

CHAPTER 2

GETTING STARTED WITH Mplus

After Mplus is installed, the program can be run from the Mplus editor. The Mplus Editor for Windows includes a language generator and a graphics module. The graphics module provides graphical displays of observed data and analysis results.

In this chapter, a brief description of the user language is presented along with an overview of the examples and some model estimation considerations.

THE Mplus LANGUAGE

The user language for Mplus consists of a set of ten commands each of which has several options. The default options for Mplus have been chosen so that user input can be minimized for the most common types of analyses. For most analyses, only a small subset of the Mplus commands is needed. Complicated models can be easily described using the Mplus language. The ten commands of Mplus are:

- TITLE
- DATA (required)
- VARIABLE (required)
- DEFINE
- ANALYSIS
- MODEL
- OUTPUT
- SAVEDATA
- PLOT
- MONTECARLO

The TITLE command is used to provide a title for the analysis. The DATA command is used to provide information about the data set to be analyzed. The VARIABLE command is used to provide information about the variables in the data set to be analyzed. The DEFINE command is used to transform existing variables and create new variables. The ANALYSIS command is used to describe the technical

CHAPTER 2

details of the analysis. The `MODEL` command is used to describe the model to be estimated. The `OUTPUT` command is used to request additional output not included as the default. The `SAVEDATA` command is used to save the analysis data, auxiliary data, and a variety of analysis results. The `PLOT` command is used to request graphical displays of observed data and analysis results. The `MONTECARLO` command is used to specify the details of a Monte Carlo simulation study.

The Mplus commands may come in any order. The `DATA` and `VARIABLE` commands are required for all analyses. All commands must begin on a new line and must be followed by a colon. Semicolons separate command options. There can be more than one option per line. The records in the input setup must be no longer than 90 columns. They can contain upper and/or lower case letters and tabs.

Commands, options, and option settings can be shortened for convenience. Commands and options can be shortened to four or more letters. Option settings can be referred to by either the complete word or the part of the word shown in bold type in the command boxes in each chapter.

Comments can be included anywhere in the input setup. A comment is designated by an exclamation point. Anything on a line following an exclamation point is treated as a user comment and is ignored by the program.

The keywords `IS`, `ARE`, and `=` can be used interchangeably in all commands except `DEFINE`, `MODEL CONSTRAINT`, and `MODEL TEST`. Items in a list can be separated by blanks or commas.

Mplus uses a hyphen (-) to indicate a list of variables or numbers. The use of this feature is discussed in each section for which it is appropriate. There is also a special keyword `ALL` which can be used to indicate all variables. This keyword is discussed with the options that use it.

Following is a set of Mplus input files for a few prototypical examples. The first example shows the input file for a factor analysis with covariates (MIMIC model).

```

TITLE:      this is an example of a MIMIC model
            with two factors, six continuous factor
            indicators, and three covariates
DATA:      FILE IS mimic.dat;
VARIABLE:  NAMES ARE y1-y6 x1-x3;
MODEL:     f1 BY y1-y3;
            f2 BY y4-y6;
            f1 f2 ON x1-x3;

```

The second example shows the input file for a growth model with time-invariant covariates. It illustrates the new simplified Mplus language for specifying growth models.

```

TITLE:      this is an example of a linear growth
            model for a continuous outcome at four
            time points with the intercept and slope
            growth factors regressed on two time-
            invariant covariates
DATA:      FILE IS growth.dat;
VARIABLE:  NAMES ARE y1-y4 x1 x2;
MODEL:     i s | y1@0 y2@1 y3@2 y4@3;
            i s ON x1 x2;

```

The third example shows the input file for a latent class analysis with covariates and a direct effect.

```

TITLE:      this is an example of a latent class
            analysis with two classes, one covariate,
            and a direct effect
DATA:      FILE IS lcax.dat;
VARIABLE:  NAMES ARE u1-u4 x;
            CLASSES = c (2);
            CATEGORICAL = u1-u4;
ANALYSIS:  TYPE = MIXTURE;
MODEL:     %OVERALL%
            c ON x;
            u4 ON x;

```

The fourth example shows the input file for a multilevel regression model with a random intercept and a random slope varying across clusters.

```

TITLE:      this is an example of a multilevel
            regression analysis with one individual-
            level outcome variable regressed on an
            individual-level background variable where
            the intercept and slope are regressed on a
            cluster-level variable
DATA:       FILE IS reg.dat;
VARIABLE:   NAMES ARE clus y x w;
            CLUSTER = clus;
            WITHIN = x;
            BETWEEN = w;
            MISSING = .;
DEFINE:     CENTER x (GRANDMEAN);
ANALYSIS:   TYPE = TWOLEVEL RANDOM;
MODEL:
            %WITHIN%
            s | y ON x;
            %BETWEEN%
            y s ON w;

```

OVERVIEW OF Mplus EXAMPLES

The next eleven chapters contain examples of prototypical input setups for several different types of analyses. The input, data, and output, as well as the corresponding Monte Carlo input and Monte Carlo output for most of the examples are on the CD that contains the Mplus program. The Monte Carlo input is used to generate the data for each example. They are named using the example number. For example, the names of the files for Example 3.1 are `ex3.1.inp`; `ex3.1.dat`; `ex3.1.out`; `mcex3.1.inp`, and `mcex3.1.out`. The data in `ex3.1.dat` are generated using `mcex3.1.inp`.

The examples presented do not cover all models that can be estimated using Mplus but do cover the major areas of modeling. They can be seen as building blocks that can be put together as needed. For example, a model can combine features described in an example from one chapter with features described in an example from another chapter. Many unique and unexplored models can therefore be created. In each chapter, all commands and options for the first example are discussed. After that, only the highlighted parts of each example are discussed.

For clarity, certain conventions are used in the input setups. Program commands, options, settings, and keywords are written in upper case. Information provided by the user is written in lower case. Note,

however, that Mplus is not case sensitive. Upper and lower case can be used interchangeably in the input setups.

For simplicity, the input setups for the examples are generic. Observed continuous and censored outcome variable names start with a y; observed binary or ordered categorical (ordinal), unordered categorical (nominal), and count outcome variable names start with a u; time-to-event variables in continuous-time survival analysis start with a t; observed background variable names start with an x; observed time-varying background variables start with an a; observed between-level background variables start with a w; continuous latent variable names start with an f; categorical latent variable names start with a c; intercept growth factor names start with an i; and slope growth factor names and random slope names start with an s or a q. Note, however, that variable names are not limited to these choices.

Following is a list of the example chapters:

- Chapter 3: Regression and path analysis
- Chapter 4: Exploratory factor analysis
- Chapter 5: Confirmatory factor analysis and structural equation modeling
- Chapter 6: Growth modeling and survival analysis
- Chapter 7: Mixture modeling with cross-sectional data
- Chapter 8: Mixture modeling with longitudinal data
- Chapter 9: Multilevel modeling with complex survey data
- Chapter 10: Multilevel mixture modeling
- Chapter 11: Missing data modeling and Bayesian analysis
- Chapter 12: Monte Carlo simulation studies
- Chapter 13: Special features

The Mplus Base program covers the analyses described in Chapters 3, 5, 6, 11, 13, and parts of Chapters 4 and 12. The Mplus Base program does not include analyses with `TYPE=MIXTURE`, `TYPE=TWOLEVEL`, `TYPE=THREELEVEL`, or `TYPE=CROSSCLASSIFIED`.

The Mplus Base and Mixture Add-On program covers the analyses described in Chapters 3, 5, 6, 7, 8, 11, 13, and parts of Chapters 4 and 12. The Mplus Base and Mixture Add-On program does not include analyses with `TYPE=TWOLEVEL`, `TYPE=THREELEVEL`, or `TYPE=CROSSCLASSIFIED`.

CHAPTER 2

The Mplus Base and Multilevel Add-On program covers the analyses described in Chapters 3, 5, 6, 9, 11, 13, and parts of Chapters 4 and 12. The Mplus Base and Multilevel Add-On program does not include analyses with TYPE=MIXTURE.

The Mplus Base and Combination Add-On program covers the analyses described in all chapters. There are no restrictions on the analyses that can be requested.