean and variance in group g, while λ_{pmg} is the loading of the $p-$th variable on the $m-$th factor. The rest of the procedure remains unchanged. The alignment of the loading parameters remains unchanged while the alignment of the intercept parameters now must account for the mean effect of all factors. Prior to version 8.8, the alignment optimization could be done for each factor separately. With the cross-loading extension the alignment optimization must be done for all factors simultaneously.

In Mplus 8.8 only the "fixed" alignment option is implemented for models with cross-loadings. The "free" alignment options allows us to estimate a few more parameters, however, it has the drawback of increasing the standard errors of the parameters, sometimes substantially. In the presence of more complex loading structures, the "free" alignment is likely to be even more susceptible to such issues. In simple loading structures, Mplus produces warnings regarding the substantial increase in standard errors due to "free" alignment. This option is, nevertheless, being misused in practical applications. Even for simple loading structures, the "fixed" alignment should always be a first-step analysis, rather than the "free" alignment. The "free" alignment works best when there is sufficient level of non-invariance. If most parameters appear to be invariant, the use of the "free" alignment is somewhat unjustified.

By default the reference group for alignment is the first group. This implies that in the first group the factor means are fixed to 0 and the factor variances are fixed to 1. The reference group can be changed using the option ALIGNMENT=FIXED(g), where g is the desired reference group. Changing the reference group does not affect the estimation of the configural model or the alignment optimization, i.e., changing the reference group does not affect the fit of the model in terms of likelihood or the alignment fit function. The alternative models are reparameterizations of the aligned model. Such post alignment reparameterization will not affect for example the ranking of the factor means, or the differences between the factor means across the groups. Changing the reference group however may affect which measurement parameters are considered invariant. Just as this happens with the ML and WLS estimators with any other re-parameterization, significance is not preserved. This mostly affects marginally significant differences in practical applications.

2.4 Extending alignment to the general structural equation model

The extension of the alignment methodology to the general SEM model is fairly simple. The first step is again the estimation of the configural SEM model. This configural model is defined as follows. All loading parameters are estimated as group-specific. The intercepts of all factor indicators are estimated as free and group specific. All factor intercepts are fixed to 0 and all residual factor variances are fixed to 1. All other parameters in the configural SEM model are specified as they are specified in the original SEM model. The parameters of the configural SEM model are estimated as well as their asymptotic variance covariance matrix. Next, the alignment procedure is used to align the loadings and intercepts of all factor indicators. In this process we obtain the factor means α_{mg} and residual variances ψ_{mg} which minimize the alignment loss function and maximize the amount of invariance in the SEM model. Note that only the configural model intercepts and loadings participate in the alignment. All other structural parameters are ignored in this stage of the estimation. The joint asymptotic variance covariance matrix of α_{mg}, ψ_{mg}, and all configural parameters is obtained by the same implicit methodology used in the simple alignment method in Asparouhov and Muthén (2014).

At this point, the parameters of the ASEM model are obtained from the parameters of the configural SEM model, α_{mg}, and ψ_{mg}. The ASEM model is a simple rescaling of the configural SEM model. The log-likelihood value and data fit is preserved. Changing the scales of the factors affects some of the SEM parameters but not all. The factor indicator intercepts and loadings are again adjusted according to equations (5) and (9). The factor covariance parameters are adjusted as follows. If $\psi_{ijg,0}$ is the configural model covariance parameter for factors η_i and η_j, the corresponding aligned parameter is computed as

$$\psi_{ijg} = \psi_{ijg,0}\sqrt{\psi_{ig}\psi_{jg}}. \tag{10}$$

If $\theta_{ijg,0}$ is the configural model covariance parameter between an observed variable Y_i and factor η_j, the corresponding aligned parameter is computed as

$$\theta_{ijg} = \theta_{ijg,0}\sqrt{\psi_{jg}}. \tag{11}$$

Regression parameters of a factor f_i on a covariate X_j in group g are adjusted as follows. The configural SEM model estimates the equation

$$f_i = ... + \beta_{ijg,0}X_j + ... \ . \tag{12}$$

The ASEM model regression parameter is computed as

$$\beta_{ijg} = \beta_{ijg,0}\sqrt{\psi_{ig}}. \tag{13}$$

The same transformation is used when the factor is regressed on a dependent variable. All other parameters remain unchanged. The above transformation equations are then used with the delta method to obtain the asymptotic variance covariance matrix for the parameters of the ASEM model.

Note also that the parameters in the transformation equations (10), (11) and (13) must be free and unequal across groups. That is because any equality constraints in the configural model will not hold for the aligned model. The same applies to parameters that are fixed to non-zero values. Note, however, that the parameters in (10), (11) and (13) can be fixed to 0 because the transformations will not alter such a constraint. This is important for example in the case of the bi-factor model where factor covariances are fixed to 0. Model parameters that are not altered by the above transformation can be held equal across groups or be fixed. Currently, the ASEM model does not support regressions of one factor on another factor. The model transformation for these regression parameters is slightly more complex and is not implemented in Mplus 8.8 but will likely be implemented in a future release.

In the ASEM model, all variables regressed on a factor are interpreted as factor indicators. Consider the situation where the model contains a distal outcome regressed on a factor. The distal outcome is not a measurement for the factor and we usually are not interested in the invariance of that regression parameter. The alignment optimization, however, will treat the distal outcome as another factor indicator. The implication of this is as follows. The alignment procedure will attempt to align the distal outcome regression parameters just as it would do so for any loading parameter, even if that is not intended. This also means that the distal outcome will have an effect on the scale of the factor. With the current implementation in Mplus, it is not possible to separate the distal outcome from the measurement indicators. In principle, however, this should not cause any estimation problems. Even when the regression parameters are group-specific, alignment will not make these regression parameters group-invariant because the alignment procedure accommodates non-invariance. Furthermore, if there are enough factor indicators that can identify the factor scale well, the addition of the distal outcome will affect the alignment only marginally. If the number of indicators is small, however, or the measurement model has a substantial amount of non-invariance, the addition of the distal outcome may negatively affect the identification of the factor scales. In such situations, the Marsh et al. (2018) two-step estimation might be preferable.

Residual covariances among the factor indicators can be included in the ASEM model and those parameters are not adjusted by the factor scales, i.e., these parameters will be identical to the configural model parameters. This also applies to the direct effects from a predictor to the factor indicators.

In summary, the alignment extension to the general SEM model is intended to make the alignment procedure an integral part of multiple-group SEM modeling. It appears that adding alignment to SEM has no serious drawbacks and it has the advantage of capturing the group effect on the factors in a way that is simple and easy to interpret.

2.5 Extending alignment to the ESEM/EFA models

The alignment extension to the ESEM/EFA models parallels the alignment extension of the SEM model. First, the configural ESEM model is estimated in every group. Such an estimation consists of two steps. First the unrotated configural ESEM model is estimated. Then, in a second step, the estimated model is rotated with an optimal rotation matrix selected by minimizing a simplicity criterion such as the geomin or quartimin criteria. Finally, the configural rotated ESEM model is aligned as if it is a regular SEM model. Such an estimation approach clearly has three distinct steps. The Mplus implementation encompasses all three steps and thus it can be viewed as a one-step approach.

The AESEM method can also be viewed as an estimation with a joint simplicity function. The simplicity rotation function is added to the alignment loss function to form a joint simplicity function. The joint simplicity function is then minimized as a function of the factor means, the factor variances and the factor rotation. The key issue in the joint simplicity function is how to weigh the two components. The approach implemented in Mplus essentially uses an infinitely large weight for the rotation part. This is to reflect the fact that we rotate the configural model without considering the alignment loss function. Conditional on these rotated results, the alignment is then conducted. It is in principle possible to combine rotation and alignment in a more equitable way. However, such an estimation will be more complex and it would still have the uncertainty about how to weigh the two components. Presumably, a more equitable rotation/alignment approach can have an MSE advantage in some situations. However, we can view the equitable rotation/alignment procedure as an estimation with an alternative simplicity function, which more or less is going to yield similar results in most situations.

3 Illustrations

3.1 Invariance depends on the theta/delta parameterization

In this section we demonstrate that the concept of invariance depends on the parameterization (delta or theta) used for models with categorical data. Consider a two-group factor analysis model with 1 factor and P binary indicators. With scalar invariance and the theta parameterization the model is defined as follows. For $i = 1, ..., P$ and $g = 1, 2$

$$P(Y_{ig} = 0|\eta_g) = \Phi(\tau_i - \lambda_i \eta_g) \tag{14}$$

where Φ is the standard probit function. Here

$$Y^*_{ig} = \lambda_i \eta_g + \varepsilon_{ig} \tag{15}$$

where $\varepsilon_{ig} \sim N(0, 1)$, $\eta_g \sim N(\alpha_g, \psi_g)$ and

$$P(Y_{ig} = 0|\eta_g) = P((Y^*_{ig} < \tau_{ig}|\eta_g) \tag{16}$$

We also set $\alpha_1 = 0$ and $\psi_g = 1$. This model is scalar invariant in the theta parameterization. If the model is scalar invariant in the delta parameterization as well then the delta parameterization loadings must be equal between the two groups

$$\frac{\lambda_i}{\sqrt{\lambda_i^2 + 1}} = \frac{\lambda_i \sqrt{\psi_2/\phi_2}}{\sqrt{\lambda_i^2 \psi_2 + 1}} \tag{17}$$

where ϕ_2 is the factor variance in the second group in the delta parameterization. Simple algebra then concludes that either $\psi_2 = \phi_2 = 1$, or $\psi_2 = \phi_2 = 0$, or all loadings λ_i must be equal across indicators and must be equal to

$$\lambda_i = \sqrt{\frac{\psi_2 - \phi_2}{\psi_2(\phi_2 - 1)}}. \tag{18}$$

We conclude that in general scale invariance in one metric does not translate into scalar invariance in another metric. When it comes to statistical significance, however, the issue is slightly more complex. Consider the case where scalar invariance holds in the theta metric. To be able to reject scalar invariance in the delta metric, all three of these hypotheses must be statistically rejected

$$\phi_2 = 1 \tag{19}$$

$$\phi_2 = 0 \tag{20}$$

$$\lambda_1 = \lambda_2 = ... = \lambda_P. \tag{21}$$

The larger the loading differences across indicators and the more distant the factor variance is from 0 and 1, the more likely the scalar invariance assumption is to break in a statistically significant way when the parameterization is changed to the delta metric.

Since alignment is attempting to approximate the scalar invariance model, we can expect that alignment results will be different across the two parameterizations as well, not just in the estimated parameter estimate, but also in the inference regarding the invariance of individual parameters. If there is no substantive reason to prefer one parameterization over another, the results of both parameterizations should be assessed and the most easily interpretable model should be preferred (which is likely to be the parameterization that yields the most invariance for the measurement model).

Next we illustrate the dependence of alignment on the parameterization with some simulated examples. Figure 2 in the Appendix shows the Mplus input for a simulation study where we generate data from a two-group one-factor scale invariant factor analysis model with 5 binary indicators using the theta parameterization. We use a very large sample size for this simulation to ensure that significance will be archived for the three hypotheses (19-21) needed to establish a difference between the two parameterizations. Smaller sample sizes will not have enough power in this small example. The data is analyzed with the theta parameterization alignment method over 100 replications and the results are presented in Table 1. The loading parameter for Y_i is denoted by λ_i and the threshold parameter by τ_i. The factor mean and variance in the first group are fixed to 0 and 1 respectively, while in the second group are estimated as α and ψ. The results indicate that the alignment estimation works very well. Next we analyze the first data set from that simulation with both the theta and the delta parameterizations with the alignment method. The Mplus input files for these analyses are given in Figures 3 and 4 in the Appendix.

The chi-square test for the two models is the same, however, only the theta parameterization concludes that the scalar invariance holds. The delta parameterization concludes that the fifth loading is not invariant. The delta parameterization output seen in Figure 1, places the fifth loading in round brackets which means that the loadings are not invariant. The details of the invariance parameter analysis can be obtained using the OUTPUT:ALIGN option.

Next we repeat this experiment by generating the data with the delta parameterization scalar invariance model. Figure 5 in the Appendix shows the Mplus input for this simulation study which also analyzes the data with the delta parameterization alignment. To properly set up the simulation study the residual variances for the dependent variables must be included here and we must make sure that the total variance for each variable in every group is 1. The results of the simulation are shown in Table 2 and the parameter estimates are recovered well. The first data set of this simulation is then analyzed with the theta and delta parameterization alignment using again the input files in Figures 3 and 4. Here we find that the delta parameterization concludes that the scalar invariance holds while the theta parameterization alignment concludes that the fifth loading is not invariant.

In summary, the delta and theta parameterization alignments are not equivalent. Furthermore, chi-square can not be used to determine which parameterization to use because both of these aligned models have identical chi-squares. Note, that this is different from the situation of the standard unaligned scalar model estimation. The scalar unaligned factor analysis models have different chi-square values, and if the scalar model is assumed, the chi-squares can be used to select the better metric for the estimation. If one of the two alignment models has a clear advantage in terms of the level of non-invariance that it reaches, it will be easy to select that metric. However, the distinction between the two metrics appears to require bigger sample sizes, which might not be feasible in practical situations. If the above simulation is repeated with half the sample size, the parameterizations are not distinguishable.

In this section we did not include non-invariant measurement parameters because we wanted to illustrate the difference between the delta and theta parameterizations. Non-invariant measurement parameters can be added to the simulation setups given in Figures 2 and 5.

Table 1: Montecarlo alignment results for a theta parameterization scalar invariant factor analysis model

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_1	.6	.00	.97
λ_2	.8	.00	.93
λ_3	1.0	.00	.94
λ_4	1.2	.00	.99
λ_5	1.4	.00	.96
τ_1	0	.00	.95
τ_2	0	.00	.94
τ_3	0	.00	.96
τ_4	0	.00	.93
τ_5	0	.00	.93
Group 2			
λ_1	.6	.00	.98
λ_2	.8	.00	.99
λ_3	1.0	.00	.95
λ_4	1.2	.00	.96
λ_5	1.4	.00	.97
τ_1	0	.00	.94
τ_2	0	.00	.92
τ_3	0	.00	.91
τ_4	0	.00	.93
τ_5	0	.00	.96
α	0	.00	.92
ψ	1.8	.00	.97

Figure 1: Output for alignment with the delta parameterization showing loading non-invariance for the fifth loading in brackets

```
APPROXIMATE MEASUREMENT INVARIANCE (NONINVARIANCE) FOR
GROUPS

 Intercepts/Thresholds
   Y1$1        1 2
   Y2$1        1 2
   Y3$1        1 2
   Y4$1        1 2
   Y5$1        1 2

 Loadings for F1
   Y1          1 2
   Y2          1 2
   Y3          1 2
   Y4          1 2
   Y5          (1) (2)
```

Table 2: Montecarlo alignment results for a delta parameterization scalar invariant factor analysis model

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_1	.4	.00	.95
λ_2	.5	.00	.95
λ_3	.6	.00	.95
λ_4	.7	.00	.95
λ_5	.8	.00	.96
τ_1	0	.00	.95
τ_2	0	.00	.94
τ_3	0	.00	.97
τ_4	0	.00	.93
τ_5	0	.00	.93
Group 2			
λ_1	.4	.00	.96
λ_2	.5	.00	.98
λ_3	.6	.00	.97
λ_4	.7	.00	.99
λ_5	.8	.00	.97
τ_1	0	.00	.95
τ_2	0	.00	.93
τ_3	0	.00	.93
τ_4	0	.00	.97
τ_5	0	.00	.91
α	0	.00	.92
ψ	.5	.00	.96

3.2 Factor analysis with cross-loadings

In this section we illustrate the performance of the alignment method for factor analysis models with cross-loadings. Here we use a 3-group, 2-factor analysis model with a total of 9 indicators. Each factor has 3 pure indicators, labeled Y_i for the first factor and Z_i for the second, and 3 of the indicators load on both factors, which are labeled as W_i. The estimated ASEM model is given by the following equations. For $i = 1, ..., 3$, $g = 1, ..., 3$, and $j = 1, 2$

$$Y_{ig} = \mu_{yig} + \lambda_{yi1g} f_1 + \varepsilon_{yi} \tag{22}$$

$$Z_{ig} = \mu_{zig} + \lambda_{zi2g} f_2 + \varepsilon_{zi} \tag{23}$$

$$W_{ig} = \mu_{wig} + \lambda_{wi1g} f_1 + \lambda_{wi2g} f_2 + \varepsilon_{wi} \tag{24}$$

$$\varepsilon_{yi} \sim N(0, \theta_{yig}), \varepsilon_{zi} \sim N(0, \theta_{zig}), \varepsilon_{wi} \sim N(0, \theta_{wig}), f_j \sim N(\alpha_{jg}, \psi_{jg}). \tag{25}$$

The factor intercept and variance in the first group are fixed to $\alpha_{j1} = 0$, $\psi_{j1} = 1$. All other parameters are free and unequal across-groups.

The input file for this simulation study is given in Figure 6 in the Appendix. The data is generated with 3 non-invariant intercepts and 3 non-invariant loadings. The lines marked with !NI in Figure 6 specify the non-invariant parameters within the group-specific statements. These statements are placed in the MODEL POPULATION and the MODEL parts of the input file. Note however, that group-specific statements are generally not needed in the MODEL part. In the simulation study, the group-specific models are used purely for starting value purposes and for the computation of the coverage of the confidence intervals. In practical applications, only the main MODEL part is needed as in Figures 2 and 3. The group-specific models are needed only in simulation studies. The model in Figure 6 also features a cross-loading present only in some of the groups. This is specified by setting "f1 by w1@0;" in group 1.

Table 3 contains the results of this simulation study for some of the model parameters. We see that the alignment procedure is able to estimate well all invariant and non-invariant measurement parameters as well as the factor means and variances.

Table 3: Alignment simulation results for a 2-factor analysis model with cross-loadings. Non-invariant measurement parameters are marked with *. Parameters existing in only some of the groups are marked with **.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_{y11}	1	.00	.96
λ_{w21}	.5	.00	1.00
λ_{w12}	.5	.00	.93
λ_{w22} *	.5	.00	.97
λ_{w32}	.5	.00	.94
μ_{y1} *	1	.01	.96
μ_{y2}	0	.01	.91
Group 2			
λ_{y11}	1	.00	.95
λ_{w11} **	.5	.00	.95
λ_{w21}	.5	.00	.93
λ_{w12}	.5	.00	.97
λ_{w22} *	1	.01	.93
λ_{w32}	.5	.01	.94
μ_{y1} *	0	.00	.96
μ_{z1} *	-1	.01	.96
α_1	0.4	.01	.99
α_2	0.6	.01	.96
ψ_1	1.4	.01	.95
ψ_2	1.2	.03	.98
Group 3			
λ_{y11}	1	.00	1.00
λ_{w11} **	.5	.01	.93
λ_{w21}	.5	.00	.98
λ_{z12}	1	.02	.98
λ_{z22} *	.5	.01	.91
λ_{z32}	1	.02	.94
μ_{w1}	0	.00	.98
μ_{w2}	0	.00	.95
μ_{w3} *	1	.01	.90
α_1	-1	.00	.93
α_2	-.5	.03	.90
ψ_1	1.4	.01	.98
ψ_2	1.2	.05	.89

3.3 Bi-factor models

In this section we illustrate the performance of the alignment methodology for the bi-factor models. The bi-factor model is technically just another factor analysis model with cross-loadings. However, the model is somewhat special because all indicators load on more than one factor. In addition, there are certain identifiability issues related to the bi-factor model that are somewhat special. The covariances among the factors must be fixed to 0 in the bi-factor model. To some extent, the bi-factor model can be considered to be an extreme version of the factor analysis model with cross-loadings.

Figure 7 in the Appendix shows the input file for a 2-group 3-factor bi-factor model with 10 indicators Y_i and Z_i, $i = 1, ..., 5$. All 10 indicators load on the general factor f_1. There are also two specific factors f_2 and f_3, f_2 is the specific factor for Y_i and f_3 is the specific factor for Z_i. The model is given by the following equations. For $i = 1, ..., 5$, $j = 1, ...3$, and $g = 1, 2$

$$Y_{ig} = \mu_{yig} + \lambda_{yi1g} f_1 + \lambda_{yi2g} f_2 + \varepsilon_{yi} \tag{26}$$

$$Z_{ig} = \mu_{zig} + \lambda_{zi1g} f_1 + \lambda_{yi3g} f_3 + \varepsilon_{zi} \tag{27}$$

$$\varepsilon_{yi} \sim N(0, \theta_{yig}), \varepsilon_{zi} \sim N(0, \theta_{zig}), f_j \sim N(\alpha_{jg}, \psi_{jg}). \tag{28}$$

The factor intercept and variance in the first group are fixed to $\alpha_{j1} = 0$, $\psi_{j1} = 1$. All other parameters are free and unequal across-groups.

We generate data with 2 non-invariant intercepts, one non-invariant general factor loading and one specific factor loading. The lines containing the non-invariant parameters in Figure 7 are marked with "!NI". One of the important identifiability conditions of the bi-factor model is that the specific factor loadings must not be proportional to the general factor loadings. If the loadings are proportional, a local non-identification condition exists. I this example, we generate data with equal general factor loadings and unequal specific factor loadings. This ensures that the parameters of the bi-factor model are not close to these local non-identifiability conditions.

In this simulation study, we use the TOLERANCE option, which refers to the ϵ value in equation (8). The default value for ϵ is 0.01. By lowering the value to 0.0001, we reduce the bias in the point estimates. However, lowering the value of ϵ is often associated with an upwards bias in the standard errors. The bi-factor model is somewhat more challenging for the alignment procedure than standard factor analysis models. Thus, the 0.0001 value is beneficial here. The bias in the standard error appears to be minimal. We do not recommend using epsilon values outside of the range 0.0001 to 0.01. Smaller values will make the computation less stable because of divisions near 0 in the derivatives of the loss function. In addition, the standard error estimation which is based on these derivatives will not be as accurate. Larger values of ϵ have the effect of spreading non-invariance beyond the source, which is also undesirable.

The results of the simulation study for a selection of the parameters are given in Table 4. The alignment method appears to work well for the bi-factor model as well.

Table 4: Alignment simulation results for a bi-factor model. Non-invariant measurement parameters are marked with *.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_{y11}	1	.00	.92
λ_{y21} *	1.5	.00	.93
λ_{y21}	1	.00	.95
λ_{z11}	1	.01	.94
λ_{z21}	1	.01	.94
λ_{y32}	1	.00	.96
λ_{y42}	1.2	.01	.95
λ_{y52} *	1.4	.00	.97
λ_{z43}	1.2	.01	.95
λ_{z53}	1.2	.01	.97
μ_{y1} *	1	.00	.94
μ_{z1} *	0	.00	.97
Group 2			
λ_{y11}	1	.01	.95
λ_{y21} *	1	.02	.94
λ_{y21}	1	.02	.97
λ_{z11}	1	.02	.96
λ_{z21}	1	.02	.98
λ_{y32}	1	.02	.98
λ_{y42}	1.2	.02	.95
λ_{y52} *	1	.02	.98
λ_{z43}	1.2	.01	.95
λ_{z53}	1.2	.01	.97
μ_{y1} *	0	.00	.96
μ_{z1} *	1	.00	.96
α_1	0.3	.01	.96
α_2	-0.4	.01	.98
α_3	0.5	.01	.98
ψ_1	1.4	.04	.95
ψ_2	1.4	.05	.92
ψ_3	1.4	.02	.95

3.4 Factor analysis with covariates

In this section we illustrate the performance of the alignment methodology for a factor analysis model with covariates. We use a 3-group analysis with one factor f and 5 ordered categorical indicators Y_i, $i = 1, ..., 5$, with 3 categories 0,1,and 2; and one covariate X. The covariate predicts the factor in every group and also directly the first indicator in the first and the third groups. A residual covariance is also estimated for the second and the third indicators in the second and the third groups. We estimate the following ASEM model. For $g = 1, ..., 3$

$$Y^*_{1g} = \lambda_{1g} f + \gamma_g X + \varepsilon_1, \tag{29}$$

and for $i = 2, ..., 5$

$$Y^*_{ig} = \lambda_{ig} f + \varepsilon_i \tag{30}$$

$$f = \alpha_g + \beta_g X + e \tag{31}$$

$$\varepsilon_i \sim N(0, 1), e \sim N(0, \psi_g), Cov(\varepsilon_2, \varepsilon_3) = \theta_{23g} \tag{32}$$

$$Y_{ig} = k \iff \tau_{i,k,g} \leq Y^*_{ig} < \tau_{i,k+1,g}. \tag{33}$$

The factor intercept and variance in the first group are fixed to $\alpha_1 = 0$ and $\psi_1 = 1$. The direct effect is not estimated in the second group, so γ_2 is also fixed to 0. Similarly, the residual covariance is not estimated in the first group so θ_{231} is also fixed to zero. The threshold parameters τ_{i0g} and τ_{i3g} are as usual set to $-\infty$ and $+\infty$. Two threshold parameters are therefore estimated for every indicator in every group: τ_{i1g} and τ_{i2g}.

We generate data according to the above model with two non-invariant factor loadings and two non-invariant thresholds. The full Mplus input for this Montecarlo simulation is given in Figure 8. We analyze the above ASEM model with the theta parameterization of the WLSMV estimator. It takes only 7 seconds to complete 100 replications of this study. The model can also be analyzed with the ML estimator if the residual covariance parameter θ_{23} is not included in the model.

The results of this simulation study for a subset of the parameters are shown in Table 5. The biases in the parameters estimates are minimal and the coverage is near the nominal level of 95%. We conclude that the alignment method can easily accommodate factor predictors as well as other features of the general SEM framework.

Table 5: Alignment simulation results for a factor analysis with a covariate. Non-invariant measurement parameters are marked with *.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_3	1	.00	.92
λ_4 *	1	.00	.95
λ_5 *	1	.00	.97
β	.7	.01	.98
γ	.4	.01	.94
τ_{11} *	-1	.01	.96
τ_{12} *	1	.00	.94
Group 2			
λ_3	1	.04	.98
λ_4 *	.6	.03	.94
λ_5 *	.6	.02	.97
β	.4	.01	.94
τ_{11} *	0	.02	.97
τ_{12} *	1.5	.04	.92
α	.4	.03	.97
ψ	.8	.04	.92
θ_{23}	.3	.01	.92
Group 3			
λ_3	1	.00	.96
λ_4 *	1	.01	.98
λ_5 *	1	.01	.99
β	.7	.00	.96
γ	.4	.00	.95
τ_{11} *	-1	.02	.98
τ_{12} *	1	.00	1.00
α	-.3	.00	.99
ψ	1.2	.02	.99
θ_{23}	.3	.01	.94

3.5 AESEM simulation study

In this section we illustrate with a simulation study the performance of the alignment methodology for a 2-factor ESEM model with a covariate. The full model is given in Figure 9 in the Appendix. The two factors are measured by a total of 6 indicators, where each of the two factors is measured by 3 main indicators. We generate the data so that factor f_1 has main(large) loadings for $Y_1, ..., Y_3$ and factor f_2 has main loadings for $Z_1, ..., Z_3$. Factor f_1 has a main loading non-invariance for Y_3 and in Groups 1 and 3 has a non-zero cross-loading for Z_3. Intercept non-invariance is introduced in group 3 for Z_2. Both factors are regressed on the covariate X.

The estimated AESEM model is given by the following equations. In group g, for $i = 1, ..., 3$

$$Y_{ig} = \mu_{yig} + \lambda_{yi1g} f_1 + \lambda_{yi2g} f_2 + \varepsilon_{yi} \tag{34}$$

$$Z_{ig} = \mu_{zig} + \lambda_{zi1g} f_1 + \lambda_{zi2g} f_2 + \varepsilon_{zi} \tag{35}$$

$$\varepsilon_{yi} \sim N(0, \theta_{yig}),\ \varepsilon_{zi} \sim N(0, \theta_{zig}) \tag{36}$$

For $j = 1, 2$ in group g

$$f_j = \alpha_{jg} + \beta_{jg} X + e_j, \tag{37}$$

where

$$e_j \sim N(0, \psi_{jg}),\ Cov(e_1, e_2) = \psi_{12g}. \tag{38}$$

The alignment method fixes the factor intercept and variance in the first group: $\alpha_{j1} = 0$, $\psi_{j1} = 1$. All other parameters are free and unequal across-groups. However, the parameters of the AESEM model are implicitly constructed from the parameters of the unrotated contextual model through the optimization of the rotation and the alignment loss functions.

For the purposes of the chi-square and the BIC computation the number of free parameters in the AESEM model is the number of free parameters in the unrotated contextual model. In this example, that model has 25 parameters in each group for a total of 75 parameters. Here is how these parameters can be counted: 6 intercepts, 12 loadings, 6 residual variance, plus 2 regression parameters for the 2 factors regressions on the covariates. From this quantity we subtract the loadings that are fixed to 0 in the unrotated model, which are the loadings above the main diagonal. In this example, just one of the loadings is fixed to 0. As usual, the unrotated model parameter specification can be found in **tech1**. As with all ESEM models, there are two sets of **tech1**. The second version of **tech1** is the actual AESEM model where all parameters are given. In this example, the additional parameters are the 2 factor variances, 1 factor covariance, 2 factor means, as well as the 1 loading that is fixed in the unrotated solution. This yields 6 additional parameters, which give a total of 31 parameters in each group and a total of 93 parameters that are reported in the output for the model results. The additional 18 parameters are not free parameters. These additional parameters are dependent parameters that are identified through the rotation and alignment loss function optimization. For the purpose of chi-square testing and BIC computations, the model has 75 parameters. The important thing to understand for the AESEM model is that the loading parameters are all free and unequal across groups so that the EFA measurement invariance is accommodated, and in addition to that the scale parameters for the factors are estimated in all but the reference group.

The results of the simulation study for some of the parameters are given in Table 6. The parameter estimates are unbiased and the coverage is near the nominal level of 95%. This includes all non-invariant loadings and intercepts as well as the factor means and variances. The average chi-square value across the replications is 25.1 and with 24 degrees of freedom this yields a rejection rate of 0.07 which is sufficiently close to the nominal rejection rate of 0.05. We conclude that the AESEM methodology works well. The entire simulation study on the 100 generated data sets takes only 8 seconds to complete. It should be noted, however, that models with more groups, indicators and factors will not be as fast to estimate. The number of parameters that must be handled in such an estimation grows linearly with the number of groups, indicators and factors. Thus, larger models are expected to have slower estimation.

For comparative purposes we also estimate the scalar ESEM model which assumes measurement model invariance. To obtain this estimation, we have to remove the alignment option from the input file given in Figure 9 as well as the group specific measurement and intercept statements in the model statement. In this model, the loading and indicator intercept parameters are held equal across groups and the rotation is group invariant. The factor means and variances are also estimated with the exception of the reference group. The results are given in Table 7. We see biases and coverage problem in almost all of the parameters. This is somewhat of a surprising result since the model has only 3 non-invariant parameters (1 main loading, 1 cross-loading and 1 intercept). Nevertheless, the non-invariance in this simulation is not small. Clearly, the larger size of the non-invariance propagates to most model parameters. The average chi-square value in this simulation is 1392.3 and with 48 degrees of freedom the model is rejected in all replications. The average CFI and TLI values are 0.93 and 0.90. Despite the larger number of parameters, the AESEM model has a much better average BIC value than the scalar ESEM model. This simulation study further emphasizes the importance of the AESEM model which can accommodate measurement non-invariance.

Table 6: Results for an AESEM model with a covariate. Non-invariant measurement parameters are marked with *.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_{y11}	1	.00	.93
λ_{y21}	1	.00	.90
λ_{y31} *	1	.00	.92
λ_{z31} *	.4	.01	.98
λ_{z32}	1	.00	.98
λ_{y12}	0	.00	.99
β_1	.3	.00	.94
β_2	-.2	.00	.94
ψ_{12}	.5	.00	.95
Group 2			
λ_{y11}	1	.01	.99
λ_{y21}	1	.01	.97
λ_{y31} *	.7	.01	.95
λ_{z31} *	0	.00	.97
β_1	-.3	.00	.95
β_2	.2	.00	.94
ψ_{12}	.3	.00	.93
α_1	.5	.01	1.00
α_2	.8	.01	.99
ψ_1	1.2	.03	.92
ψ_2	1.5	.02	.98
Group 3			
λ_{y11}	1	.00	1.00
λ_{y21}	1	.00	.98
λ_{y31} *	1	.00	.96
λ_{z31} *	.4	.01	.95
α_1	-.5	.01	.98
α_2	.3	.01	.92
ψ_1	1.5	.01	.99
ψ_2	1.2	.00	.98
μ_{z2} *	1	.00	.94

Table 7: Scalar ESEM results. Non-invariant measurement parameters are marked with *.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_{y11}	1	.03	.79
λ_{y31} *	1	.07	.09
λ_{z11}	0	.01	.00
λ_{z21}	0	.19	.00
λ_{z31} *	.4	.11	.00
λ_{z12}	1	.00	.96
λ_{z22}	1	.13	.00
λ_{z32}	1	.00	.98
λ_{y12}	0	.01	.99
β_1	.3	.00	.95
β_2	-.2	.03	.66
ψ_{12}	.5	.07	.11
μ_{z1}	0	.07	.12
μ_{z2} *	0	.18	.00
μ_{z3}	0	.13	.00
θ_{y3}	1	.05	.66
θ_{z1}	1	.04	.82
θ_{z2}	1	.01	.94
θ_{z3}	1	.08	.50
Group 2			
β_1	-.3	.04	.39
β_2	.2	.00	.93
ψ_{12}	.3	.06	.45
α_1	.5	.06	.52
α_2	.8	.05	.73
ψ_1	1.2	.32	.00
ψ_2	1.5	.15	.37
Group 3			
β_1	.3	.00	.98
β_2	.2	.02	.86
ψ_{12}	.4	.13	.06
α_1	-.5	.06	.65
α_2	.3	.28	.00

3.6 Models with large amount of non-invariance

Asparouhov and Muthén (2014) present several simulation studies to evaluate the alignment methodology with different levels of non-invariance. Up to 20% non-invariance has been used in these simulation studies. Subsequently, the question has arisen regarding the level of non-invariance that the alignment methodology can handle. In this section we illustrate the performance of the alignment method with a large amount of non-invariance.

Multiple factors play a role in the quality of the alignment estimates. The most important of these is whether the generating measurement parameters is the most invariant pattern that represents the data. The interpretation of the "most invariant" concept in this regard is that the set of generating parameters represents a minimum of the alignment loss function (7). If the generating parameters doesn't not represent a minimum of the alignment loss function, then alignment will replace the generating parameters with a set of parameters that represents a minimum. This essentially means that the alignment procedure will align the parameters better to a set of parameters that has greater non-invariance than the generating parameters. Thus, if the generating parameters is not an alignment loss function minimum, we can not expect the generating set to be recovered by alignment. This concept is similar to what happens in EFA. If a set of generating loading parameters does not represent a minimum of the rotation function, we can not expect the EFA procedure to reproduce the these parameters. The parameters will be rotated further to an actual minimum that yields simpler loading structure.

If the configural set of parameters is estimated by the ML/WLS estimators without any error, the alignment estimates of α_g and ψ_g are considered the "most invariant" if these yield the smallest loss function (7), i.e., they minimize the loss function. When the amount of non-invariance is small, i.e., less than 20%, this usually ends up to be the case, especially when the non-invariance is spread around the groups. However, there are exceptions to this rule. Consider the situation where there are G groups and the first $G-1$ groups are fully invariant while the last group is not invariant in any of the parameters. Such a situation represents $1/G$ non-invariance, i.e., with $G = 10$ groups, the amount of non-invariance would be 10%. In that case, however, the factor mean and variance in the last group can be adjusted to make at least two (one intercept and one loading) parameter invariant and likely reducing the loss function. In this hypothetical case, the true parameters do not represent the most invariant pattern, even though there is only 10% non-invariance, and the alignment procedure is not expected to recover the estimates even with large sample size.

In every situation where the alignment optimization has converged and there is a sufficient number of random starts used in the procedure (Mplus default is 30), one can be assured that the reported alignment results represent the most invariant pattern, i.e., a loss function minimum. Suppose that a real data alignment model has been estimated successfully. The natural question in this case is to see if the reported solution can be recovered in a simulation study. The answer to that question is that the aligned solution will be recovered if the simulation study is conducted with a large sample size. The large sample size will ensure that the configural parameters are the same as the configural parameters in the real data. At that point the alignment optimization between the real data and the simulated data will be the same, thereby producing the same outcome. When the sample size is small the configural estimates in the simulated data and the real data will differ. Depending on the circumstances, some replications may be aligned to the same "most invariant" solution and some may be aligned to a different "most invariant" solution, thereby creating the illusion that the parameter estimates are not recovered. Such a result however is not a reason to disregard the real data analysis. Even if the sample size is small, the aligned solution still represents the "most invariant" choice.

When the amount of non-invariance is larger, it would not be easy to verify that the true values represent the "most invariant" solution. However, if the solution is obtained from a real data run, then by definition the aligned solution is "most invariant". In principle, there is no upper limit on the amount of non-invariance that the alignment methodology can handle successfully. Here we provide a simulation study where 50% of the measurement parameters are not invariant. The input file for this simulation study is given in Figure 10. The model has one factor measured by 6 indicators, Y_i and Z_i for $i = 1, ..., 3$ and predicted by a covariate X. The indicators Y_i have fully invariant measurement parameters while the indicators Z_i are not invariant. Thus, 50% of the measurement parameters are non-invariant. The estimated ASEM model is given by the following equations. For $i = 1, ..., 3$

$$Y_{ig} = \mu_{yig} + \lambda_{yig} f + \varepsilon_{yi} \tag{39}$$

$$Z_{ig} = \mu_{zig} + \lambda_{zig} f + \varepsilon_{zi} \quad (40)$$

$$f = \alpha_g + \beta_g X + e \quad (41)$$

$$\varepsilon_{yi} \sim N(0, \theta_{yig}), \ \varepsilon_{zi} \sim N(0, \theta_{zig}), e \sim N(0, \psi_g) \quad (42)$$

where in the first group $\alpha_1 = 0$ and ψ_1. The parameters μ_{yig} and λ_{yig} are invariant, while μ_{zig} and λ_{zig} are not invariant. In the ASEM estimation, all parameters are group specific. The results of the simulation study for a selection of the parameters are given in Table 8. We see here that the alignment method performed well even when the amount of non-invariance is substantial.

We conclude this section with the following observation. The result of the alignment optimization is that the invariant group specific measurement parameters are typically very close across the groups, but they are not identical. The small differences across groups are generally not of practical significance. For situations with large sample sizes, however, these practically insignificant differences may become statistically significant. In that regard, parameters that for all practical purposes can be considered invariant may be reported in the Mplus output as having some non-invariance. This should be taken into consideration. Large sample sizes are often the source of non-invariant results. Substantive judgement should be exercised in such situation to properly depict the amount of non-invariance. The simulation study described in this section is not of this type because the differences in the parameters across groups is substantial. When the sample size is large, the amount of non-invariance reported in the Mplus alignment output is likely overestimated and some subjective assessment must be done to obtain a more realistic account for the level of non-invariance.

Table 8: Alignment results with large amount of non-invariance. Non-invariant measurement parameters are marked with *.

Parameter	True Value	Abs. Bias	Coverage
Group 1			
λ_{y1}	1	.01	.93
λ_{y2}	1	.01	.93
λ_{y3}	1	.01	.93
λ_{z1} *	1	.01	.92
λ_{z2} *	1	.00	.95
λ_{z3} *	1	.00	.93
β	.4	.00	.91
μ_{y1}	0	.00	.93
μ_{y2}	0	.00	.94
μ_{y3}	0	.00	.92
μ_{z1} *	0	.00	.91
μ_{z2} *	0	.00	.94
μ_{z3} *	0	.00	.96
Group 2			
λ_{y1}	1	.01	.96
λ_{y2}	1	.01	.95
λ_{y3}	1	.01	.96
λ_{z1} *	.5	.01	.97
λ_{z2} *	.8	.01	.95
λ_{z3} *	1.2	.01	.94
β	.7	.01	.94
μ_{y1}	0	.01	.94
μ_{y2}	0	.01	.98
μ_{y3}	0	.01	.97
μ_{z1} *	.3	.01	1.00
μ_{z2} *	.5	.01	.96
μ_{z3} *	.3	.01	.95
Group 3			
λ_{y1}	1	.01	.94
λ_{y2}	1	.01	.95
λ_{y3}	1	.01	.96
λ_{z1} *	.8	.00	.94
λ_{z2} *	1.2	.01	.96
λ_{z3} *	.5	.00	.99
β	.2	.00	.95

4 Empirical example

In this section we illustrate the AESEM model with an empirical example. We utilize the student background questionnaire data of PISA 2006, which contains eight scales measuring a variety of motivational and engagement constructs in science as described in Marsh et al. (2018). The data contains nationally representative samples of 15-year-old students from 30 OECD countries/groups with a total sample size $N = 249,840$. Sampling weights are included in this data as well as the school identification, which is also the primary sampling unit. Complex sampling methodology is therefore utilized in the analysis, see Asparouhov (2005). Nagengast and Marsh (2013) establish a fairly well-defined eight factor CFA model based on 44 indicators and pure loading structure. For this illustration we will use only the first 4 factors which are measured by 22 indicators. The four factors we consider here are Enjoyment, Instrumental motivation, Future-oriented motivation, and Self-efficacy. The data includes also 3 covariates that are used to predict the 4 factors: gender, socioeconomic status, and science achievement.

We consider 4 different models. The first model is the scalar measurement invariance SEM model which does not use alignment and utilizes pure loading structure, i.e., no EFA rotation/cross-loadings. The second model is the ASEM model which uses alignment, i.e., doesn't assume scalar measurement invariance, but relies on the pure loading structure, i.e., doesn't use EFA rotation/cross-loadings. The third model is the ESEM model which assumes scalar measurement invariance but utilizes EFA rotation/cross-loadings, i.e., it doesn't assume pure loading structure. The fourth model is the AESEM model which utilizes both alignment and EFA rotation/cross-loadings, i.e., it doesn't assume scalar measurement invariance or pure loading structure.

Model fit comparison is given in Table 9. The various measures of fit show that the AESEM model fits best for this data, followed by the ASEM model, followed by he ESEM model. The added parameters, needed for the groups-specific aligned measurement model, and the added cross-loadings parameters, obtained with the EFA rotation, are well justified for this data. Despite the substantial increase in the number of parameters, the AESEM model yields the best BIC. The Mplus input file used for the estimation of the AESEM model is given in Figure 11. The only difference between the input file for the

Table 9: Model fit comparison for PISA empirical example: 22 indicators, 4 factors, and 3 covariates

Model	SEM	ESEM	ASEM	AESEM
Number of parameters	1,476	1,530	2,520	4,140
Chi-Square	162,449	145,919	85,308	62,943
Degrees of Freedom	8,754	8,700	7,710	6090
BIC	10,404,759	10,376,631	10,286,767	10,268,659
CFI	0.92	0.93	0.96	0.97
TLI	0.92	0.93	0.95	0.96
SRMR	0.049	0.045	0.029	0.019
RMSEA	0.046	0.044	0.035	0.034

ESEM and AESEM model is the specification of the alignment option: ALIGNMENT=FIXED. The DEFINE statement in this input is needed to ensure that the school identification numbers are different across all the countries.

Overall, the results across the models are fairly consistent. For example, if we compare the 30 factor means for the AESEM and the SEM models, the correlation between these estimates for the four factors are 0.997, 0.997, 0.925 and 0.994. If we compare the rankings of groups by their factor means, in the AESEM and ASEM models, more than 50% of the groups preserved their rankings and those that did not, changed ranking only slightly. A notable exception is a ranking change for one of the groups for the third factor where the ranking changed from third highest to 24-th highest. The percentage of non-invariant factor loadings in the AESEM model for the four factors is 17%, 14%, 17%, and 12%. The percentage of non-invariant intercept parameters is at 50%. The most invariant indicator has invariant intercepts in 23 of the groups while the least invariant indicator has invariant intercepts in only 5 groups. The cross-loadings in the AESEM model are generally small and the pure structure holds up quite well. In 20 of the 30 groups there are no cross-loadings with Z-score above 10, in 7 of the groups there is only 1 such cross-loading, and in 3 of the groups there are 2 such cross-loadings. The size of these more substantial cross-loadings amounts to about 25% of the main loadings by size.

5 The alignment R-squared

After the alignment estimation has been completed, Mplus produces a detailed invariance analysis for all measurement parameters. The differences between every pair of parameters is evaluated for significance. A subset of the parameters is declared as an invariant set and the remaining parameters are declared as non-invariant. Verbose details of this analysis can be obtained with the option OUTPUT:ALIGN. The computation is outlined in Asparouhov and Muthén (2014). This procedure is somewhat ad-hoc because of the large number of comparisons that are analyzed, however, it appears to work fairly well for most situations.

For every measurement parameter, Mplus also produces an R-squared value, which is meant to be interpreted as the amount of variation in the parameter explained by the alignment. This R-squared value is reported under the title: "R-square/Explained variance/Invariance index". The R-squared value is between 0 and 1. Values close to 1 are associated with invariant parameters, while values closer to 0 are generally associated with non-invariant parameters. In practical applications, however, the R-squared value appears to frequently deviate from these general guidelines and it becomes a source of confusion. In this section, we elaborate somewhat on the properties of this quantity and show how it is computed for the general ASEM model.

For the loading parameters, the computation of R^2 remains unchanged and is as in formula (14) in Asparouhov and Muthén (2014)

$$R^2_{\lambda_{pm}} = 1 - Var(\lambda_{0,pmg} - \sqrt{\psi_{mg}}\,\lambda_{pm})/Var(\lambda_{0,pmg}). \qquad (43)$$

Here $\lambda_{0,pmg}$ is the configural loading for the $p-$th indicator and the $m-$th factor in group g, λ_{pm} is the average aligned loading across the groups, ψ_{mg} is the alignment factor variance in group g. The configural loading $\lambda_{0,pmg}$ is interpreted as the observed value. The quantity $\sqrt{\psi_{mg}}\,\lambda_{pm}$ is the predicted configural loading in group g, while $\lambda_{0,gpm} - \sqrt{\psi_{mg}}\,\lambda_{pm}$ is the residual. If the aligned loadings are identical, then the observed and the predicted values will be identical and the R^2 value will be 1, i.e., for a fully invariant loading parameter, the R^2 value is 1.

For the intercept parameter, the R^2 value is computed as follows

$$R^2_{\nu_p} = 1 - Var(\nu_{0,pg} - \nu_p - \sum_{m=1}^{M} \alpha_{mg}\,\lambda_{pm})/Var(\nu_{0,pg}). \qquad (44)$$

Here $\nu_{0,pg}$ is the configural intercept of the $p-$th variable in group g, ν_p is the average across groups aligned intercept, λ_{pm} is the average across groups aligned loading for the $m-$th factor, α_{mg} is the mean of the $m-$th factor in group g. In this case, $\nu_{0,pg}$ is the observed value, $\nu_p + \sum_{m=1}^{M} \alpha_{mg}\ \lambda_{pm}$ is the predicted value, and $\nu_{0,pg} - \nu_p - \sum_{m=1}^{M} \alpha_{mg}\ \lambda_{pm}$ is the residual. Under scalar invariance, the R^2 value is 1. The further away R^2 is from 1, the less invariant the intercept parameter is.

The R^2 measure has a number of caveats that should be taken into account when the value is used in practical applications. First, the R^2 value is based on variance estimates. It is somewhat difficult to rely on such estimates when the number of groups is small. Second, the R^2 value does not reflect statistical significance. It is a common occurrence to see a low R^2 value and at the same time the parameter estimates across groups to be invariant because the parameters are not statistically significant from each other. This can occur for example in those situations where the sample size is small and the power to establish statistical significance is low. Third, the R^2 value is not mathematically constrained to be above 0, as in regression analysis. It is not uncommon for one reason or another to obtain estimates for which the variance of the "predicted" or the "residual" values are bigger than that of the configural parameters. In that case R^2 is simply reported as 0. The fourth instance in which the R^2 is not a meaningful measure of invariance is the siltation where all the variances in the computation are very small. This would occur when the variation across groups in the factor means is small and when the variation in the configural intercepts is very small. Similar issues can occur also for the loading parameters. If the variation in the factor variance across groups is negligible and the variation in the configural loadings is negligible, it would be difficult to make some inference from the 0/0 ratio that the R^2 is based on. The final issue with R^2 we want to mention here is the fact that for an intercept parameter, the R^2 computation also involves the loading parameters. This implies that even if the intercept is invariant, but the loading parameters are not, the R^2 may not be 1.

In summary, the R^2 invariance index is a rough measure for how far we are from a scalar model on the level of individual parameters. Low R^2 occurs for a specific reason, but that reason can not be universally identified. In general, we do not recommend using the R^2 measure as a criterion for which parameter can be considered invariant. This should be properly done as in the pairwise comparisons that Mplus uses. The proper interpretation of the R^2 measure is that this is the proportion of variation that can be explained by the variation in the factors. If R^2 is small, the parameter is somewhat different from what can be observed for scalar invariance model parameters. As a general rule of thumb, we can expect that high R^2 are usually obtained for invariant parameters and low R^2 are usually obtained for non-invariant parameters. The five situations we listed above, however, are all exceptions to this rule of thumb.

6 Practical Guidelines

6.1 Constructing a well fitting ASEM model

The fit of the ASEM model is the same as the fit of the configural SEM model. If the substantively suggested configural SEM model does not provide an acceptable fit, the model can be modified. This should be done before the alignment procedure is used. If the number of groups is small, it may be feasible to use modification indices for the configural SEM model to guide in the model adjustments needed for an acceptable fit. Analyzing every group separately or as a part of the configural model, we can add residual correlations and cross-loading parameters suggested by the modification indices output. The modifications need not be the same in every group.

If the number of groups is large, for example, more than 5, such a process may not be realistic and may compromise the replicability of the statistical analysis. In that situation, we recommend to settle the SEM model on the population as a whole (using a single group analysis). Analyzing the entire population as one group, the SEM model can be modified with residual correlations and cross-loadings suggested by the modification indices, until the fit of the model is acceptable. In this step, we can use either an approximate fit criteria such as CFI/TLI or an exact fit criteria such as the chi-square test of fit. Given that the number of groups is large, however, which is like associated with a large total sample size, the approximate fit criteria are likely more appropriate. Once the SEM model is settled, the alignment procedure can then be used to estimate the group-specific ASEM model.

6.2 The 2-step ASEM model

If there are just a few groups in the sample, a two-step approach can be used to conduct the multiple group analysis. In the first step, alignment can be used to discover which parameters are invariant and which are not. In the second step, we can construct a CFA model without alignment where the invariant parameters are held equal across groups and those that are not invariant are free and unequal across groups. Since a large portion of the parameters will be held equal across group, the factor means and variances can be estimated in all but the first group as in the scalar invariance and the ASEM models. We call this model the 2-step ASEM model. The 2-step ASEM model, may not be feasible in the case where there is a large number of groups or a large number of non-invariant parameters. This is simply because the input file would be too tedious to write down and interpret. In principle, however, the 2-step ASEM model is valuable even if the number of groups is large.

There are 3 competing models that essentially purport to do the same thing: estimate a SEM model that includes group-specific factor means and variances. The three models are: ASEM, the 2-step ASEM, and the scalar invariance model. The ASEM model will have a much larger number of parameters as compared to the 2-step ASEM and the scalar invariance models. The question then arises, whether this large number of parameters is supported by the data. A formal chi-square can be conducted between the three models since the models are nested within each other. Alternatively, the three models can be compared with the BIC criterion. The 2-step ASEM model is generally expected to have a better BIC than the scalar invariance model if there are statistically significant non-invariant parameters. Furthermore, if the group sizes are small, the ASEM model is expected to lose in the BIC comparison because of the much larger number of parameters relative to the sample size. Therefore, the 2-step ASEM model is expected to have the best BIC in most situations.

6.3 Considering the size of the model

In the AESEM model, a large portion of the analysis is automated. In the absence of covariates for example, the only decision that must be made is how many factors are in the model. The model statement itself is a single line. Because of this simplicity, it may become increasingly tempting to estimate bigger and bigger AESEM models without much consideration. This in turn may lead to a variety of problems.

An AESEM model with large number of variables and groups will have a very large number of parameters and may become difficult to estimate in terms of computational time. A number of preliminary steps can be taken to better understand the effect of the number of variables and factors on the computational time. For example, the model can be estimated with smaller number of variables or with smaller number of groups.

Furthermore, the power of the model may be reduced because of the large number of parameters. A reasonable approach that could address this issue is to convert the AESEM model to a SEM model where non-significant loadings are removed and the invariant parameters are held equal across the groups, thereby obtaining a 2-step AESEM model.

Finally, a large AESEM model may lead to convergence problems. A number of preliminary steps can be taken to resolve such issues. For example, the configural model can be estimated in each group separately. Convergence problems will surely be easier to resolve in a single group analysis then in the full AESEM model. More generally, the AESEM model can be estimated with smaller number of variables and/or with smaller number of groups. Such an approach may provide a path to identifying and resolving convergence problems that occur in the full AESEM model.

6.4 Factor permutation and sign

The sign of the factor is generally an unidentified quantity. A factor f provides the same model fit as the factor $-f$ when the loadings are also reversed. The alinement procedure must have the same factor direction in all groups. Otherwise the alignment becomes meaningless. This is achieved by constraining the sum of all loadings to be positive for every factor and group. Such a constraint is also implemented for ESEM models. Therefore for AESEM, ESEM and ASEM models, the uncertainty of the sign of the factors is technically removed. However, that is not the case for SEM models. SEM models are more general than AESEM/ESEM/ASEM models. For example, factor loadings can be fixed to 1 in SEM models but they can not be fixed to any value in models with rotation or alignment. Being more general, however, also leads to the fact that the sign of the factor is not completely removed from the SEM model (as that is not possible in the general

SEM model). This becomes important if the parameter estimates are compared between the SEM model and the AESEM/ESEM/ASEM models. It is necessary to manually check that the SEM model has positive sum of loadings in all groups and factors. If random starting values have not been used, this is generally assured by the starting values of the SEM model. With random starting values the picture may become a little more complex and multiple runs might be necessary with various manually entered starting values.

If an alignment model contains factors that have both positive and negative large loadings, it may happen that the sum of the loadings is near zero. In that case the direction of the factor may switch between the groups and the alignment may become invalid. To avoid this problem, the factor indicators showing large negative loadings can be reversed.

The ESEM model and the AESEM model also have to deal with factor permutation. Reordering of the EFA factors yields equivalent models. This model non-identification is resolved by ordering the factors according to the average index of their large loadings, see Appendix D, Asparouhov and Muthén (2009). This becomes somewhat of a critical issue in AESEM. We have to make sure that all groups yield the same EFA factor order. If they don't, the alignment again becomes meaningless as we would be aligning the loadings of different factors. A simple way to ensure that the order stays the same in all groups is to order the factor indicators according to the desired order of the factors. The primary indicators of the first factor should be placed first in the USEVAR option, the primary indicators of the second factor should be placed next, etc. If we don't know what the primary indicators are, it may be necessary to estimate an ESEM model for all groups combined as one and use such a model to determine which the primary indicators are for every factor.

7 Conclusion

Expanding the alignment methodology to the general SEM and ESEM models provides a much broader application area for the multiple group analysis. Multiple groups SEM and ESEM models can now be estimated without assuming measurement invariance. Adopting the methodology to the WLS type estimators, in addition to the ML estimator, completes the availability of alignment for the most commonly used SEM and ESEM frameworks. Alignment models with continuous, binary and ordinal dependent variables, and any number of latent variables can be estimated with the WLS estimators. Improvements in the Mplus language and implementation also facilitates greater ease of use. The difference between the scalar model specification and the ASEM specification is only in the addition of the alignment option. Also, test of fit statistics and modifications indices are now obtained within the alignment estimation, which simplifies the overall multiple group analysis.

The alignment procedure can also be used with panel/longitudinal CFA models where measurement invariance doesn't necessarily hold across time. Such models cannot be formulated as multiple group models because the variables are correlated across time/groups. Future release of Mplus will include the possibility to use the alignment method in panel CFA models.

References

[1] Asparouhov T. (2005). Sampling weights in latent variable modeling. Structural Equation Modeling, 12, 411–434.

[2] Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. Structural Equation Modeling, 16, 397–438

[3] Asparouhov, T. & Muthén, B. (2014). Multiple-group factor analysis alignment. Structural Equation Modeling: A Multidisciplinary Journal, 21, 495-508.

[4] Flake, J.K. & McCoach, D. B. (2018). An investigation of the alignment method with polytomous indicators under conditions of partial measurement invariance. Structural Equation Modeling: A Multidisciplinary Journal, 25, 56-70.

[5] Lomazzi, V. (2018). Using Alignment Optimization to Test the Measurement Invariance of Gender Role Attitudes in 59 Countries. methods, data, analyses, [S.l.], v. 12, n. 1, p. 27, ISSN 2190-4936.

[6] Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., & Dicke, T. (2018). What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. Psychological Methods, 23(3), 524–545.

[7] Munck, I., Barber, C., & Torney-Purta, J. (2018). Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009: The alignment method applied to IEA CIVED and ICCS. Sociological Methods & Research, 47, 687-728.

[8] Muthén, B. and Asparouhov, T. (2014). IRT studies of many groups: the alignment method. Front. Psychol. 5:978.

[9] Muthén, B. and Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus Web Notes: No. 4, Version 5, December 9, 2002. `http://www.statmodel.com/download/webnotes/catNGLong.pdf`.

[10] Muthén, B., & Asparouhov, T. (2018). Recent methods for the study of measurement invariance with many groups: Alignment and random effects. Sociological Methods & Research, 47, 637-664.

[11] Nagengast, B., & Marsh, H. W. (2013). Motivation and engagement in science around the globe: testing measurement invariance with multigroup SEMs across 57 countries using PISA 2006. In L. Rutkowski, M. von Davier, D. Rutkowski (Eds.). A Handbook of International Large-Scale Assessment Data Analysis (pp. 317-344). Chapman & Hall, CRC Press

[12] Rudnev, M. (2019). Elements of cross-cultural research. `https://maksimrudnev.com/2019/05/01/alignment-tutorial/`

8 Appendix

Figure 2: Generating data with theta parameterization scalar invariance

```
MONTECARLO:
      NAMES = y1-y5;
      NGROUPS = 2; NOBSERVATIONS = 2(10000); NREPS = 100;
      generate=y1-y5(1); categorical=all;
      SAVE=1.dat;

ANALYSIS: alignment=fixed; parameterization=theta;

MODEL POPULATION: f1 BY y1*0.6 y2*0.8 y3*1 y4*1.2 y5*1.4;
                  f1*1;
MODEL POPULATION-G2: f1*1.8;

MODEL: f1 BY y1*0.6 y2*0.8 y3*1 y4*1.2 y5*1.4;
       f1*1;
MODEL G2: f1*1.8;
```

Figure 3: Input file for alignment with the theta parameterization

```
VARIABLE: NAMES = y1-y5 g; categorical=all;
          grouping=g(1-2);
DATA: file=1.dat;
ANALYSIS: alignment=fixed; parameterization=theta;
MODEL: f1 BY y1-y5;
OUTPUT: align;
```

Figure 4: Input file for alignment with the delta parameterization

```
VARIABLE: NAMES = y1-y5 g; categorical=all;
          grouping=g(1-2);
DATA: file=1.dat;
ANALYSIS: alignment=fixed;
MODEL: f1 BY y1-y5;
OUTPUT: align;
```

Figure 5: Generating data with delta parameterization scalar invariance

```
MONTECARLO:
    NAMES = y1-y5;
    NGROUPS = 2; NOBSERVATIONS = 2(10000); NREPS = 100;
    generate=y1-y5(1); categorical=all;
    SAVE=1.dat;

ANALYSIS: alignment=fixed;

MODEL POPULATION:
    f1 BY y1*0.4 y2*0.5 y3*.6 y4*.7 y5*.8; f1*1;
    y1*0.84 y2*0.75 y3*0.64 y4*0.51 y5*0.36;

MODEL POPULATION-G2:
    f1*0.5;
    y1*0.92 y2*0.875 y3*0.82 y4*0.755 y5*0.68;

MODEL:
    f1 BY y1*0.4 y2*0.5 y3*.6 y4*.7 y5*.8; f1*1;

MODEL G2:
    f1*0.5;
```

Figure 6: Alignment simulation for a 2-factor analysis model with cross-loadings

```
MONTECARLO:
    NAMES = y1-y3 z1-z3 w1-w3;
    NOBSERVATIONS = 3(1000);
    NGROUPS = 3; NREPS = 100;

ANALYSIS: alignment=fixed;

MODEL POPULATION:
    y1-w3*1;
    f1 BY y1-y3*1 w1-w3*0.5;
    f2 BY z1-z3*1 w1-w3*0.5;

MODEL POPULATION-G1:
    f1-f2*1; f1 with f2*0.4;
    f1 by w1@0; [y1*1]; !NI

MODEL POPULATION-G2:
    f1*1.4 f2*1.2; f1 with f2*-0.4;
    [f1*0.4 f2*0.6];
    f2 by w2*1; [z1*-1]; !NI

MODEL POPULATION-G3:
    f1*1.4 f2*1.2; f1 with f2*0.3;
    [f1*-1 f2*-0.5];
    f2 by z2*0.5; [w3*1]; !NI

MODEL:
    y1-w3*1; [y1-w3*0];
    f1 BY y1-y3*1 w1-w3*0.5;
    f2 BY z1-z3*1 w1-w3*0.5;

MODEL G1:
    f1-f2*1; f1 with f2*0.4;
    f1 by w1@0; [y1*1]; !NI

MODEL G2:
    f1*1.4 f2*1.2; f1 with f2*-0.4;
    [f1*0.4 f2*0.6];
    f2 by w2*1; [z1*-1]; !NI

MODEL G3:
    f1*1.4 f2*1.2; f1 with f2*0.3;
    [f1*-1 f2*-0.5];
    f2 by z2*0.5; [w3*1]; !NI
```

Figure 7: Alignment simulation for a bi-factor model

```
MONTECARLO:
      NAMES = y1-y5 z1-z5;
      NOBSERVATIONS = 2(2000);
      NGROUPS = 2;
      NREPS = 100;

ANALYSIS: alignment=fixed; tolerance=0.0001;

MODEL POPULATION:
      f1 by y1-y5*1 z1-z5*1;
      f2 BY y1*0.4 y2*0.6 y3*1 y4*1.2 y5*1.4;
      f3 by z1-z2*0.6 z3-z5*1.2;
      y1-z5*1;

MODEL POPULATION-G1:
      f1-f3*1;
      [y1*1]; f1 by y2*1.5; ! NI

MODEL POPULATION-G2:
      f1-f3*1.4;
      [f1*0.3 f2*-0.4 f3*0.5];
      [z1*1]; f2 by y5*1; ! NI

MODEL:
      f1 by y1-y5*1 z1-z5*1;
      f2 BY y1*0.4 y2*0.6 y3*1 y4*1.2 y5*1.4;
      f3 by z1-z2*0.6 z3-z5*1.2;
      y1-z5*1;
      f2 with f3@0;
      f1 with f2-f3@0;

MODEL G1:
      f1-f3*1; [y1-z5*0];
      [y1*1]; f1 by y2*1.5; ! NI

MODEL G2:
      f1-f3*1.4; [y1-z5*0];
      [f1*0.3 f2*-0.4 f3*0.5];
      [z1*1]; f2 by y5*1; ! NI
```

Figure 8: Alignment simulation for a factor analysis with a covariate

```
MONTECARLO:
      NAMES = y1-y5 x;
      NOBSERVATIONS =3(1000);
      NGROUPS = 3; NREPS = 100;
      generate=y1-y5(2);
      categorical=y1-y5

ANALYSIS: alignment=fixed; parameterization=theta;
          tolerance=0.0001;

MODEL POPULATION:
      f1 by y1-y5*1; x*1;
      f1 on x*0.7;
      [y1$1-y5$1*-1];
      [y1$2-y5$2*1];
      y1 on x*0.4;
      y2 with y3*0.3;

MODEL POPULATION-G1:
      f1*1;
      y2 with y3@0;

MODEL POPULATION-G2:
      [f1*0.4]; f1*0.8;
      f1 on x*0.4;
      y1 on x@0;
      f1 by y4-y5*0.6; ! NI
      [y1$1*0 y1$2*1.5] !NI

MODEL POPULATION-G3:
          f1*1.2; [f1*-0.3];

MODEL:
      f1 by y1-y5*1;
      f1 on x*0.7;
      [y1$1-y5$1*-1];
      [y1$2-y5$2*1];
      y1 on x*0.4;
      y2 with y3*0.3;

MODEL G1:
      f1*1;
      y2 with y3@0;

MODEL G2:
      [f1*0.4]; f1*0.8;
      f1 on x*0.4;
      y1 on x@0;
      f1 by y4-y5*0.6; ! NI
      [y1$1*0 y1$2*1.5] !NI

MODEL G3:
          f1*1.2; [f1*-0.3];
```

Figure 9: AESEM model with a covariate

```
MONTECARLO:
      NAMES = y1-y3 z1-z3 x;
      NGROUPS = 3; NOBSERVATIONS = 3(3000);
      NREPS = 100;

ANALYSIS: alignment = fixed; tolerance=0.0001;

MODEL POPULATION:
      f1 BY y1-y3*1 z1-z3*0 (*1);
      f2 BY z1-z3*1 y1-y3*0 (*1);
      y1-z3*1; x*1;
      f1 on x*0.3; f2 on x*-.2;

Model population-G1:
      f1-f2*1; f1 with f2*0.5;
      f1 by z3*0.4;

Model population-G2:
      [f1*0.5 f2*0.8]; f1*1.2 f2*1.5;  f1 with f2*0.3;
      f1 by y3*0.7;
      f1 on x*-0.3; f2 on x*.2;

Model population-G3:
      [f1*-0.5 f2*0.3]; f1*1.5 f2*1.2; f1 with f2*0.4;
      [z2*1];  f1 by z3*0.4;
      f1 on x*0.3; f2 on x*.2;

MODEL:
      f1 BY y1-y3*1 z1-z3*0 (*1);
      f2 BY z1-z3*1 y1-y3*0 (*1);
      y1-z3*1;
      f1 on x*0.3; f2 on x*-.2;

Model G1:
      f1 with f2*0.5;
      f1 by z3*0.4;

Model G2:
      [f1*0.5 f2*0.8]; f1*1.2 f2*1.5;  f1 with f2*0.3;
      f1 by y3*0.7;
      f1 on x*-0.3; f2 on x*.2;

Model G3:
      [f1*-0.5 f2*0.3]; f1*1.5 f2*1.2; f1 with f2*0.4;
      [z2*1];  f1 by z3*0.4;
      f1 on x*0.3; f2 on x*.2;
```

Figure 10: Large amount of non-invariance

```
MONTECARLO:
    NAMES = y1-y3 z1-z3 x;
    NGROUPS = 3; NOBSERVATIONS = 3(1000);
    NREPS = 100;

ANALYSIS:
    alignment = fixed; tolerance=0.0001;

MODEL POPULATION:
    f1 BY y1-z3*1;
    [y1-z3*0]; y1-z3*.5;
    f1 on x*0.4; x*1;

MODEL POPULATION-G1:
    [f1*0]; f1*1;

MODEL POPULATION-G2:
    [f1*0.3]; f1*1.5;
    f1 on x*0.7;
    [z1*-0.3 z2*0.3 z3*0.5];
    f1 BY z1*0.5 z2*0.8 z3*1.2;

MODEL POPULATION-G3:
    [f1*0.3]; f1*1.2;
    f1 on x*0.2;
    [z1*0.3 z2*0.5 z3*-0.3];
    f1 BY z1*0.8 z2*1.2 z3*0.5;

MODEL:
    f1 BY y1-z3*1;
    [y1-z3*0]; y1-z3*.5;
    f1 on x*0.4;

MODEL G1:
    [f1*0]; f1*1;

MODEL G2:
    [f1*0.3]; f1*1.5;
    f1 on x*0.7;
    [z1*-0.3 z2*0.3 z3*0.5];
    f1 BY z1*0.5 z2*0.8 z3*1.2;

MODEL G3:
    [f1*0.3]; f1*1.2;
    f1 on x*0.2;
    [z1*0.3 z2*0.5 z3*-0.3];
    f1 BY z1*0.8 z2*1.2 z3*0.5;
```

Figure 11: AESEM input file for PISA empirical example

```
DATA: FILE = pisa06_alignment_fiml_data_r.dat;

VARIABLE:

NAMES = schoolid stidstd country oecd w_fstuwt st16q01-st16q05
st17q01-st17q08 st18q01-st18q10 st19q01-st19q06 st21q01-st21q08
st29q01-st29q04 st35q01-st35q05 st37q01-st37q06
pv1scie gender ses cntgen zpv1scie zgender zses;

USEVARIABLES = st16q01-st16q05 st35q01-st35q05 st29q01-st29q04
st17q01-st17q08 zgender zses zpv1scie;

WEIGHT = w_fstuwt;
CLUSTER = schoolid;
MISSING=.;
GROUPING = country(30);

DEFINE: schoolid=(country*10000)+schoolid;

ANALYSIS: TYPE= COMPLEX; ALIGNMENT=FIXED;

MODEL:
f1-f4 BY st16q01-st17q08 (*1);
f1-f4 ON zgender zses zpv1scie;
```