

Mplus Survival Guide

(Exerts from the Mplus User's Guide)

The Nine Command(ment)s of Mplus:

TITLE
DATA
VARIABLE
DEFINE
ANALYSIS
MODEL
OUTPUT
SAVEDATA
MONTECARLO

Example of an Mplus input file:

Title: Filename is mexample.inp
This is an example input file

Data:
file is c:\mydata.dat;

Variable:
Names are y1 y2 y3 x;

Analysis:
Type is General;

Model:

f by y1-y3;
f on x;

Output:
Sampstat Tech1;

Savedata:

File is c:\mydata.out;
Save = fscores;

TITLE:

You can write whatever you want here. You can have a whole paragraph. There is no required syntax in this command block, e.g., you don't need “;” at the end of each line. It is a good practice to include the actual file name in this command block along with brief description of your data set and the model you are estimating

DATA:

This command block is required for all analyses.

The following are the options for the DATA command (defaults are in bold):

DATA:	
File is	filename;
Format is	format statement; Free;
Type is	Individual; Covariance; Correlation; Correlation; Fullcov; Means; Stddeviations;
Nobservations are	number of observations;
Ngroups =	number of groups;

“File is” identifies the path and name of the ASCII file that contains your data set.

“Format is” describes the format of the data set. For space-, tab-, or comma-delimited data, the default of “free” can be used. For data in fixed format, a FORTRAN-like format statement is required.

“Type is” identifies the type of data in the specified data set. The default “individual” is a data matrix where the rows represent observations and the columns represent variables.

“Nobservations are” is required when data are not “individual”.

VARIABLE:

This command block is required for all analyses.

The following are select options for the VARAIBLE command (defaults are in bold):

VARIABLE:	
Names are	names of the variables in the order in which they appear in the data set;
Useobservations are	conditional statement to subset data;
Usevariables are	variables to be analyzed;
Missing are	variable (#) or . or * ;
Categorical are	names of categorical dependent variables;
Grouping is	name of grouping variable;

“Names are” assigns names to the variables in the data set. Names can be separated by blanks are commas and can be up to 8 characters long. Names must begin with a letter. Mplus will generate a variable list, e.g., “Names are y1-y3;” creates the variables y1, y2, y3.

“Useobservations are” option selects a subset of the data for analysis (only available for data of type “individual”). Only those observations the satisfy the conditional statement are selected.

Conditional statements:

AND
OR
NOT
EQ (= =)
NE (/=
GE, GT (>=, >)
LE, LT (<=, <)

For example, “Useobservations are gender EQ 1 AND ethnicity EQ 2;”

“Usevariables are” selects variables for the analysis. If this statement is excluded, all variables will be included which mean they will appear in the sample var-cov matrix even if you do not use them explicitly in your model. *All original variables found in the data file plus all new variables created using the DEFINE command to be used in the analysis should appear under this option.*

“Missing are” specific the values in the data that are to be considered missing or invalid. For example, “Missing are all(999);”, “Missing are .;”, “Missing are ethnicity(9 99);”

“Categorical are” identifies the dependent variables in the analysis that have been measured using a categorical scales. The program recoded these variables with zero as the lowest value, one as the second lowest value, etc.

“Grouping are” is used for multiple group analysis and names the variable in the data that contains the information on group membership. A grouping variable can be created in the DEFINE command.

DEFINE:

The following are the type of transformations available for the DEFINE command:

DEFINE:

variable = mathematical expression;

IF (conditional statement) THEN variable = expression;

CUT variable (cutpoints);

This command allows you to transform existing variables and create new variable within Mplus.

Transformations of existing variables will only be used in the analysis—the changes are *not* made to your original data.

New variables that are created in DEFINE that are to be used in the analysis *must* be listed in the USEVARIABLES list *after* the original variables that will be used.

Arithmetic Operators:

+	addition
-	subtraction
*	multiplication
/	division
**	exponentiation

Some functions:

Log	base e log
Exp	exponential
Sqrt	square root
Abs	absolute value
Sin	sine
Cos	cosine
Tan	tangent

Examples:

kilos = 0.454 * pounds;

Creates a new variable, kilos, that is a transformation of pounds

weight = 0.454 * weight;

Transforms the existing variable, weight, from pounds to kilos

abuse = item1 + item2 + item3;

Creates a new variable, abuse, that is an aggregate of several existing vars

IF (gender EQ 0 AND ses EQ 3) THEN genses = 5;

IF (gender EQ 1 AND ses EQ 3) THEN genses = 6;

Creates a new variable, genses, and assigns values based on combined values of the existing variables, gender and ses.

The CUT option categorizes a variable using a set of cutpoints.

y1 CUT (30 40);

Transforms the existing variable, y1, into the follow:

y1 = 0	when y1 <= 30
y1 = 1	when 30 < y1 <= 40
y1 = 2	when y1 > 40

Missing values in the expression statement will be carried over to any new or transformed variables.

ANALYSIS:

The ANALYSIS command describes the technical details of the analysis including the type of analysis, the statistical estimator, the matrix to be analyzed, and the specifics of the computer algorithm.

The following are select options for the ANALYSIS command (defaults are in bold):

ANALYSIS:	
Type is	Basic; General Meanstructure Missing H1 Mcohort Mixture Complex Twolevel; EFA # # ; Logistic;
Estimators is	ML; MLM; . . .
Matrix is	Covariance ; Correlation;

“Type is” option is used to describe the type of analysis that is to be performed.

“Basic” allows you to obtain sample statistics, etc., without estimating a model.

“General” encompasses all models containing relationships among continuous and categorical observed variables and continuous and categorical latent variables, including linear regression, probit regression, path analysis, CFA, SEM, multilevel modeling, growth modeling, general mixture modeling.

“EFA # #” provides EFA for categorical and/or continuous dependent variables. The two numbers following “EFA” are the lower and upper limits on the number of factors to be extracted.

“Logistic” provides logistic regression for binary and ordered categorical dependent variables.

“Estimator is” option specifies the estimator to be used in the analysis. The default estimator varies according to the type of analysis.

“Matrix is” identifies the matrix to be analyzed.

MODEL:

The MODEL command describes the specific model to be estimated.

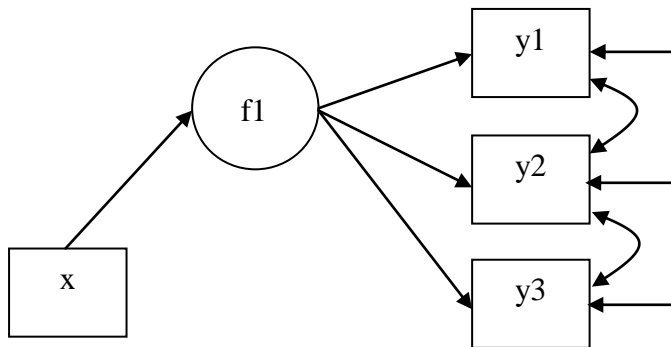
The following are select options for the MODEL command:

MODEL:	
BY	Short for “measured by” Defines latent variables
ON	Short for “regressed on” Defines regression relationships
WITH PWITH	Short for “correlated with” Defines correlational relationships
list of vars	Refers to variances and residual variances
[list of vars]	Refers to means, intercepts, thresholds

*	Frees a parameter to be estimated
@	Fixes a parameter at a specific value
variable\$number	Label for a variable threshold
variable#number	Label for a latent class

“BY” statements define and name continuous latent variables.

Consider the following path diagram:



This model corresponds the following MODEL specification:

f1 by y1 y2 y3;

y1 with y2;

y2 with y3;

f1 on x;

Note that you can name your factors whatever you wish. For examples, “f1” could just as well be “attitude”.

If you wanted to fix all the y-intercepts to zero, you would use the following statement:

[y1@0 y2@0 y3@0];

If you wanted to fix the variance of your factor to one, you would use the following statement:

```
f1@1;
```

If you wanted to factor variance to be estimated and instead fix the loading on your first factor indicator to one, you could make the following statements:

```
f1 by y1@1 y2 y3;
```

```
f1*;
```

OUTPUT:

The OUTPUT command allows you to request extra output not included by default.

The following are options for the OUTPUT command:

OUTPUT:

```
Sampstat  
Modindices (number)  
Standardized  
Residuals  
Cinterval  
H1SE  
H1Tech3  
Patterns  
FSCoefficient  
FSDeterminacy  
Tech1  
Tech2  
Tech3  
Tech3  
Tech4  
Tech5  
Tech6  
Tech7  
Tech8  
Tech9;
```

This is not a required command.

By default, the input setup, summary of the analysis specifications, and summary of analysis results are all printed by default.

“Sampstat” produces samples statistics for the data being analyzed.

“Modindices” produces modification indices with the minimum expected drop in the chi-square given in parentheses.

“Standardized” produces standardized coefficients.

“Residual” produces model estimated means, variances, and covariances for the observed variables and the residuals that represent the differences between the observed sample statistics and their model estimated values.

“Cinterval” produces 95% and 99% confidence intervals for all of the parameter estimates.

“Tech1” produces the arrays containing parameter specification and starting values for all free parameters in the model.

SAVEDATA:

The SAVEDATA command allow you to save the analysis data and/or a variety of model results in ASCII files for future use.

The following are options for the SAVEDATA command (defaults are in bold):

SAVEDATA:	
File is	filename;
File (sample) is	filename;
File (sigb) is	filename;
File (results) is	filename;
File (tech3) is	filename;
File (tech4) is	filename;
Format is	format statement; Free;

Type (sample) is	Covariance; Correlation;
Type (sigb) is	Covariance; Correlation;
Recordlength is	characters per record;
Save is	Fscores; Cprobabilities;

“File” names the ASCII file to which the saved data will be written.

MISCELLANEOUS:

“IS”, “ARE”, and “=” can be used interchangeably in all the command blocks except in DEFINE (where you must use “=”).

Command options are NOT case-sensitive.

You may add a comment to any part of you input file by proceeding it with “!”

In the Mplus environment, the File → New will open a new text window in which you can type your input file. Then use File → SaveAs to save your input file. The convention is to use “.inp” for the input file extension. Click on the “RUN” button or use Mplus → Run Mplus to submit your input file. Your output file will then appear on top of your input file in the text window. The output file is automatically saved and the name of the output file is the same as the input file with the extension “.out”. The default directory is the one in which your input file resides.