Multilevel Modeling Guide

Du Toit, S.H.C. & Du Toit, M. (2021). *LISREL: Multilevel Modeling Guide*. Chapel Hill, NC: Scientific Software International, Inc.

Please download the examples from our website at https://ssicentral.com/index.php/products/lisrel/lisrel-examples and unzip them into a convenient folder location. The examples shown in the manual uses 'C:\LISREL Examples' and you are more than welcome to use the same or a different location. Please note that the actual location may be different on your machine.

Table of Contents

1 Overview

Social science research often entails the analysis of data with a hierarchical structure. A frequently cited example of multilevel data is a dataset containing measurements on children nested within schools, with schools nested within education departments.

The need for statistical models that take account of the sampling scheme is well recognized and it has been shown that the analysis of survey data under the assumption of a simple random sampling scheme may give rise to misleading results.

Multilevel models are particularly useful in the modeling of data from complex surveys. Cluster or multi-stage samples designs are frequently used for populations with an inherent hierarchical structure. Ignoring the hierarchical structure of data has serious implications. The use of alternatives such as aggregation and disaggregation of information to another level can induce an increase in collinearity among predictors and large or biased standard errors for the estimates. In order to address concerns regarding the appropriate analyses of survey data, the LISREL multilevel module features an option for users to include design weights on levels 1, 2 , 3, 4 or 5 of the hierarchy.

2 Multilevel models

2.1 Introduction

There has been a growing interest in recent years in fitting models to data collected from longitudinal surveys that use complex sample designs. This interest reflects expansion in requirements by policy makers and researchers for in-depth studies of social processes over time.

Although structural equation modeling allows for a tremendous flexibility in modeling error structures, it is in general not straightforward to analyze nested data structures with it. This, on the other hand, is a strong point of multilevel modeling, which is also more flexible than structural equation modeling when repeated measurement occasions vary between individuals. In order to address concerns regarding the appropriate analyses of survey data, the LISREL multilevel module features an option for users to include design weights on levels 1, 2, 3, 4 or 5 of the hierarchy. Correct parameter estimates and robust standard error estimates, using a Taylor linearization approach, are produced.

Section 2.2 is a brief overview of the graphical user interface (GUI) for the linear multilevel linear modeling module implemented in LISREL. Section 2.3 gives the multilevel linear model syntax that is generated via the dialog boxes. For advanced users, there are additional syntax specifications presently not available via the interface dialog boxes. A number of examples are given in Section 2.4. In Section 2.5 we provide evaluation and simulation studies. Section 2.6 describes the general weighting strategy of Pfeffermann *et al.* (1997), followed by a more rigorous theoretical treatment.

Additional examples to illustrate various aspects of hierarchical linear modeling are contained in the **Multilevel Examples** folder. A list of these examples is given in the table below.

LIST OF MULTILEVEL EXAMPLES

Mouse data: Variance decomposition Data File: MOUSE.LSF Input File: MOUSE1.PRL	**Mouse data: Modeling linear growth** Data File: MOUSE.LSF Input Files: MOUSE2.PRL, MOUSE3.PRL
Mouse data: Modeling non-linear growth Data File: MOUSE.LSF Input File: MOUSE4.PRL	**Mouse data: Adding a covariate** Data File: MOUSE.LSF Input File: MOUSE5.PRL
Mouse data: Non-homogeneous level-1 variation Data File: MOUSE.LSF Input File: MOUSE6.PRL	**Air traffic data: Variance decomposition** Data File: KANFER.LSF Input File: KANFER1.PRL
Air traffic data: Non-linear model Data File: KANFER.LSF Input File: KANFER2.PRL	**Air traffic data: Adding additional predictors** Data File: KANFER.LSF Input File: KANFER3.PRL
Educational data: 3-Level Model. Variance decomposition Data File: JSP.LSF Input File: JSP1.PRL	**Educational data: 3-Level Model. Adding additional predictors** Data File: JSP.LSF Input File: JSP2.PRL
CPC data: 3-level model, all data Data File: INCOME.LSF Input File: INCOME1.PRL, INCOME2.PRL, INCOME3.PRL	**CPC data: 3-level model, education sector** Data File: EDUC.LSF Input File: EDUC.PRL
CPC data: 3-level model, construction sector Data File: CONS.LSF Input File: CONS.PRL	**Simulated data: 4-level model** Data File: SIM_LEV4.LSF Input File: SIM1_LEV4.PRL
Simulated data: 5-level model Data File: SIM_LEV5.LSF Input File: SIM1_LEV5.PRL	**Assessment data: 4-level model** Data File: THERAPIS_L4.LSF Input File: THERAPIS1_L4.PRL
Survey design data Data File: SURVEYHLM.LSF Input File: SURVEYHLM1.PRL	**Survey design data (Saturated model)** Data File: SURVEYHLM.LSF Input File: SURVEYHLM2.PRL

2.2 Graphical User Interface

2.2.1 The Multilevel Models menu

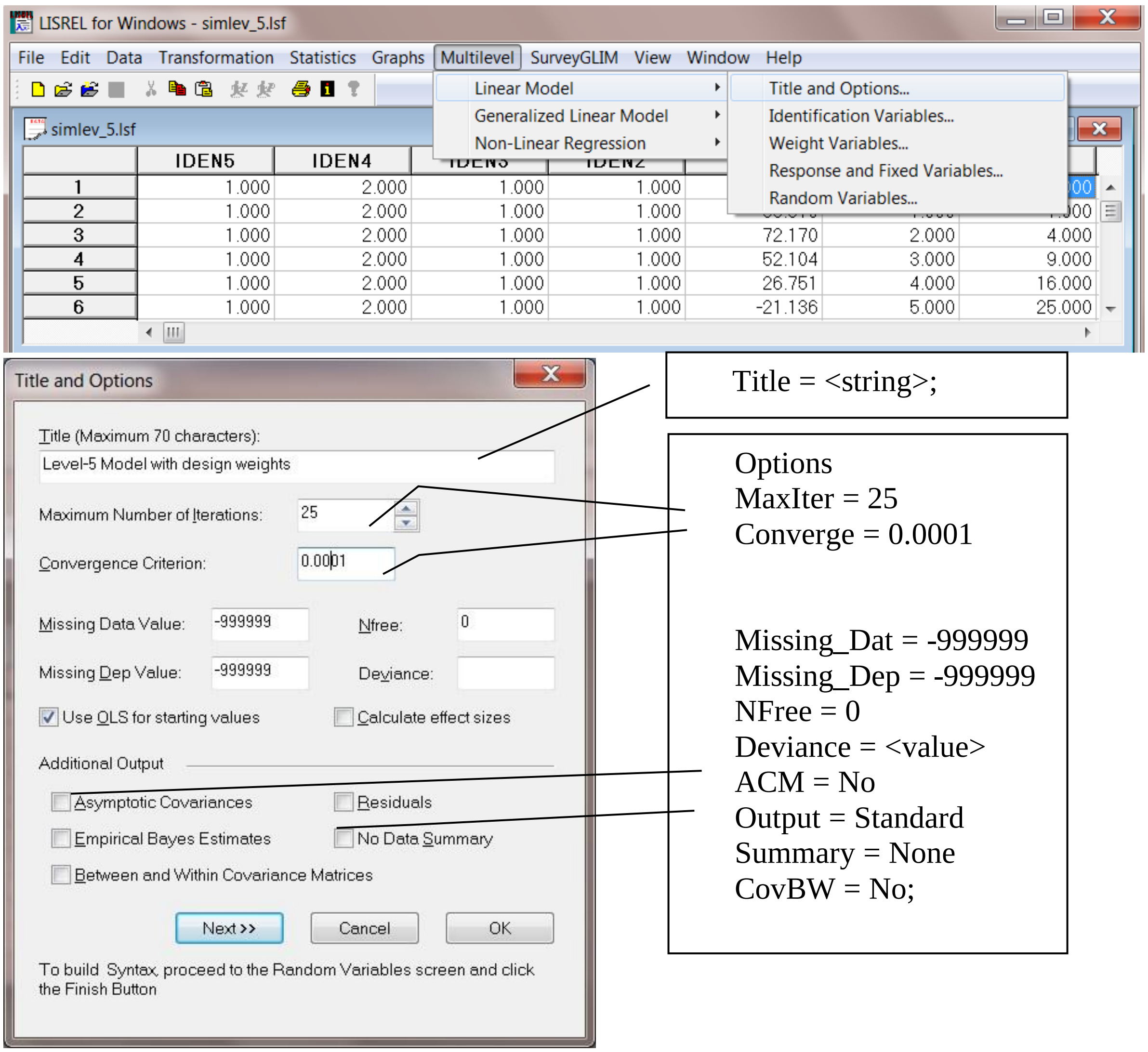
The **Multilevel** menu provides you access to three options: **Linear Model**, **Generalized Linear Model** and **Non-Linear Regression**. In this chapter, the **Linear Model** option is introduced. This menu as shown below provides you access to a sequence of five dialog boxes that can be used to create a basic syntax file interactively. It is located on the LSF (LISREL System File) window of LISREL which is used to display, manipulate and process raw data. In other words, this menu is only available when a LISREL system data file (*.**lsf**) is opened. To illustrate this, the LSF window for the file **simlev_5.lsf** is shown below with the **Multilevel**, **Linear Model** menu expanded.

Presently, the **Multilevel**, **Linear Model** menu has five dialog boxes that can be used to perform basic multilevel analyses. Advanced options that enable the user to specify more complex models must be typed in once a syntax file has been generated. These options are described in this section.

The typical first step for using the **Multilevel**, **Linear Model** menu would be to click on the **Title and Options** option to activate that dialog box (see Section 2.2.2). However, you can click on other options to go directly to the **Identification Variables** (see Section 2.2.3), **Weight Variables** (see Section 2.2.4), **Response and Fixed Variables** (see Section 2.2.5) or **Random Variables** dialog box (see Section 2.2.6).

2.2.2 The Title and Options dialog box

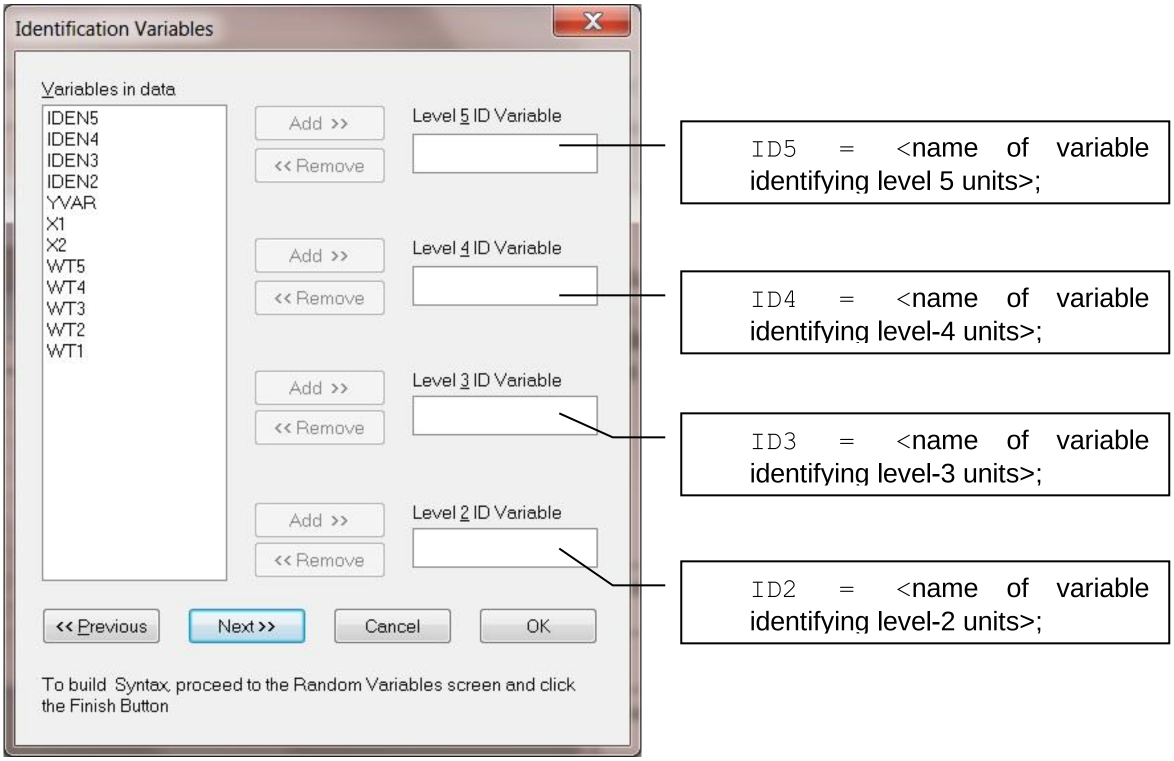
The **Title and Options** dialog box is used to provide a title for the analysis and keywords concerning the iterative procedure. The image below shows the default settings for this dialog box and, to the right, the corresponding syntax commands. See the alphabetical list of syntax commands for details on the options available other than the defaults settings: TITLE command (section 2.3.16); OPTIONS command, including the MAXITER, NFREE and OUTPUT keywords (section 2.3.12); and MISSING_DAT (section 2.3.10).



The **Next** button takes you to the **Identification of Variables** dialog box.

2.2.3 The Identification Variables dialog box

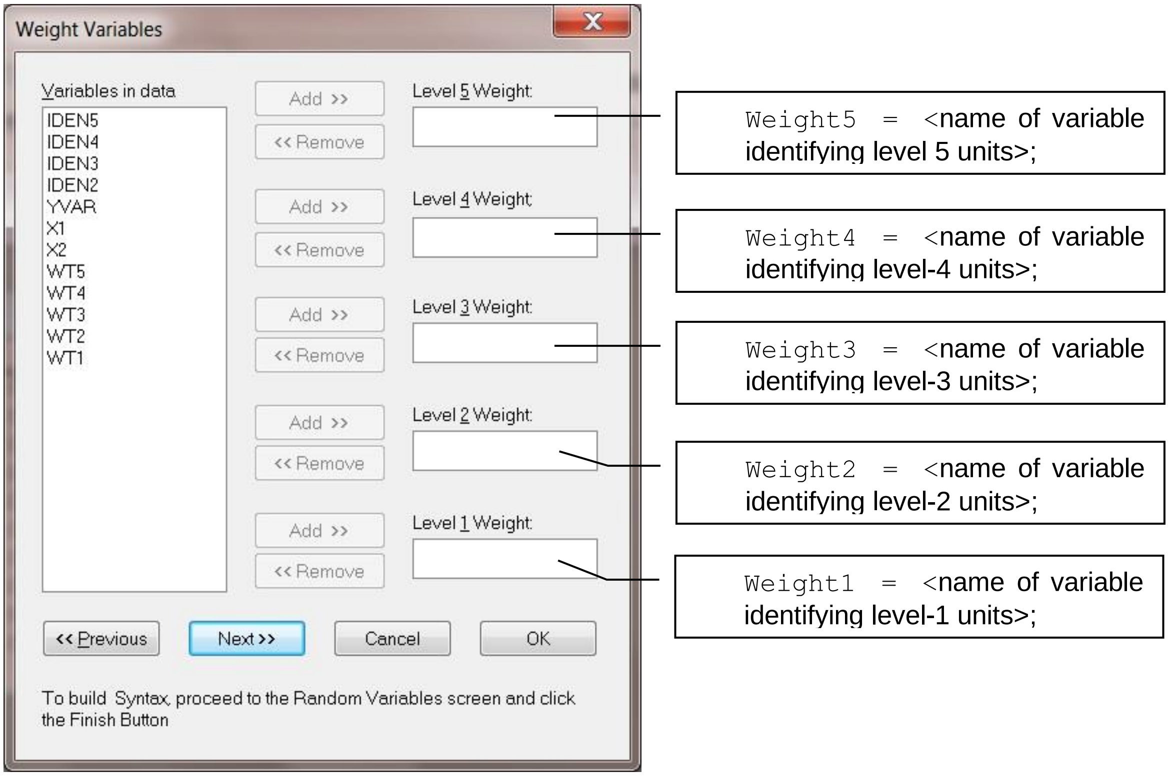
The **Identification of Variables** dialog box is used to select the variables in the LISREL data file (***.lsf**) that identify the various levels of the hierarchy. The image below shows the default settings for this dialog box and, to the right, the corresponding syntax commands. See the alphabetical list of syntax commands for details on the options available other than the default settings: IDn command (section 2.3.9).



The **Next** button provides access to the **Weight Variables** dialog box.

2.2.4 The Weight Variables dialog box

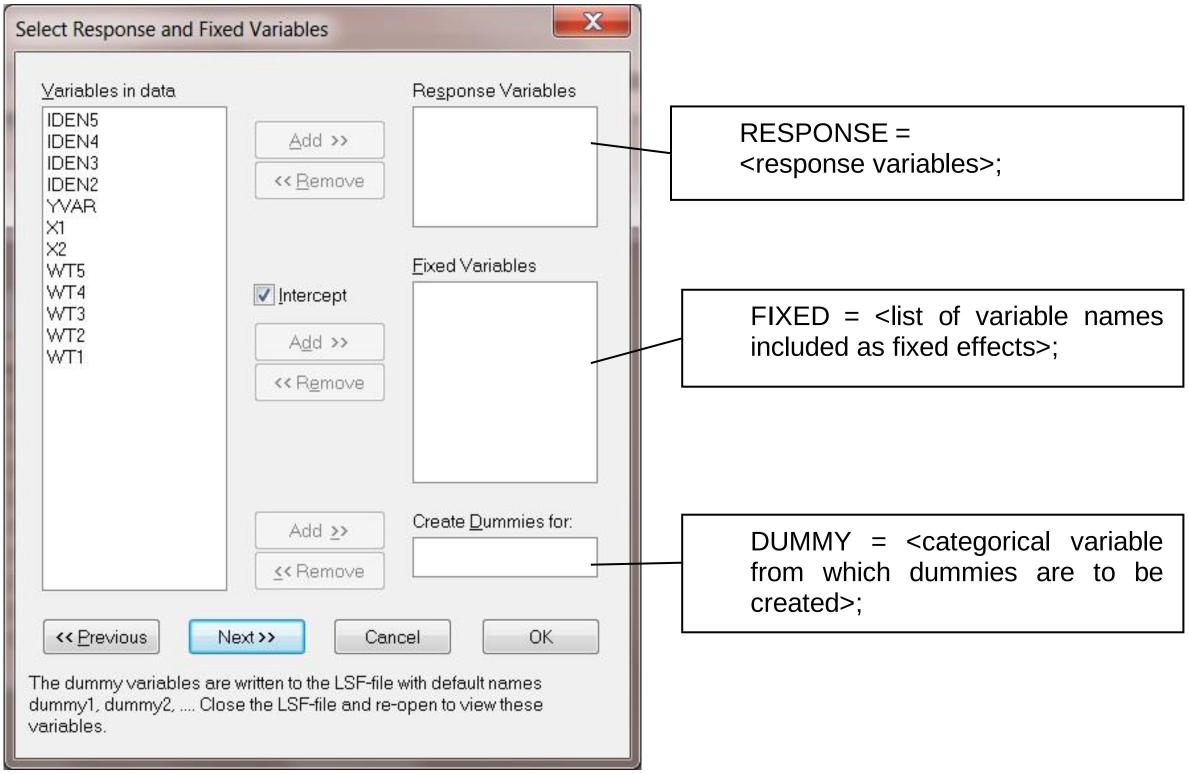
The **Weight Variables** dialog box is used to select the weight variables in the LISREL system data file for each level of the model. The image below shows the default settings for this dialog box and, to the right, the corresponding syntax commands. See the alphabetical list of syntax commands for details on the options available other than the default settings: WEIGHTn command (section 2.3.17).



The **Next** button provides access to the **Select Response and Fixed Variables** dialog box.

2.2.5 The Response and Fixed Variables dialog box

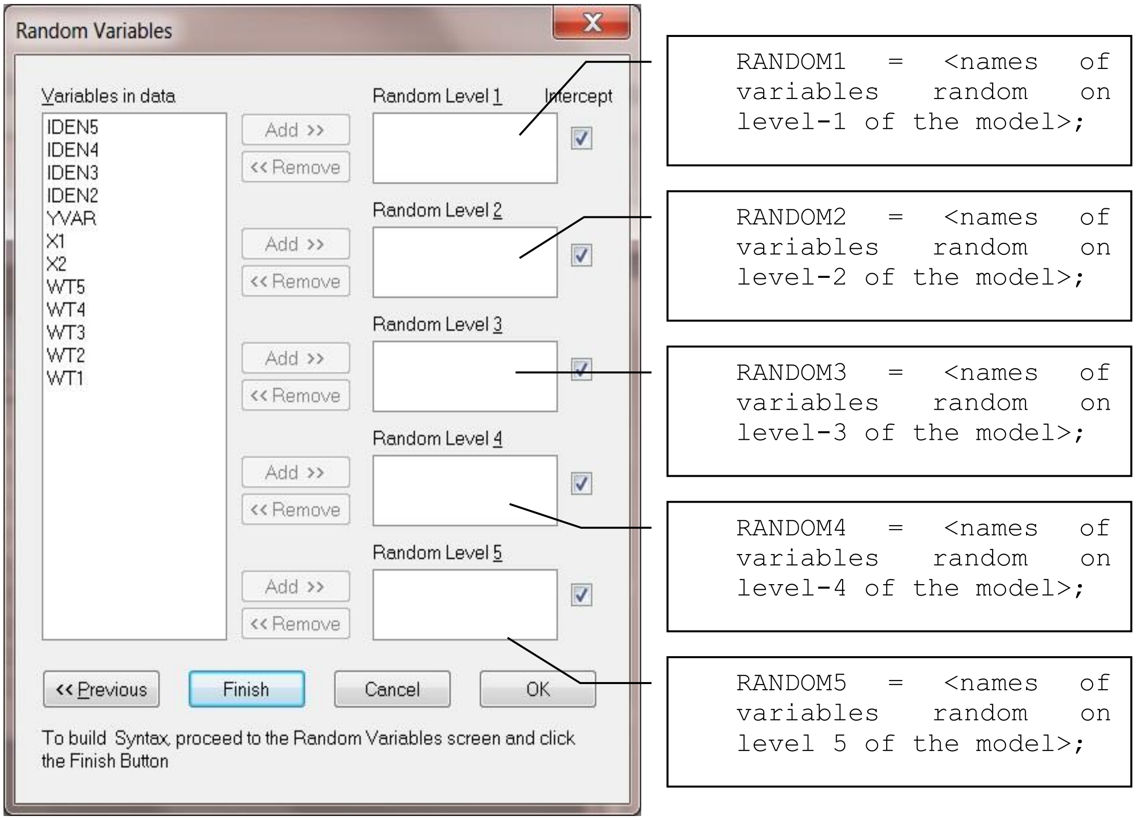
The **Select Response and Fixed Variables** dialog box is used to select the outcome and fixed variables to be included in the model from the LISREL system data file (*.**lsf**). The image below shows the default settings for this dialog box and, to the right, the corresponding syntax commands. See the alphabetical list of syntax commands for details on the option available other than the default settings: RESPONSE command (section 2.3.14), FIXED command (section 2.2.6) and DUMMY command (section 2.3.5).



The **Next** button takes you to the **Random Variables** dialog box.

2.2.6 The Random Variables dialog box

The **Random Variables** dialog box is used to select the variables for which coefficients are assumed to be random from the LISREL system data file (*.**lsf**). Default settings for this dialog box are shown in the image below. To the right, the corresponding syntax commands are given. See the alphabetical list of syntax commands for details on the options available other than the default settings: RANDOMn command (section 2.3.13).



Once all the options are set as desired, click the **Finish** button to generate the syntax.

2.3 Syntax

2.3.1 The structure of the syntax file

The basic structure of the syntax file is as given below, and the **required** commands are indicated. In this section, commands appear in the order in which they are used in the syntax file. In the sections to follow, the commands are listed in alphabetical order.

OPTIONS;	Required
SY = name of LISREL system file;	Required
IDn = name of variable identifying level n units;	Required
WEIGHT = label;	Optional
MISSING_DAT = real value;	Optional
MISSING_DEP = real value;	Optional
RESPONSE = name(s) of response variables(s);	Required
FIXED = names of variables included as fixed effects in the model ;	Required
RANDOMn = names of variables included as random effects on level n of the model ;	Required
TITLE = job title;	Optional
CONTRAST = name of contrast file;	Optional
COVnVAL = starting values for level n random coefficient covariance matrix;	Optional
COVnPAT = pattern for level n random coefficient covariance matrix;	Optional
FIXVAL = starting values for fixed effect parameters;	Optional
FIXPAT = pattern for fixed effect parameters ;	Optional
DUMMY = categorical variable from which dummies are to be created;	Optional
SUBPOP = names of variables to be used to construct subpopulations;	Optional

Guidelines for constructing or changing the syntax file:

When syntax is generated through the interface, the commands generated and saved to a ***.prl** file will automatically conform to the syntax rules given in the next section. When the syntax file is constructed or edited outside the interface, the following guidelines should be kept in mind:

- All commands start with a keyword and conclude with a semi-colon.
- There is no specific required order in which commands have to be given, with the exception of the OPTIONS command, which must always be the first line in the syntax file.
- Lines may be left blank between commands.
- Multilevel commands and keywords are not case-sensitive, but variable names are.
- Not all of the available commands have to be included in the syntax file.

In LISREL, the CONTRAST, COVnPAT, COVnVAL, FIXVAL and FIXPAT commands are not available via the graphical user interface. These commands are typically used in more advanced applications and can be added by editing the syntax file, or by writing an syntax file in a text editor. The separate commands are discussed, in alphabetical order, in the next 16 sections.

2.3.2 CONTRAST command

The CONTRAST command is used to specify the path to and name of the syntax file containing information on any contrast(s) between the fixed effects in the model to be tested. This is an **optional** command.

Syntax

```
CONTRAST = <filename>;
```

where <filename> denotes the complete name (including drive and folder names) of the file containing information on the fixed effects contrasts to be tested.

Examples:

Specifying the filename:

```
CONTRAST= C:\MLEVEL\EXAMPLES\MLEVEL.CTR;
```

The drive and folder names may be omitted if the contrast file and the syntax file are in the same folder.

Specifying the contrasts between fixed effects in a ***.ctr** file:

Suppose that there are six fixed effects in a particular model, these being INTERCEPT, GENDER, MATHS, READING, SCIENCE and WRITING.

If, for example, one wishes to test

$$H_0 : \beta_{READING} - \beta_{WRITING} = 0;$$
$$\beta_{MATHS} - \beta_{SCIENCE} = 0;$$

this can be tested by specifying

$$H_0 : \mathbf{C\beta} = 0$$

where

$$C = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$$

and

$$\beta = [\beta_{INTERCEPT} \quad \beta_{GENDER} \quad \beta_{MATHS} \quad \beta_{READING} \quad \beta_{SCIENCE} \quad \beta_{WRITING}]$$

Note that each row of **C** has six elements, corresponding to the six fixed effects. The first contrast between fixed effects is in the first row. Since the fourth element in the first equals 1, and the sixth element is -1, this denotes a contrast between the $\beta_{READING}$ and $\beta_{WRITING}$ effects.

The contrast file will have the following form:

```
2
0  0  0  1  0 -1
0  0  1  0 -1  0
```

The first row indicates the number of contrasts and the second and third rows the actual contrasts to be tested.

If the contrast file is specified as

```
1
0  0  0  1  0 -1
1
0  0  0  0 -1  0,
```

two separate contrast tests are performed as opposed to a simultaneous test for two contrasts.

2.3.3 COVnPAT command

The COVnPAT commands are used to place constraints on the covariance matrices of random coefficients on the different levels of the model. We denote these covariance matrices by $\Phi_{(1)}$, $\Phi_{(2)}$, $\Phi_{(3)}$, $\Phi_{(4)}$ and $\Phi_{(5)}$ or, in general, by $\Phi_{(n)}$, $n = 1, 2, 3, 4, 5$.

One COVnPAT command is allowed for each level of the hierarchy. If, for instance, a 5-level model with random components on all five levels of the hierarchy is to be fitted, up to five COVnPAT commands may be included in the syntax file.

Note that on level-1, only structures pertaining to the diagonal elements of the level-1 random effects covariance matrix are permissible. The use of COVnPAT commands is **optional**.

Syntax

```
COVnPAT= <keywords>;
```

Valid keywords are as follows:

DIAG In this case the covariance matrix of random parameters on level *n* of the model will be constrained to be a diagonal matrix.

TOEPLITZ The covariance matrix on levels 2 or 3 will be constrained to be of the form of a so-called Toeplitz matrix, that is

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & & \\ \gamma_1 & \gamma_0 & \gamma_1 & \ddots & \\ \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \gamma_2 \\ & \ddots & \ddots & \ddots & \gamma_1 \\ & & \gamma_2 & \gamma_1 & \gamma_0 \end{bmatrix}$$

INTRA The covariance matrix of random parameters on levels 2 or 3 will be constrained to have an intra-class structure, that is

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \alpha & \beta & \cdots & \cdots & \beta \\ \beta & \alpha & \beta & \cdots & \vdots \\ \vdots & \beta & \alpha & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \beta \\ \beta & \cdots & \cdots & \beta & \alpha \end{bmatrix}$$

MA1 Constrains the covariance matrix on level *n* to be similar to that of a time series process of order MA1. The form of the covariance matrix will then be

$$\mathbf{\Phi}_{(n)} = \begin{bmatrix} \gamma & \beta & 0 & \cdots & 0 \\ \beta & \gamma & \beta & \ddots & \vdots \\ 0 & \beta & \gamma & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \beta \\ 0 & \cdots & 0 & \beta & \gamma \end{bmatrix}$$

The following conventions apply to the use of the COVnPAT command:

- Any line of input may not exceed 127 characters. Thus, if a large COVnPAT matrix is entered, care should be taken that no row of this matrix exceeds this limit. If a row of the matrix is too long, it may simply be continued on the next line of the syntax file.
- If elements of the covariance matrix to be estimated are constrained to be equal in value, the number assigned to those elements must be the same.
- As with all other commands in the syntax file, the command should end with a semi-colon that may be placed directly after the last element of the matrix as specified or on the next line of the syntax file.
- The matrix specified must have the same number of elements as implied by the RANDOMn command. That is, if there are p variables listed in the RANDOMn command, a total number of $\frac{1}{2}\,p\,(p+1)$ elements must be entered.
- In order to assign initial values to elements of the covariance matrix on level-n or to set fixed elements of the matrix to user specified values, the COVnPAT command must be used in conjunction with the COVnVAL command.

User-specified values

To constrain the elements of the covariance matrix to be of a form other than those discussed above, you may specify this required structure with the COVnPAT command. This can be done by entering a lower-triangular matrix with the required structure on the COVnPAT command. If, for example, the covariance matrix corresponding to the RANDOMn command

```
RANDOMn = X1 X2 X3 X4;
```

is to be constrained, it can be accomplished by following a row-wise numbering convention for the lower triangular elements of the covariance matrix as shown below.

```
1
2  3
4  5  6
7  8  9  10
```

The elements to be fixed are then replaced with "0." If, for example, the matrix is constrained to be diagonal, the command to be used is as follows:

```
COVnPAT = 1
          0  3
          0  0  6
          0  0  0  10;
```

The structure as specified indicates that there are four parameters to be estimated (*i.e.* numbers 1, 3, 6 and 10, corresponding to the variances) and six fixed parameters (corresponding to the covariances), indicated by 0. The values which the fixed parameters are to be set equal to can be supplied using the COVnVAL command. If the COVnVAL command is omitted, the fixed parameters will be constrained to be equal to zero, as the initial structure of all covariance matrices are assumed to be diagonal at the start of the iterative procedure.

Examples:

In the case of an MA1 process, for example, the command will be as follows:

```
COVnPAT = 1
        2 1
        0 2 1
        0 0 2 1;
```

From this structure it follows that there are only two parameters to be estimated (numbers 1 and 2) while all other parameters are constrained to be equal to zero, unless otherwise specified using the COVnVAL command.

It is permissible to constrain diagonal elements of the level-n covariance matrix to be fixed through the use of the COVnVAL command.

The following commands, for example, are permissible:

```
COVnPAT = 1
        2 0
        4 2 0
        0 0 2 0;
```

```
COVnPAT = 0
        2 0
        4 2 0
        0 0 2 0;
```

Notes:

- 0-values indicate that the corresponding elements remain fixed at the values specified in the COVnVAL paragraphs.
- No line of input may exceed 80 charcters. Thus, if a large COVnPAT matrix is entered, care should be taken that no row of this matrix exceeds this limit. If a row of a matrix is too long, it may simply be continued on the next line of the input file.
- The matrix specified by the user must have the same number of elements as implied by the RANDOMn command. That is, if there are p variables listed in the RANDOMn command, a total number of $\frac{1}{2}p(p+1)$ elements must be entered by the user.
- In order to assign initial values to elements of the covariance matrix at level n or to set fixed elements of the matrix to user specified values, the COVnPAT command should be used in conjunction with the COVnVAL command.

2.3.4 COVnVAL command

COVnVAL commands may be used to provide either initial values for elements of the covariance matrix on level n of the model or to provide values for elements fixed through the use of keywords of the COVnPAT command. Note that the use of COVnVAL commands is **optional**.

One COVnVAL command is allowed for each level of the hierarchy. If, for instance, a 5-level model with random coefficients on all five levels of the hierarchy is to be fitted, up to five COVnVAL commands may be included in the syntax file.

The values to be used for the elements of the covariance matrix must be entered in the form of a lower-triangular matrix. The number of values entered must be the same as the number of elements implied by the relevant RANDOMn command. If there are p variables listed in the RANDOMn command, $\frac{1}{2} p (p + 1)$ values must be entered. If a large number of values is entered, a row of the lower-triangular matrix may be continued on the next line of the syntax file if the number of characters in that row of the matrix exceeds 127 characters. The command must end with a semi-colon, which may be entered on the last line of the values given or on the next line of the syntax file.

Syntax

```
COVnVAL = <values specified by user>;
```

Examples:

Providing values for the elements of the covariance matrix:

Continuing with the example used to illustrate the use of the COVnPAT command to obtain a user specified covariance structure, the following command illustrates how you may provide values for the elements of the covariance matrix (n):

```
COVnVAL = 1.00
            0.32  0.85
            0.63  0.62  0.78
            0.19  0.00  0.25  0.99;
```

If an accompanying COVnPAT command is not used, these values will function as starting values for the level-n covariance matrix. When good starting values for the elements of this covariance matrix are known, the use of the command as shown above together with the use of the keyword OLS = NO in the OPTIONS command could decrease the number of iterations required to obtain convergence.

Specifying a diagonal structure for a covariance matrix:

When the command

```
COVnPAT = DIAG;
```

is used together with the COVnVAL command given in the previous example, the values specified on the diagonal of the lower-triangular matrix will be used as initial values for the parameters which are to be estimated. The off-diagonal elements of the covariance matrix will then be constrained to be equal to the values of off-diagonal elements of the matrix given above.

2.3.5 DUMMY command

The DUMMY command is used to create dummy variables for a selected variable. Names for the dummy variables are denoted by dummy1, dummy2, …, dummyk, where k equals the number of distinct values of the selected variable. Use of the DUMMY command is **optional**.

Syntax

```
DUMMY = <varname>;
```

Example:

```
DUMMY = TIME;
```

Note:

- If the variable TIME has 4 distinct values, 0, 1, 2, and 3, then the command above will result in the creation of four dummy variables: dummy1, dummy2, dummy3, and dummy4.
- You can change the default names of the dummy variables by the inclusion of the PREFIX keyword on the DUMMY command. For example:

  ```
  DUMMY = TIME PREFIX = TIM;
  ```

  In this case, the dummy variable names will be TIM1, TIM2, TIM3, and TIM4.

2.3.6 FIXED command

The FIXED command is used to identify the fixed effects for the model to be analyzed. When the syntax file is created through the interface, the FIXED command is automatically generated. If, however, the file is edited manually, the guidelines below should be followed. Identification of the fixed effects are done on the **Select Response and Fixed Variables** dialog box (see below). The FIXED command is a **required** command, and must appear in any syntax file.

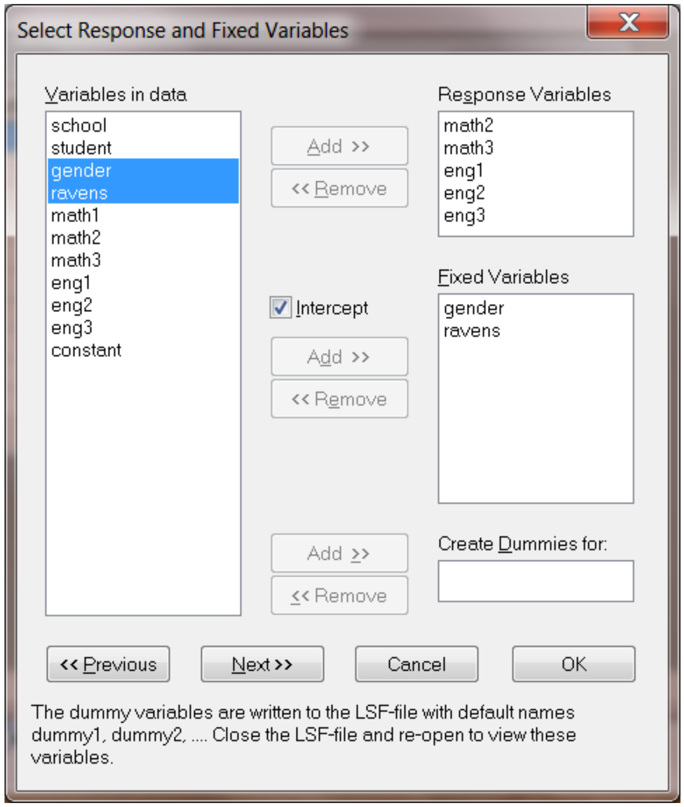
Syntax

```
FIXED = <list of variables names to be included as fixed effects>;
```

The fixed effects may be all of the predictor variables contained in the raw data file or any subset of these predictors and may be specified in any order. Variable names are case sensitive and thus spelling of the names must correspond to the spelling used in the data spreadsheet (*.**lsf** file).

If a covariate is included in the analysis, this should be reflected in the FIXED command. The format in which the covariate should be entered is as follows:

```
FIXED = intcept1 var1 var2 . . . varn covariate covariate*var1 covariate*var2 . . . covariate*varn;
```



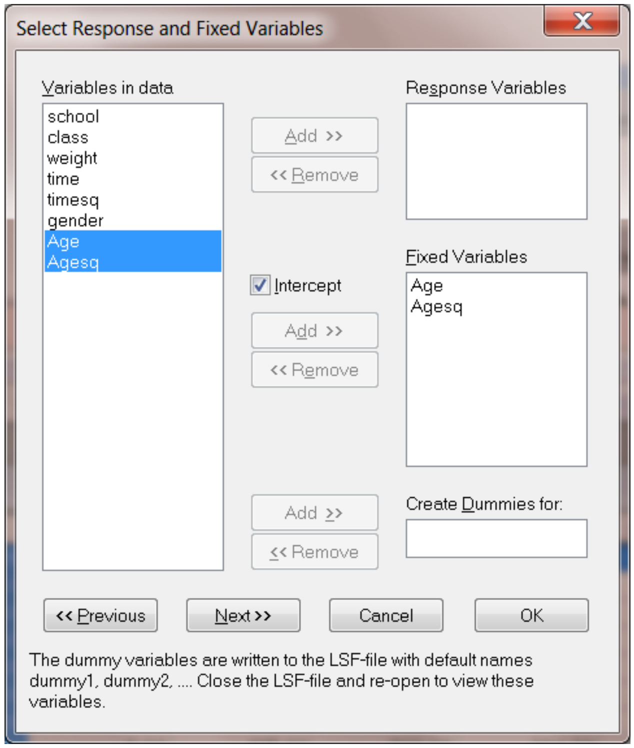
The covariate can be used in combination with any of the predictors listed in the FIXED paragraph. Note, however, that the multilevel module accepts a FIXED command of the form FIXED = var1*var2; The specification var1*var2 cannot be generated by the interface, only by manually editing the syntax file. A specification of the form var1*var2*var3 is not presently allowed.

Initial estimates for the fixed effects may be provided manually. This is done through use of the **optional** FIXVAL command that will be discussed in Section 2.3.8. See also Section 2.3.7 for a description of the FIXPAT command.

Examples:

```
FIXED = INTCEPT AGE AGESQ;
FIXED = Dummy1:Dummy6;
```

or any other similar command. The first command shown corresponds to the settings in the **Select Response and Fixed Variables** dialog box shown below.



If the variable GENDER is to be included as covariate in (1), the appropriate FIXED command is as follows:

FIXED = intcept AGE AGESQ GENDER GENDER*AGE GENDER*AGESQ;

2.3.7 FIXPAT command

To specify a patterned structure for the vector of fixed parameters, the FIXPAT command may be used, with or without an additional FIXVAL command (see Section 2.3.8). Use of this command is **optional**.

Syntax

FIXPAT = <list of numbers>;

where <list of numbers> denotes a list of positive integers separated by blank spaces. The number of values entered must equal the number of predictors in the model.

Examples:

1. Constraining fixed effects to be equal:

FIXPAT = 1 1 3 3 5 6;

This statement specifies that the vector of six parameters in the fixed part of the model are constrained as follows: BETA1 = BETA2; BETA3 = BETA4 while BETA5 and BETA6 are estimated freely. In the command shown above, the actual numbers correspond to the order of the parameter in question: "1" denotes the first parameter, "3" the third and "5" and "6" the fifth and sixth of the parameters in the fixed part of the model.

2. Fixing fixed effects to user-specified values:

```
FIXPAT = 0 0 3;
```

If '0' values are in the list of numbers, then the FIXPAT command should be used in conjunction with the FIXVAL command. If, for example, FIXVAL = 10 2.5 0.15; then BETA1 and BETA2 are fixed at their initial values (10 and 2.5 respectively) while BETA3 is estimated freely.

2.3.8 FIXVAL command

It is also possible to provide initial values for the fixed parameters in the model to be analyzed. This may be achieved with the FIXVAL command, which allows you to provide starting values for these parameters. The use of the FIXVAL command and the OLS = NO keyword of the OPTIONS command may be particularly effective when good starting values of these parameters are available. Use of this command is **optional**.

Syntax

```
FIXVAL = <as specified by user>;
```

The number of values entered by you using this command must be equal to the number of fixed parameters to be estimated. There is no specific format in which the values have to be entered.

Example

The commands

```
FIXVAL = 0.151 0.355 0.654;
FIXVAL = 0.151
         0.355
         0.654;
```

and

```
FIXVAL = 0.151  0.355
          0.654
          ;
```

are all permissible. If the first of these commands is used and the number of characters in the user specified string exceeds 127 characters, the next line of the syntax file should be used.

2.3.9 IDn command

The ID command(s) are used to indicate the variable(s) identifying the units on the different levels of the hierarchy. ID command(s) are **required** command(s).

If the model specified is a 2-level model, the command ID2 is required. Likewise, if a level-5 model is to be considered, the ID2, ID3, ID4, and ID5 commands are required in the syntax file.

Variables listed in the ID commands must be included in the data file (*.**lsf** file). Variable names are case sensitive; therefore the spelling and case in which they are given need to correspond to that given in the spreadsheet.

Syntax

IDn = <variable name identifying level-n units>;

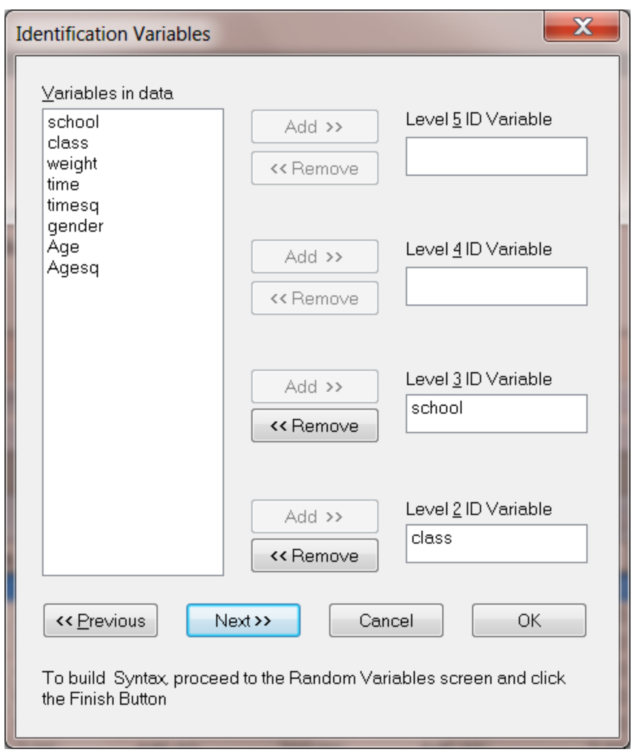
where n denotes a positive integer, (1, 2, 3, 4 or 5).

Example

Suppose the raw data file contains information on the test scores, age and gender of pupils belonging to classes within schools, and the variables school, class, pupil, age, gender and score are contained in the data file. The following ID commands may be used to identify the levels of the hierarchical structure:

```
ID3 = school;
ID2 = class;
```

The **Identification Variables** dialog box shown below shows the settings needed to obtain these commands.



2.3.10 MISSING_DAT command

The MISSING_DAT and MISSING_DEP commands may be used when missing data are present in the raw data file. The MISSING_DAT command allows you to specify a numeric value, which will represent a missing value on any of the variables used in the analysis. This command may also be used in conjunction with the MISSING_DEP command, as described in Section 2.3.11. Note that use of the MISSING_DAT command is **optional**.

Syntax

```
MISSING_DAT = <value>;
```

Any positive or negative value may be used. Only one value is allowed in this command. All records with data values equal to the code specified in this command will subsequently be removed from the analysis.

Default

-999999.0.

Examples

Valid examples of the use of this command include the following:

```
MISSING_DAT = 99;
MISSING_DAT = -998.0;
MISSING_DAT = 0;
```

2.3.11 MISSING_DEP command

The MISSING_DEP command may be used to specify a code assigned to missing values on the response variables only. The consequence of using the MISSING_DEP command is that only records with response variable values equal to the code assigned through the MISSING_DEP command will be removed from the analysis. Note that use of this command is **optional**.

The MISSING_DEP command is recommended for use in the case of multivariate analysis. If only one of the response variables to be used in the multivariate analysis has a missing response, only that particular response will be considered missing while the remaining responses will still be used.

Syntax

```
MISSING_DEP = <value>;
```

Any positive or negative value may be used. Only one value is allowed in this command. All records with response variable values equal to the code specified in this command will subsequently be removed from the analysis.

Default

-999999.0.

Example

Consider the observations

Response variables	Predictor variables
4.0 5.3 1.7 99	1 10 14.5 999
3.2 4.4 99 7.7	3 12 13.7 53.2

and the command

```
MISSING_DEP = 99;
```

If the code 99 is identified as the code for missing data values on the dependent variables, this will imply that the analysis of this record will use the first three response values and disregard the fourth one in the case of the first observation. The third response variable will be omitted for the second observation.

If, additionally, the code 999 is specified (MISSING_DAT = 999) as the code for missing data values on all the variables included in the analysis, the whole first record as given above will be deleted from the data set to be analyzed. The second observation will be retained with the exception of the third response variable value.

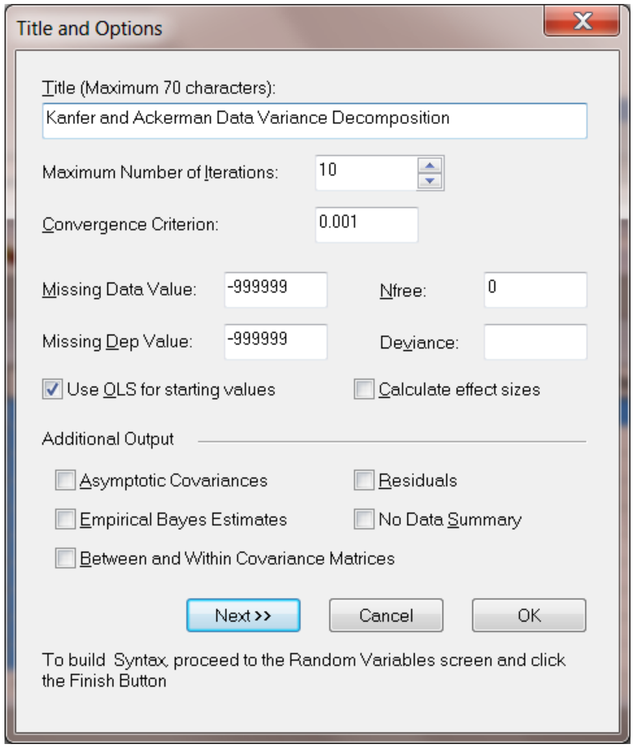
This is accomplished by using both the MISSING_DEP and MISSING_DAT commands as follows:

```
MISSING_DEP = 99;
MISSING_DAT = 999;
```

Note that if only the MISSING_DEP command is used for the two observations given above, the value of 999 for the last predictor variable on the first observation will be considered valid data and will be used as such in the analysis.

2.3.12 OPTIONS command

Each problem for a multilevel analysis starts with an OPTIONS command. The keywords of the OPTIONS command are used to control the estimation procedure and the amount of output to be written at convergence of the iterative procedure. All keywords are set via the **Title and Options** dialog box through selection of the **Title and Options** option from the **Multilevel Models** menu on the main menu bar. Inclusion of an OPTIONS command in a syntax file is **required**.



Syntax

OPTIONS <keywords>;

The keywords and options that may be used with the OPTIONS command are listed below. Details on each keyword, option, and its default value are provided below. In the generated syntax file, keywords may occur in any order. If any OPTIONS keywords are not given, the default values will be used. Also see the examples of OPTIONS commands on p. 31 of this guide.

ACM	Requests printing of asymptotic covariance matrices
Converge	Sets a value for the test for convergence made at the end of each iteration
CovBW	Requests printing of the within- and between-clusters covariance matrices
Deviance	Provides the value of $-2\ln L$ as reported in a previous analysis, with the purpose to obtain a chi-square test statistic for comparing two nested models
Effect	Estimates and print indirect effects of coefficients in the fixed part of the model Maxiter. Indicates the maximum number of iterations to be performed
Nfree	Indicates the number of free parameters as reported in a previous analysis
OLS	Indicates whether OLS estimates are to be calculated during the first iteration
Output	Sets the amount and type of output required
Summary	Requests printing of summary table containing sample sizes of units
Maxiter	The maximum number of iterations to be performed.

ACM keyword

The ACM keyword is used to print the large-sample covariance matrices of the estimated parameters in the fixed part and random part of the model. This keyword is controlled from the **Title and Options** dialog box.

Standard errors of the estimated parameters are equal to the square roots of the diagonal elements. The non-duplicated elements of these asymptotic covariance matrices are written to external files with the following default names:

<Outputfilename>_random.acm
<Outputfilename>_random.acm

If the output file name is, for example, **kanfer1.out**, then the large-sample covariance matrices are saved to the files **kanfer1_fixed.acm** and **kanfer1_random.acm**.

Syntax

ACM = Yes/No

Default

No: asymptotic covariance matrices will not be printed.

Example

In the OPTIONS command below, the ACM keyword is used to request the printing of the asymptotic covariance matrices at convergence. A convergence criterion of 0.0001 is set as the requirement for convergence, and 30 iterations is indicated as the maximum number of iterations to be performed.

```
OPTIONS MAXITER = 30 CONV = 0.0001 ACM = Yes;
```

CONVERGE Keyword

A test for convergence is made at the end of each iteration. If the absolute difference between the estimated parameters and their previous values are all smaller than the convergence criterion, convergence is said to have been reached. In the **Title and Options** dialog box shown above, the default value is automatically shown next to **Convergence Criterion**. To change the value, click in the box and enter the required convergence criterion.

Syntax

CONVERGE = <value>

Default

0.001 (10^{-2}).

Example

In order to use a value of, for example, 0.0001 as convergence criterion, the keyword CONVERGE = 0.0001 must be included as part of the OPTIONS command, as shown in the following example:

```
OPTIONS MAXITER = 10 CONVERGE = 0.0001;
```

The iterative procedure will terminate if this requirement is met, or if 10 iterations (set with MAXITER the keyword described below) have been performed without meeting this requirement.

COVBW keyword

The COVBW keyword is used to request printing of the within-clusters and between-clusters matrices of the random effects. The non-duplicated elements of these matrices are written to external files with the following default names:

<Outputfilename> _between.cov
<Outputfilename> _within.cov

For example, consider **jsp1.prl** in the **Multilevel Examples** folder. For this syntax file, the default names are **jsp1_between.cov** and **jsp1_within.cov**. This keyword is applicable to multivariate response models only and is controlled from the **Title and Options** dialog box (see example above).

Syntax

```
COVBW = Yes/No
```

Default

No: the within- and between-clusters covariance matrices will not be printed.

Example

In the OPTIONS command below, the COVBW keyword is used to request the printing of the within- and between-cluster matrices at convergence, for which a convergence criterion of 0.0001 is set as the requirement for convergence, to be attained within a maximum number of 30 iterations.

```
OPTIONS MAXITER = 30 CONVERGE = 0.0001 COVBW = Yes;
```

DEVIANCE keyword

The DEVIANCE keyword is used to provide the value of –2 log likelihood as reported in a previous analysis, in order to obtain a χ^2 test statistic for comparing two nested models. The χ^2 statistic is defined as the difference in the deviance statistics for the two models and has as associated degrees of freedom the difference in the number of parameters estimated in the models compared. It must be accompanied by the NFREE keyword, which is used to indicate the number of parameters estimated in the previous model. The DEVIANCE and NFREE keywords are controlled from the **Title and Options** dialog box (see earlier example of a **Title and Options** dialog box).

Syntax

```
DEVIANCE = <value>
```

where value equals the deviance (- 2 log L) value at convergence printed to the output file of the previous analysis.

Default

None: no –2 log likelihood value provided.

Example

In the OPTIONS command below, the DEVIANCE keyword indicates that a – 2 log likelihood value of 22735.524 was obtained in the previous analysis, and that 44 parameters were estimated (NFREE = 44). See Section 2.4.2 for a detailed example.

```
OPTIONS NFREE = 44 DEVIANCE = 22735.524;
```

EFFECTS keyword

The EFFECTS keyword is used to estimate and print indirect effects of coefficients in the fixed part of the model. This keyword is controlled from the **Title and Options** dialog box.

Syntax

```
EFFECTS = Yes/No
```

where Yes indicates that indirect effects will be computed and listed for all the predictors in the model.

Default

No: Indirect effects will not be computed.

Example

When the OPTIONS command shown below, with EFFECTS keyword set to Yes, is used in combination with accompanying FIXED command, indirect effects will be computed and listed for the predictors INTCEPT, AGE, and AGESQ.

```
OPTIONS EFFECTS = Yes;
FIXED = INTCEPT AGE AGESQ;
```

Typical output generated in the case of EFFECTS = Yes is shown below.

```
------------------------------------------------------------------------
COEFFICIENTS       LEVEL   T-SQUARED      APPROX DF      EFFECT SIZE
------------------------------------------------------------------------
Ability            1        50.83568      839            0.23902
intcept            2         4.51051      138            0.17791
time               2       520.05345      138            0.88898
timesq             2       284.90950      138            0.82079

Effect sizes are obtained by replacing the Z-values reported in the fixed part of the model by t-
values with DF as listed above.

Effect Size = sqrt[tsq/(DF + tsq)]

Level = 2: DF equals the number of level2 units – the number of level2 random coefficients – the
number of level2 covariates associated with level2 random coefficients.
```

MAXITER keyword

The keyword MAXITER is used to indicate the maximum number of iterations to be performed. The value of the keyword is set on the **Title and Options** dialog box (see Section 2.3.2). To change the value, click in the box and enter the required maximum number of iterations.

Syntax

```
MAXITER = <value>
```

Default

10.

The default number of iterations should be sufficient for convergence to be reached in most cases. If, however, a more stringent convergence criterion is used or previous experience with a particular data set indicates slow convergence, this keyword may be used to increase the maximum number of iterations. If, on the other hand, you wish to obtain only the OLS estimates calculated in the first iteration, MAXITER may be set equal to 1.

Example

In the OPTIONS command below, MAXITER is set to 30, indicating that a maximum of 30 iterations should be performed. The iterative procedure may terminate before this number is reached if the convergence criterion of 0.0001 (CONVERGE = 0.0001) is met.

```
OPTIONS MAXITER = 30 CONVERGE = 0.0001;
```

NFREE keyword

The NFREE keyword is used to denote the number of free parameters as reported in a previous analysis, in order to obtain a χ^2 test statistic for comparing two nested models. The χ^2 statistic is defined as the difference in the deviance statistics for the two models and has as associated degrees of freedom the difference in the number of parameters estimated in the models compared. It must be accompanied by the DEVIANCE keyword, which is used to provide the value of – 2 log likelihood as reported in the previous analysis. The DEVIANCE and NFREE keywords are controlled from the **Title and Options** dialog box.

Syntax

```
NFREE = <number>;
```

where number is the number of free parameters, that is, the total number of parameters estimated during the previous analysis, as reported in the output file.

Default

None: number of parameters for previous model not given.

Example

In the OPTIONS command below, the NFREE keyword indicates that 44 parameters were estimated in the previous model, with a – 2 log likelihood value of 22735.524 (DEVIANCE = 22735.524).

```
OPTIONS NFREE = 44 DEVIANCE = 22735.524;
```

See Section 2.4.2 for a detailed example.

OLS keyword

OLS estimates of the fixed effects are calculated as a first step of the iterative procedure unless otherwise specified. The OLS keyword is used to indicate whether the OLS estimates are to be calculated during the first iteration. On the **Title and Options** dialog box, the default value for this keyword is shown for **Use OLS for Starting Values**.

Syntax

```
OLS = <Yes/No>;
```

If starting values (see the FIXVAL command described in Section 2.3.8) are provided, use the OLS = NO option. To set OLS to NO, use the check box on the **Title and Options** dialog box.

Default

Yes: OLS estimates will be calculated during the first iteration.

Example

As starting values are provided for the fixed effects on the FIXVAL command, the OLS keyword is set to NO on the OPTIONS command below.

```
OPTIONS OLS = No;
FIXVAL = 0.151 0.355 0.654;
```

OUTPUT keyword

The OUTPUT keyword determines the amount of output produced. The output options are controlled from the **Title and Options** dialog box. To get more than the default output, check one or both of the boxes next to **Residuals** or **Empirical Bayes Estimates** (see description of the valid options listed below).

Syntax

```
OUTPUT = <option>;
```

where the valid options are as follows:

STANDARD The default output only

BAYES	The default output and empirical Bayes estimates
RESIDUAL	The default output and residuals
ALL	The default output, residuals and empirical Bayes estimates.

Details on each of these options are given below.

Default output (STANDARD):

The following information is written to the default output file:

1. Input specifications as supplied by you in the syntax file.
2. A summary of the hierarchical structure of the raw data.
3. Details of the iterative procedure at iteration 1 and at convergence, or MAXITER if convergence was not attained. For each iteration, aside from the first iteration, these details include the estimates, their standard errors, *z*-values and exceedance probabilities.
4. The covariance and correlation matrices of the random parameters on the different levels of the model.
5. The value of –2 log likelihood (deviance) at each iteration and number of parameters estimated.
6. The CPU time for completion of the iterative procedure and writing of required results to the output file.

Empirical Bayes estimates (BAYES):

If OUTPUT = BAYES is specified, (1) to (6) are written to the output file. One, or in the case of a 3-level model, two additional output files are also created.

The empirical Bayes estimates on levels 2 and 3 of the model are calculated and, along with their variance and relevant variable codes, are written to the files *.**ba2** and *.**ba3**, where these file names refer to the second and third level of the hierarchy respectively. The filename and path are the same as for the **.out** file.

Residuals (RESIDUAL):

If OUTPUT = RESIDUAL is specified, (1) to (6) are written to the output file. An additional file, *.**res**, is created, and contains the residuals as at convergence. The following information is provided:

- the residuals $(y_{ijk} - \mathbf{x}'_{(f)ijk}\hat{\boldsymbol{\beta}})$,
- the expected value $(\tilde{y}_{ijk})$, and
- observed value (y_{ijk}) for each observation in the raw data set.

All output (ALL):

All of the above files are created.

Example

By using the OPTIONS command shown below (without keywords), the convergence criterion will be 0.001, a maximum number of 10 iterations will be carried out and partial output will be written to the output file *.**out**. OLS estimates will be calculated during the first iteration.

```
OPTIONS;
```

Use of the command shown below will exclude the calculation of the OLS estimates during the first iteration. The convergence criterion is 0.0001 and the maximum number of iterations is 20. Lack of convergence will be noted in the default output file. All output files (standard output, Empirical Bayes estimate and residual files) will be created based on the solution obtained at termination.

```
OPTIONS OLS=NONE MAXITER=20 OUTPUT=ALL CONVERGE=0.0001;
```

SUMMARY keyword

The SUMMARY keyword is used to suppress the printout of the data summary table. This keyword is controlled from the **Title and Options** dialog box. In the example of a **Title and Options** dialog box (see above), the **No data summary** check box is not checked, indicating that the SUMMARY keyword is not used.

Syntax

```
SUMMARY = Yes/No
```

Default

Yes: the summary table containing sample sizes of units within the various levels of the hierarchy is printed.

Example

The OPTIONS command below request use of the default values for the OLS, MAXITER, and CONVERGE keywords, along with suppression of the printing of the summary table, as indicated by the absence of the SUMMARY keyword.

```
OPTIONS OLS=YES MAXITER=10 CONVERGE=0.001;
```

2.3.13 RANDOMn command

The RANDOMn command is used to identify those variables whose coefficients are allowed to vary randomly over a given level of the hierarchy. One RANDOM command is allowed for each level of the hierarchy. When the syntax file is created through the interface, the RANDOM command(s) are automatically generated. Variables listed, except for the variable intcept (intercept), must be included in the data spreadsheet (*.**lsf** file). The spelling and case in which they are given need to correspond to that given in the spreadsheet. By default, the intercept is automatically included as a random effect at each level of the model. To exclude the intercept term at any level, the corresponding **Intercept** check box (see the **Random Variables** dialog box below) must be unchecked. At least one RANDOMn command is **required** if a 2-level model is fitted. For a 3-level model, at least two RANDOMn commands must be included in the syntax file. For a 4-level model, at least three RANDOMn commands must be included in the syntax file. At least four RANDOMn commands must be included in the syntax file for a level-5 model.

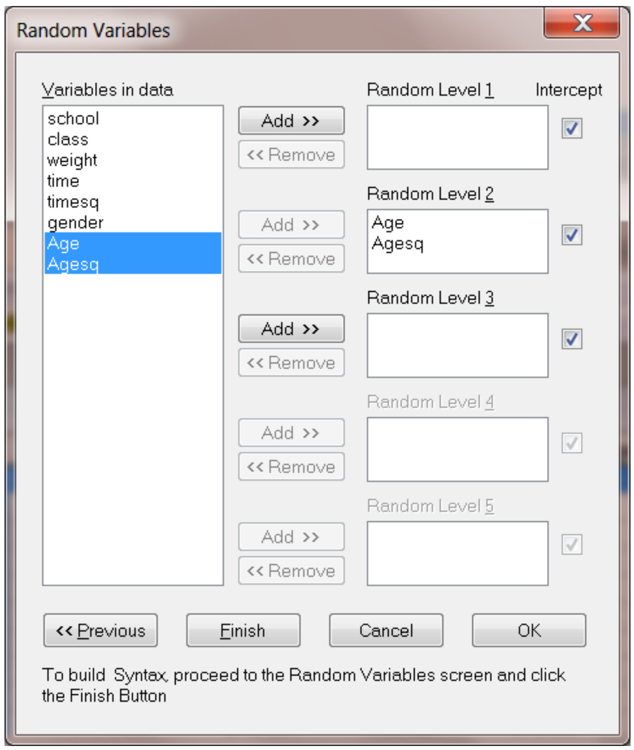
Syntax

RANDOMn = <list of variables names to be included as random effects on level n> ;

Example

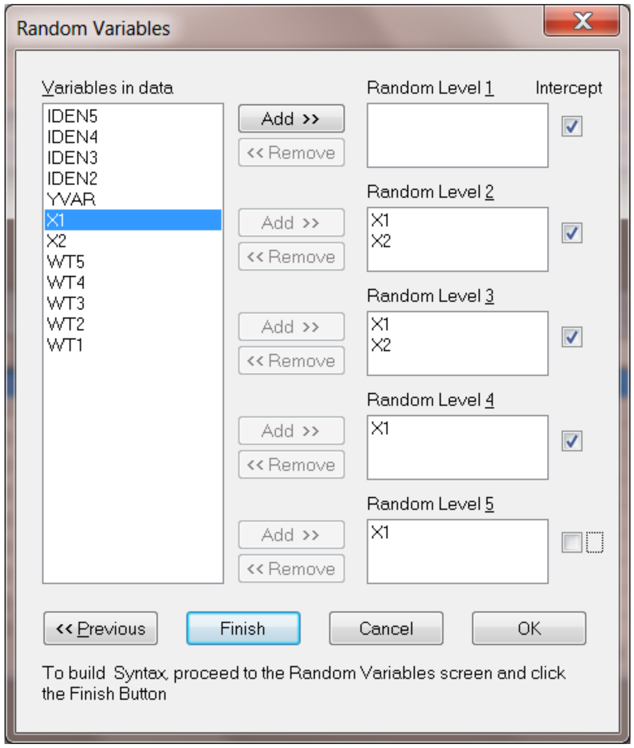
The **Random Variables** dialog box shown below corresponding to the commands

```
RANDOM1 = intcept;
RANDOM2 = intcept Age Agesq;
RANDOM3 = intcept;
```



The settings corresponding to the following commands are shown on the **Random Variables** dialog box below:

```
RANDOM5 = X1;
RANDOM4 = intcept X1;
RANDOM3 = intcept X1 X2;
RANDOM2 = intcept X1 X2;
RANDOM1 = intcept;
```



From this hypothetical example the following can be seen:

- The random variables may be listed in any order.
- Any or all of the possible predictors may be included in a RANDOM command at any level of the model.

The RANDOM1 command may be omitted in the case of a multivariate model or if a model with no random component on level-1 of the hierarchy is to be fitted. Thus the following set of commands may be used:

```
ID3 = iden3;
ID2 = iden2;
RANDOM3 = X1:X4 ;
RANDOM2= X3:X4 ;
```

It is possible to place constraints on elements of the random coefficient covariance matrices. Information on the constraints permitted and on the provision of initial values for elements of these matrices are discussed elsewhere (see Sections 2.3.3 and 2.3.4 for the COVnPAT and COVnVAL commands respectively).

2.3.14 RESPONSE command

The RESPONSE command contains information on the response variable(s) to be used in the analysis. When the syntax file is created using the interface dialogs, the RESPONSE command is automatically generated. This command is entered in the **Select Response and Fixed Variables** dialog box, which follows the **Title and Options** dialog box. Since variable names are case sensitive, spelling, etc. of the names of the response variables must be the same as those used in the data spreadsheet (*.**lsf** file). The RESPONSE command is a **required** command.

Syntax

RESPONSE = <response variable(s)>;

In the case of a multivariate model, more than one response variable may be listed in the RESPONSE command. Response variables may be entered in any order.

Example

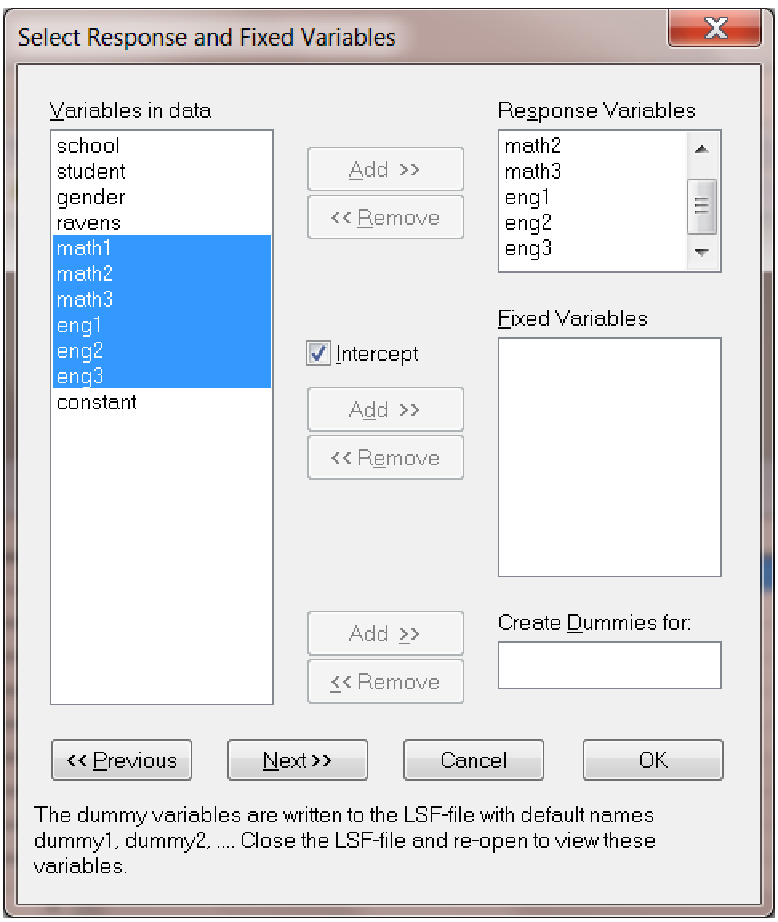
In the RESPONSE command below, the response variable is indicated as the variable Y1:

RESPONSE = Y1;

The RESPONSE command for a multivariate model, in which 6 response variables are listed,

RESPONSE = Math1 Math2 Math3 Eng1 Eng2 Eng3;

corresponds to the selection shown on the **Select Response and Fixed Variables** dialog box shown below.



2.3.15 SY command

The SY command is used to specify the LISREL System File (LSF) to be analyzed, and is automatically generated if the multilevel model specifications are built via the dialog boxes. The SY command is a **required** command.

Syntax

```
SY = <filename>;
```

where <filename> denotes the complete name (including folder name) of the LSF. The folder name may be omitted if the LSF and multilevel syntax file are in the same folder.

Example

The command shown below is used to open the LSF file **kanfer.lsf**.

```
SY = 'kanfer.lsf';
```

2.3.16 TITLE command

The TITLE command allows you to provide a description of the analysis to be performed. This command, like all commands excluding the OPTIONS command, can be placed anywhere in the syntax file. When generating syntax via the interface, the TITLE command corresponds to the first entry on the **Title and Options** dialog box. The maximum permissible length of this **optional** command is 70 characters.

Syntax

```
TITLE = <title as provided by the user>;
```

Default

No title.

Example

The TITLE command shown below corresponds to the screen shot discussed in Section 2.2.2.

```
TITLE = Level-5 model with design weights;
```

2.3.17 WEIGHTn command

The WEIGHT command is used to specify design weights for each level of the multilevel model. One WEIGHT command for each level of the hierarchy may be included in the syntax file. For a 2-level model, either or both level-1 and level-2 weights, if available, can be used. Likewise, any combination of weights can be selected for a 5-level model. Use of the command is **optional**.

Syntax

```
WEIGHTn = <name>;
```

where n denotes a positive integer, (1, 2, 3, 4 or 5), for the weight level and <name> denotes the case sensitive name of the weight variable.

Default

No weights.

Example

The WEIGHT command shown below indicates the use of the level-4 weighting variable SPWT.

```
WEIGHT4 = SPWT;
```

2.3.18 SUBPOP command

When categorical data are to be analyzed, subpopulations may be created through use of the SUBPOP command. Use of the command is **optional**.

Syntax

```
SUBPOP = <names of variables used to create subpopulations>;
```

Default

No subpopulations.

Example

Consider the two variables GENDER and AGE. If GENDER has two possible outcomes, for example 1 = Male and 2 = Female and AGE has three outcomes, 1 = less than 20 years old, 2 = 21 – 40 years old, and 3 = 41+ years old, the use of the SUBPOP command

```
SUBPOP = GENDER AGE;
```

will induce the creation of six subpopulations for the combination (GENDER; AGE), namely:

```
(GENDER = 1; AGE = 1)  (GENDER = 1; AGE = 2)  (GENDER = 1; AGE = 3)
(GENDER = 2; AGE = 1)  (GENDER = 2; AGE = 2)  (GENDER = 2; AGE = 2)
```

2.4 Examples

The analysis of data with a hierarchical structure is known in the literature as, amongst others, hierarchical modeling, random coefficient modeling, latent curve modeling, growth curve modeling, or multilevel modeling. Here we opt to use "multilevel modeling" to describe models exhibiting nested hierarchical structures.

The basic idea is that units, be it patients or measurements, are nested within units at a higher level of the hierarchy. For example, multiple blood pressure measurements may be “nested” within patients, where patients form the next, higher, level of the hierarchy. Alternatively, duration of stay within a hospital for each individual may form measurements nested within a hospital. Here the individuals are the lower-level units, nested within the hospitals that serve as the higher-level units. No matter which of these structures applies, the outcome measured at the lowest level may be described using regression coefficients at some or all the levels of the hierarchy. Variance components at different levels of the hierarchy can be included for study. This allows the researcher to evaluate the variation in outcome at various levels of the hierarchy, while inclusion of any moderating effects is optional. In addition, the dependence of repeated measurements belonging to one experimental unit in a typical growth curve analysis, for example, is taken into account with this approach. Multilevel models are also suited to the analysis of unbalanced data, and thus estimates can be obtained for units for which a very limited amount of information is available.

Multilevel models are particularly useful in the modeling of data from complex surveys. Cluster or multi-stage samples designs are frequently used for populations with an inherent hierarchical structure. Ignoring the hierarchical structure of data has serious implications. The use of alternatives such as aggregation and disaggregation of information to another level can induce an increase in collinearity among predictors and large or biased standard errors for the estimates.

Multilevel models may be fitted to complex survey data or to data from a simple random sample by using the options on the **Multilevel Models** menu. This feature is illustrated by fitting models to both real and simulated data in the sections to follow.

2.4.1 Three-level analysis of health expenditure data

2.4.1.1 The data

The data set used here is the same as that used in Section 4.2 of the *Generalized Linear Modeling Guide,* and forms part of the data library of the Medical Expenditure Panel Survey (MEPS). Collected in 1999, these data from a longitudinal national survey were used to obtain regional and national estimates of health care use and expenditure based on the health expenditures of a sample of U.S. civilian non-institutionalized participants. The survey sample design utilized stratification, clustering, multiple stages of selection, and disproportionate sampling. The sample was drawn from 143 strata, divided into 460 PSUs. Information on 23,565 participants included positive person-level weights and forms the data set used here, excluding the 1,053 participants in the original data with zero person-level weights. Data for the first 10 participants on most of the variables used in this section are shown below in the form of a LISREL spreadsheet file, named **meps2.lsf**.

meps2.lsf

	VARSTR99	VARPSU99	PERWT99F	TOTEXP99	RACE	GENDER
2	1.00	1.00	4495.33	4.20	0.00	1.00
3	1.00	1.00	7689.24	4.98	0.00	0.00
4	1.00	2.00	11069.78	5.18	0.00	1.00
5	1.00	2.00	9288.98	8.11	0.00	0.00
6	1.00	2.00	12651.61	10.09	1.00	0.00
7	1.00	2.00	6156.33	6.64	1.00	1.00
8	1.00	2.00	31359.46	8.29	0.00	0.00
9	1.00	2.00	29082.10	7.85	0.00	1.00
10	1.00	2.00	14474.96	8.47	0.00	1.00

The variables of interest are:

- VARSTR99 is the stratum identification variable (143 strata in total).
- VARPSU99 is the PSU identification variable (460 PSUs in total).
- PERWT99F represents the final sample weight, with weights ranging between 307.16 and 80061.61, correcting for both non-response and adjustments to population control totals from the Current Population Survey.
- TOTEXP99 is the natural logarithm of the total health expenditure of a respondent in 1999, ranging between 0 and 12.24 and representing actual expenditure of between $0 and $206,721.
- RACE is an ethnicity indicator, with a value of 1 indicating white respondents, and 0 denoting all other ethnic groups as well as respondents for which ethnicity is not known. This variable was recoded from the original MEPS variable RACEX.
- GENDER is a gender indicator, with a value of 0 indicating a male participant and 1 a female participant; recoded from the original MEPS variable RGENDER.
- INSCOV is an indicator of the level of insurance coverage, where 0 indicates private coverage any time during 1999, and 1 indicates public coverage or no insurance at all during 1999.
- RPOVC991 to RPOVC995 are five indicator variables, each associated with a category of the original MEPS variable RPOVC99 which was constructed by dividing family income by the applicable poverty line (selection of which depended on family size and composition), expressed as a percentage.

Income is a variable that is often transformed using its natural log. Doing so in effect causes the impact of each additional dollar to decrease as income increases. Logarithmic transformation is also useful in lessening the influence of outliers, as the natural logarithm of a variable is much less sensitive to extreme observations than is the variable itself.

The original MEPS variable RPOVC99 assumed a value of 1 for a family with "high" income level where family income was equal to or greater than 400% of the applicable poverty line, and a value of 2 for those with a "low income" level (associated with 125% to 200% of the poverty line). Families with "middle income", "near poor" and "negative or poor" levels of income relative to poverty line income were coded 3, 4 and 5 respectively. For the "middle income" category, the ratio (as percentage) of family income to poverty line was 200% to less than 400%. In the case of "near poor" families, the percentages ranged between 100% and 125%, and for "negative or poor," the family income was less than 100% of the relevant poverty line. Thus, a value of 1 on the indicator variable RPOVC991 indicates a family with income at the "high " level, while a value of 1 on the variable RPOVC995 indicates a family with "negative or poor" income level. The variables RPOVC992, RPOVC993, and RPOVC994 are associated with the categories "low income", "middle income" and "near poor" respectively.

Note that as each of the five indicator variables for categories of RPOVC99 is coded 1 if a participant responded in that category and 0 otherwise, only four of the five indicator variables can be used in a model where an intercept is included. Indicator variables of this type can easily be created by using the **Create Dummies for** option on the **Select Response and Fixed variables** dialog box as described in Section 2.2.5. Here, we opted to create them prior to analysis as illustration of that feature is not relevant to the example at hand.

2.4.1.2 The model

The multilevel model does not make provision for the specification of design related variables such as stratum or PSU. Instead, these design variables are used to define the hierarchical structure of the data. In this example, the stratum identification variable VARSTR99 is used as the level-3 identifier and the PSU identification variable VARPSU99 serves to identify level-2 units (*i.e.*, PSUs) nested within a given stratum. We thus use the design variables to define a three-level hierarchical structure, with participants as level-1 observations nested within PSUs, in turn nested within strata. While not explicitly acknowledging the survey design or offering a conventional design effect estimate to measure the difference in estimates obtained when implementing this design compared to estimates obtained under a simple random sample, a multilevel model offers the advantage of estimating the variation in total health care expenditure within and between PSUs.

A general three-level model for a response variable y depending on a set of r predictors $x_1, x_2, \ldots, x_r$ can be written in the form

$$y_{ijk} = \mathbf{x}'_{(f)ijk}\boldsymbol{\beta} + \mathbf{x}'_{(3)ijk}\mathbf{v}_i + \mathbf{x}'_{(2)ijk}\mathbf{u}_{ij} + \mathbf{x}'_{(1)ijk}\mathbf{e}_{ijk}$$

where $i = 1, 2, \ldots, N$ denotes the level-3 units, $j = 1, 2, \ldots, n_i$ the level-2 units, and $k = 1, 2, \ldots, n_{ij}$ the level-1 units. In this context, y_{ijk} represents the response of individual k, nested within level-2 unit j and level-3 unit i. The model shown here consists of a fixed and a random part. The fixed part of the model is represented by the vector product $\mathbf{x}'_{(f)ijk}\boldsymbol{\beta}$, where $\mathbf{x}'_{(f)ijk}$ is a typical row of the design matrix of the fixed part of the model with, as elements, a subset of the r predictors. The vector $\boldsymbol{\beta}$ contains the fixed, but unknown parameters to be estimated. The vector products $\mathbf{x}'_{(3)ijk}\mathbf{v}_i$, $\mathbf{x}'_{(2)ijk}\mathbf{u}_{ij}$, and $\mathbf{x}'_{(1)ijk}\mathbf{e}_{ijk}$ denote the random part of the model at levels 3, 2, and 1 respectively. For example, $\mathbf{x}'_{(3)ijk}$ represents a typical row of the design matrix of the random part at level-3, and $\mathbf{v}_i$ the vector of random level-3 coefficients to be estimated. The products $\mathbf{x}'_{(2)ijk}\mathbf{u}_{ij}$ and $\mathbf{x}'_{(1)ijk}\mathbf{e}_{ijk}$ serve the same purpose at levels 2 and 1 respectively. It is assumed that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are independently and identically distributed (i.i.d.) with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(3)}$. Similarly, $\mathbf{u}_{i1}, \mathbf{u}_{i2}, \ldots, \mathbf{u}_{in_i}$ are assumed i.i.d., with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(2)}$, and $\mathbf{e}_{ij1}, \mathbf{e}_{ij2}, \ldots, \mathbf{e}_{ijn_{ij}}$ are assumed i.i.d., with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(1)}$.

Within this hierarchical framework, the model fitted to the data uses the participant's gender, ethnicity, type of health insurance cover, and measure of income relative to poverty level to predict the total expenditure on health care in 1999, the latter transformed to the natural logarithm of the actual expenses incurred.

$$\begin{aligned} \text{TOTEXP99}_{ijk} = \beta_0 + \beta_1 * \text{GENDER}_{ijk} + \beta_2 * \text{RACE}_{ijk} + \beta_3 * \text{INSCOV}_{ijk} + \\ \beta_4 * \text{RPOVC991}_{ijk} + \beta_5 * \text{RPOVC992}_{ijk} + \beta_6 * \text{RPOVC993}_{ijk} + \\ \beta_7 * \text{RPOVC994}_{ijk} + v_{i0} + u_{ij0} + e_{ijk} \end{aligned}$$

where β_0 denotes the average expected total expenditure on health care in 1999, and $\beta_1, \beta_2, \ldots, \beta_7$ indicate the estimated coefficients associated with the fixed part of the model which contains the predictor variables GENDER, RACE, INSCOV and the indicator variables for categories of income relative to the poverty level. The random part of the model is represented by v_{i0}, u_{ij0} and e_{ijk}, which denote the variation in average total health related expenditure over strata, between PSUs (or, in other words, over PSUs nested within strata) and between participants at the lowest level of the hierarchy.

2.4.1.3 Multilevel analysis with sampling weights

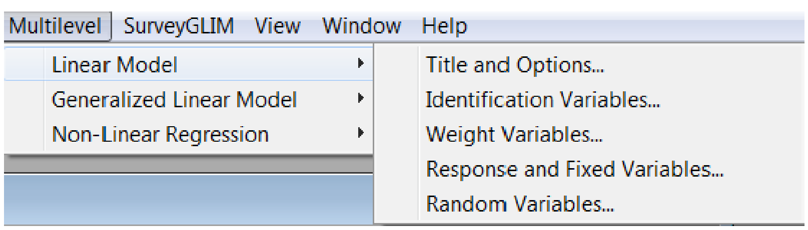
1.2 Setting up the analysis

The model is fitted to the data in **meps2.lsf** by using the sequence of four dialog boxes accessed via the **Multilevel**, **Linear Model** option from the main menu bar in LISREL. Note that options such as **Multilevel** and **SurveyGLIM** are only available on the main menu bar when a ***.lsf** file is open.

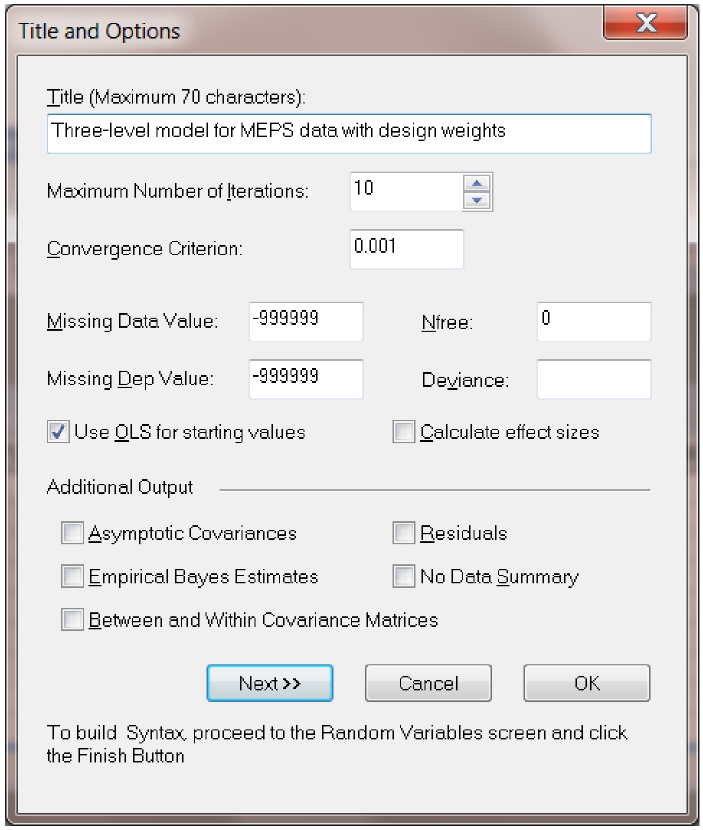
The first step is to open the LSF shown above, which is accomplished as follows:

- Use the **File**, **Open** option to activate the display of an **Open** dialog box.
- Set the **Files of type** drop-down list box to **LISREL System Data (*.lsf)** and browse for the file **meps2.lsf**.
- Select the file and click the **Open** button to return to the main LISREL window, where the contents of the LSF are displayed.

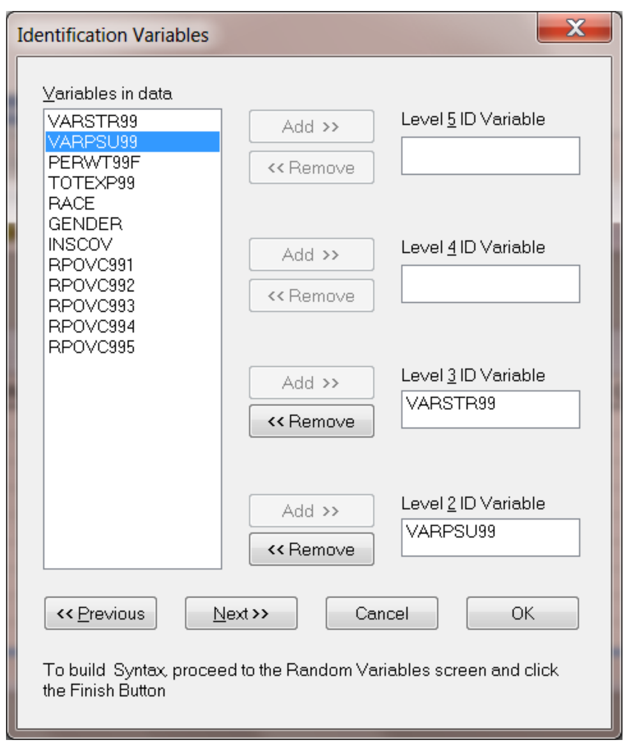
The next step is to describe the model to be fitted using the multilevel module in LISREL. From the main menu bar, select the **Multilevel** option. Here we limit our discussion to linear models, and thus the **Linear Model** option will be used throughout.

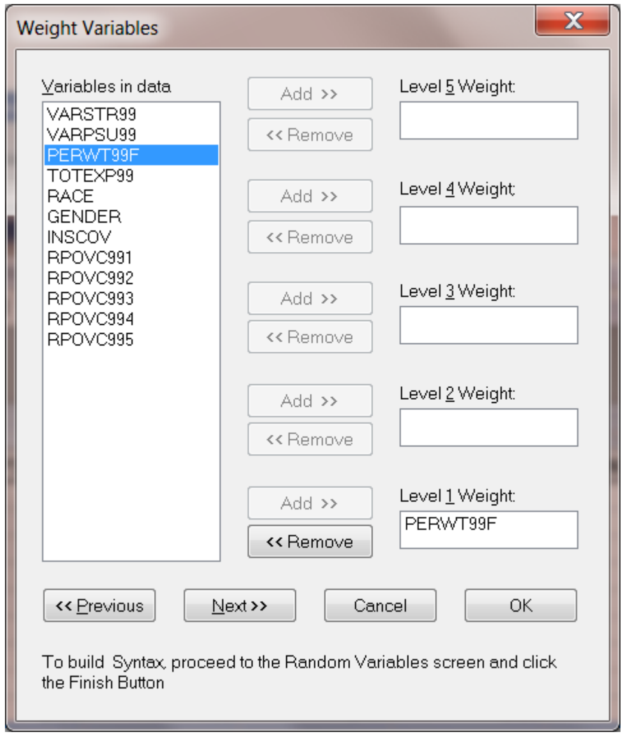


The first of the four options on the pop-up menu provide access to the **Title and Options** dialog box discussed in Section 2.2.2. Start by providing a title for the analysis in the **Title** field. In this example, default settings for all other options associated with this dialog box are used. Click the **Next** button to go to the **Identification Variables** dialog box.



On the **Identification Variables** dialog box, enter the variables defining the hierarchical structure as ID variables (see Section 2.3.9 for detailed information on the ID command). As mentioned before, the stratum identification variable is used to indicate the level-3 units in the hierarchical structure, and the PSU identification variable serves a similar purpose at level-2. Select the variables VARSTR99 and VARPSU99 as Level-3 ID variable and Level-2 ID variable respectively by clicking on the variable names in the **Variables in data** field at the left of the dialog box. Add them to the ID variable fields by clicking the appropriate **Add** button for each.



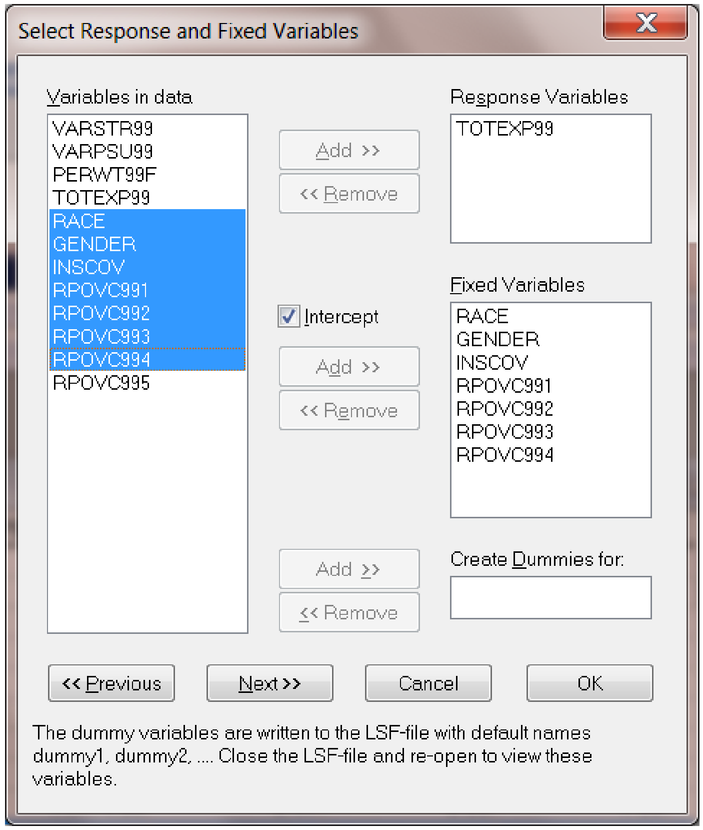


The **Weight Variables** dialog box is used to provide information on weight variables, if any. In our case, only one weight, denoted by the variable PERWT99F, is available. Select this variable from the **Variables in data** field, and add it to the **Level-1 weight** field as shown below. As all available information is now entered on this dialog box, click the **Next** button to proceed to the **Select Response and Fixed Variables** dialog box.

The **Select Response and Fixed Variables** dialog box, described in detail in Section 2.2.5, is used to identify the outcome variable and predictor variables, if any. Select and add the outcome variable TOTEXP99 to the **Response Variables** field in the same way as described for the previous dialog box. Next, select the variables starting from RACE to RPOVC994 by dragging the mouse over them and click the **Add** button next to the **Fixed Variables** field to include these variables as predictors in the model. This completes the specification of the response and fixed variables.

Before moving to the next dialog box, two other options available on this dialog box are worth noting.

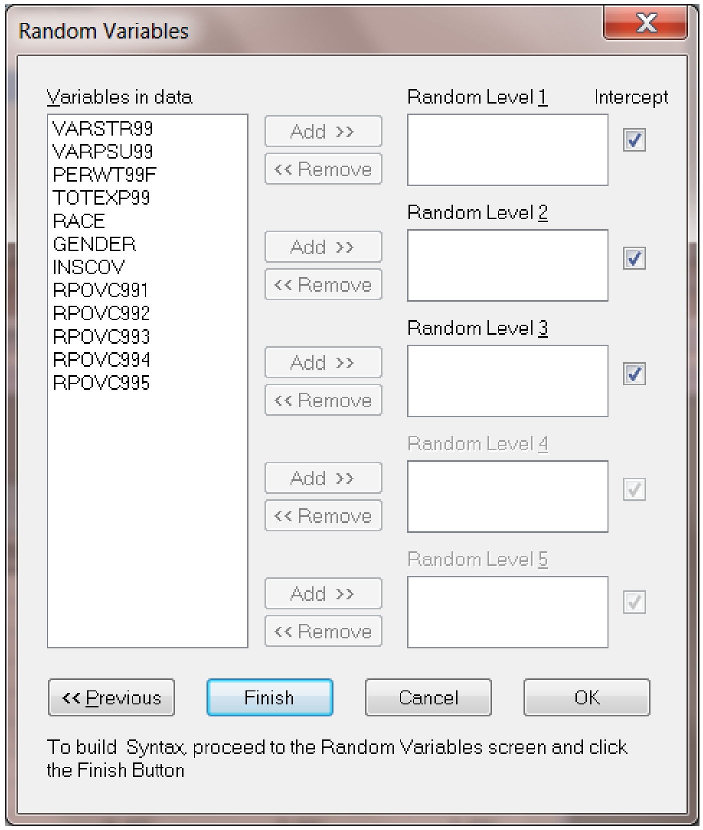
As previously discussed, the indicator variable associated with the highest level of income relative to the poverty line income is not selected for inclusion as the model fitted to the data has an intercept. Because of the intercept term, inclusion of all five indicator variables would lead to a design matrix of less than full rank and is bound to cause problems during the iterative procedure. An alternative approach would be to use all five indicator variables in a model without an intercept term. This can be achieved by deselecting the intercept term by unchecking the box next to Intercept.



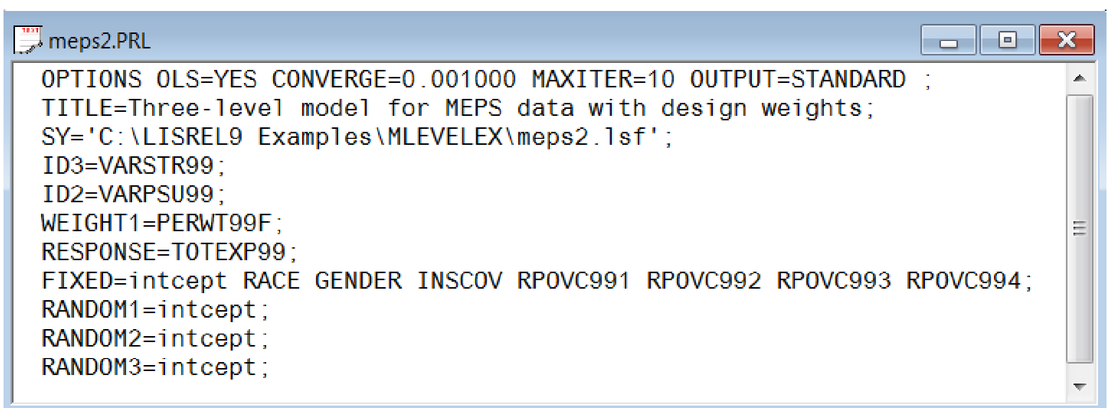
The **Create Dummies for** option available on the **Select Response and Fixed** dialog box can be used to create indicator variables for the categories of a categorical variable such as RPOVC99. In fact, the indicator variables RPOVC991 to RPOVC995 were created in precisely this way for inclusion in the present analysis, and simply renamed from their default names of DUMMY1 to DUMMY5 using the **Define Variable** option from the **Data** menu accessed from the main menu bar in LISREL.

That said; proceed to the **Random Variables** dialog box by clicking the **Next** button.

The **Random Variables** dialog box shown below displays the default settings associated with this dialog box. In the current model, only the intercept coefficients are allowed to vary randomly at the various levels of the hierarchy. As this corresponds to the default settings shown on the dialog box, click the **Finish** button to generate the syntax for the model.



The syntax shown below corresponds to the information entered via the dialog boxes above. Run the model by clicking the **Run Prelis** icon on the main menu bar.



```
OPTIONS OLS=YES CONVERGE=0.001000 MAXITER=10 OUTPUT=STANDARD ;
TITLE=Three-level model for MEPS data with design weights;
SY='C:\LISREL9 Examples\MLEVELEX\meps2.lsf';
ID3=VARSTR99;
ID2=VARPSU99;
WEIGHT1=PERWT99F;
RESPONSE=TOTEXP99;
FIXED=intcept RACE GENDER INSCOV RPOVC991 RPOVC992 RPOVC993 RPOVC994;
RANDOM1=intcept;
RANDOM2=intcept;
RANDOM3=intcept;
```

Discussion of results – Multilevel model with sampling weights

Portions of the output file **meps2.out** are shown below.

meps2.OUT

```
                              +--------------+
                              | DATA SUMMARY |
                              +--------------+

NUMBER OF LEVEL 3 UNITS :       143
NUMBER OF LEVEL 2 UNITS :       460
NUMBER OF LEVEL 1 UNITS :     23564

ID3  :      1      2      3      4      5      6      7      8
N2   :      2      2      2      2      2      3      2      2
N1   :     29     85     55     86    159     48     48     78

ID3  :      9     10     11     12     13     14     15     16
N2   :      2      2      2     11      2      2      2      3
N1   :    114     23     68    408     68     62    168     59

ID3  :     17     18     19     20     21     22     23     24
N2   :      2      5      2      2      2      2      2      2
N1   :    364    215     69     44     40     59    171    151

ID3  :     25     26     27     28     29     30     31     32
N2   :      3      2      2      2      2      2     10      2
N1   :    225     18     86     25     58     26    417     26
```

In the first section of the output file a description of the hierarchical structure is provided in the **Data Summary** section. A total of 143 strata, 460 PSUs and information from 23,564 individual participants were included at levels 3, 2 and 1 of the multilevel model. This corresponds to the survey design described earlier. In addition, a summary of the number of PSUs and participants nested within each stratum is provided. For stratum number 1 (ID3: 1), data are available from only 29 participants nested within 2 primary sampling units (N2: 2). By contrast, for stratum number 12 (ID3: 12), data are available from 408 participants (N1: 408) nested within 11 primary sampling units (N2: 11).

meps2.OUT

```
Three-level model for MEPS data with design weights

ITERATION NUMBER     6

                        +-----------------------+
                        | FIXED PART OF MODEL   |
                        +-----------------------+

------------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.       Z-VALUE       PR > |Z|
------------------------------------------------------------------------------
 intcept                   4.39189       0.11481       38.25193       0.00000
 RACE                      0.94272       0.08693       10.84511       0.00000
 GENDER                    0.91059       0.03677       24.76122       0.00000
 INSCOV                   -0.65101       0.07624       -8.53904       0.00000
 RPOVC991                  0.35725       0.11520        3.10110       0.00193
 RPOVC992                 -0.13848       0.10523       -1.31593       0.18820
 RPOVC993                  0.07022       0.11782        0.59595       0.55121
 RPOVC994                 -0.32929       0.14183       -2.32166       0.02025

                        +-----------------------+
                        |   -2 LOG-LIKELIHOOD   |
                        +-----------------------+

 DEVIANCE= -2*LOG(LIKELIHOOD) =    119135.095120823
 NUMBER OF FREE PARAMETERS =                11

 CHI-SQUARE SCALE FACTOR =      0.25580
```

The output describing the estimated **fixed effects** after convergence is shown next. The estimates are shown in the column with heading BETA-HAT and correspond to the coefficients $\beta_0, \beta_2, \ldots, \beta_7$ in the model specification. From the z-values and associated exceedance probabilities, we see that the coefficients associated with gender, ethnicity and insurance coverage type were all highly significant. Recall that a value of 1 for the ethnicity indicator variable RACE indicated that a participant was white, with a value of 0 assigned to participants from all other ethnic groups. The positive estimated coefficient for this variable indicates an increase of 0.94298 units in the logarithm of total health expenditure, holding all other predictors constant. Similarly, female participants (coded "1" on the gender indicator GENDER), are expected to have a total health expenditure 0.91057 higher than male participants if all other variables are held constant. In contrast, participants with public coverage or no coverage have a lower expected total expenditure, as indicated by the negative estimated coefficient -0.65109.

Turning to the indicator variables associated with income relative to the poverty line income, it can be seen that only two of the indicator variables, RPOVC991 and RPOVC994, have estimated coefficients that are significantly different from zero at a 5% level of significance. In the case of families with a "high" income, the estimate of 0.35750 for RPOVC991 indicates an expected increase in expenditure, while for "near poor" families, the estimate of -0.32939 indicates an expected decrease in expenditure, holding all other variables constant.

Estimated outcomes for different groups

To evaluate the expected effect of the measure of a family's income to the corresponding poverty line income, suppose that the variables RACE, GENDER, and INSCOV are held at zero, as would be the case for a nonwhite male participant with private insurance coverage. If such a participant originates from a family with "high" income, the logarithm of total health expenditure is expected to be

$$\begin{aligned}&\beta_0 + \beta_4(\text{RPOVC991}) + \beta_5(\text{RPOVC992}) + \beta_6(\text{RPOVC993}) + \beta_7(\text{RPOVC994}) \\ &= \beta_0 + \beta_4 \\ &= 4.39123 + 0.35750 \\ &= 4.74873\end{aligned}$$

which translates to a projected total expenditure of $e^{4.74873} = \$115.437$. In contrast, for a participant with similar demographic background and coverage from a "near poor" family, we obtain a projected total expenditure of

$$\begin{aligned}&e^{\beta_0+\beta_7} \\ &= e^{4.39123-0.32929} \\ &= \$58.086\end{aligned}$$

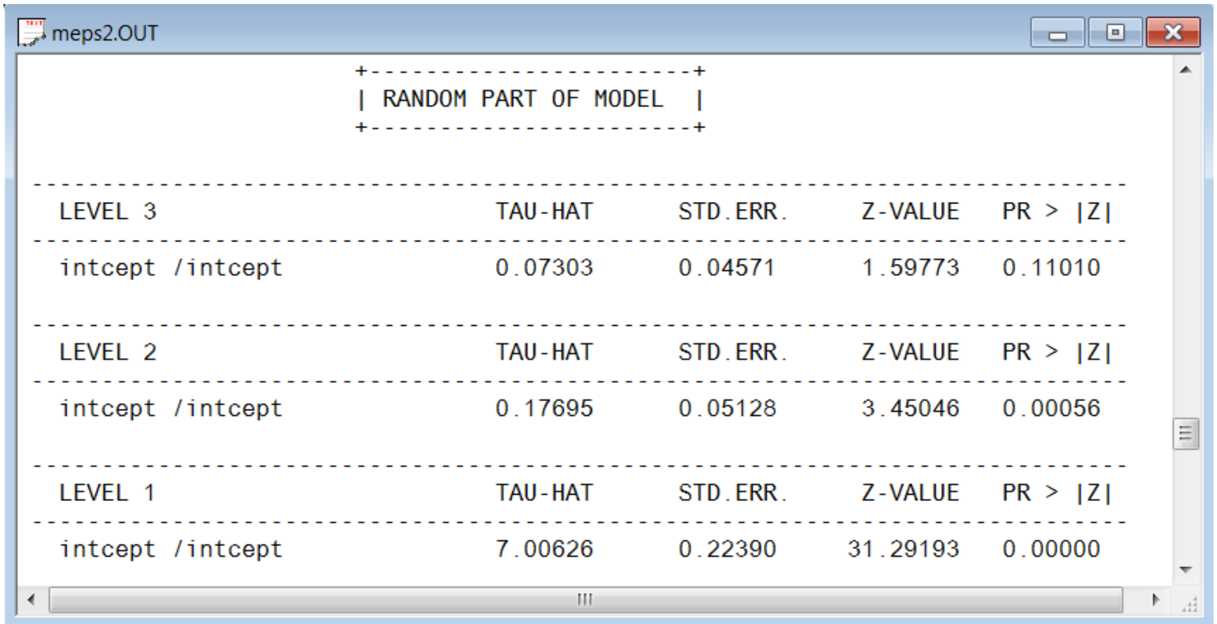
The predicted total expenditure (as natural logarithm) for similar participants from "low", "middle" or "negative or poor" families are similarly obtained by calculating $e^{\beta_0+\beta_5}$, $e^{\beta_0+\beta_6}$ and e^{β_0} respectively.

Table 2.1: Predicted total health expenditure for various subgroups

Respondents with high family income (RPOVC991 = 1)	**Male** (GENDER = 0)		**Female** (GENDER = 1)	
	Insurance coverage:		**Insurance coverage:**	
	Private (INSCOV=0)	**Public/none** (INSCOV = 1)	**Private** (INSCOV=0)	**Public/none** (INSCOV = 1)
Nonwhite (RACE = 0)	\$115	\$60	\$287	\$150
White (RACE = 1)	\$296	\$155	\$737	\$384
Respondents with near poor income (RPOVC994 = 1)				
Nonwhite (RACE = 0)	\$58	\$30	\$145	\$75
White (RACE = 1)	\$149	\$78	\$370	\$193

In Table 2.1, the predicted total health expenditure is given for respondents with high or near poor family income, for each of the subpopulations formed by gender, ethnicity and insurance coverage. For purposes of the comparison, results are expressed in U.S. dollars, rather than in the natural logarithmic units of the outcome variable TOTEXP99. Respondents from families with high income consistently outspend their near poor counterparts by approximately 100%, regardless of gender, ethnicity or level of insurance coverage. In families with high income, female respondents spent more in 1999 than their male counterparts, regardless of ethnicity. This is generally also true for near poor respondents. It is also apparent that the total health expenditure in 1999 was higher for respondents with private insurance than for respondents with public or no coverage, and that white respondents spent more than respondents from other ethnic groups, regardless of gender or the level of family income. From exploratory analyses, we know that the outcome variable TOTEXP99 is highly skewed, with median

1999 expenditure of $ 377.41. When this is taken in account, we can conclude that, generally speaking, white females spent more on health in 1999 than 50% of all respondents in the sample.



LEVEL 3	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	0.07303	0.04571	1.59773	0.11010

LEVEL 2	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	0.17695	0.05128	3.45046	0.00056

LEVEL 1	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	7.00626	0.22390	31.29193	0.00000

The output for the **random part** of the model follows, and is shown in the image above. There is significant variation in the average estimated total health expenditure at all levels, with the most variation over the participants (level-1), and the least variation over strata (level-3).

An estimate of the level-2 cluster effect, for example, is obtained as

$$\frac{0.17706}{0.07305+0.17706+7.00628}\times 100\% = 2.41\%$$

indicating that only 2.41% of the total variance explained is at level-2 of the model.

2.4.1.4 Multilevel analysis without sampling weights

To evaluate the effect on the estimated coefficients if the sampling weights are ignored for data known to come from a disproportionally sampled survey, we fit the same model without a WEIGHT command.

Setting up the analysis

To fit the unweighted model, the syntax file from the previous analysis can be edited by simply deleting the WEIGHT1 command from the syntax file. Alternatively, the **Level-1 Weight** field on the **Weight Variables** dialog box can be cleared by clicking on this field and then clicking the **Remove** button next to this field.

Clicking **Next** on this and the next two dialog boxes, followed by clicking the **Finish** button on the **Random Variables** dialog box will generate a revised syntax file.

Discussion of results – Multilevel model without sampling weights

After running the analysis by clicking the **Run Prelis** icon on the main menu bar, the following output is obtained for the fixed and random parts of the unweighted model.

meps2.OUT

```
 ITERATION NUMBER     4

                           +----------------------+
                           | FIXED PART OF MODEL  |
                           +----------------------+

 ------------------------------------------------------------------------------
  COEFFICIENTS             BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
 ------------------------------------------------------------------------------
  intcept                   4.45841      0.08350      53.39310       0.00000
  RACE                      0.68364      0.04971      13.75292       0.00000
  GENDER                    0.93063      0.03581      25.98628       0.00000
  INSCOV                   -0.61785      0.04571     -13.51696       0.00000
  RPOVC991                  0.49302      0.06489       7.59791       0.00000
  RPOVC992                 -0.15390      0.06677      -2.30502       0.02117
  RPOVC993                  0.10053      0.06193       1.62342       0.10450
  RPOVC994                 -0.34592      0.08978      -3.85321       0.00012

                           +----------------------+
                           |  -2 LOG-LIKELIHOOD   |
                           +----------------------+
  DEVIANCE= -2*LOG(LIKELIHOOD) =    115546.006821173
  NUMBER OF FREE PARAMETERS =              11
```

meps2.OUT

```
                           +---------------------+
                           | RANDOM PART OF MODEL |
                           +---------------------+

 ------------------------------------------------------------------------
  LEVEL 3                          TAU-HAT      STD.ERR.     Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------
  intcept /intcept                 0.14587      0.03648      3.99824    0.00006

 ------------------------------------------------------------------------
  LEVEL 2                          TAU-HAT      STD.ERR.     Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------
  intcept /intcept                 0.17442      0.02767      6.30265    0.00000

 ------------------------------------------------------------------------
  LEVEL 1                          TAU-HAT      STD.ERR.     Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------
  intcept /intcept                 7.46282      0.06938    107.56310    0.00000
```

In Table 2.2, the predicted total health expenditure is given for respondents with high or near poor family income, for each of the subpopulations formed by gender, ethnicity and insurance coverage. When compared to Table 2.1, where similar results were given for the weighted analysis, no difference in the overall pattern of expenditure is detected. Note, however, that the predicted expenditure for Nonwhite respondents (RACE = 0) are consistently higher in Table 2.2 than was the case in Table 2.1. For white respondents, the unweighted results shown in Table 2.2 are consistently lower than the corresponding results in Table 2.1. If sample weights are not used in the analysis, it may lead to a consistent, although small, overestimation of the health expenditure of nonwhite respondents, and to an underestimation of the health expenditures of their white counterparts.

Table 2.2: Predicted total health expenditure for various subgroups

Respondents with high family income (RPOVC991 = 1)	**Male** (GENDER = 0)		**Female** (GENDER = 1)	
	Insurance coverage:		**Insurance coverage:**	
	Private (INSCOV=0)	**Public/none** (INSCOV = 1)	**Private** (INSCOV=0)	**Public/none** (INSCOV = 1)
Nonwhite (RACE = 0)	\$141	\$76	\$359	\$193
White (RACE = 1)	\$280	\$151	\$710	\$383
Respondents with near poor income (RPOVC994 = 1)				
Nonwhite (RACE = 0)	\$61	\$33	\$154	\$83
White (RACE = 1)	\$120	\$65	\$304	\$164

Results for the two models (weighted and unweighted) are summarized in Table 2.3. While results for the models fitted in this case are not dramatically different, we observe that while some coefficients are larger for the unweighted model (for example, the estimates for intcept, GENDER, INSCOV, and most markedly for RPOVC991), coefficients for RPOVC992 and RACE are larger for the weighted model. The largest difference observed is in the case of ethnicity, where an estimated increase of 0.94 in expenditure is associated with a white respondent under the weighted model, compared to only 0.68 for a white respondent in the unweighted model (holding all other variables constant). As this translates to a difference of $e^{0.26} = 1,296$ in total health expenditure for 1999, this difference is more important than it seems at first glance. In addition, the models are sufficiently different in that coefficients statistically significant in one model are no longer significant in the other, as illustrated by the estimated coefficients for the indicator variable RPOVC992. In the weighted model, low income respondents are not expected to have a significantly different expected total expenditure, while the estimated coefficient under the unweighted model indicates a statistically significant decrease of -0.15 units in the total expected expenditure.

Table 2.3: Results of weighted and unweighted level-3 models for the MEPS data

Coefficient	Estimate (weighted)	Estimate (unweighted)
intcept	4.39123	4.45841
RACE	0.94298	0.68364
GENDER	0.91057	0.93063
INSCOV	-0.65109	-0.61785
RPOVC991	0.35750	0.49302
RPOVC992	-0.13832*	-0.15390
RPOVC993	0.07036*	0.10053*
RPOVC994	-0.32929	-0.34592
Level-1 variance	7.00628	7.46282
Level-2 variance	0.17706	0.17442
Level-3 variance	0.07305	0.14587

* Not significant at a 5% level of significance.

Comparison with SurveyGLIM model

A similar model was fitted to the data using the SurveyGLIM module (see Section 2.1 of the *Generalized Linear Modeling Guide)* and a Normal-Identity model. Results are summarized in Table 2.4. In general, results obtained for the two models are similar.

Table 2.4: Results of weighted multilevel and SurveyGLIM models for the MEPS data

Coefficient	Multilevel model	SurveyGLIM model
intcept	4.39123	4.2771
RACE	0.94298	0.9393
GENDER	0.91057	0.9204
INSCOV	-0.65109	-0.6952
RPOVC991	0.35750	0.4319
RPOVC992	-0.13832*	-0.1415*
RPOVC993	0.07036*	0.1186*
RPOVC994	-0.32929	-0.3433

* Not significant at a 5% level of significance.

We conclude that, where weight variables are available for survey data, these should be included in the model as neglecting to do so can have a definite impact on the estimated coefficients. In the current example, results for the two models were not dramatically different, but comparison of predicted expenditure indicated the risk of consistently over- or underestimating the total health expenditure for groups with different levels of family poverty. From the results it seems reasonable to assume that it included a component to adjust for the over/undersampling of ethnic and gender groups, a procedure commonly used in survey design to ensure representativeness. This is in agreement with the fact that, according to the MEPS HC-054: 1999 report, Hispanic and black households were oversampled at rates of approximately 2 and 1.5 times the rate of remaining households.

2.4.2 Three-level analysis of simulated data

Unlike real data sets, simulated data sets have the advantage that the true population parameters are known. Consequently, it is possible to evaluate how closely a particular model approaches these values.

2.4.2.1 The data

A linear growth curve model with two dummy-coded covariates (Lang1 and Lang2) is fitted to a simulated dataset **surveyhlm.lsf** in the **Multilevel Examples** folder. It is assumed that the level-3 units are institutions. Within each of 100 institutions, 10 patients are selected on the basis of their initial achievement in a test of short term memory (Score1) and measurements were repeated over six time intervals for five patients from each institution and over 4 time intervals for the remaining 5. In the table below, (Weight3) shows the level-3 weight calculations based on standardized initial scores. See Section 2.5.3 for additional information on the weight calculations.

Interval	Lower	Upper	% Expected	% Selected	Weight3
1	-Inf	-1.00	15.87	10.00	1.587
2	-1.00	-0.70	8.33	10.00	0.833
3	-0.70	-0.20	17.88	10.00	1.788
4	-0.20	0.00	7.93	10.00	0.793
5	0.00	0.30	11.79	10.00	1.179
6	0.30	1.00	22.34	10.00	2.234
7	1.00	1.30	6.19	10.00	0.619
8	1.30	1.80	6.09	10.00	0.609
9	1.80	2.30	2.52	10.00	0.252
10	2.30	Inf	1.07	10.00	0.107

Ten patients were selected from each institution as follows:

- Four from ethnic group 1 with Weight2 = 7.0/4.0
- Three from ethnic group 2 with Weight2 = 2.0/3.0
- Three from ethnic group 3 with Weight2 = 1.0/3.0

The first 10 records of the dataset in **surveyhlm.lsf** are shown below.

surveyhlm.lsf

	Institut	Patient	Score	Time	Lang1	Lang2	WT3	WT2
1	1.00	1.00	-1.84	0.00	0.00	0.00	1.59	1.75
2	1.00	1.00	-0.89	1.00	0.00	0.00	1.59	1.75
3	1.00	1.00	-1.21	2.00	0.00	0.00	1.59	1.75
4	1.00	1.00	-3.24	3.00	0.00	0.00	1.59	1.75
5	1.00	1.00	-1.16	4.00	0.00	0.00	1.59	1.75
6	1.00	1.00	-1.30	5.00	0.00	0.00	1.59	1.75
7	1.00	2.00	-0.07	0.00	0.00	0.00	1.59	1.75
8	1.00	2.00	2.99	1.00	0.00	0.00	1.59	1.75
9	1.00	2.00	0.92	2.00	0.00	0.00	1.59	1.75
10	1.00	2.00	3.63	3.00	0.00	0.00	1.59	1.75

Note that the data were simulated in such a way that odd-numbered patients have six score measurements at time points 0, 1, 2, 3, 4, 5. The even-numbered patients have only four score measurements.

2.4.2.2 The model

The three-level model used here is similar to that described in the previous section. Within that hierarchical framework, 500 data sets, **surveyhlm.lsf** being the first, were simulated (see Section 2.5.3) according to the following hypothetical model

$$Score_{ijk} = \beta_0 + \beta_1 * Time + \gamma_1 * Lang1 + \gamma_2 * Lang2$$
$$+ v_{i0} + Time * v_{i1} + u_{ij0} + Time * u_{ij1} + e_{ijk}$$

where i denotes institution i, (i = 1, 2, …, 100), ij patient j (j =1, 2, …, 10) in institution i and ijk the k-th measurement (k = 1, 2, …, 6) on patient j in institution i. The outcome variable Score denotes a patient's measurement on some test of interest, Time the time of measurement, and Lang1 and Lang2 are indicator variables indicating a patient's first or home language as being English or another language.

In this model, β_0 denotes the average expected score, while β_1 indicates the estimated coefficients associated with the time of measurement as represented by the fixed effect Time. The fixed part of the model also includes the predictor variables Lang1 and Lang2. The random part of the model is represented by v_{i0}, u_{ij0} and e_{ijk}, which denote the variation in score over institutions, between patients (or, in other words, over patients nested within institutions) and between measurements at the lowest level of the hierarchy.

The data were simulated under the assumption that

$$\begin{pmatrix}\beta_0\\ \beta_1\end{pmatrix}=\begin{pmatrix}1.0\\ 0.5\end{pmatrix},\quad \begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix}=\begin{pmatrix}0.5\\ -1.0\end{pmatrix}$$

$$\phi_2 = Cov(u_{ij0}, u_{ij1}) = \begin{pmatrix}1 & \\ 0.3 & 0.2\end{pmatrix}$$

$$\phi_3 = Cov(v_{i0}, v_{i1}) = \begin{pmatrix}1 & \\ 0.3 & 0.2\end{pmatrix}$$

and

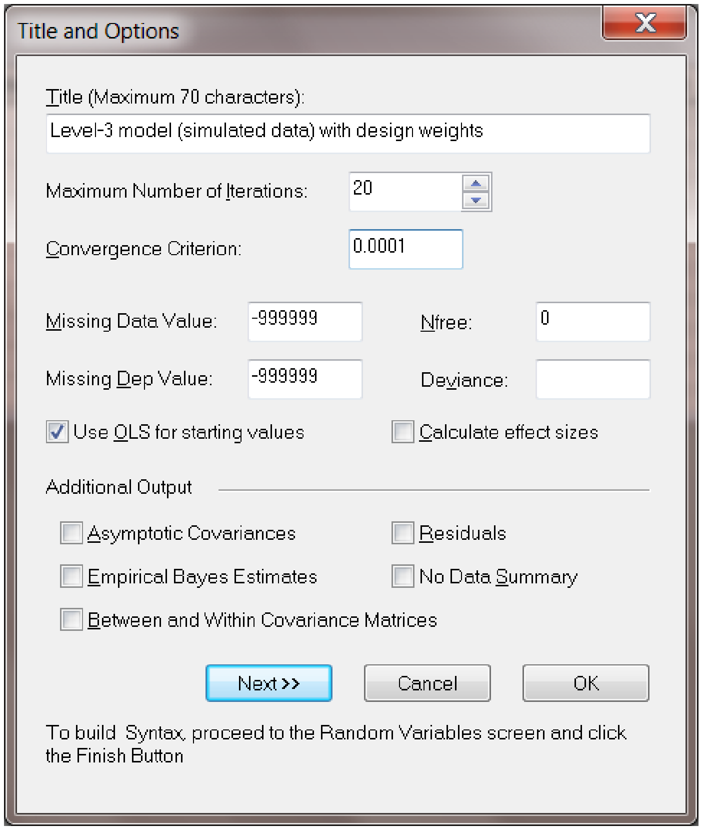
$$\sigma^2 = Var(e_{ijk}) = 1.0.$$

2.4.2.3 Fitting the model

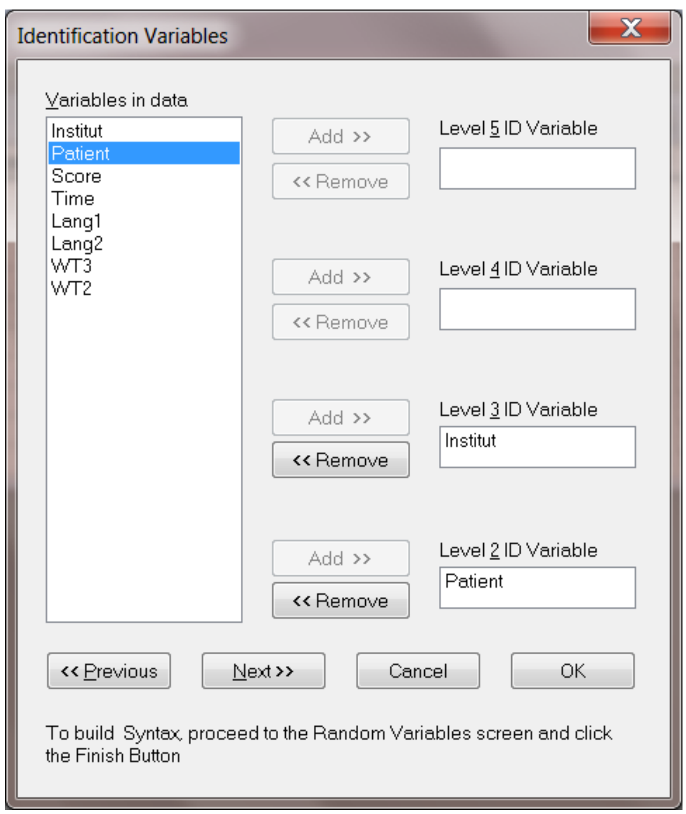
The first step is to open the LSF shown above, which is accomplished as follows:

- Use the **File**, **Open** option to activate the display of an **Open** dialog box.
- Set the **Files of type** drop-down list box to **LISREL System Data (*.lsf)** and browse for the file **surveyhlm.lsf**.
- Select the file and click the **Open** button to return to the main LISREL window, where the contents of the LSF are displayed.

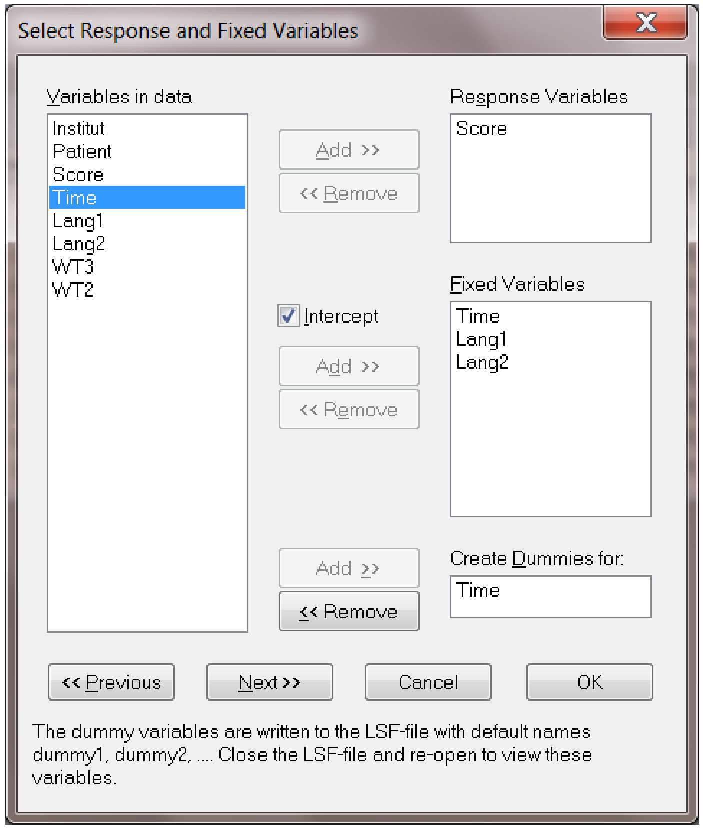
The next step is to describe the model to be fitted using the multilevel module in LISREL. To fit a growth model to the data, we proceed as follows. From the main menu bar, select **Multilevel, Linear Model, Title and Options**. Type in a title, and change the number of iterations to 20 and the convergence criterion to 0.0001 as shown below.



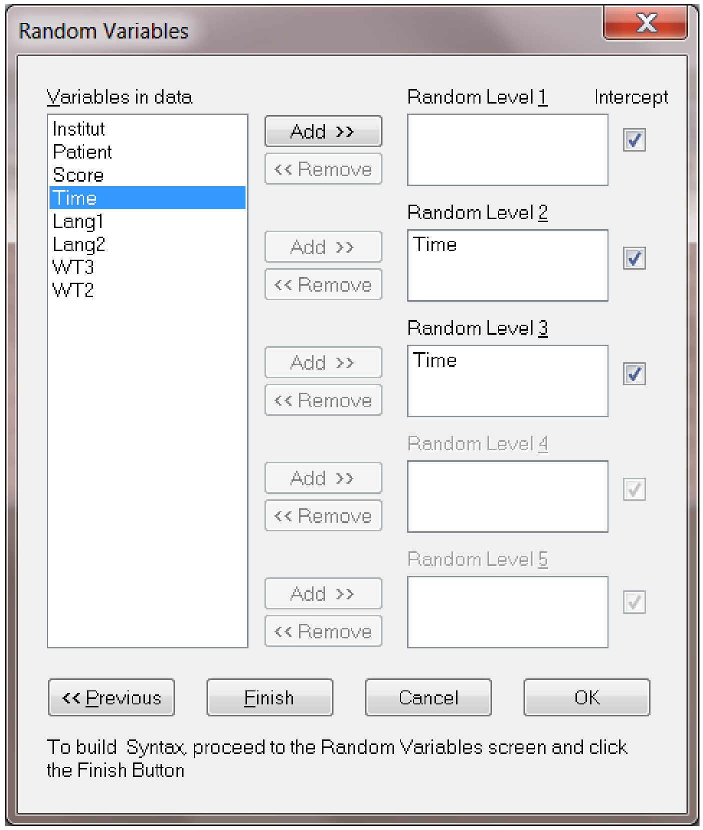
Click the **Next** button to activate the **Identification Variables** dialog box. Add the level-3 and level-2 identification variables (Institut and Patient). Use the **Weight Variables** dialog box to select the variables WT3 and WT2 as the level-3 and level-2 weights respectively. To continue to the **Response and Fixed variables** dialog box click **Next**.



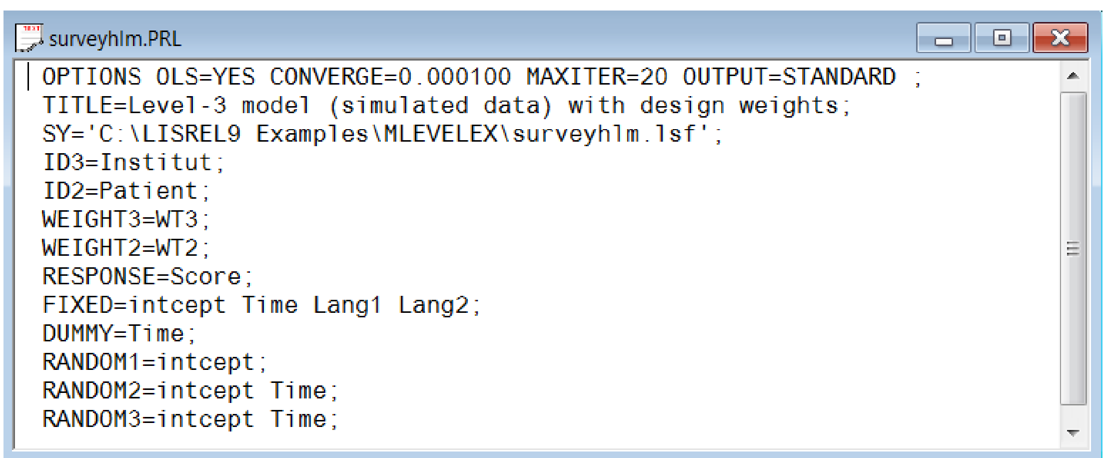
Select Score as the dependent (response) variable and Time, Lang1 and Lang2 as the fixed variables (predictors). Note that an intercept term is automatically included unless the **Intercept** check box is unchecked. For illustrative purposes, Time was added to the **Create Dummies for:** text box.



Click **Next** to go to the **Random Variables** dialog box and add Time as level-2 and level-3 random components (the variances are denoted by $Var(u_{ij1}) = \phi_{(2),22}$ and $Var(v_{i1}) = \phi_{(3),22}$ respectively. Note that by default, intercept terms are included at the different levels of the hierarchy. The level-1, level-2 and level-3 variance components for the intercept are denoted by σ_e^2, $\phi_{(2),11}$ and $\phi_{(3),11}$ respectively.



When done, click the **Finish** button to obtain the PRELIS syntax file **surveyhlm.prl**. Click the **Run PRELIS** icon button to invoke the multilevel module.




```
OPTIONS OLS=YES CONVERGE=0.000100 MAXITER=20 OUTPUT=STANDARD ;
TITLE=Level-3 model (simulated data) with design weights;
SY='C:\LISREL9 Examples\MLEVELEX\surveyhlm.lsf';
ID3=Institut;
ID2=Patient;
WEIGHT3=WT3;
WEIGHT2=WT2;
RESPONSE=Score;
FIXED=intcept Time Lang1 Lang2;
DUMMY=Time;
RANDOM1=intcept;
RANDOM2=intcept Time;
RANDOM3=intcept Time;
```

2.4.2.4 Discussion of results

The output for the **fixed part** of the model is given first, as shown below.

```
surveyhlm.OUT

Level-3 model (simulated data) with design weights

ITERATION NUMBER     4

                         +-----------------------+
                         | FIXED PART OF MODEL   |
                         +-----------------------+

--------------------------------------------------------------------------------
 COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
--------------------------------------------------------------------------------
 intcept                    0.92865       0.12985       7.15145       0.00000
 Time                       0.51411       0.05250       9.79254       0.00000
 Lang1                      0.39340       0.11381       3.45676       0.00055
 Lang2                     -1.03430       0.14107      -7.33172       0.00000

                         +-----------------------+
                         |   -2 LOG-LIKELIHOOD   |
                         +-----------------------+

 DEVIANCE= -2*LOG(LIKELIHOOD) =    23121.9339178070
 NUMBER OF FREE PARAMETERS =              11

 CHI-SQUARE SCALE FACTOR =      0.45957
```

Recall that the "true" values of the intcept, Time, Lang1 and Lang2 parameters were 1.0, 0.5, 0.5, and -1.0 respectively. To obtain 95% confidence intervals for these estimates, we calculate

$$Estimate \pm 1.96(std.error)$$

and find that the confidence intervals for the estimated intcept, Time, Lang1 and Lang2 parameters are (0.7009; 1.1565), (0.4220; 0.6062), (0.1937; 0.5931) and (-1.2818; -0.7868) respectively. In all four cases, the confidence intervals include the "true" values of the corresponding parameter. Note that a χ^2 scale factor of 0.45957 is reported. This value is used to obtain a corrected χ^2-statistic for testing one model against another model, as will be shown in the next example.

The output for the **random part** of the model is given next. Note that the parameter estimates reported in the output are generally close to the population values which were used to simulate the data. The "true" values for both the level-3 and level-2 variance-covariance components are 1.0, 0.3, and 0.2 respectively. The standard error estimates shown have been corrected as described in the theoretical section (see Section 2.6).

surveyhlm.OUT

```
                    +----------------------+
                    | RANDOM PART OF MODEL |
                    +----------------------+

 ------------------------------------------------------------------------------
  LEVEL 3                         TAU-HAT     STD.ERR.    Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------------
  intcept /intcept                0.93174     0.20698     4.50164    0.00001
  Time    /intcept                0.25568     0.07490     3.41361    0.00064
  Time    /Time                   0.17510     0.04203     4.16637    0.00003

 ------------------------------------------------------------------------------
  LEVEL 2                         TAU-HAT     STD.ERR.    Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------------
  intcept /intcept                0.96301     0.23237     4.14424    0.00003
  Time    /intcept                0.36079     0.06511     5.54100    0.00000
  Time    /Time                   0.20039     0.04226     4.74169    0.00000

 ------------------------------------------------------------------------------
  LEVEL 1                         TAU-HAT     STD.ERR.    Z-VALUE    PR > |Z|
 ------------------------------------------------------------------------------
  intcept /intcept                1.02326     0.04431    23.09147    0.00000
```

For the level-3 variance components, 95% confidence intervals can be obtained as shown previously. The confidence intervals corresponding to intcept/intcept, Time/intcept, and Time/Time are (0.6002; 1.2633), (0.1367; 0.3747) and (0.1083; 0.2419) respectively. Again, the "true" values fall within these intervals. This conclusion also holds for confidence intervals for the level-1 and level-2 variance-covariance components, which are calculated in the same way.

Note that the spreadsheet presentation of **surveyhlm.lsf** will only show the variables Institut, Patient, …, WT2, although dummy variables corresponding to the six measurement occasions were written to the actual LSF file. To see these dummy variables, close the LSF file (**without** saving it) and then use **File**, **Open** to display the modified LSF file.

surveyhlm.lsf

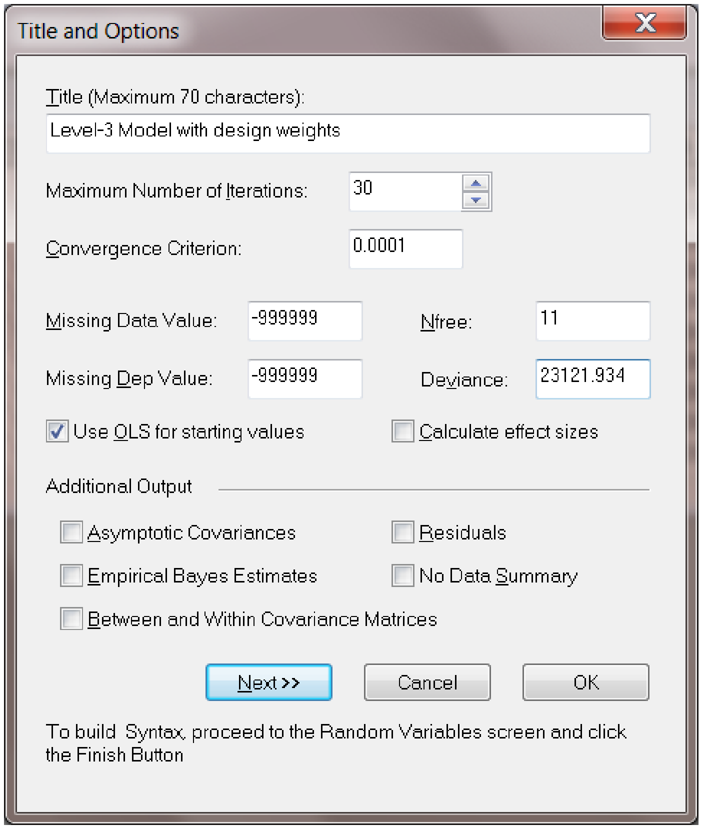
	Lang2	WT3	WT2	dummy1	dummy2	dummy3	dummy4	dummy5	dummy6
1	0.00	1.59	1.75	1.00	0.00	0.00	0.00	0.00	0.00
2	0.00	1.59	1.75	0.00	1.00	0.00	0.00	0.00	0.00
3	0.00	1.59	1.75	0.00	0.00	1.00	0.00	0.00	0.00
4	0.00	1.59	1.75	0.00	0.00	0.00	1.00	0.00	0.00
5	0.00	1.59	1.75	0.00	0.00	0.00	0.00	1.00	0.00
6	0.00	1.59	1.75	0.00	0.00	0.00	0.00	0.00	1.00
7	0.00	1.59	1.75	1.00	0.00	0.00	0.00	0.00	0.00
8	0.00	1.59	1.75	0.00	1.00	0.00	0.00	0.00	0.00
9	0.00	1.59	1.75	0.00	0.00	1.00	0.00	0.00	0.00
10	0.00	1.59	1.75	0.00	0.00	0.00	1.00	0.00	0.00

2.4.3 Three-level saturated model for simulated data

Using the same simulated data described in Section 2.4.2, a "saturated model" using the dummy variables created previously is now fitted to the data. This model illustrates the use of the NFREE and DEVIANCE keywords to obtain a chi-square statistic for testing two nested models. A model of particular interest is the so-called saturated model, which is obtained by estimating the population means and the covariance matrices for both level-3 and level-2 at the six measurement occasions. The dummy variables created in the previous section, each corresponding to a specific measurement occasion, are used for this purpose.

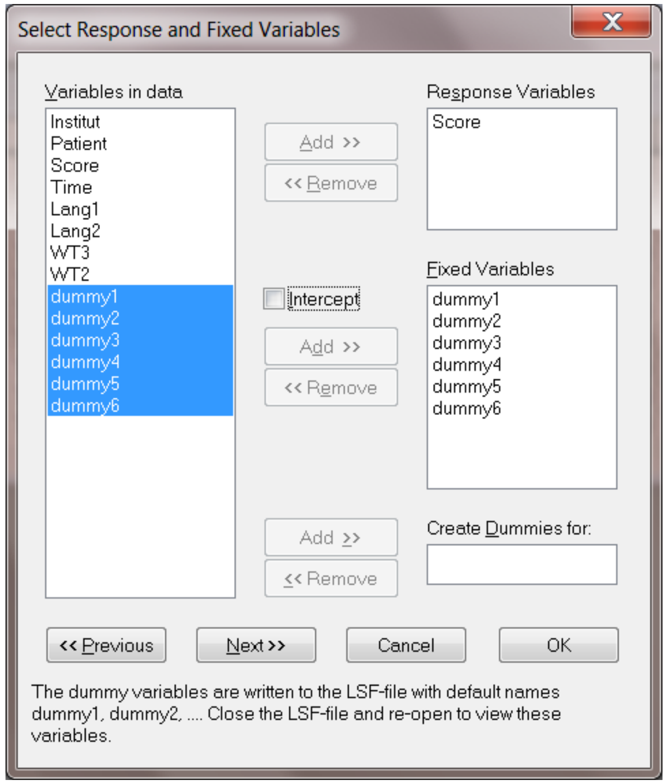
2.4.3.1 Fitting the model

The model is fitted using the same sequence of dialog boxes shown in the previous Section. In order to compare the fit of the saturated model with the fit of the model described in Section 2.4.2, the deviance statistic and number of estimated parameters from the first model are used. In the previous model, 11 parameters (4 fixed and 7 random) were estimated and a deviance statistic (-2 log*L*) of 23121.934 was obtained. Enter these values in the **Nfree** and **Deviance** fields of the **Title and Options** dialog box. Click **Next** to display the **Identification Variables** dialog box.

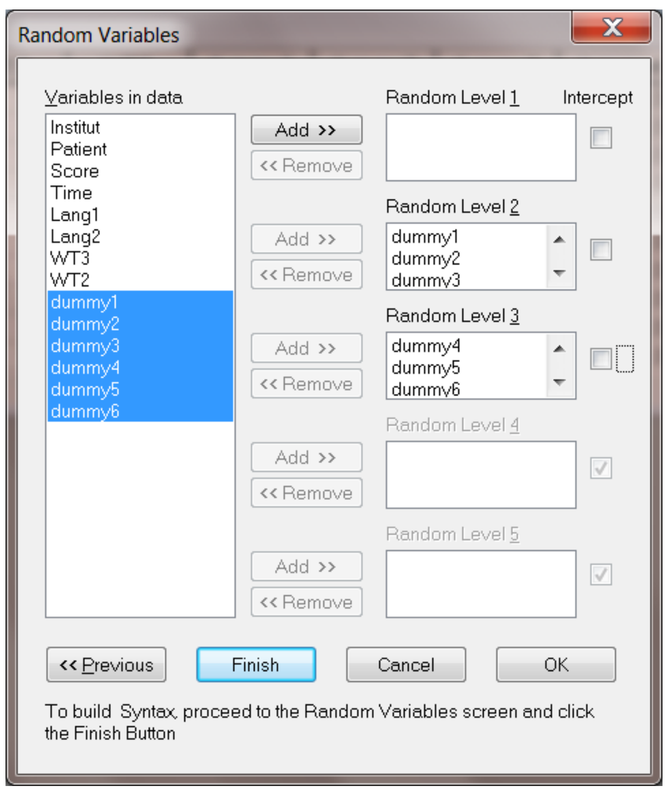


As no changes to the hierarchical structure or weight specification entered previously on the **Identification Variables** dialog box, click **Next** to load the **Response and Fixed Variables** dialog box.

Note that one cannot add an intercept term to the fixed part of the model when dummy1 to dummy6 are selected as predictors. If the intercept term is not unchecked, then the fixed parameter coefficients cannot be estimated, since the fixed-effect design matrix containing intcept, dummy1, …, dummy6 will not be of full rank. Click **Next** to proceed to the **Random Variables** dialog box.



On the **Random Variables** dialog box, the intercept terms for the random effects are unchecked and dummy variables one to six are only added at levels 2 and 3.



Click the **Finish** button to produce the PRELIS syntax file (which was subsequently saved as **surveyhlm2.prl**).

```
surveyhlm2.prl
OPTIONS OLS=YES NFREE=11 DEVIANCE=23121.934 CONVERGE=0.000100 MAXITER=30 OUTPUT=STANDARI
TITLE=Level-3 Model with design weights;
SY='C:\LISREL9 Examples\MLEVELEX\surveyhlm.lsf';
ID3=Institut;
ID2=Patient;
WEIGHT3=WT3;
WEIGHT2=WT2;
RESPONSE=Score;
FIXED=dummy1 dummy2 dummy3 dummy4 dummy5 dummy6;
RANDOM2=dummy1 dummy2 dummy3 dummy4 dummy5 dummy6;
RANDOM3=dummy1 dummy2 dummy3 dummy4 dummy5 dummy6;
```

2.4.3.2 Discussion of results

The portions of the output below summarize the estimated parameter values for the fixed part of the model and the goodness of fit χ^2 statistic. The $\hat{\beta}$ values are estimates of the population mean scores at each of the six measurement occasions, after controlling for the within institution and within patient variation. Note that the difference in the -2 log(L) values is 2832.814. The χ^2 value of 1312.7329 was obtained by multiplying 2832.814 with the scale factor (0.46) obtained when design weights are included.

```
surveyhlm2.OUT
                          +----------------------+
                          | FIXED PART OF MODEL  |
                          +----------------------+

 ------------------------------------------------------------------------------
   COEFFICIENTS              BETA-HAT      STD.ERR.      Z-VALUE       PR > |Z|
 ------------------------------------------------------------------------------
   dummy1                    0.86268       0.11765       7.33286       0.00000
   dummy2                    1.47507       0.16289       9.05572       0.00000
   dummy3                    1.92424       0.19218      10.01261       0.00000
   dummy4                    2.43611       0.24712       9.85820       0.00000
   dummy5                    2.98651       0.29087      10.26742       0.00000
   dummy6                    3.40673       0.34567       9.85530       0.00000


                          +----------------------+
                          |   -2 LOG-LIKELIHOOD  |
                          +----------------------+

   DEVIANCE= -2*LOG(LIKELIHOOD) =    20289.1201028391
   NUMBER OF FREE PARAMETERS =              48

        |-----------------------------------------------|
        | Chi-square Statistic for Testing the Fit of   |
        | the Current Model versus an Alternative Model |
        |-----------------------------------------------|

       -2Log(L)=      20289.1201 with     48  Free Parameters (Current Model)
       -2Log(L)=      23121.9340 with     11  Free Parameters (Alternative Model)
        Chi-Square=     1312.7329, df=     37, p-value=  0.00000
```

Results for the variance components (random part of the model) are shown below. The $\hat{\tau}$ values are estimates of the population variances/covariances at level-3 (institutions) and level-2 (patients). An inspection of the output shows that, in general, there is greater variation in scores at each time point within patients than is the case within institutions.

surveyhlm2.OUT

LEVEL 3		TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
dummy1	/dummy1	0.91628	0.16213	5.65142	0.00000
dummy2	/dummy1	1.14733	0.21999	5.21535	0.00000
dummy2	/dummy2	1.74251	0.39431	4.41920	0.00001
dummy3	/dummy1	1.37995	0.24954	5.52995	0.00000
dummy3	/dummy2	2.03690	0.43315	4.70253	0.00000
dummy3	/dummy3	2.54612	0.49287	5.16596	0.00000
dummy4	/dummy1	1.67425	0.32598	5.13604	0.00000
dummy4	/dummy2	2.61631	0.55997	4.67225	0.00000
dummy4	/dummy3	3.22516	0.62377	5.17039	0.00000
dummy4	/dummy4	4.18454	0.81608	5.12759	0.00000
dummy5	/dummy1	1.84738	0.33817	5.46284	0.00000
dummy5	/dummy2	2.84893	0.61342	4.64437	0.00000
dummy5	/dummy3	3.58711	0.68160	5.26281	0.00000
dummy5	/dummy4	4.69860	0.89993	5.22107	0.00000
dummy5	/dummy5	5.52001	1.03310	5.34316	0.00000
dummy6	/dummy1	2.17376	0.40571	5.35786	0.00000
dummy6	/dummy2	3.46369	0.77193	4.48704	0.00001
dummy6	/dummy3	4.38163	0.85862	5.10311	0.00000
dummy6	/dummy4	5.67540	1.11176	5.10487	0.00000
dummy6	/dummy5	6.58911	1.27135	5.18275	0.00000
dummy6	/dummy6	8.03923	1.59318	5.04602	0.00000

surveyhlm2.OUT

LEVEL 2		TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
dummy1	/dummy1	2.14267	0.25276	8.47722	0.00000
dummy2	/dummy1	1.45763	0.22475	6.48547	0.00000
dummy2	/dummy2	3.05585	0.43746	6.98541	0.00000
dummy3	/dummy1	1.93046	0.29417	6.56248	0.00000
dummy3	/dummy2	2.57313	0.40916	6.28886	0.00000
dummy3	/dummy3	4.32406	0.58250	7.42334	0.00000
dummy4	/dummy1	2.19516	0.31774	6.90869	0.00000
dummy4	/dummy2	3.11563	0.50019	6.22893	0.00000
dummy4	/dummy3	3.99632	0.57238	6.98198	0.00000
dummy4	/dummy4	5.92608	0.83054	7.13521	0.00000
dummy5	/dummy1	2.60166	0.38782	6.70850	0.00000
dummy5	/dummy2	3.88162	0.59735	6.49803	0.00000
dummy5	/dummy3	4.89266	0.68106	7.18386	0.00000
dummy5	/dummy4	6.06821	0.84521	7.17954	0.00000
dummy5	/dummy5	8.14698	1.05293	7.73741	0.00000
dummy6	/dummy1	2.88799	0.41279	6.99622	0.00000
dummy6	/dummy2	4.37310	0.74994	5.83124	0.00000
dummy6	/dummy3	5.73908	0.83557	6.86847	0.00000
dummy6	/dummy4	7.07375	0.99921	7.07933	0.00000
dummy6	/dummy5	8.61156	1.16767	7.37497	0.00000
dummy6	/dummy6	11.07436	1.57773	7.01916	0.00000

2.4.4 Four-level model for assessment data

2.4.4.1 The data

The data set used here consists of four levels of nesting. Information is available on repeated measurements made on 1,192 participants at three occasions. In the case of some of the participants, measurements were made on only one or two occasions. Data for the first 10 participants on most of the variables used in this section are shown below in the form of a LISREL spreadsheet file, named **therapist_L4.lsf**.

Therapis_L4.lsf

	site	therapis	particip	assesmt	gender	occasion	thera1	thera2	thera3	thera4
1	1.0	1.0	2.0	14.0	1.0	0.0	1.0	0.0	0.0	0.0
2	1.0	1.0	2.0	14.0	1.0	1.0	1.0	0.0	0.0	0.0
3	1.0	1.0	4.0	24.0	1.0	0.0	1.0	0.0	0.0	0.0
4	1.0	1.0	4.0	29.0	1.0	1.0	1.0	0.0	0.0	0.0
5	1.0	1.0	4.0	32.0	1.0	2.0	1.0	0.0	0.0	0.0
6	1.0	1.0	6.0	19.0	0.0	0.0	1.0	0.0	0.0	0.0
7	1.0	1.0	6.0	26.0	0.0	1.0	1.0	0.0	0.0	0.0
8	1.0	1.0	6.0	11.0	0.0	2.0	1.0	0.0	0.0	0.0
9	1.0	1.0	7.0	22.0	1.0	0.0	1.0	0.0	0.0	0.0
10	1.0	1.0	7.0	25.0	1.0	1.0	1.0	0.0	0.0	0.0

The variables of interest are:

- site is the level-4 identification variable (49 units in total).
- therapis is the level-3 identification variable (187 units in total).
- particip is the level-2 identification variable (1192 units in total).
- assesmt is a score assigned by a therapist to a particular participant on occasion 0, 1 or 2.
- gender is a gender indicator, with a value of 0 indicating a male participant and 1 a female participant.
- occasion is a predictor variable coded 0, 1 and 2.
- thera1 - thera4 are dummy coded variables indicating four types of therapy.

2.4.4.2 The model

In this example, the identification variable site is used as the level-4 identifier, therapis is used as the level-3 identifier, and particip serves to identify level-2 units.

A general four-level model for a response variable y depending on a set of r predictors $x_1, x_2, \ldots, x_r$ can be written in the form

$$y_{ijkl} = \mathbf{x}'_{(f)ijkl}\boldsymbol{\beta} + \mathbf{x}'_{(4)ijkl}\mathbf{v}_i + \mathbf{x}'_{(3)ijkl}\mathbf{v}_{ij} + \mathbf{x}'_{(2)ijkl}\mathbf{u}_{ijk} + \mathbf{x}'_{(1)ijkl}\mathbf{e}_{ijkl}$$

where $i = 1, 2, \ldots, N$ denotes the level-4 units, $j = 1, 2, \ldots, n_i$ the level-3 units, $k = 1, 2, \ldots, n_{ij}$ the level-3 units, and $l = 1, 2, \ldots, n_{ijk}$ the level-1 units. In this context, y_{ijkl} represents the response of individual l, nested within level-2 unit k, level-3 unit j and level-4 unit i. The model shown here consists of a fixed and a random part. The fixed part of the model is represented by the vector product $\mathbf{x}'_{(f)ijkl}\boldsymbol{\beta}$, where $\mathbf{x}'_{(f)ijkl}$ is a typical row of the design matrix of the fixed part of the model with, as elements, a subset of the r predictors. The vector $\boldsymbol{\beta}$ contains the fixed, but unknown parameters to be estimated. The vector products $\mathbf{x}'_{(4)ijkl}\mathbf{v}_i$, $\mathbf{x}'_{(3)ijkl}\mathbf{v}_{ij}$, $\mathbf{x}'_{(2)ijkl}\mathbf{u}_{ijk}$, and $\mathbf{x}'_{(1)ijkl}\mathbf{e}_{ijkl}$ denote the random part of the model at levels 4, 3, 2, and 1 respectively. For example, $\mathbf{x}'_{(3)ijkl}$ represents a typical row of the design matrix of the random part at level-3, and $\mathbf{v}_{ij}$ the vector of random level-3 coefficients to be estimated. The products $\mathbf{x}'_{(2)ijkl}\mathbf{u}_{ijk}$ and $\mathbf{x}'_{(1)ijkl}\mathbf{e}_{ijkl}$ serve the same purpose at levels 2 and 1 respectively. It is assumed that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are independently and identically distributed (i.i.d.) with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Phi}_{(4)}$ and that $\mathbf{v}_{i1}, \mathbf{v}_{i2}, \ldots, \mathbf{v}_{in_i}$ are independently and identically distributed (i.i.d.) with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Phi}_{(3)}$. Similarly, $\mathbf{u}_{ij1},\mathbf{u}_{ij2},\ldots,\mathbf{u}_{ijn_{ij}}$ are assumed i.i.d., with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Phi}_{(2)}$, and $\mathbf{e}_{ijk1},\mathbf{e}_{ijk2},\ldots,\mathbf{e}_{ijkn_{ijk}}$ are assumed i.i.d., with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Phi}_{(1)}$.

Within this hierarchical framework, the model fitted to the data uses the participant's gender and type of therapy treatment to predict the assessment scores on three occasions.

$$\begin{aligned}\text{assessment}_{ijkl} = {} & \beta_1 * \text{gender}_{ijkl} + \beta_2 * \text{occasion}_{ijkl} + \beta_3 * \text{thera1}_{ijkl} + \\ & \beta_4 * \text{thera2}_{ijkl} + \beta_5 * \text{thera3}_{ijkl} + \beta_6 * \text{thera4}_{ijkl} + \\ & + v_{i0} + v_{ij0} + u_{ijk0} + u_{ijk1} * \text{occasion}_{ijkl} + e_{ijkl}\end{aligned}$$

where $\beta_1, \beta_2, \ldots, \beta_6$ indicate the estimated coefficients associated with the fixed part of the model that contains the predictor variables gender, occasion, and the four indicator variables. The random part of the model is represented by v_{i0}, v_{ij0}, u_{ijk0}, u_{ijk1} and e_{ijkl}, which denote the variation in average assessment across sites, therapists, participants and measurement occasions.

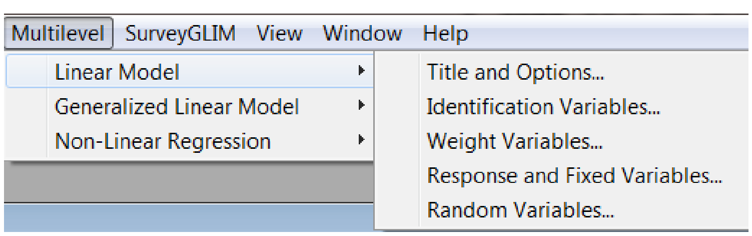
2.4.4.3 Setting up the analysis

The model is fitted to the data in t**herapist_L4.lsf** by using the sequence of five dialog boxes accessed via the **Multilevel**, **Linear Model** option from the main menu bar in LISREL. Note that options such as **Multilevel** and **SurveyGLIM** are only available on the main menu bar when a ***.lsf** file is open.

The first step is to open the LSF shown above, which is accomplished as follows:

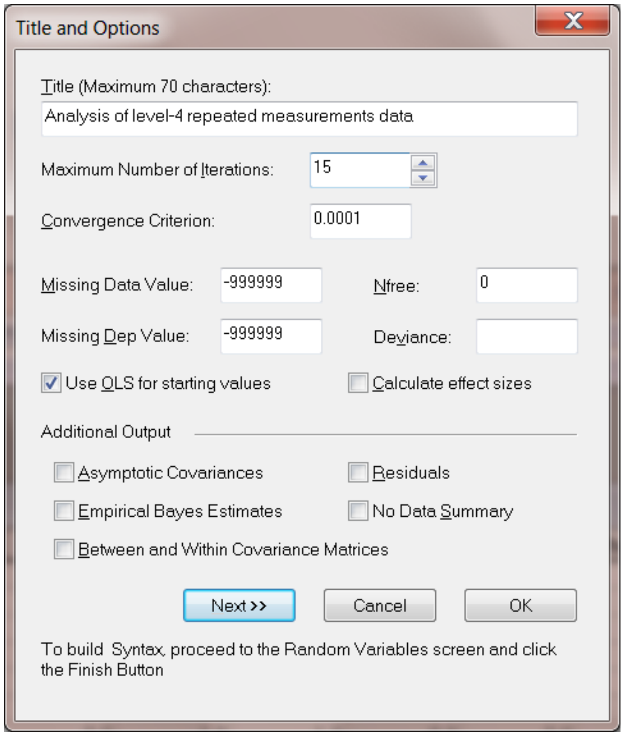
- Use the **File**, **Open** option to activate the display of an **Open** dialog box.
- Set the **Files of type** drop-down list box to **LISREL System Data (*.lsf)** and browse for the file **therapist_L4.lsf**.
- Select the file and click the **Open** button to return to the main LISREL window, where the contents of the LSF are displayed.

The next step is to describe the model to be fitted using the multilevel module in LISREL. From the main menu bar, select the **Multilevel** option. Here we limit our discussion to linear models, and thus the **Linear Model** option will be used throughout.

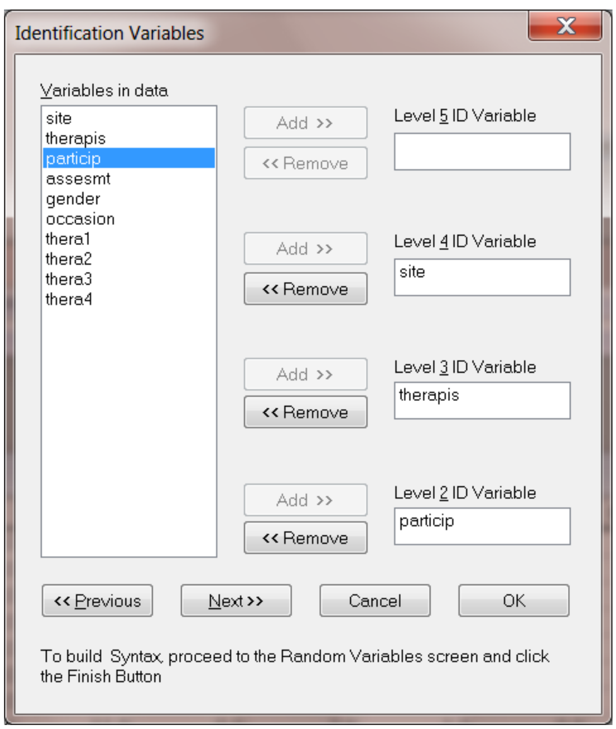


The first of the five options on the pop-up menu provide access to the **Title and Options** dialog box. Start by providing a title for the analysis in the **Title** field. In this example, we set the maximum number of iterations to

15 and the convergence criterion to 0.0001. Default settings for all other options associated with this dialog box are used. Click the **Next** button to go to the **Identification Variables** dialog box.

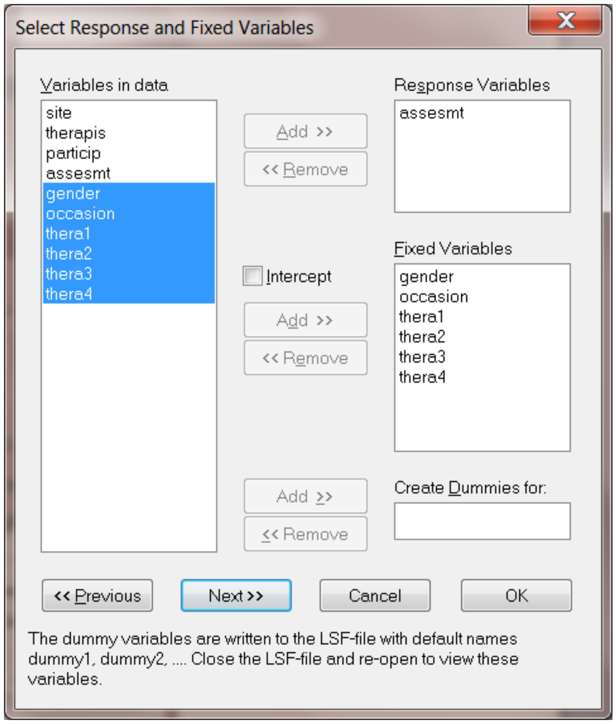


On the **Identification Variables** dialog box, enter the variables defining the hierarchical structure as ID variables.

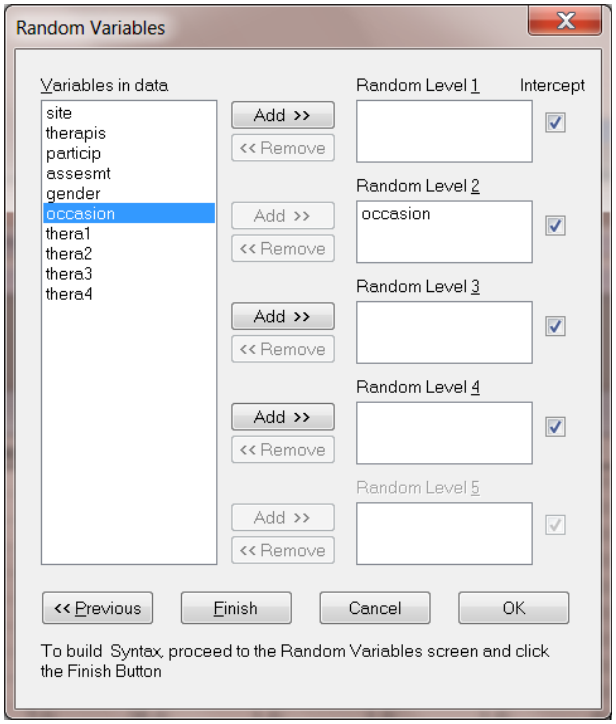


The next dialog box is used to provide information on weight variables, if any. In our case, weights are not available. Therefore, click the **Next** button to proceed to the **Select Response and Fixed Variables** dialog box.

The **Select Response and Fixed Variables** dialog box is used to identify the outcome variable and predictor variables, if any. Select and add the outcome variable assesmt to the **Response Variables** field. Next, select the variables starting from gender to thera4 by dragging the mouse over them and click the **Add** button next to the **Fixed Variables** field to include these variables as predictors in the model. Note that the **Intercept** box should be unchecked. This completes the specification of the response and fixed variables. Click the **Next** button to proceed to the **Random Variables** dialog box.

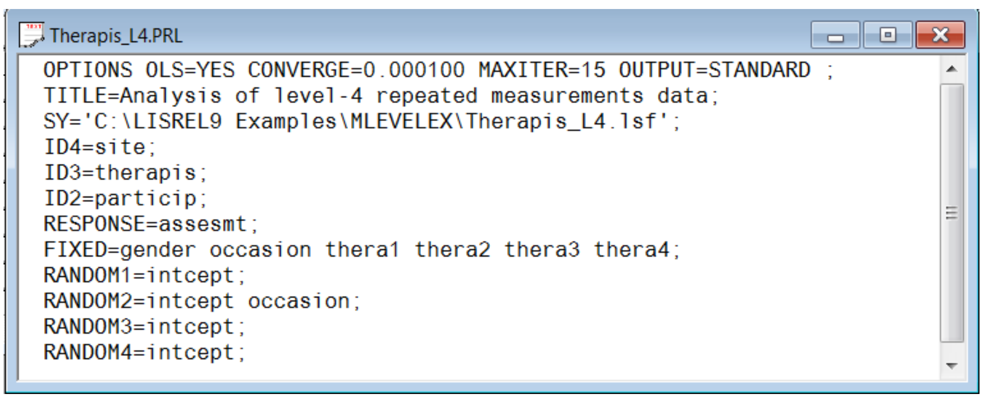


The **Random Variables** dialog box shown below displays the default settings associated with this dialog box. In the current model, only intercept coefficients are allowed to vary randomly at the various levels of the hierarchy, except at level-2 where we add the predictor occasion as a random effect. Once this is done, click the **Finish** button to generate the syntax for the model.



The syntax shown below corresponds to the information entered via the dialog boxes above. Run the model by clicking the **Run Prelis** icon on the main menu bar.

2.4.4.4 Discussion of results




```
OPTIONS OLS=YES CONVERGE=0.000100 MAXITER=15 OUTPUT=STANDARD ;
TITLE=Analysis of level-4 repeated measurements data;
SY='C:\LISREL9 Examples\MLEVELEX\Therapis_L4.lsf';
ID4=site;
ID3=therapis;
ID2=particip;
RESPONSE=assesmt;
FIXED=gender occasion thera1 thera2 thera3 thera4;
RANDOM1=intcept;
RANDOM2=intcept occasion;
RANDOM3=intcept;
RANDOM4=intcept;
```

Only selected parts of the output are shown. The output describing the estimated **fixed effects** after convergence is shown next. The estimates are shown in the column with heading BETA-HAT and correspond to the coefficients $\beta_0, \beta_2, \ldots, \beta_6$ in the model specification. From the z-values and associated exceedance probabilities, we see that except for the coefficient associated with gender, the remaining coefficients were all highly significant.

Therapis_L4.OUT

ITERATION NUMBER 9

FIXED PART OF MODEL

COEFFICIENTS	BETA-HAT	STD.ERR.	Z-VALUE	PR > |Z|
gender	-0.68445	0.31043	-2.20484	0.02747
occasion	2.52207	0.08571	29.42622	0.00000
thera1	18.46583	0.52985	34.85093	0.00000
thera2	22.81821	0.39635	57.57015	0.00000
thera3	27.19290	0.39524	68.80062	0.00000
thera4	30.63948	0.45528	67.29788	0.00000

-2 LOG-LIKELIHOOD

DEVIANCE= -2*LOG(LIKELIHOOD) = 19795.3841910891
NUMBER OF FREE PARAMETERS = 12

Therapis_L4.OUT

RANDOM PART OF MODEL

LEVEL 4	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	2.25830	0.74765	3.02055	0.00252

LEVEL 3	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	0.36906	0.53853	0.68532	0.49314

LEVEL 2	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	22.09563	1.57712	14.01007	0.00000
occasion/intcept	-0.50019	0.64561	-0.77476	0.43848
occasion/occasion	0.21014	0.47079	0.44636	0.65534

LEVEL 1	TAU-HAT	STD.ERR.	Z-VALUE	PR > |Z|
intcept /intcept	14.27640	0.63112	22.62080	0.00000

A study of the random part of the model shows that all the intercept effects are highly significant, except for the level-3 (therapists) intercept. From this, we conclude that intercept estimates vary significantly over sites, but not over therapists. Finally, we note that there is no significant occasion effect at level-2. This fact can be further substantiated by fitting an intercepts only model to the data. In this case, we obtain the following results:

DEVIANCE= -2*LOG(LIKELIHOOD)= 19795.9780720184
NUMBER OF FREE PARAMETERS= 10

Since the difference between the deviance statistics equals 0.58, this leads to the conclusion (at 12 – 10 = 2 degrees of freedom) that the level-2 random occasion effect is not significant.

2.5 Evaluation

2.5.1 Introduction

A feature of many sampling surveys is that the probability of selection is unequal. This can be the result of stratified sampling, cluster sampling, subpopulation oversampling, designed unequal probability sampling, etc. If the unequal probability of selection is not incorporated in the analysis a substantial bias in the parameter estimates may arise. This bias is commonly known as the selection bias. If the probability of selection is known and incorporated in the analysis the selection bias can be eliminated.

In the next section we compare the performance of the methods implemented in the four statistical software packages LISREL, HLM, Mplus, and MLWiN. In all the tables to follow, the abbreviation MLevel is used to denote the multilevel module in LISREL.

2.5.2 Comparison of results using two-level simulated data

Asparouhov (2004) selected a linear growth model for continuous outcomes as the basis for a simulation study. This model can be estimated by all the different statistical packages for hierarchical linear modeling. An unbalanced design, consisting of 500 univariate observations that are clustered within 100 level two units, was used. Half of the level-2 units have four observations and the other half have six observations. The times of the observations are equally spaced starting at 0 and ending with 3 for the clusters with 4 observations and ending with 5 for the clusters with 6 observations. The linear growth model has random intercept and slope coefficients.

The observed variable Y_{ij} for level-2 unit i at time j is given by

$$Y_{ij} = b_{0i} + b_{1i} + \varepsilon_{ij}$$

where ε_{ij} is a zero-mean, normally distributed residual with variance σ^2. The random effects b_{0i} and b_{1i} have means β_0 and β_1, variances ϕ_{11} and ϕ_{22} respectively, and covariance ϕ_{21}.

The selection model is defined by the initial status in the growth model Y_{i0}, namely $P(b_{0i} = 1) = 1/(1 + \exp\{-Y_{i0}\})$, *i.e.*, level-2 units with higher initial status have been oversampled. The analysis was replicated 500 times. An example of a few records for the first of the 500 data sets, in the form of a LSF, is shown below.

Simul1.LSF

	Level2ID	Y	X	WT
1	10001.000	-0.484	0.000	2.623
2	10001.000	0.968	1.000	2.623
3	10001.000	2.165	2.000	2.623
4	10001.000	2.512	3.000	2.623
5	10002.000	0.256	0.000	1.774
6	10002.000	1.617	1.000	1.774
7	10002.000	0.555	2.000	1.774
8	10002.000	2.479	3.000	1.774
9	10002.000	4.189	4.000	1.774
10	10002.000	1.656	5.000	1.774

Table 2.5 shows the bias in the parameter estimates as well as the coverage rates for the 95% confidence intervals computed by LISREL, HLM 6, Mplus 3 and MLWiN 2. Note that this table contains updated HLM results when compared to Asparouhov (2004), where the previous version of HLM was used. In addition, results obtained with LISREL have been added.

Table 2.5: Bias and Coverage in LISREL, HLM, MLWiN and HLM

Parameter	True Value	Bias				Coverage			
		LISREL	HLM	MLWiN	Mplus	LISREL	HLM	MLWiN	Mplus
β_0	0.5	0.019	0.016	0.017	0.017	0.906	0.908	0.782	0.908
β_1	0.1	0.001	0.03	0.002	0.002	0.948	0.938	0.888	0.942
ϕ_{11}	1	-0.029	-0.012	-0.024	-0.024	0.840	-	0.758	0.848
ϕ_{22}	0.2	-0.006	-0.001	-0.006	-0.006	0.878	-	0.848	0.902
ϕ_{21}	0.3	-0.005	-0.008	-0.005	-0.006	0.938	-	0.846	0.940
σ^2	1	-0.005	-0.012	-0.008	-0.008	0.946	-	0.878	0.910

The bias shown in subsequent tables is the difference between the mean of the estimated parameters over the 500 simulated data sets and the true value for that parameter as used in the actual simulation. For example, the first value for LISREL in the body of Table 2.5, *i.e.* 0.019, indicates that the average of the estimates of β_0 with this program was 0.5 + 0.019 = 0.519.

The coverage reported was calculated by determining the lower and upper bounds of a 95% confidence interval for each of the parameters for each of the simulated data sets. If the true value for the parameter fell within the confidence interval, an indicator variable was assigned a value of 1; if not, the indicator variable was coded 0. The mean value of the indicator variable over all 500 data sets, expressed as a percentage, is the coverage as reported in the tables to follow and indicates the percentage of data sets where the confidence interval based on the estimates obtained included the simulated or "true" value. The SAS code for calculation of coverage for the intercept is given below.

```
title 'coverage of intercept';
*Upper and lower limits of interval;
upper=intcept+1.96*serror;
lower=intcept-1.96*serror;
*Determine inclusion of true value;
if lower<=0.5<=upper then include=1;
else include=0;
proc means;
var include;
```

The biases produced by LISREL, HLM, Mplus and MLWiN are virtually identical. While the difference in the point estimation between the three methods is very small, the difference in the variance estimation (where available) is not.

Table 2.6 shows the effect of ignoring the design weight, as computed with LISREL. There is large bias present in the estimation of the intercept coefficient (β_0) and the variance of the level-2 intercept error term (ϕ_{11}) when weights are omitted. This conclusion is substantiated by the low coverage for these parameters.

Table 2.6: Bias and Coverage in LISREL without inclusion of a weight variable

Parameter	True Value	Without weights		With weights	
		Bias	Coverage	Bias	Coverage
β_0	0.5	-0.482	0.60	0.019	0.906
β_1	0.1	0.031	0.906	0.001	0.948
ϕ_{11}	-1.0	-0.312	0.592	-0.029	0.840
ϕ_{22}	0.2	-0.002	0.932	-0.006	0.878
ϕ_{21}	0.3	-0.020	0.922	-0.005	0.938
σ^2	1	-0.005	0.936	-0.005	0.946

2.5.3 Comparison of results using three-level simulated data

In this section we discuss the results of a simulation study for the evaluation of a 3-level model with level-2 and level-3 weights. Five hundred datasets were simulated according to the following hypothetical model

$$Score_{ijk} = \beta_0 + \beta_1 * Time + \gamma_1 * Lang1 + \gamma_2 * Lang2$$
$$+ v_{i0} + Time * v_{i1} + u_{ij0} + Time * u_{ij1} + e_{ijk}$$

where i denotes institution i, (i = 1, 2, …, 100), ij patient j (j =1, 2, …, 10) in institution i and ijk the k-th measurement (k = 1, 2, …, 6) on patient j in institution i. The outcome variable Score denotes a patient's measurement on some test of interest, TIME the time of measurement, and Lang1 and Lang2 are indicator variables indicating a patient's first or home language. The data were simulated under the assumption that

$$\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} 1.0 \\ 0.5 \end{pmatrix}, \quad \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -1.0 \end{pmatrix}$$
$$\phi_2 = Cov(u_{ij0}, u_{ij1}) = \begin{pmatrix} 1 & \\ 0.3 & 0.2 \end{pmatrix}$$
$$\phi_3 = Cov(v_{i0}, v_{i1}) = \begin{pmatrix} 1 & \\ 0.3 & 0.2 \end{pmatrix}$$

and

$$\sigma^2 = Var(e_{ijk}) = 1.0.$$

Note that the data were simulated in such a way that odd-numbered patients have six score measurements at time points 0, 1, 2, 3, 4, 5. The even-numbered patients have only four score measurements.

Level-3 weights

To incorporate design weights, the simulated initial scores were standardized to a normal (0,1) distribution and an equal number of "institutions" were subsequently drawn from each of the 10 score intervals ($(-\infty, -1)$, $(-1, -0.7)$, …, $(2.3, \infty)$) as shown in the table below. It can easily be verified that for a standardized normal variable z, $P(0.30 \le z \le 1.00) = 22.34\%$. In the simulation study, patients (cases) were selected from 10 institutions if their standardized scores fell within the interval (0.30, 1.00). To correct for this undersampling, a weight of 22.34/10.0 = 2.234 was assigned to each of those institutions. In a similar fashion, 10 institutions were selected according to the remaining nine score intervals as shown below.

Interval	Lower	Upper	% Expected	% Selected	Weight3
1	-Inf	-1.00	15.87	10.00	1.587
2	-1.00	-0.70	8.33	10.00	0.833
3	-0.70	-0.20	17.88	10.00	1.788
4	-0.20	0.00	7.93	10.00	0.793
5	0.00	0.30	11.79	10.00	1.179
6	0.30	1.00	22.34	10.00	2.234
7	1.00	1.30	6.19	10.00	0.619
8	1.30	1.80	6.09	10.00	0.609
9	1.80	2.30	2.52	10.00	0.252
10	2.30	Inf	1.07	10.00	0.107

Level-2 weights

In order to incorporate level-2 weights, it was further assumed that the actual percentages of patients in each of three ethnic groups are 70%, 20% and 10%. However, in each institution four patients were drawn from the first ethnic groups, and three from each of the second and third ethnic groups. To compensate for this unequal probability of selection, level-2 ("patient") weights were assigned as follows:

- Four from ethnic group 1 with Weight2 = 7.0/4.0
- Three from ethnic group 2 with Weight2 = 2.0/3.0
- Three from ethnic group 3 with Weight2 = 1.0/3.0

The first 10 records of the dataset in **surveyhlm.lsf** are shown below.

surveyhlm.lsf

	Institut	Patient	Score	Time	Lang1	Lang2	WT3	WT2
1	1.00	1.00	-1.84	0.00	0.00	0.00	1.59	1.75
2	1.00	1.00	-0.89	1.00	0.00	0.00	1.59	1.75
3	1.00	1.00	-1.21	2.00	0.00	0.00	1.59	1.75
4	1.00	1.00	-3.24	3.00	0.00	0.00	1.59	1.75
5	1.00	1.00	-1.16	4.00	0.00	0.00	1.59	1.75
6	1.00	1.00	-1.30	5.00	0.00	0.00	1.59	1.75
7	1.00	2.00	-0.07	0.00	0.00	0.00	1.59	1.75
8	1.00	2.00	2.99	1.00	0.00	0.00	1.59	1.75
9	1.00	2.00	0.92	2.00	0.00	0.00	1.59	1.75
10	1.00	2.00	3.63	3.00	0.00	0.00	1.59	1.75

The model was fitted to 500 simulated data sets using HLM 6.08 (Bryk & Raudenbush, 2004) and LISREL. Table 2.7 shows the bias and coverage for estimates obtained with weighted analyses using LISREL and HLM. Results are very similar. HLM 6.8 does not provide estimates of the variance-covariance components in the case of weighted models. Due to this, coverage for the HLM results could not be calculated.

Table 2.8 shows the effect of ignoring the design weight, as computed with LISREL. For the unweighted analyses it was found that the estimates of the fixed parameters (β_0, β_1, γ_1 and γ_2) were strongly biased as is also reflected by the low coverage (0.042 in the case of the intercept coefficient). As was the case for the similar comparison shown in Table 2.6, both bias and coverage for the weighted model yield closer approximations to the theoretical expected values for bias and coverage (0 and 0.95).

Table 2.7: Bias and Coverage for simulated three-level data (weighted analyses)

Parameter	True Value	LISREL		HLM	
		Bias	Coverage	Bias	Coverage
β_0	1.0	-0.002	0.986	-0.002	0.986
β_1	0.5	0.001	0.948	-0.004	0.948
γ_1	0.5	-0.004	0.950	0.001	0.950
γ_2	-1.0	-0.004	0.936	-0.004	0.936
σ^2	1.0	0.001	1.000	0.010	-
$\phi_{(2)11}$	1.0	-0.072	0.856	-0.075	-
$\phi_{(2)22}$	0.2	-0.003	0.970	-0.000	-
$\phi_{(2)21}$	0.3	-0.011	0.906	-0.013	-
$\phi_{(3)11}$	1.0	-0.049	0.930	0.049	-
$\phi_{(3)22}$	0.2	-0.002	0.922	-0.003	-
$\phi_{(3)21}$	0.3	0.008	0.922	0.008	-

Table 2.8: Bias and Coverage for simulated three-level data (unweighted analyses)

Parameter	True Value	Bias	Coverage
β_0	1.0	0.418	0.042
β_1	0.5	0.090	0.548
γ_1	0.5	-0.115	0.756
γ_2	-1.0	-0.122	0.764
σ^2	1.0	0.004	0.920
$\phi_{(2)11}$	1.0	0.064	0.908
$\phi_{(2)22}$	0.2	0.000	0.966
$\phi_{(2)21}$	0.3	0.006	0.958
$\phi_{(3)11}$	1.0	0.256	0.850
$\phi_{(3)22}$	0.2	0.017	0.954
$\phi_{(3)21}$	0.3	0.070	0.914

2.5.4 Comparison of results using a 3-level model for the MEPS data

The model fitted in Section 2.2.1 using the multilevel module (MLevel) of LISREL was also fitted using HLM 6, MLWiN 2 and the SurveyGLIM (GLIM) module of LISREL. Table 2.9 below contains estimates obtained with these four procedures for both weighted and unweighted models. Standard error estimates, where available, are given below the estimates in parentheses. Simplified two-level models using the same data and a wider array of statistical software are given in the next section.

Results of the GLIM analyses are not directly comparable to those obtained using the multilevel analysis programs HLM, MLevel and MLWiN, but are presented here to demonstrate the effect of the different modeling assumptions. The standard errors reported for the SurveyGLIM module were obtained under the assumption of stratification and clustering, using a Taylor linearization approach to the asymptotic covariance matrix (see Section 6 of the *Generalized Linear Modeling Guide*). In the multilevel programs, the stratum and cluster variables define the hierarchical structure by serving as level-3 and level-2 identifiers, and it is assumed that the intercept coefficients vary randomly across the level-3 and level-2 units.

Table 2.9: Comparison of results from 4 procedures for model fitted to MEPS data

Coefficient	Weighted HLM	MLevel	MLWiN	GLIM	Unweighted HLM	MLevel	MLWiN	GLIM
intercept	4.360 (0.122)	4.391 (0.114)	4.336 (0.108)	4.277 (0.125)	4.458 (0.115)	4.458 (0.084)	4.459 (0.084)	4.282 (0.025)
race	0.944 (0.088)	0.943 (0.086)	0.939 (0.082)	0.880 (0.100)	0.684 (0.096)	0.684 (0.050)	0.684 (0.050)	0.623 (0.017)
GENDER	0.904 (0.038)	0.911 (0.036)	0.920 (0.039)	0.932 (0.041)	0.931 (0.037)	0.931 (0.036)	0.931 (0.036)	0.945 (0.013)
inscov	-0.616 (0.081)	-0.651 (0.076)	-0.695 (0.074)	-0.630 (0.086)	-0.618 (0.084)	-0.618 (0.046)	-0.618 (0.046)	-0.733 (0.016)
rpovc991	0.363 (0.118)	0.358 (0.114)	0.432 (0.103)	0.439 (0.109)	0.493 (0.095)	0.493 (0.065)	0.493 (0.065)	0.668 (0.023)
rpovc992	-0.122 (0.110)	-0.138 (0.104)	-0.142 (0.106)	-0.114 (0.107)	-0.154 (0.097)	-0.154 (0.067)	-0.154 (0.067)	-0.128 (0.024)

rpovc993	0.088 (0.111)	0.070 (0.117)	0.119 (0.097)	0.130 (0.116)	0.101 (0.096)	0.101 (0.062)	0.101 (0.062)	0.205 (0.022)
rpovc994	-0.318 (0.152)	-0.329 (0.140)	-0.343 (0.150)	-0.263 (0.141)	-0.346 (0.109)	-0.346 (0.090)	-0.346 (0.090)	-0.336 (0.032)
variance (level-1)	6.682	7.006 (0.196)	7.233 (0.166)	N/A	7.463 (0.069)	7.463 (0.069)	7.463 (0.069)	N/A
variance (level-2)	0.190	0.177 (0.037)	0.200 (0.046)	N/A	0.175 (0.028)	0.174 (0.028)	0.174 (0.028)	N/A
variance (level-3)	0.079	0.073 (0.029)	0.101 (0.048)	N/A	0.145 (0.037)	0.146 (0.0365)	0.146 (0.036)	N/A
deviance	112067	118256	114698		114663	114663	114663	
parameters	11	11	11		11	11	11	

Parameter estimates for the unweighted analyses with HLM, MLevel and MLWiN are identical. The HLM standard error estimates for the fixed effects are generally larger than those reported for MLevel and MLWiN. The reason for this is that the robust standard error estimates produced by HLM were reported. By including the commands

```
WEIGHT2 = intcept;
WEIGHT3 = intcept;
```

in the MLevel syntax file, results similar to those produced by HLM can be obtained with the MLevel module (see Section 2.6.4 for a further discussion of this topic). The parameter and standard error estimates for the HLM, MLevel and MLWiN procedures are very similar. Note that HLM 6 does not produce standard errors for the variance components.

Turning to the GLIM results, we note relatively large differences in parameter and standard error estimates for both the weighted and unweighted models. More research may be required to provide users with guidelines if a choice has to be made between fitting a multilevel or a generalized linear model to complex survey data.

2.5.5 Comparison of results using a 2-level model for the MEPS data

In order to expand the comparison of results for weighted models using all the software packages at our disposal, we fitted a series of two-level models to the MEPS data. While this implies ignoring the survey sample design to some extent, doing so was necessary in order to obtain results for SAS PROC MIXED and Mplus. In the case of PROC MIXED, fitting 3-level models is computationally intensive and thus not an option when a large number of models is to be fitted, while Mplus cannot presently accommodate level-3 models.

In the first set of models, it was assumed that respondents were nested within the 143 strata only, and no distinction was made in terms of the PSU they were drawn from. Three models were introduced, each using a different subset of the predictors used in Section 2.4.1. For each of the models, results for both weighted and unweighted analyses are given in Tables 2.10, 2.11 and 2.12.

An inspection of Tables 2.10 to 2.12 shows that the GLIM estimates and standard errors differ from those obtained using the multilevel procedures. This result is to be expected, since the multilevel approach allows for all or a subset of the regression coefficients to vary randomly over the different levels of the hierarchical structure, while the GLIM approach assumes fixed regression coefficients and uses stratification and clustering variables to produce appropriate standard errors.

Table 2.10: Comparison of results for first model fitted to MEPS data

Model 1: Results for unweighted analyses

Coefficients	**MLevel**	**HLM**	**SAS**	**MLWiN**	**GLIM**	**Mplus**
intercept	4.652	4.652	4.652	4.652	4.571	4.654
	(0.066)	(0.092)	(0.066)	(0.066)	(0.078)	(0.093)
race	0.715	0.715	0.715	0.725	0.685	0.713
	(0.048)	(0.097)	(0.048)	(0.048)	(0.117)	(0.097)
GENDER	0.919	0.919	0.919	0.929	0.924	0.920
	(0.036)	(0.036)	(0.036)	(0.036)	(0.039)	(0.036)
inscov	-0.838	-0.838	-0.838	-0.838	-1.015	-0.840
	(0.040)	(0.081)	(0.040)	(0.040)	(0.124)	(0.081)
variance (level-1)	7.632	7.633	7.632	7.632	N/A	7.634
	(0.071)			(0.071)		(0.181)
variance (level-2)	0.221	0.224	0.222	0.333	N/A	0.223
	(0.036)			(0.036)		(0.036)

Model 1: Results for weighted analyses

Coefficients	**MLevel**	**HLM**	**SAS**	**MLWiN**	**GLIM**	**Mplus**
intercept	4.527	4.528	4.514	4.541	4.470	4.527
	(0.092)	(0.093)	(0.062)	(0.091)	(0.091)	(0.093)
race	0.970	0.970	0.966	0.920	0.985	0.970
	(0.082)	(0.082)	(0.048)	(0.084)	(0.086)	(0.082)
GENDER	0.902	0.902	0.902	0.906	0.901	0.902
	(0.036)	(0.036)	(0.035)	(0.040)	(0.037)	(0.036)
inscov	-0.822	-0.822	-0.825	-0.832	-0.899	-0.824
	(0.078)	(0.078)	(0.041)	(0.095)	(0.099)	(0.078)
variance (level-1)	7.157	6.337	7.158	7.357	N/A	7.161
	(0.243)			(0.190)		(0.168)
variance (level-2)	0.139	0.148	0.161	0.172	N/A	0.138
	(0.019)			(0.027)		(0.020

Table 2.11: Comparison of results for second model fitted to MEPS data

Model 2: Results for unweighted analyses

Coefficients	**MLevel**	**HLM**	**SAS**	**MLWiN**	**GLIM**	**Mplus**
intercept	4.335 (0.067)	4.335 (0.085)	4.335 (0.067)	4.335 (0.067)	4.160 (0.089)	4.338 (0.086)
race	0.821 (0.048)	0.821 (0.101)	0.821 (0.048)	0.821 (0.048)	0.810 (0.122)	0.819 (0.101)
GENDER	0.904 (0.036)	0.904 (0.035)	0.904 (0.036)	0.904 (0.036)	0.906 (0.038)	0.905 (0.035)
variance (level-1)	7.765 (0.072)	7.765	7.765	7.765 (0.072)	N/A	7.768 (0.197)
variance (level-2)	0.275 (0.042)	0.224	0.274	0.275 (0.042)	N/A	0.276 (0.048)

Model 2: Results for weighted analyses

Coefficients	MLevel	HLM	SAS	MLWiN	GLIM	Mplus
intercept	4.210 (0.085)	4.211 (0.085)	4.184 (0.062)	4.220 (0.082)	4.119 (0.091)	4.209 (0.085)
race	1.108 (0.080)	1.108 (0.080)	1.105 (0.048)	1.044 (0.086)	1.134 (0.084)	1.108 (0.081)
GENDER	0.894 (0.037)	0.894 (0.037)	0.894 (0.035)	0.896 (0.040)	0.892 (0.037)	0.894 (0.037)
variance (level-1)	7.275 (0.240)	6.442	7.276	7.482 (0.207)	N/A	7.279 (0.177)
variance (level-2)	0.161 (0.023)	0.148	0.191	0.205 (0.034)	N/A	0.160 (0.024)

Table 2.12: Comparison of results for third model fitted to MEPS data

Model 3: Results for unweighted analyses

Coefficients	MLevel	HLM	SAS	MLWiN	GLIM	Mplus
intercept	5.143 (0.064)	5.144 (0.095)	5.144 (0.064)	5.144 (0.063)	5.066 (0.078)	5.148 (0.096)
race	0.692 (0.049)	0.692 (0.095)	0.692 (0.049)	0.692 (0.049)	0.660 (0.116)	0.689 (0.095)
inscov	-0.817 (0.041)	-0.817 (0.082)	-0.817 (0.041)	-0.817 (0.041)	-0.995 (0.124)	-0.819 (0.081)
variance (level-1)	7.843 (0.073)	7.844	7.843	7.843 (0.072)	N/A	7.843 (0.184)
variance (level-2)	0.221 (0.036)	0.223	0.221	0.221 (0.036)	N/A	0.221 (0.036)

Model 3: Results for weighted analyses

Coefficients	MLevel	HLM	SAS	MLWiN	GLIM	Mplus
intercept	5.000 (0.089)	5.001 (0.090)	4.987 (0.060)	5.012 (0.090)	4.945 (0.088)	4.999 (0.090)
race	0.952 (0.081)	0.952 (0.081)	0.948 (0.048)	0.905 (0.083)	0.967 (0.086)	0.952 (0.081)
inscov	-0.810 (0.079)	-0.810 (0.079)	-0.813 (0.042)	-0.818 (0.096)	-0.887 (0.100)	-0.812 (0.079)
variance (level-1)	7.361 (0.249)	6.517	7.361	7.562 (0.192)	N/A	7.365 (0.169)
variance (level-2)	0.137 (0.019)	0.146	0.160	0.172 (0.027)	N/A	0.136 (0.020)

For all the weighted analyses, a comparison of the multilevel (MLevel) and Mplus results show that the estimated parameters and standard errors (given in parentheses) are almost identical, with the exception of the standard error estimate for the level-1 variance component.

2.5.6 Comparison of results for simulated 4-level data

Table 2.13: Bias and Coverage for simulated four-level data (weighted analysis)

Parameter	True Value	0% missing		10% missing		20% missing	
		Bias	Coverage(%	Bias	Coverage(%	Bias	Coverage(%
β_1	150	0.837	85.92	0.874	87.84	1.069	89.42
β_2	1.5	0.233	93.27	0.209	92.27	0.229	94.99
β_3	-0.5	0.018	93.27	-0.038	95.00	0.003	94.85
σ^2	400	1.821	93.37	2.330	94.43	3.273	93.45
$\phi_{2(11)}$	970	8.454	94.82	8.146	95.57	10.81	96.24
$\phi_{2(21)}$	270	11.36	92.13	12.10	93.52	11.17	94.01
$\phi_{2(22)}$	850	12.47	92.75	15.24	94.89	12.97	94.85
$\phi_{2(31)}$	200	2.446	94.93	2.533	95.80	1.291	94.43
$\phi_{2(32)}$	300	6.244	93.79	7.023	94.32	6.672	94.99
$\phi_{2(33)}$	910	20.41	88.82	17.74	90.80	21.16	92.20
$\phi_{3(11)}$	50	0.169	94.82	-0.197	95.11	0.973	97.63
$\phi_{3(21)}$	-10	0.351	94.51	0.157	94.20	1.040	94.99
$\phi_{3(22)}$	35	0.399	95.03	0.778	97.50	2.931	98.61
$\phi_{3(31)}$	9	-0.257	94.41	0.266	94.89	-0.213	95.82
$\phi_{3(32)}$	25	0.596	95.96	1.013	95.45	1.507	96.80
$\phi_{3(33)}$	40	0.054	95.45	0.993	94.09	0.795	96.94
$\phi_{4(11)}$	15	-0.309	94.51	0.448	98.98	2.008	100.00
$\phi_{4(21)}$	8	-0.467	95.55	0.211	95.80	0.882	97.49
$\phi_{4(22)}$	19	-1.069	91.51	-1.201	91.48	0.001	96.80
$\phi_{4(31)}$	6	-0.117	96.27	0.305	96.82	1.163	96.10
$\phi_{4(32)}$	12	-0.626	90.99	-0.249	93.86	0.374	95.96
$\phi_{4(33)}$	24	-1.032	91.30	-0.582	91.02	-0.496	93.45

A comparison of the results for the unweighted analyses reveals differences in standard error estimates. In this case, the MLevel, SAS, and MLWiN standard error estimates are in close agreement, while those produced by HLM and Mplus are the same. In Section 2.6.4 it is shown that robust standard error estimates can be obtained in the unweighted case if the command WEIGHT1 = intcept is included in the syntax file. In doing so, the LISREL MLevel method yields the standard errors reported in the Mplus column.

Table 2.14: Bias and Coverage for simulated four-level data (unweighted analysis)

Parameter	True Value	0% missing		10% missing		20% missing	
		Bias	Coverage(%	Bias	Coverage(%	Bias	Coverage(%
β_1	150	1.185	76.95	1.246	79.86	1.444	82.27
β_2	1.5	0.025	93.98	-0.008	93.63	0.023	95.57
β_3	-0.5	0.047	93.25	-0.009	94.99	0.037	94.60
σ^2	400	8.625	63.24	8.910	76.45	9.629	80.89
$\phi_{2(11)}$	970	19.484	92.11	20.081	93.63	22.782	95.01
$\phi_{2(21)}$	270	6.363	94.08	6.987	94.65	6.292	93.91
$\phi_{2(22)}$	850	17.228	91.38	20.532	93.40	18.980	93.49
$\phi_{2(31)}$	200	2.938	95.12	3.210	95.79	2.152	94.46
$\phi_{2(32)}$	300	5.539	94.29	6.153	94.31	5.917	94.88
$\phi_{2(33)}$	910	20.562	88.68	18.058	90.67	21.454	91.97
$\phi_{3(11)}$	50	0.428	94.81	-0.035	94.88	1.221	97.23
$\phi_{3(21)}$	-10	0.155	94.70	-0.047	94.08	1.050	94.74
$\phi_{3(22)}$	35	0.591	95.33	0.793	97.16	2.836	98.75
$\phi_{3(31)}$	9	-0.283	94.60	0.163	95.22	-0.225	96.26
$\phi_{3(32)}$	25	0.586	96.16	0.973	95.45	1.553	96.68
$\phi_{3(33)}$	40	0.014	95.43	1.031	94.08	0.828	96.81
$\phi_{4(11)}$	15	-0.209	94.60	0.516	98.75	2.099	100.00
$\phi_{4(21)}$	8	-0.493	95.53	0.203	95.79	0.773	97.51
$\phi_{4(22)}$	19	-1.022	91.80	-1.111	91.47	0.057	97.37
$\phi_{4(31)}$	6	-0.131	96.16	0.337	96.81	1.105	96.12
$\phi_{4(32)}$	12	-0.608	91.07	-0.204	93.74	0.303	96.12
$\phi_{4(33)}$	24	-1.025	91.38	-0.518	91.24	-0.613	93.21

2.5.7 Comparison of results for simulated 5-level data

Table 2.15: Bias and Coverage for simulated five-level data (weighted analysis)

Parameter	True Value	0% missing		10% missing		20% missing	
		Bias	Coverage(%	Bias	Coverage(%	Bias	Coverage(%
β_1	150	-0.238	93.73	-0.077	94.69	-0.076	93.63
β_2	1.5	-0.001	93.33	0.007	94.80	-0.033	93.36
β_3	-0.5	0.025	94.74	0.037	92.78	-0.050	93.09
σ^2	400	1.720	90.60	2.493	91.93	3.166	93.89
$\phi_{2(11)}$	975	10.184	93.02	12.767	94.69	10.737	93.89
$\phi_{2(21)}$	175	1.637	95.85	2.685	93.84	1.566	96.02
$\phi_{2(22)}$	850	1.437	95.45	3.241	94.27	3.980	94.42
$\phi_{2(31)}$	120	0.271	95.96	-1.103	94.90	-0.968	95.09
$\phi_{2(32)}$	100	0.924	94.94	1.087	95.01	1.060	94.02
$\phi_{2(33)}$	910	1.859	95.05	2.450	95.97	4.336	94.95
$\phi_{3(11)}$	50	-2.059	93.93	-2.135	95.12	1.223	96.81
$\phi_{3(21)}$	-10	-0.164	94.94	0.205	94.27	1.621	94.56
$\phi_{3(22)}$	35	0.332	96.66	-0.947	95.33	0.229	98.54
$\phi_{3(31)}$	9	-0.485	94.94	0.252	94.27	1.261	94.16
$\phi_{3(32)}$	15	-0.104	94.03	0.135	96.28	0.050	95.48
$\phi_{3(33)}$	40	-0.098	95.15	0.343	94.06	-0.100	95.88
$\phi_{4(11)}$	30	0.039	95.35	0.368	94.37	-1.686	93.49
$\phi_{4(21)}$	14	-0.016	94.74	-0.125	94.37	0.162	96.68
$\phi_{4(22)}$	15	-0.267	94.74	-0.331	95.65	1.321	98.67
$\phi_{4(31)}$	9	-0.200	94.34	0.143	95.44	-0.100	95.48
$\phi_{4(32)}$	16	-0.320	94.14	0.183	93.84	-0.536	94.29
$\phi_{4(33)}$	62	-0.167	93.83	0.026	94.90	-1.289	94.02
$\phi_{5(11)}$	15	-1.006	92.52	-0.230	94.90	0.338	99.20
$\phi_{5(21)}$	13	-0.676	91.71	0.114	93.63	0.194	95.75
$\phi_{5(22)}$	19	-0.692	90.39	-0.701	89.92	-1.135	90.44
$\phi_{5(31)}$	8	-0.282	94.64	0.335	95.44	-0.138	94.69
$\phi_{5(32)}$	8	-0.187	93.12	-0.157	92.57	-0.875	94.95
$\phi_{5(33)}$	24	-1.049	90.39	-1.813	89.49	-0.756	94.42

Table 2.16: Bias and Coverage for simulated five-level data (unweighted analysis)

Parameter	True Value	0% missing		10% missing		20% missing	
		Bias	Coverage(%	Bias	Coverage(%	Bias	Coverage(%
β_1	150	1.224	73.84	1.388	74.95	1.346	82.46
β_2	1.5	0.030	93.54	0.047	95.20	0.010	92.93
β_3	-0.5	0.017	95.05	0.042	92.75	-0.071	93.59
σ^2	400	9.139	23.74	9.818	41.58	10.26	66.36
$\phi_{2(11)}$	975	37.39	67.37	40.33	75.80	36.98	87.70
$\phi_{2(21)}$	175	5.385	94.14	6.391	92.22	5.496	95.68
$\phi_{2(22)}$	850	18.98	82.83	20.83	87.74	21.38	90.31
$\phi_{2(31)}$	120	2.632	95.25	1.379	94.88	1.176	94.76
$\phi_{2(32)}$	100	3.028	94.24	3.153	94.99	3.086	94.24
$\phi_{2(33)}$	910	20.39	77.58	20.91	85.71	22.74	87.70
$\phi_{3(11)}$	50	0.051	93.84	0.617	96.16	4.250	97.51
$\phi_{3(21)}$	-10	-0.453	95.45	0.042	94.24	1.323	94.50
$\phi_{3(22)}$	35	1.039	96.67	-0.365	96.27	0.630	98.17
$\phi_{3(31)}$	9	-0.282	95.15	0.599	94.67	1.615	94.76
$\phi_{3(32)}$	15	0.234	94.65	0.462	96.59	0.170	95.29
$\phi_{3(33)}$	40	0.807	95.45	1.180	94.88	0.548	96.07
$\phi_{4(11)}$	30	0.745	95.35	0.642	94.78	-1.193	94.50
$\phi_{4(21)}$	14	0.283	94.65	0.165	95.52	0.491	96.60
$\phi_{4(22)}$	15	0.067	94.75	0.079	96.48	1.637	98.69
$\phi_{4(31)}$	9	-0.015	94.65	0.171	95.20	0.202	96.07
$\phi_{4(32)}$	16	-0.011	94.55	0.507	93.82	-0.142	95.55
$\phi_{4(33)}$	62	1.049	94.85	1.245	95.84	0.165	93.59
$\phi_{5(11)}$	15	-0.688	92.93	0.226	95.42	0.589	99.61
$\phi_{5(21)}$	13	-0.438	92.12	0.385	94.03	0.303	95.94
$\phi_{5(22)}$	19	-0.304	91.21	-0.400	90.51	-0.836	91.75
$\phi_{5(31)}$	8	-0.121	94.75	0.561	95.52	-0.090	94.63
$\phi_{5(32)}$	8	-0.021	93.43	-0.026	93.18	-0.853	94.90
$\phi_{5(33)}$	24	-0.567	91.11	-1.300	91.04	-0.509	94.76

2.6 Theory

2.6.1 Introduction

In Section 2.6.2, we outline a general procedure for the implementation of weights in level-2 and level-3 models. A more rigorous theoretical treatment of these results are presented in Section 2.6.3. In Section 2.6.4 we provide some results for standard error estimation and fit statistics.

2.6.2 A general weighting procedure

Under the assumption that the sampling weights at a specific level are independent of the random effects at that level, Pfeffermann *et. al.* (1997) adopted the following procedure. Consider the case of a 2 level model. Denote by w_i the weight attached to the *i*-th level-2 unit and by $w_{j|i}$ the weight attached to the *j*-th level-1 unit within the *i*-th level-2 unit such that

$$\sum_j w_{j|i} = n_i, \sum_i w_i = I$$

where I is the total number of level-2 units and $N = \sum_i n_i$ the total number of level-1 units. That is, the lower level weights within each immediate higher level unit are scaled to have a mean of unity, and likewise for higher levels. For each level-1 unit we now form the final, or composite, weight

$$w_{ji} = Nw_{j|i}w_i \Big/ \sum_j \sum_i w_{j|i}w_i = Nw_{j|i}w_i \Big/ \sum_i n_i w_i .$$

Denote by $\mathbf{z}_u$ and $\mathbf{z}_e$ respectively the sets of explanatory variables defining the level-2 and level-1 random coefficients and form

$$\mathbf{z}_u^* = \mathbf{W}_i \mathbf{z}_u, \mathbf{W}_i = Diag\left\{w_i^{-0.5}\right\}$$
$$\mathbf{z}_e^* = \mathbf{W}_{ji} \mathbf{z}_e, \mathbf{W}_{ji} = Diag\left\{w_{ji}^{-0.5}\right\}.$$

We now carry out a standard estimation but using $\mathbf{z}_u^*$ and $\mathbf{z}_e^*$ as the random coefficient explanatory variables. For a 3 level model, with an obvious extension to notation, we have the following

$$\sum_j w_{j|ik} = n_{ik}, \sum_i w_{i|k} = I_k, \sum_k w_k = K, N = \sum_i \sum_k n_{ik}, I = \sum_k I_k$$
$$w_{jik} = Nw_{j|ik}w_{i|k}w_k \Big/ \sum_j \sum_i \sum_k w_{j|ik}w_{i|k}w_k , w_{ik} = Iw_{i|k}w_k \Big/ \sum_i \sum_k w_{i|k}w_k .$$

Goldstein (1995) also pointed out that in survey work analysts often have access only to the final level-1 weights w_{ji}. In this case, say for a 2-level model, we can obtain the w_i by computing $w_i^{'} = W_i I \Big/ \sum_i W_i , W_i = \left(\sum_j w_{ji} \right) \Big/ n_i$.

For a 3-level model the procedure is carried out for each level-3 unit and the resulting $w_{ik}^{'}$ are transformed analogously.

2.6.3 Weights in multilevel models

Let

$$\mathbf{y}_i = \mathbf{X}_{(f)i}\boldsymbol{\beta} + \mathbf{X}_{(2)i}\mathbf{u}_i + \sum_{j=1}^{n_i} \mathbf{Z}_{(1)ij}\mathbf{e}_{ij}, \quad i = 1, 2, ..., I$$

with $\mathbf{y}_i : (n_i \times 1)$, $\mathbf{X}_{(f)i} : n \times p$, $\mathbf{X}_{(2)i} : n_i \times q$, and $\mathbf{Z}_{(1)ij} : n_i \times r$. It is further assumed that $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_I$ are independently and identically distributed (i.i.d.) with $E(\mathbf{u}_i) = \mathbf{0}, \quad Cov(\mathbf{u}_i) = \boldsymbol{\Phi}_{(2)}$. Also, $\mathbf{e}_{11}, ..., \mathbf{e}_{1n_1}, \mathbf{e}_{21}, ..., \mathbf{e}_{In_I}$ are i.i.d. with $E(\mathbf{e}_{ij}) = \mathbf{0}, \quad Cov(\mathbf{e}_{ij}) = \boldsymbol{\Phi}_{(1)}$. Note further that

$$Z_{(1)ij} = \begin{bmatrix} \mathbf{0}' \\ \mathbf{0}' \\ \mathbf{x}'_{(1)ij} \\ \mathbf{0}' \\ \vdots \\ \mathbf{0}' \end{bmatrix}.$$

Example (r = 1):

Suppose that $\mathbf{x}'_{(1)ij} = 1, i = 1, 2, ..., I, \ j = 1, 2, ..., n_i$. In this case, $Cov\left[\sum_{j=1}^{n_i} \mathbf{Z}_{(1)ij}\mathbf{e}_{ij}\right] = \sigma^2 \mathbf{I}_{n_i}$, where $\sigma^2 = \Phi_{(1)}$, a scalar. From the distributional assumptions given above, it follows that

$$E(\mathbf{y}_i) = \mathbf{X}_{(f)i}\boldsymbol{\beta}, \ Cov(\mathbf{y}_i, \mathbf{y}'_i) = \boldsymbol{\Sigma}_i$$

where

$$\boldsymbol{\Sigma}_i = \mathbf{X}_{(2)i}\boldsymbol{\Phi}_{(2)}\mathbf{X}'_{(2)i} + \sum_{j=1}^{n_i} \mathbf{Z}_{(1)ij}\boldsymbol{\Phi}_{(1)}\mathbf{Z}'_{(1)ij}.$$

Consider the case where r = 1, then $\Phi_{(1)} = \sigma_e^2$ and $\sum_{j=1}^{n_i} \mathbf{Z}_{(1)ij}\boldsymbol{\Phi}_{(1)}\mathbf{Z}'_{(1)ij} = \mathbf{D}_i\sigma^2$.

If $\mathbf{Z}_{(1)ij} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_{(1)ij} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, then $\mathbf{Z}_{(1)ij}\mathbf{Z}_{(1)ij}^{'} = \begin{bmatrix} 0 & & & & \\ & \ddots & & & \\ & & x_{(1)ij}^2 & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix}$,

and hence

$$\sum_{j=1}^{n_i} \mathbf{Z}_{(1)ij}\sigma^2\mathbf{Z}_{(1)ij}^{'} = \sigma^2 diag(x_{(1)i1}^2, \ldots, x_{(1)in_i}^2) = \mathbf{D}_i\sigma^2 .$$

Let $\mathbf{V}_i$ be a provisional estimate of $\mathbf{\Sigma}_i$, then

$$\boldsymbol{\beta} \mid \mathbf{V}_i = \left[\sum_{i=1}^{I} \mathbf{X}_{(f)i}^{'}\mathbf{V}_i^{-1}\mathbf{X}_{(f)i}\right]^{-1}\left[\sum_{i=1}^{I} \mathbf{X}_{(f)i}^{'}\mathbf{V}_i^{-1}\mathbf{y}_i\right],$$

where

$$\begin{aligned} \mathbf{V}_i &= \mathbf{X}_{(2)i}\hat{\mathbf{\Phi}}_{(2)}\mathbf{X}_{(2)i}^{'} + \mathbf{D}_i\hat{\sigma}^2 \\ &= \hat{\sigma}^2\left(\mathbf{X}_{(2)i}\frac{\hat{\mathbf{\Phi}}_{(2)}}{\hat{\sigma}^2}\mathbf{X}_{(2)i}^{'} + \mathbf{D}_i\right). \end{aligned}$$

Using a well-known result for matrix inversion,

$$\left[\mathbf{B\Omega B}^{'} + \mathbf{\Lambda}\right]^{-1} = \mathbf{\Lambda}^{-1} - \mathbf{\Lambda}^{-1}\mathbf{B}\left[\mathbf{\Omega}^{-1} + \mathbf{B}^{'}\mathbf{\Lambda}^{-1}\mathbf{B}\right]^{-1}\mathbf{B}^{'}\mathbf{\Lambda}^{-1},$$

it follows that

$$\mathbf{V}_i^{-1} = \left(\hat{\sigma}^2\right)^{-1}\left\{\mathbf{D}_i^{-1} - \mathbf{D}_i^{-1}\mathbf{X}_{(2)i}\left(\hat{\sigma}^2\mathbf{\Phi}_{(2)}^{-1} + \mathbf{X}_{(2)i}^{'}\mathbf{D}_i^{-1}\mathbf{X}_{(2)i}\right)^{-1}\mathbf{X}_{(2)i}^{'}\mathbf{D}_i^{-1}\right\}.$$

Hence

$$\mathbf{X}_{(f)i}^{'}\mathbf{V}_i^{-1}\mathbf{X}_{(f)i} = \left(\hat{\sigma}^2\right)^{-1}\left\{\mathbf{T}_{1i} - \mathbf{T}_{2i}\left[\hat{\sigma}^2\hat{\mathbf{\Phi}}_{(2)}^{-1} + \mathbf{T}_{3i}\right]^{-1}\mathbf{T}_{2i}^{'}\right\},$$

where

$$\begin{aligned} \mathbf{T}_{1i} &= \mathbf{X}_{(f)i}^{'}\mathbf{D}_i^{-1}\mathbf{X}_{(f)i} = \sum_{j=1}^{n_i}\mathbf{x}_{(f)ij}\mathbf{x}_{(f)ij}^{'} / x_{(1)ij}^2 \\ \hat{\mathbf{T}}_{1i} &= \sum_{j=1}^{n_i} w_{j|i}\mathbf{x}_{(f)ij}\mathbf{x}_{(f)ij}^{'} / x_{(1)ij}^2 \\ \hat{\mathbf{T}}_1 &= \sum_{i=1}^{I} w_i \sum_{j=1}^{n_i} w_{j|i}\mathbf{x}_{(f)ij}\mathbf{x}_{(f)ij}^{'} / x_{(1)ij}^2 . \end{aligned}$$

Since $w_{ij} = w_i \cdot w_{j|i}$, it follows that

$$\hat{\mathbf{T}}_i = \sum_{i=1}^{I} \sum_{j=1}^{n_i} \mathbf{x}_{(f)ij} \mathbf{x}'_{(f)ij} / x^{*2}_{(1)ij},$$

where $x^*_{(1)ij} = w_{ij}^{-1/2} x_{(1)ij}$.

$$\mathbf{T}_{2i} = \sum_{j=1}^{n_i} \mathbf{x}_{(f)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij}$$

$$\hat{\mathbf{T}}_{2i} = \sum_{j=1}^{n_i} w_{j|i} \mathbf{x}_{(f)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij}$$

$$\mathbf{T}_{3i} = \sum_{j=1}^{n_i} \mathbf{x}_{(2)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij}$$

$$\begin{aligned}\hat{\mathbf{T}}_{3i} &= \sum_{j=1}^{n_i} (w_i \cdot w_i^{-1}) w_{j|i} \mathbf{x}_{(2)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij} \\ &= \sum_{j=1}^{n_i} w_i^{-1} w_{ij} \mathbf{x}_{(2)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij} \\ &= \sum_{j=1}^{n_i} \mathbf{x}^*_{(2)ij} \mathbf{x}^{*'}_{(2)ij} / x^{*2}_{(1)ij}\end{aligned}$$

where $\mathbf{x}^*_{(2)ij} = w_i^{-1/2} \mathbf{x}_{(2)ij}$.

Let

$$\hat{\mathbf{C}}_{(2)i} = \left[\hat{\sigma}^2 \hat{\mathbf{\Phi}}^{-1}_{(2)} + \hat{\mathbf{T}}_{3i} \right]^{-1}$$

then $\sum_{i=1}^{I} \mathbf{T}_{2i} \mathbf{C}_{(2)i} \mathbf{T}'_{2i}$ is estimated by

$$\begin{aligned}\sum_{i=1}^{I} w_i \hat{\mathbf{T}}_{2i} \hat{\mathbf{C}}_{(2)i} \hat{\mathbf{T}}'_{2i} = \sum_{i=1}^{I} w_i^{1/2} \left[\sum_{j=1}^{n_i} w_i^{-1/2} w_i^{-1/2} w_i w_{j|i} \mathbf{x}_{(f)ij} \mathbf{x}'_{(2)ij} / x^2_{(1)ij} \right] \times \\ \hat{\mathbf{C}}_{(2)i} w_i^{1/2} \left[\sum_{j=1}^{n_i} w_i^{-1/2} w_i^{-1/2} w_i w_{j|i} \mathbf{x}_{(2)ij} \mathbf{x}'_{(f)ij} / x^2_{(1)ij} \right].\end{aligned}$$

Since $w_i w_{j|i} = w_{ij}$, it follows that $\sum_{i=1}^{I} \mathbf{T}_{2i} \mathbf{C}_{(2)i} \mathbf{T}'_{2i}$ is estimated by

$$\sum_{i=1}^{I} \left[\sum_{j=1}^{n_i} \mathbf{x}_{(f)ij} \mathbf{x}^{*'}_{(2)ij} / x^{*2}_{(1)ij} \right] \hat{\mathbf{C}}_{(2)i} \left[\sum_{j=1}^{n_i} \mathbf{x}^*_{(2)ij} \mathbf{x}'_{(f)ij} / x^{*2}_{(1)ij} \right].$$

Similarly

$$\mathbf{X}'_{(f)i}\mathbf{V}_i^{-1}\mathbf{y}_i = \left(\hat{\sigma}^2\right)^{-1}\left[\mathbf{X}'_{(f)i}\mathbf{D}_i^{-1}\mathbf{y}_i - \mathbf{X}'_{(f)i}\mathbf{D}_i^{-1}\mathbf{X}_{(2)i}\mathbf{C}_{(2)i}\mathbf{X}'_{(2)i}\mathbf{D}_i^{-1}\mathbf{y}_i\right]$$
$$= \left(\hat{\sigma}^2\right)^{-1}\left[\mathbf{t}_{4i} - \mathbf{T}_{2i}\mathbf{C}_{(2)i}\mathbf{t}_{5i}\right]$$

where

$$\mathbf{t}_{4i} = \mathbf{X}'_{(f)i}\mathbf{D}_i^{-1}\mathbf{y}_i = \sum_{j=1}^{n_i}\mathbf{x}_{(f)ij}y_{ij} / x^2_{(1)ij},$$
$$\hat{\mathbf{t}}_{4i} = \sum_{j=1}^{n_i} w_i^{-1}w_i w_{i|j}\mathbf{x}_{(f)ij}y_{ij} / x^2_{(1)ij} = w_i^{-1}\sum_{j=1}^{n_i}\mathbf{x}_{(f)ij}y_{ij} / x^{*2}_{(1)ij}.$$
$$\mathbf{t}_{5i} = \mathbf{X}'_{(2)i}\mathbf{D}_i^{-1}\mathbf{y}_i = \sum_{j=1}^{n_i}\mathbf{x}_{(2)ij}y_{ij} / x^2_{(1)ij}$$
$$\hat{\mathbf{t}}_{5i} = w_i^{-1/2}\sum_{j=1}^{n_i}\mathbf{x}^*_{(2)ij}y_{ij} / x^{*2}_{(1)ij}.$$

It can then be shown that $\sum_{i=1}^{I}\mathbf{X}'_{(f)i}\mathbf{V}_i^{-1}\mathbf{y}_i$ is estimated by $\left(\hat{\sigma}^2\right)^{-1}\left[\hat{\mathbf{t}}_{4i} - \sum_{i=1}^{I}\hat{\mathbf{q}}_i\right]$ where

$$\hat{\mathbf{q}}_i = \left[\sum_{j=1}^{n_i}\mathbf{x}_{(f)ij}\mathbf{x}^{*'}_{(2)ij} / x^{*2}_{(1)ij}\right]\cdot\hat{\mathbf{C}}_{(2)i}\cdot\left[\sum_{j=1}^{n_i}\mathbf{x}^*_{(2)ij}y_{ij} / x^{*2}_{(1)ij}\right]$$

and

$$\hat{\mathbf{t}}_4 = \sum_{i=1}^{I} w_i\mathbf{t}_{4i} = \sum_{i=1}^{I}\sum_{j=1}^{n_i}\mathbf{x}_{(f)ij}y_{ij} / x^{*2}_{(1)ij}.$$

2.6.4 Standard errors and fit statistics

The method used in LISREL to calculate standard error estimates for multilevel models depends on whether design weights are included in the analysis or not. We first consider the case where design weights are used.

Let $\hat{\boldsymbol{\gamma}}$ denote the vector of estimated parameters. In Section 8 of the *Complex Survey Sampling Guide* it was shown that an approximate expression for the asymptotic covariance matrix of $\hat{\boldsymbol{\gamma}}$ is given by

$$Cov(\hat{\boldsymbol{\gamma}}) \approx \mathbf{I}_n^{-1}(\boldsymbol{\gamma})\mathbf{G}\mathbf{I}_n^{-1}(\boldsymbol{\gamma})$$

where

$$E\left[\frac{\partial^2 \ln L}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}\right] = -\mathbf{I}_n(\boldsymbol{\gamma}).$$

As an estimate of $\mathbf{G}$ we use

$$\mathbf{G} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{g}_i\mathbf{g}_i',$$

where $\mathbf{g}_i$ denotes the i-th contribution to the gradient vector $\mathbf{g} = \frac{\partial \ln L}{\partial\boldsymbol{\gamma}}$.

Standard error estimates of the unknown parameters are obtained by taking the square roots of the diagonal elements of $Cov\left(\hat{\boldsymbol{\gamma}}\right)$.

If no weighting variable is specified, it is assumed that $Cov\left(\hat{\boldsymbol{\gamma}}\right)$ equals the inverse of the information matrix, that is $Cov\left(\hat{\boldsymbol{\gamma}}\right) = \mathbf{I}_n^{-1}$. In the case where no weight is specified, so-called robust standard error estimates of the estimated parameters may be obtained by using the asymptotic covariance matrix for the weighted case. This is accomplished by adding the command WEIGHT1 = intcept; to the multilevel syntax file.

Likelihood ratio tests

Test of a null hypothesis against a restricted alternative hypothesis can be constructed, provided that two conditions are met. Firstly, the models under H_0 and H_1 should be estimable and secondly, the parameter space $\boldsymbol{\Omega}_0$ for H_0 must be a subset of the parameter space $\boldsymbol{\Omega}$ for H_1.

Use is made of the likelihood ratio test statistic

$$\lambda = \frac{L_0\left(\hat{\gamma}_0\right)}{L_1\left(\hat{\gamma}_1\right)}$$

where L_0 and L_1 denote the likelihood functions under H_0 and H_1 respectively. For a large N (see, for example, Anderson, 1984), $-2\ln\lambda = -2\ln L_0 - (-2\ln L_1)$ has an approximate $\chi^2_{(v)}$ distribution where the number of degrees of freedom v is the difference in the number of parameters estimated under H_0 and the number of parameters estimated under H_1. The statistic $-2\ln L$ is called the deviance.

Contrasts

Consider a clinical trial in which two types of drugs are administered to 400 obese adults. Adults are randomly assigned to four groups:

- Group 1, Drug A, low dosage (10 mg/day)
- Group 2, Drug A, high dosage (50 mg/day)
- Group 3, Drug B, low dosage (10 mg/day)
- Group 4, Drug B, high dosage (50 mg/day)

Let y_{ij} denote weight loss of subject *i* on occasion t_j, *i* = 1, 2, …, 400 and *j* = 1, 2, …, n_i, and let

$$\begin{aligned} y_{ij} &= \beta_1 AL + \beta_2 AH + \beta_3 BL + \beta_4 BH + \beta_5 TIJ + \beta_6 AGE \\ &+ \beta_7 GENDER + \beta_8 INITW + u_{1i} + TIJ \times u_{2i} + e_{ij} \end{aligned}$$

where AL, AH, BL and BH are dummy variables, coded as follows

	AL	**AH**	**BL**	**BH**
Drug A, low dosage	1	0	0	0
Drug A, high dosage	0	1	0	0
Drug B, low dosage	0	0	1	0
Drug B, high dosage	0	0	0	1

In the above model $\beta_1, \beta_2, \beta_3,$ and β_4 represent the average group loss (or gain) in weight over the study period if we control for a subject's age (AGE), gender (GENDER), weight at the onset of the trial (INITW), and time (TIJ) at which the weight loss (y_{ij}) measurement was made.

Visual inspection of the estimated β-coefficients may point to significant differences between the different treatments. The construction of contrasts or linear functions of the parameters is a useful statistical analysis tool and enables the researcher to perform hypothesis testing concerning the equality of subsets of parameters.

In the example above, the fixed part of the model has 8 parameters $\beta_1, \beta_2, \ldots, \beta_8$. We may want to test the following 3 hypotheses:

$$H_{01}: \beta_1 = \beta_2$$
$$H_{02}: \beta_1 = \beta_3$$
$$H_{03}: \beta_1 = \beta_4.$$

Each of these hypotheses can alternatively be written as

$$H_{01}: 1\beta_1 - 1\beta_2 + 0\beta_3 + 0\beta_4 + 0\beta_5 + 0\beta_6 + 0\beta_7 + 0\beta_8 = 0$$
$$H_{02}: 1\beta_1 + 0\beta_2 - 1\beta_3 + 0\beta_4 + 0\beta_5 + 0\beta_6 + 0\beta_7 + 0\beta_8 = 0$$
$$H_{03}: 1\beta_1 + 0\beta_2 + 0\beta_3 - 1\beta_4 + 0\beta_5 + 0\beta_6 + 0\beta_7 + 0\beta_8 = 0$$

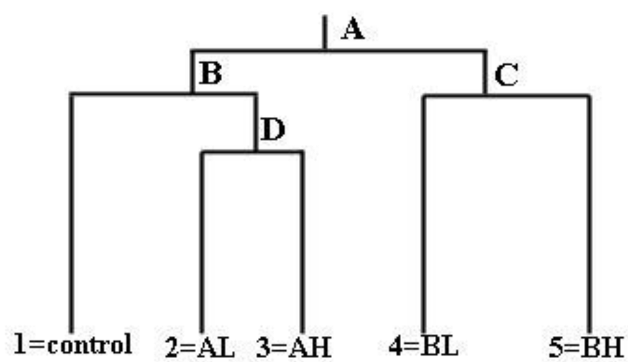
or, in matrix notation,

$$H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0},$$

where

$$\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Suppose that an additional 100 subjects (the control group) are also assigned to the experiment, but each subject from this group receives a placebo. Suppose further that the 5 treatments are hypothesized to be related as described by the tree diagram



Here we can form the orthogonal contrasts:

Contrast		**Treatments** 1	2	3	4	5	TIJ	AGE	GENDER	INITW
	A	1/3	1/3	1/3	-1/2	-1/2	0	0	0	0
	B	1	-1/2	-1/2	0	0	0	0	0	0
	C	0	0	0	1	-1	0	0	0	0
	D	0	1	-1	0	0	0	0	0	0

A complex hypothesis about several elements of the vector of fixed coefficients $\boldsymbol{\beta}$ can be tested if use is made of a $p \times m$ contrast matrix **C**, with *p* the number of contrasts and *m* the number of fixed coefficients. The hypothesis is written in the form

$$\mathbf{C}\boldsymbol{\beta} = \mathbf{k}$$

where **k** is a known vector, usually **k** = **0**.

For large samples (see *e.g.* du Toit, 1993), $\mathbf{C}\hat{\boldsymbol{\beta}}$ has an approximate $N\left(\mathbf{C}\boldsymbol{\beta}, \mathbf{C}\boldsymbol{\Gamma}^{-1}\mathbf{C}'\right)$ distribution, where $\boldsymbol{\Gamma} = Cov(\hat{\boldsymbol{\beta}})$. If the hypothesis $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{k}$ is true, it follows (see, *e.g.* Anderson (2003)), that

$$U = \left(\mathbf{C}\boldsymbol{\beta} - \mathbf{k}\right)' \left[\mathbf{C}\boldsymbol{\Gamma}^{-1}\mathbf{C}'\right]^{-1} \left(\mathbf{C}\boldsymbol{\beta} - \mathbf{k}\right)$$

follows an approximate χ^2-distribution with *p* degrees of freedom.

A set of $100\,(1-\alpha)\%$ simultaneous confidence intervals for the *p* elements of $\mathbf{C}\boldsymbol{\beta}$ is given by the *p* intervals

$$\mathbf{c}_i'\hat{\boldsymbol{\beta}} \pm \left[\mathbf{c}_i'\boldsymbol{\Gamma}^{-1}\mathbf{c}_i\chi^2_{m,\alpha}\right]^{0.5},$$

where $p \le m$, $\mathbf{c}_i'$ denotes the *i*-th row of **C** and $\chi^2_{m,\alpha}$ is the critical value of the χ^2 distribution with *m* degrees of freedom.

3 References

Anderson, T.W. (2003). *An introduction to multivariate statistical analysis*. Third edition, New York: Wiley & Sons.

Asparouhov, T. (2004). Weighting for Unequal Probability of Selection in Latent Variable Modeling. Mplus Web Notes: No. 7.

Du Toit, S.H.C. (1993). *Analysis of multilevel models. Part 1: Theoretical aspects*. Pretoria, South Africa: Human Sciences Research Council.

Du Toit, M. & du Toit, S.H.C. (2001). *Interactive LISREL guide.* Lincolnwood, IL: Scientific Software International.

Goldstein, H. (1995). *Multilevel Statistical Models*. London, Edward Arnold: New York, Wiley.

Jöreskog, K.G. & Sörbom, D. (2004). *LISREL for Windows [Computer Software]*. Lincolnwood, IL: Scientific Software International, Inc.

Jöreskog, K.G., Sörbom, D., du Toit, S.H.C. & du Toit, M. (2000). *LISREL 8: New Statistical Features*. Lincolnwood, IL: Scientific Software International, Inc.

Pfeffermann, D., Skinner, C. J., Holmes, D., Goldstein, H., *et al*. (1997). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society, B.*, **60**, 23-40.