

VERSION 5.1

EXAMPLES ADDENDUM

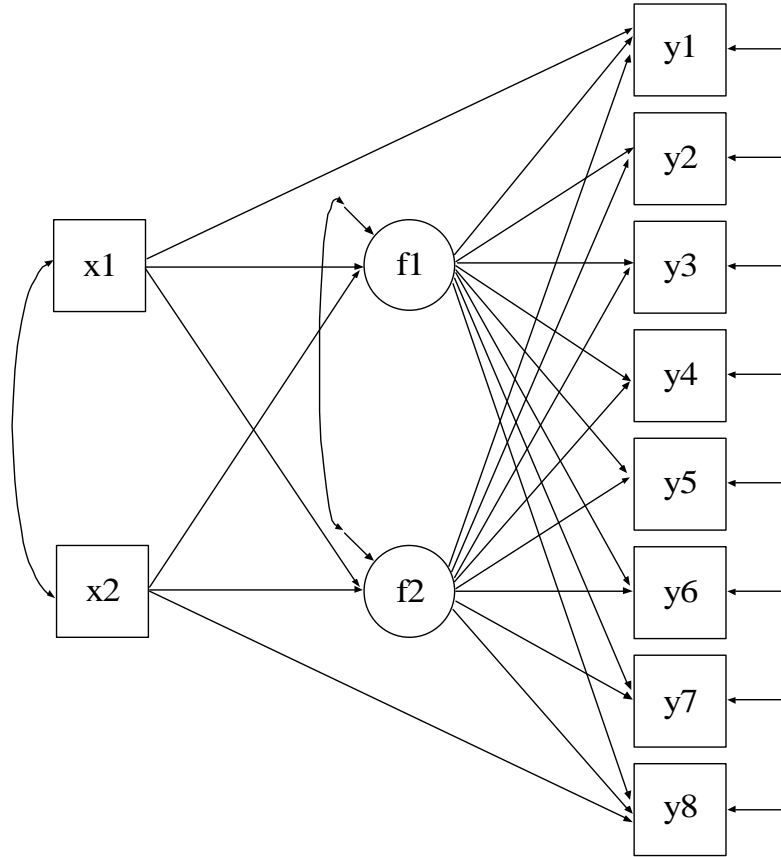
This addendum contains examples of new features included in Version 5.1. Following is the set of examples included in this addendum:

- 1: EFA with covariates (MIMIC) with continuous factor indicators and direct effects
- 2: SEM with EFA and CFA factors with continuous factor indicators
- 3: EFA at two timepoints with factor loading invariance and correlated residuals across time
- 4: Multiple-group EFA with continuous factor indicators
- 5: Monte Carlo simulation study for a multiple-group EFA with continuous factor indicators with measurement invariance of intercepts and factor loadings
- 6: GMM for a count outcome using a negative binomial model *

* Example uses numerical integration in the estimation of the model. This can be computationally demanding depending on the size of the problem.

EXAMPLE 1: EFA WITH COVARIATES (MIMIC) WITH CONTINUOUS FACTOR INDICATORS AND DIRECT EFFECTS

```
TITLE:      this is an example of an EFA with
             covariates (MIMIC) with continuous factor
             indicators and direct effects
DATA:      FILE IS example1.dat;
VARIABLE:  NAMES ARE y1-y8 x1 x2;
MODEL:    f1-f2 BY y1-y8(*1);
           f1-f2 ON x1-x2;
           y1 ON x1;
           y8 ON x2;
OUTPUT:    TECH1;
```



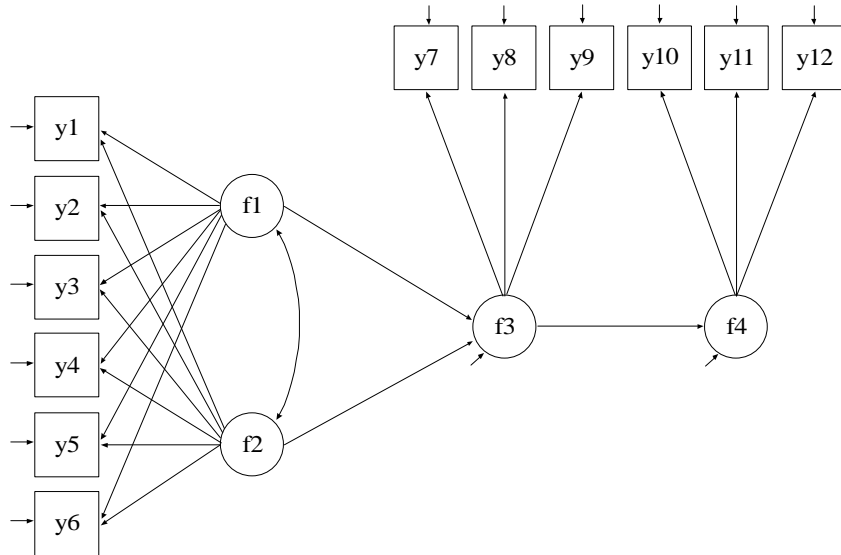
In this example, the EFA with covariates (MIMIC) with continuous factor indicators and direct effects shown in the picture above is estimated. The factors f_1 and f_2 are EFA factors which have the same factor indicators (Asparouhov & Muthén, 2008). Unlike CFA, no factor loadings are fixed at zero. Instead, the four restrictions on the factor loadings, factor variances, and factor covariances necessary for identification are imposed by rotating the factor loading matrix and fixing the factor residual variances at one.

In the MODEL command, the BY statement specifies that the factors f1 and f2 are measured by the continuous factor indicators y1 through y8. The label 1 following an asterisk (*) in parentheses following the BY statement is used to indicate that f1 and f2 are a set of EFA factors. When no rotation is specified using the ROTATION option of the ANALYSIS command, the default oblique GEOMIN rotation is used. The intercepts and residual variances of the factor indicators are estimated and the residuals are not correlated as the default. The residual variances of the factors are fixed at one as the default. The residuals of the factors are correlated under the default oblique GEOMIN rotation. The first ON statement describes the linear regressions of f1 and f2 on the covariates x1 and x2. The second and third ON statements describe the linear regressions of y1 on x1 and y8 on x2. These regressions represent direct effects used to test for measurement non-invariance.

The default estimator for this type of analysis is maximum likelihood. The ESTIMATOR option of the ANALYSIS command can be used to select a different estimator. An explanation of the other commands can be found in the example chapters of the Mplus User's Guide.

EXAMPLE 2: SEM WITH EFA AND CFA FACTORS WITH CONTINUOUS FACTOR INDICATORS

```
TITLE:      this is an example of a SEM with EFA and
            CFA factors with continuous factor
            indicators
DATA:      FILE IS example2.dat;
VARIABLE:  NAMES ARE y1-y12;
MODEL:    f1-f2 BY y1-y6 (*1);
           f3 BY y7-y9;
           f4 BY y10-y12;
           f3 ON f1-f2;
           f4 ON f3;
```



In this example, the SEM with EFA and CFA factors with continuous factor indicators shown in the picture above is estimated. The factors f1 and f2 are EFA factors which have the same factor indicators (Asparouhov & Muthén, 2008). Unlike CFA, no factor loadings are fixed at zero. Instead, the four restrictions necessary for identification, factor variances, and factor covariances necessary for identification are imposed by rotating the factor loading matrix and fixing the factor variances at one. The factors f3 and f4 are CFA factors.

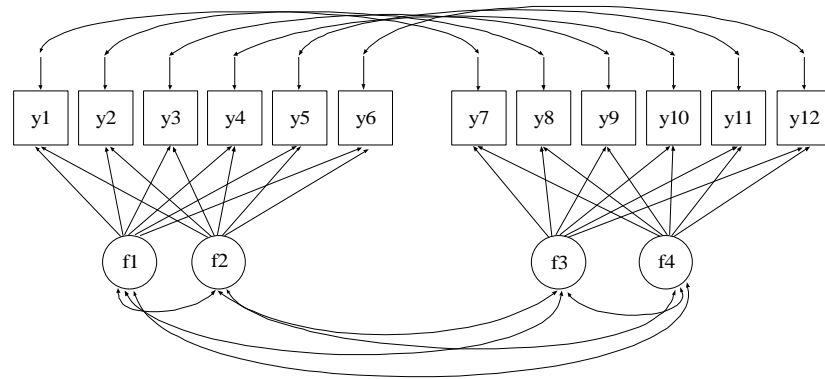
In the MODEL command, the first BY statement specifies that the factors f1 and f2 are measured by the continuous factor indicators y1 through y6. The label 1 following an asterisk (*) in parentheses following the BY statement is used to indicate that f1 and f2 are a set of EFA factors. When no rotation is specified using the ROTATION option of the ANALYSIS command, the default oblique GEOMIN rotation is used. For EFA factors, the intercepts and residual variances of the factor indicators are estimated and the residuals are not correlated as the default. The variances of the factors are fixed at one as the default. The factors are correlated under the default oblique GEOMIN rotation. The second BY statement specifies that f3 is measured by y7, y8, and y9. The third BY statement specifies that f4 is measured by y10, y11, and y12. The metric of the factors is set automatically by the program by fixing the first factor loading in each BY statement to 1.

This option can be overridden. The intercepts and residual variances of the factor indicators are estimated and the residual are not correlated as the default. The residual variances of the factors are estimated as the default.

The first ON statement describes the linear regression of f3 on the set of EFA factors f1 and f2. The second ON statement describes the linear regression of f4 on f3. The default estimator for this type of analysis is maximum likelihood. The ESTIMATOR option of the ANALYSIS command can be used to select a different estimator. An explanation of the other commands can be found in the example chapters of the Mplus User's Guide.

EXAMPLE 3: EFA AT TWO TIMEPOINTS WITH FACTOR LOADING INVARIANCE AND CORRELATED RESIDUALS ACROSS TIME

```
TITLE:      this is an example of an EFA at two
            timepoints with factor loading invariance
            and correlated residuals across time
DATA:      FILE IS example3.dat;
VARIABLE:  NAMES ARE y1-y12;
MODEL:     f1-f2 BY y1-y6 (*t1 1);
            f3-f4 BY y7-y12 (*t2 1);
            f3-f4 WITH f1-f2;
            y1-y6 PWITH y7-y12;
OUTPUT:    TECH1 STANDARDIZED;
```



In this example, the EFA at two timepoints with factor loading invariance and correlated residuals across time shown in the picture above is estimated. The factor indicators y1 through y6 and y7 through y12 are the same variables measured at two timepoints. The factors f1 and f2 are one set of EFA factors which have the same factor indicators and the factors f3 and f4 are a second set of EFA factors which have the same factor indicators (Asparouhov & Muthén, 2008). Unlike CFA, no factor loadings are fixed at zero in either set. Instead, for each set, the four restrictions on the factor loadings, factor variances, and factor covariances necessary for identification are imposed by rotating the factor loading matrix and fixing the factor variances at one.

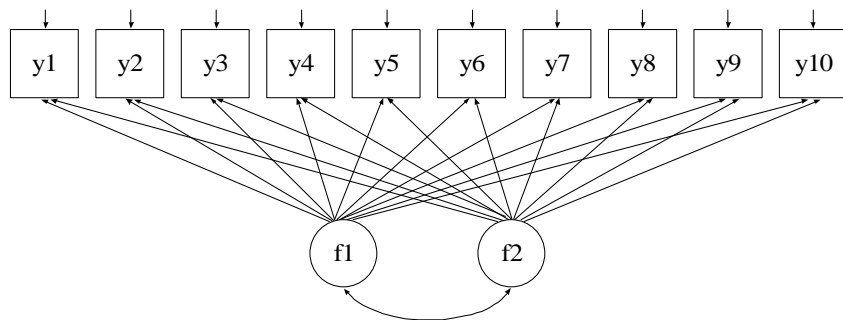
In the MODEL command, the first BY statement specifies that the factors f1 and f2 are measured by the continuous factor indicators y1 through y6. The label t1 following an asterisk (*) in parentheses following the BY statement is used to indicate that f1 and f2 are a set of EFA factors. The second BY statement specifies that the factors f3 and f4 are measured by the continuous factor indicators y7 through y12. The label t2 following an asterisk (*) in parentheses following the BY statement is used to indicate that f3 and f4 are a set of EFA factors. The number 1 following the labels t1 and t2 specifies that the factor loading matrices for the two sets of EFA factors are held equal. When no rotation is specified using the ROTATION option of the ANALYSIS command, the default oblique GEOMIN rotation is used. For EFA factors, the intercepts and residual variances of the factor indicators are estimated and the residuals are not correlated as the default. The

variances of the factors are fixed at one as the default. The factors are correlated under the default oblique GEOMIN rotation.

The WITH statement specifies that the factors are correlated over time. The PWITH statement specifies that the residuals for each factor indicator are correlated over time. The ESTIMATOR option of the ANALYSIS command can be used to select a different estimator. An explanation of the other commands can be found in the example chapters of the Mplus User's Guide.

EXAMPLE 4: MULTIPLE-GROUP EFA WITH CONTINUOUS FACTOR INDICATORS

```
TITLE:      this is an example of multiple-group EFA
            with continuous factor indicators with no
            measurement invariance
DATA:      FILE IS example4.dat;
VARIABLE:  NAMES ARE y1-y10 group;
            GROUPING IS group (1 = g1 2 = g2);
MODEL:     f1-f2 BY y1-y10 (*1);
            [f1-f2@0];
MODEL g2:  f1-f2 BY y1-y10 (*1);
            [y1-y10];
OUTPUT:    TECH1;
```



In this example, the multiple-group EFA with continuous indicators shown in the picture above is estimated. The factors f1 and f2 are EFA factors which have the same factor indicators (Asparouhov & Muthén, 2008). Unlike CFA, no factor loadings are fixed at zero. Instead, for the first group the four restrictions on the factor loadings, factor variances,

and factor covariances necessary for identification are imposed by rotating the factor loading matrix and fixing the factor variances at one. For other groups, factor variances are free to be estimated. Factor means are fixed at zero in the first group and are free to be estimated in the other groups as the default. The first model in this example imposes no restrictions on the model parameters across the two groups. Four subsequent models impose varying degrees of invariance on the model parameters.

In the MODEL command, the BY statement specifies that the factors f1 and f2 are measured by the continuous factor indicators y1 through y10. The label 1 following an asterisk (*) in parentheses following the BY statement is used to indicate that f1 and f2 are a set of EFA factors. When no rotation is specified using the ROTATION option of the ANALYSIS command, the default oblique GEOMIN rotation is used. The intercepts and residual variances of the factor indicators are estimated and the residuals are not correlated as the default. The variances of the factors are fixed at one as the default. The factors are correlated under the default oblique GEOMIN rotation. The bracket statement specifies that the factor means are fixed at zero in both groups to override the default of the factor means being fixed at zero in the first group and being free in the other group.

In the group-specific MODEL command for g2, the BY statement relaxes the default equality constraint on the factor loading matrices in the two groups. The bracket statement relaxes the default equality constraint on the intercepts of the factor indicators y1 through y10 in the two groups. The default estimator for this type of analysis is maximum likelihood. The ESTIMATOR option of the ANALYSIS command can be used to select a different estimator. An explanation of the other commands can be found in the example chapters of the Mplus User's Guide.

Following is the second part of the example where equality of factor loading matrices across the two groups is imposed.

```
MODEL:      f1-f2 BY y1-y10 (*1);
            [f1-f2@0];
MODEL g2:   [y1-y10];
```


Equality of factor loading matrices is accomplished by removing the BY statement from the group-specific MODEL command for g2. Equality of factor loadings matrices is the default.

Following is the third part of the example where equality of factor loading matrices and intercepts of the factor indicators across the two groups is imposed.

```
MODEL:      f1-f2 by y1-y10 (*1);
```

Equality of factor indicator intercepts is accomplished by removing the bracket statement for y1 through y10 from the group-specific MODEL command for g2. Equality of factor indicator intercepts is the default. This specification is the default setting in multiple group analysis, specifying measurement invariance of the intercepts of the factor indicators and the factor loading matrices.

Following is the fourth part of the example where equality of factor variances and the factor covariance is imposed in addition to measurement invariance of the intercepts and factor loading matrices.

```
MODEL:      f1-f2 by y1-y10 (*1);  
            f1 WITH f2 (1);  
            f1-f2@1;
```

In the MODEL command, the number one in parentheses following the WITH statement specifies that the covariance between f1 and f2 is held equal across the two groups. The default in multiple group EFA is that the factor variances are fixed to one in the first group and are free to be estimated in the other groups. The third statement in the MODEL command specifies that the factor variances are fixed at one in both groups.

Following is the fifth part of the example where in addition to equality of factor variances and the factor covariance, equality of the factor means is imposed in addition to measurement invariance of the intercepts and factor loading matrices.

```
MODEL:      f1-f2 by y1-y10 (*1);  
            f1 WITH f2 (1);  
            f1-f2@1;  
            [f1-f2@0];
```

The default in multiple group EFA is that the factor means are fixed to zero in the first group and are free to be estimated in the other groups. The bracket statement in the MODEL command specifies that the factor means are fixed at zero in both groups.

EXAMPLE 5: MONTE CARLO SIMULATION STUDY FOR A MULTIPLE-GROUP EFA WITH CONTINUOUS INDICATORS WITH MEASUREMENT INVARIANCE OF INTERCEPTS AND FACTOR LOADINGS

```
TITLE:      this is an example of a Monte Carlo
             simulation study for a multiple-group EFA
             with continuous factor indicators with
             measurement invariance of intercepts and
             factor loadings
MONTECARLO:
             NAMES ARE y1-y10;
             NOBSERVATIONS = 500 500;
             NREPS = 1;
             NGROUPS = 2;
MODEL POPULATION:
             f1 BY y1-y5*.8 y6-y10*0;
             f2 BY y1-y5*0 y6-y10*.8;
             f1-f2@1;
             f1 WITH f2*.5;
             y1-y10*1;
             [y1-y10*1];
             [f1-f2@0];
MODEL POPULATION-g2:
             f1*1.5 f2*2;
             f1 WITH f2*1;
             y1-y10*2;
             [f1*.5 f2*.8];
MODEL:      f1 BY y1-y5*.8 y6-y10*0 (*1);
             f2 BY y1-y5*0 y6-y10*.8 (*1);
             f1-f2@1;
             f1 WITH f2*.5;
             y1-y10*1;
             [y1-y10*1]; [f1-f2@0];
MODEL g2:  f1*1.5 f2*2;
             f1 WITH f2*1;
             y1-y10*2; [f1*.5 f2*.8];
OUTPUT:    TECH9;
```

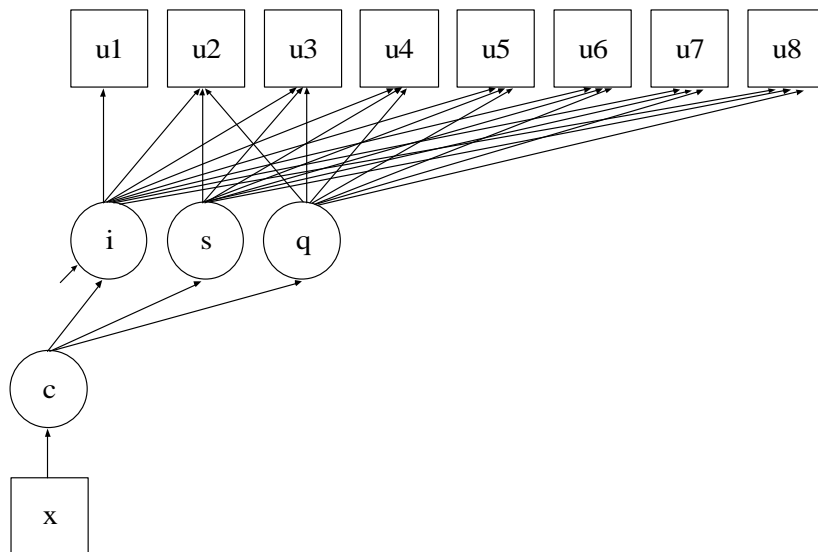
In this example, a Monte Carlo simulation study for the model estimated in Example 4 is shown for the case of measurement invariance across groups of the factor indicator intercepts and factor loading matrices. One difference between the MODEL command when EFA factors are involved rather than CFA factors is that the values given using the asterisk (*) are used only for coverage. Starting values are not allowed for the factor loading and factor covariance matrices for EFA factors. An explanation of the other commands can be found in the example chapters of the Mplus User's Guide.

EXAMPLE 6: GMM FOR A COUNT OUTCOME USING A NEGATIVE BINOMIAL MODEL

```

TITLE:      this is an example of a GMM for a count
            outcome using a negative binomial model
DATA:      FILE IS example6.dat;
VARIABLE:  NAMES ARE u1-u8 x;
            CLASSES = c(2);
            COUNT = u1-u8(NB);
ANALYSIS:  TYPE = MIXTURE;
            ALGORITHM = INTEGRATION;
MODEL:
            %OVERALL%
            i s q | u1@0 u2@.1 u3@.2 u4@.3 u5@.4 u6@.5
            u7@.6 u8@.7;
            [u1-u8@0];
            s-q@0;
            i WITH s-q@0;
            q WITH s@0;
            c ON x;
OUTPUT:    TECH1 TECH8;

```



In the example above, a growth mixture model (GMM) for a count outcome using a negative binomial model is estimated. The difference between this example and Example 8.5 in the Mplus User's Guide is that a negative binomial model is estimated instead of a zero-inflated Poisson model. The negative binomial model estimates a dispersion parameter for each of the outcomes.

The COUNT option is used to specify which dependent variables are treated as count variables in the model and its estimation and which type of model is estimated. The NB in parentheses following u1-u8 indicates that a negative binomial model will be estimated. The dispersion parameters for each of the outcomes are held equal across the latent classes as the default. By specifying ALGORITHM=INTEGRATION, a maximum likelihood estimator with robust standard errors using a numerical integration algorithm will be used. Note that numerical integration becomes increasingly more computationally demanding as the number of factors and the sample size increase. In this example, one dimension of integration is used with 15 integration points. The ESTIMATOR option of the ANALYSIS command can be used to select a different estimator. An explanation of the other commands can be found in Example 8.5 in the Mplus User's Guide.

REFERENCES

Asparouhov, T. & Muthén, B. (2008). Exploratory structural equation modeling. Technical report. www.statmodel.com.