

VERSIONS 8.9 AND 8.10 MPLUS LANGUAGE ADDENDUM

In this addendum, changes introduced in Versions 8.9 and 8.10 are described. They include corrections to minor problems that have been found since the release of Version 8.8 in April 2022 and new features. Following are the new features in Version 8.9:

- Automatic testing of measurement invariance for single-group longitudinal models is now available using the MODEL option of the ANALYSIS command.
- Alignment is now available for single-group longitudinal models (Asparouhov & Muthén, 2023a, Section 5.3).
- Multiple group alignment has been extended by allowing SEM and giving standardized results (Asparouhov & Muthén, 2022).
- New setting for ALIGNMENT=FIXED to choose the optimal reference group or time point. The setting is the asterisk (*) symbol.
- A new method called Penalized Structural Equation Modeling (PSEM) is available. PSEM uses an Alignment Loss Function (ALF) prior for maximum likelihood estimation that can improve estimation of models like EFA, SEM, multiple group alignment, and longitudinal alignment (Asparouhov & Muthén, 2023a).
- Random correlations are available for continuous, binary, ordered categorical (ordinal), or combinations of these variable types for TYPE=TWOLEVEL and TYPE=CROSSCLASSIFIED for ESTIMATOR=BAYES (Asparouhov & Muthén, 2023b, Section 9.1).
- New DSEM features are available. They include random correlations and changes to TINTERVAL, SAVEDATA, MONTECARLO and plots. See Mplus Web Talk No. 6.

Following are the new features in Version 8.10:

- Addition of residual covariances, residual auto-regressions, and random intercepts to automatic testing of longitudinal measurement invariance.
- PSEM for categorical variables using weighted least-squares estimation (Asparouhov & Muthén, (2023a).
- New PSEM penalty functions: LASSO and GEOMIN.
- Additional output for DIFF priors when ALIGNMENT is used.
- Hat (^) language allowed with ESEM.
- Between-level histogram for cluster-specific random effects with colors denoting significance.
- SAVEDATA for cluster-specific random effects.

TESTING FOR MEASUREMENT INVARIANCE

The MODEL option of the ANALYSIS command is used to automatically test multiple group models and single-group longitudinal models for measurement invariance. For multiple group models, the GROUPING option or the KNOWNCLASS option is used. Measurement invariance testing is available for CFA and ESEM models for continuous variables using the maximum likelihood and Bayes estimators; for censored variables using the weighted least squares and maximum likelihood estimators; for binary and ordered categorical (ordinal) variables using the

weighted least squares, maximum likelihood, and Bayes estimators; and for count variables using the maximum likelihood estimator. It is not available for censored-inflated, count-inflated, nominal, continuous-time survival, negative binomial variables, or combinations of variable types. The MODEL command can contain only BY statements for first-order factors. The metric for the factors can be set by fixing a factor loading to one in each group/time point or by fixing the factor variance to one in one group/time point. No partial measurement invariance is allowed. The configural, metric, and scalar models used are described in the next section.

The MODEL option has three settings for testing for measurement invariance: CONFIGURAL, METRIC, and SCALAR. These settings can be used alone to set up a particular model or together to test the models for measurement invariance. Chi-square difference testing is carried out automatically using scaling correction factors for MLM, MLR, and WLSM and using the DIFFTEST option for WLSMV and MLMV. The settings cannot be used together for ESTIMATOR=BAYES and for Monte Carlo analyses. Full analysis results are printed along with a summary of the difference testing. The CONFIGURAL setting produces a model with the same number of factors and the same set of zero factor loadings in all groups/time points. The METRIC setting produces a model where factor loadings are held equal across groups/timepoints. The SCALAR setting produces a model where factor loadings and intercepts/thresholds are held equal across groups/timepoints. For multiple group models, when the factor variance is fixed to one in one group, it is the first group when the GROUPING option is used and the last class when the KNOWNCLASS option is used.

The MODEL option for testing measurement invariance is specified as follows:

```
MODEL = CONFIGURAL METRIC SCALAR;
```

which specifies that configural, metric, and scalar models will be estimated and difference testing of the models will be done.

For testing longitudinal measurement invariance, time-specific MODEL commands and an overall MODEL command are used to describe the analysis model. The time-specific MODEL commands are used to specify the factor model to be tested for measurement invariance. A time-specific MODEL command must be specified for each time point and the order of the factor indicators must be the same for each time-specific MODEL command. The time-specific MODEL commands are labelled t1 for the first time point, t2 for the second time point, etc. Following is an example of how to specify the time-specific MODEL commands for a factor measured at three time points. MODEL t1 shows the model at the first time point with factor indicators y11, y12, and y13. MODEL t2 shows the model at the second time point with factor indicators y21, y22, and y23. MODEL t3 shows the model at the third time point with factor indicators y31, y32, and y33.

```
MODEL t1:  
F1 BY y11 y12 y13;  
MODEL t2:  
F2 BY y21 y22 y23;  
MODEL t3:  
F3 BY y31 y32 y33;
```

The overall MODEL command is used to specify relationships across time such as residual covariances, residual auto-regressions, and random intercepts. Following is an example of how to specify residual covariances across time.

MODEL:

y11 WITH y21; y21 WITH y31;
y12 WITH y22; y22 WITH y32;
y13 WITH y23; y23 WITH y33;

MODELS FOR TESTING MEASUREMENT INVARIANCE

Following is a description of the models used by the MODEL option to test for multiple group and longitudinal measurement invariance for various variable types and estimators.

MODELS FOR CONTINUOUS, CENSORED, AND COUNT VARIABLES

Following is a set of models that can be considered for measurement invariance of continuous, censored, and count variables. They are listed from least restrictive to most restrictive. Both maximum likelihood and Bayes estimators are available for continuous variables. Only the maximum likelihood estimator is available for censored and count variables.

For continuous, censored, and count variables, the configural model has factor loadings, intercepts, and residual variances free across groups/time points and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The metric model has factor loadings constrained to be equal across groups/time points, intercepts and residual variances free across groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

The scalar model has factor loadings and intercepts constrained to be equal across groups/time points, residual variances free across groups/time points, and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

MODELS FOR BINARY VARIABLES

Following is a set of models that can be considered for measurement invariance of binary variables. They are listed from least restrictive to most restrictive. For binary variables and weighted least squares estimation, only the configural and scalar models are considered. The metric model is not identified because scale factors or residual variances are allowed to vary across groups/time points. For binary variables and maximum likelihood estimation, the configural, metric, and scalar models are considered. The metric model is identified because residual variances are implicitly fixed at one in all groups/time points.

WEIGHTED LEAST SQUARES ESTIMATION USING THE DELTA PARAMETERIZATION

For binary variables using weighted least squares estimation and the Delta parameterization, the configural model has factor loadings and thresholds free across groups/time points, scale factors fixed at one in all groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points, scale factors fixed at one in one group/time point and free in the other groups/time points, and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

WEIGHTED LEAST SQUARES ESTIMATION USING THE THETA PARAMETERIZATION

For binary variables using weighted least squares estimation and the Theta parameterization, the configural model has factor loadings and thresholds free across groups/time points, residual variances fixed at one in all groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points, residual variances fixed at one in one group/time point and free in the other groups/time points, and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

MAXIMUM LIKELIHOOD AND BAYES ESTIMATION

For binary variables and maximum likelihood estimation, the configural model has factor loadings and thresholds free across groups/time points and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The metric model has factor loadings constrained to be equal across groups/time points, thresholds free across groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

MODELS FOR ORDERED CATEGORICAL (ORDINAL) VARIABLES

Following is a set of models that can be considered for measurement invariance of ordered categorical (ordinal) variables. They are listed from least restrictive to most restrictive. For ordered categorical (ordinal) variables and weighted least squares estimation, only the configural and scalar models are considered. The metric model is not identified (Wu & Estabrook, 2016). For ordered categorical (ordinal) variables and maximum likelihood estimation, the configural, metric, and scalar models are considered. The metric model is identified because residual variances are implicitly fixed at one in all groups/time points.

WEIGHTED LEAST SQUARES ESTIMATION USING THE DELTA PARAMETERIZATION

For ordered categorical (ordinal) variables using weighted least squares estimation and the Delta parameterization, the configural model has factor loadings and thresholds free across groups/time points, scale factors fixed at one in all groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points, scale factors fixed at one in one group/time point and free in the other groups/time points, and factor means fixed at zero in one group/time point and free in the other groups/time points.

If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

WEIGHTED LEAST SQUARES ESTIMATION USING THE THETA PARAMETERIZATION

For ordered categorical (ordinal) variables using weighted least squares estimation and the Theta parameterization, the configural model has factor loadings and thresholds free across groups/time points, residual variances fixed at one in all groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points, residual variances fixed at one in one group/time point and free in the other groups/time points, and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

MAXIMUM LIKELIHOOD AND BAYES ESTIMATION

For ordered categorical variables and maximum likelihood estimation, the configural model has factor loadings and thresholds free across groups/time points and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings and fixing the factor variance to one, the factor variance is fixed at one in all groups/time points.

The metric model has factor loadings constrained to be equal across groups/time points, thresholds free across groups/time points, and factor means fixed at zero in all groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

The scalar model has factor loadings and thresholds constrained to be equal across groups/time points and factor means fixed at zero in one group/time point and free in the other groups/time points. If the metric of a factor is set by fixing a factor loading to one, factor variances are free across groups/time points. If the metric of a factor is set by freeing all factor loadings within a group/time point and fixing the factor variance to one, the factor variance is fixed at one in one group/time point and is free in the other groups/time points.

ALIGNMENT

The ALIGNMENT option is used with multiple group and longitudinal models to assess measurement invariance and compare factor means, variances, and structural parameters across groups (Asparouhov & Muthén, 2014, 2022) and time (Asparouhov & Muthén, 2023a, Section 5.3). It is most useful when there are many groups or time points as seen in country comparisons of achievement like the Programme for International Student Assessment (PISA), the Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS) as well as in cross-cultural studies like the International Social Survey Program (ISSP) and the European Social Survey (ESS). It is available for multiple group and longitudinal models when all variables are continuous or binary with the ML, MLR, MLF, and BAYES estimators and when all variables are ordered categorical (ordinal) with the ML, MLR, MLF, and WLSMV estimators. It is available for regular and Monte Carlo analyses using TYPE=MIXTURE and TYPE=COMPLEX MIXTURE in conjunction with the KNOWNCLASS option for real data and the NGROUPS option for Monte Carlo analyses. The MODEL command can contain only BY statements for first-order factors.

The alignment optimization method consists of three steps:

1. Analysis of a configural model with the same number of factors and same pattern of zero factor loadings in all groups/time points.
2. Alignment optimization of the measurement parameters, factor loadings and intercepts/thresholds according to a simplicity criterion that favors few non-invariant measurement parameters.
3. Adjustment of the factor means, variances, and structural parameters in line with the optimal alignment.

The ALIGNMENT option has two settings: FIXED and FREE. There is no default. In the FIXED setting, a factor mean is fixed at zero in the reference group. In the FREE setting, all factor means are estimated. FREE is the most general approach. FIXED is recommended when there is little factor loading non-invariance which may occur when there is a small number of groups. The ALIGNMENT option has two subsettings, one for specifying the reference group and another for specifying the type of model used in the first step of the alignment optimization. The default for the reference group is the first group/time point. There is also an asterisk (*) setting which selects the optimal reference group (Asparouhov & Muthén, 2022a). The asterisk (*) setting is not available for longitudinal alignment. The default for the type of model is CONFIGURAL. The alternative setting is BSEM where approximate invariance of measurement parameters is specified using Bayes priors (Muthén & Asparouhov, 2013). The BSEM setting is not available for longitudinal alignment.

The subsettings are specified in parentheses following the FIXED or FREE settings. Following is an example of how to specify the ALIGNMENT option:

```
ALIGNMENT = FREE;
```

where the default reference group is the first group/time point. The default model is CONFIGURAL.

Following is an equivalent way of specifying this:

ALIGNMENT = FREE (1 CONFIGURAL);

PENALIZED STRUCTURAL EQUATION MODELING

A new method called Penalized Structural Equation Modeling (PSEM) uses a prior to impose a penalty on model estimation. The following four penalty functions are available for PSEM: Normal (N), Alignment Loss Function (ALF), LASSO, and GEOMIN. They can be used with the ML, MLR, and WLSMV estimators to improve estimation of models like EFA, SEM, growth models, multiple group alignment, and longitudinal alignment (Asparouhov & Muthén, 2023a).

MODEL PRIORS is used to specify the N, ALF, and LASSO penalties by giving two arguments that describe the penalty. The first argument (m) is the target. A typical value is zero. The second argument defines a penalty that is a function of the inverse of v. A typical value of v is one. As v increases, the penalty decreases and the prior has less influence on the analysis. Conversely, as v decreases, the penalty increases and the prior has more influence on the analysis.

Following are the MODEL PRIOR specifications for N, ALF, and LASSO where p is the parameter label from the MODEL command:

$p \sim N(0, v)$ for which the penalty is p^2/v .

$p \sim ALF(0, v)$ for which the penalty is $\sqrt{|p|}/v$.

$p \sim LASSO(0, v)$ for which the penalty is $|p|/v$.

The GEOMIN prior is used when the model involves EFA or ESEM to specify the GEOMIN rotation function as the penalty. It is a multivariate prior. All loading parameters must be specified together. Following is the MODEL PRIOR specification for GEOMIN where l1 – ln are the parameters labels for the factor loadings from the MODEL command.

$l1-ln \sim GEOMIN(m, v, eps)$

The first argument refers to the number of factors in the factor loading matrix that is to be rotated and determines the loading matrix containing the n loading parameters. The dimensions of the loading matrix is n/m by m. The second argument v is used to determine the penalty in line with the other penalty priors. The third argument eps is optional and is the epsilon of the GEOMIN rotation function. Its default is 0.01.

The order of parameters specified in the GEOMIN prior is important. They should be given column wise starting with the first column of the factor loading matrix. Following is an example of how this is specified:

MODEL:
f1-f3 BY y1-y10 (11-130);
MODEL PRIOR:
11-130 ~ GEOMIN (3,1);

Alternatively, the loadings can be given as

MODEL:
f1 BY y1-y10 (11-110);
f2 BY y1-y10 (111-120);
f3 BY y1-y10 (121-130);
MODEL PRIOR:
11-130 ~ GEOMIN (3, 1);

The second specification allows for starting values which are on the rotated scale. The unrotated model is not estimated in PSEM. Such starting values cannot be given in ESEM or EFA.

RANDOM CORRELATIONS

The | symbol is used in conjunction with TYPE=RANDOM and ESTIMATOR=BAYES to name and define the random correlation variables in the model. Random correlations are available for TYPE=TWOLEVEL and TYPE=CROSSCLASSIFIED. The name on the left-hand side of the | symbol names the random correlation variable. The variables on the right-hand side of the | symbol specify the variables that will have a random correlation. The Fisher z transformation of the random correlation is used in the model (Asparouhov & Muthén, 2023b, Section 9.1). The asterisk (*) or @ symbols may not be used on the right-hand side of the | symbol. The means and the variances of the random correlation variables are free as the default. Covariances among random correlation variables are fixed at zero as the default. Covariances among random correlation variables and growth factors, latent variables defined using BY statements, and observed variables are fixed at zero as the default. Following is an example of how to specify a random correlation using the | symbol.

c | y1 WITH y2;

where c is the random correlation for the variables y1 and y2. Examples are shown in Muthén & Asparouhov (2023).

TINTERVAL

The TINTERVAL option is used in time series analysis to specify the time interval that is used to create a time variable when times of measurement are not the same across people, for example, due to random measurement occasions (Asparouhov, Hamaker, & Muthén, 2018; Muthén & Asparouhov (2023).

The TINTERVAL option has been changed in Version 8.9. Initial missing values for an individual are no longer dropped for TYPE=TWOLEVEL. It works the same way for

TYPE=TWOLEVEL and TYPE=CROSSCLASSIFIED. The old approach can be obtained by adding DROP as the third argument in parentheses after TINTERVAL. The original time variable is no longer changed. Instead, a new time variable is created. This variable can be named by the user or the default name `_BINT` is used.

Following is an example of how to specify the TINTERVAL option:

```
TINTERVAL = hours (2 time);
```

where hours is the time variable in the data set, 2 specifies a time interval of two, and time is the time variable created by the program and used in the analysis. The variable time has values of 1, 2, 3, etc. For further details, see Asparouhov, Hamaker, and Muthén (2018) and Mplus Web Talk No. 6.

REFERENCES

Asparouhov, T. & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling*, 21, 495-508.

Asparouhov, T. & Muthén, B. (2022). Multiple group alignment for exploratory and structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, DOI: 10.1080/10705511.2022.2127100

Asparouhov, T. & Muthén, B. (2023a). Penalized structural equation models. Submitted for publication.

Asparouhov, T. & Muthén, B. (2023b). Bayesian analysis using Mplus: Technical implementation. Technical Report. Los Angeles: Muthén & Muthén.

Asparouhov, T., Hamaker, E.L. & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25:3, 359-388, DOI: 10.1080/10705511.2017.1406803

Muthén, B. & Asparouhov, T. (2013). BSEM measurement invariance analysis. *Mplus Web Notes*: No. 17. www.statmodel.com.

Muthén, B. & Asparouhov, T. (2023). Mplus Web Talk No. 6. www.statmodel.com/MplusWebTalks.shtml.

Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes." *Psychometrika*, 81, 1014–45.