HLM 8

SSI SCIENTIFIC SOFTWARE INTERNATIONAL

HIERARCHICAL LINEAR & NONLINEAR MODELING

Stephen Raudenbush Anthony Bryk Richard Congdon

HLM 8

Hierarchical Linear and Nonlinear Modeling

Stephen W. Raudenbush
University of Chicago
Anthony S. Bryk
The Carnegie Foundation
Yuk Fai Cheong
Emory University
Richard T. Congdon, Jr.
Harvard University
Mathilda du Toit
Scientific Software International, Inc.

SSI SCIENTIFIC SOFTWARE INTERNATIONAL

Published by:

Scientific Software International, Inc.
P.O. Box 4728,
Skokie, IL 60076-4728
Tel: +1.847.675.0720
Fax: +1.847.675.2140
URL: http://www.ssicentral.com

Preface

New Program Features in HLM 8 for Windows

Estimating HLM from incomplete data

- A completely automated approach that generates and analyzes multiply imputed data sets from incomplete data.
- The model is fully multivariate and enables the analyst to strengthen imputation through auxiliary variables.

Here the user specifies the HLM; the program automatically searches the data to discover which variables have missing values; it then estimates a multivariate hierarchical linear model "imputation model" in which all variables having missing values are regressed on all variables having complete data; it then uses the resulting parameter estimates to generate *M* imputed data sets; it then analyzes each of these in turn and combines the results using the "Rubin rules."

Flexible Combinations of Fixed Intercepts and Random Coefficients

- Included in HLM2, HLM3, HLM4, HCM2, and HCM3.
- Two-level examples: a) a longitudinal study with fixed child effects and random treatment effect; and b) a study in which children are randomly assigned to treatments within pre-school centers with fixed center intercepts and a random coefficient for treatment.
- A three-level study in which children are nested within classrooms within schools; we have fixed school intercepts and a randomly varying treatment effect and randomly varying classroom intercepts.

A concern that can arise in multilevel causal studies is that random effects may be correlated with treatment assignment. For example, suppose that treatments are assigned non-randomly to students who are nested within schools. Estimating a two-level model with random school intercepts will generate bias if the random intercepts are correlated with treatment effects. The conventional strategy is to specify a fixed effects model for schools. However, this approach assumes homogeneous treatment effects, leading possibly to biased estimates of the average treatment effect, incorrect standard errors, and inappropriate interpretations. Our tools allow the analyst to combine fixed intercepts with random coefficients in models that address these problems and to facilitate a richer summary including an estimate of the variation of treatment effects and empirical Bayes estimates of unit-specific treatment effects. This approach was proposed in Bloom, Raudenbush, Weiss, and Porter (2017).

Contents

1 Conceptual and Statistical Background for Two-Level Models

Behavioral and social data commonly have a nested structure. For example, if repeated observations are collected on a set of individuals and the measurement occasions are not identical for all persons, the multiple observations are properly conceived as nested within persons. Each person might also be nested within some organizational unit such as a school or workplace. These organizational units may in turn be nested within a geographical location such as a community, state, or country. Within the hierarchical linear model, each of the levels in the data structure (*e.g.*, repeated observations within persons, persons within communities, communities within states) is formally represented by its own sub-model. Each sub-model represents the structural relations occurring at that level and the residual variability at that level.

This manual describes the use of the HLM computer program for the statistical modeling of two-, three- and four-level data structures, respectively. It should be used in conjunction with the text *Hierarchical Linear Models: Applications and Data Analysis Methods* (Raudenbush, S.W. & Bryk, A.S., 2002: Newbury Park, CA: Sage Publications)[1]. The HLM programs have been tailored so that the basic program structure, input specification, and output of results closely coordinate with this textbook. This manual also cross-references the appropriate sections of the textbook for the reader interested in a full discussion of the details of parameter estimation and hypothesis testing. Many of the illustrative examples described in this manual are based on data distributed with the program and analyzed in the Sage text.

We begin by discussing the two-level model below and the use of the HLM2 program in Chapter 2. Building on this framework, Chapters 3 and 4 introduce the three-level model and the use of the HLM3 program. The four-level model and the use of the HLM4 program are discussed in Chapters 5 and 6. Chapters 7 and 8 discuss use of hierarchical modeling for non-normal level-1 errors. Chapters 9 and 10 consider multivariate models that can be estimated from incomplete data. Chapter 11 describes several special features of HLM2 and HLM3, including analyses involving latent variables, multiply-imputed data, and known level-1 variances, as well as the procedure for graphing data and equations. Chapters 12 and 13 introduce two-level cross-classified random effects that are applicable for analyses of models that do not have a strictly hierarchical data structure, and Chapters 14 and 15 discuss three-level cross-classified random effects models. Hierarchical linear models with cross-classified random effects are considered in Chapters 16 and 17. Chapter 18 illustrates HLM's ability to produce data- and model-based graphs. Flexible combinations of Fixed Intercepts and Random Coefficients (FIRC) is introduced in Chapter 19. In Chapter 20, a completely automated approach that generates and analyzes multiply imputed data sets from incomplete data is discussed.

[1] Also available from SSI.

1.1 The general two-level model

As the name implies, a two-level model consists of two submodels at level 1 and level 2. For example, if the research problem consists of data on students nested within schools, the level-1 model would represent the relationships among the student-level variables and the level-2 model would capture the influence of school-level factors. Formally, there are $i = 1,...,n_j$ level-1 units (*e.g.*, students) nested within $j = 1,...,J$ level-2 units (*e.g.*, schools).

1.1.1 Level-1 model

We represent in the level-1 model the outcome for case i within unit j as:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{Qj}X_{Qij} + r_{ij} \\ &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj}X_{qij} + r_{ij}, \end{aligned} \quad (0.0)$$

where

β_{qj} ($q = 0,1,...,Q$) are *level-1 coefficients*;

X_{qij} is the *level-1 predictor* q for case i in unit j;

r_{ij} is the level-1 random effect; and

σ^2 is the variance of r_{ij}, that is the level-1 variance.

Here we assume that the random term $r_{ij} \sim N(0,\sigma^2)$.

1.1.2 Level-2 model

Each of the level-1 coefficients, β_{qj}, defined in the level-1 model becomes an outcome variable in the level-2 model:

$$\begin{aligned} \beta_{qj} &= \gamma_{q0} + \gamma_{q1}W_{1j} + \gamma_{q2}W_{2j} + \cdots + \gamma_{qS_q}W_{S_q j} + u_{qj} \\ &= \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs}W_{sj} + u_{qj}, \end{aligned} \quad (1.0)$$

where

γ_{qs} ($q = 0,1,...,S_q$) are *level-2 coefficients*;

W_{sj} is a level-2 predictor; and

u_{qj} is a level-2 random effect.

We assume that, for each unit *j*, the vector $\left(u_{0j}, u_{1j}, ..., u_{Qj}\right)'$ is distributed as multivariate normal, with each element of u_{qj} having a mean of zero and variance of

$$Var(u_{qj}) = \tau_{qq}. \quad (0.0)$$

For any pair of random effects q and q',

$$Cov(u_{qj}, u_{q'j}) = \tau_{qq'}. \tag{0.0}$$

These *level-2 variance* and *covariance components* can be collected into a dispersion matrix, $\mathbf{T}$, whose maximum dimension is $(Q+1)\times(Q+1)$.

We note that each level-1 coefficient can be modeled at level-2 as one of three general forms:

1. a *fixed level-1 coefficient*; *e.g.*,

$$\beta_{qj} = \gamma_{q0,} \tag{0.0}$$

2. a *non-randomly varying level-1 coefficient*, *e.g.*,

$$\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj}, \tag{1.0}$$

3. a *randomly varying level-1* coefficient, *e.g.*,

$$\beta_{qj} = \gamma_{q0} + u_{qj} \tag{1.0}$$

or a level-1 coefficient with both non-random and random sources of variation,

$$\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj} + u_{qj} \tag{1.0}$$

The actual dimension of $\mathbf{T}$ in any application depends on the number of level-2 coefficients specified as randomly varying. We also note that a different set of level-2 predictors may be used in each of the $Q+1$ equations of the level-2 model.

1.2 Parameter estimation

Three kinds of parameter estimates are available in a hierarchical linear model: empirical Bayes estimates of randomly varying level-1 coefficients; generalized least squares estimates of the level-2 coefficients; and maximum-likelihood estimates of the variance and covariance components.

1.3 Empirical Bayes ("EB") estimates of randomly varying level-1 coefficients, β_{qj}

These estimates of the level-1 coefficients for each unit j are optimal composites of an estimate based on the data from that unit and an estimate based on data from other similar units. Intuitively, we are borrowing strength from all of the information present in the ensemble of data to improve the level-1 coefficient estimates for each of the J units. These "EB" estimates are also referred to as "shrunken estimates" of the level-1 coefficients. They are produced by HLM as part of the residual file output (see Section 2.5.4, *Model checking based on the residual file*). (For further discussion see *Hierarchical Linear Models*, pp. 45-51; 85-95.)

1.4 Generalized least squares (GLS) estimates of the level-2 coefficients, γ_{qs}

Substitution of the level-2 equations for β_{qj} into their corresponding level-1 terms yields a single-equation linear model with a complex error structure. Proper estimation of the regression coefficients of this model (*i.e.*, the γ's) requires that we take into account the differential precision of the information provided by each of the *J* units. This is accomplished through generalized least squares. (For further discussion see *Hierarchical Linear Models*, pp. 38-44.)

1.5 Maximum likelihood estimates of variance and covariance components

Because of the unbalanced nature of the data in most applications of hierarchical linear models (*i.e.*, n_j varies across the J units and the observed patterns on the level-1 predictors also vary), traditional methods for variance-covariance component estimation fail to yield efficient estimates. Through iterative computing techniques such as the EM algorithm and Fisher scoring, maximum-likelihood estimates for σ^2 and $\mathbf{T}$ can be obtained. (For further discussion, see *Hierarchical Linear Models*, pp. 51-56; also Chapters 13, 14).

1.6 Some other useful statistics

Based on the various parameter estimates discussed above, HLM2 and HLM3 also compute a number of other useful statistics. These include:

1. *Reliability* of $\hat{\beta}_{qj}$.

The program computes an overall or average reliability for the least squares estimates of each level-1 coefficient across the set of J level-2 units. These are denoted in the program output as RELIABILITY ESTIMATES and are calculated according to Equation 3.58 in *Hierarchical Linear Models*, p. 49.

2. *Least squares residuals,*($\hat{u}_{qj}$).

These residuals are based on the deviation of an ordinary least squares estimate of a level-1 coefficient, $\hat{\beta}_{qj}$, from its predicted or "fitted" value based on the level-2 model, *i.e.*,

$$\hat{u}_{qj} = \hat{\beta}_{qj} - \left(\hat{\gamma}_{q0} + \sum_{s=1}^{S_q} \hat{\gamma}_{qs} W_{sj} \right). \tag{0.0}$$

These ordinary least square residuals are denoted in HLM residual files by the prefix OL before the corresponding variable names.

3. *Empirical Bayes residuals* (u^*_{qj})

These residuals are based on the deviation of the empirical Bayes estimates, β^*_{qj}, of a randomly varying level-1 coefficient from its predicted or "fitted" value based on the level-2 model, *i.e.*,

$$u^*_{qj} = \beta^*_{qj} - \left(\hat{\gamma}_{q0} + \sum_{s=1}^{S_q} \hat{\gamma}_{qs} W_{sj} \right). \tag{0.0}$$

These are denoted in the HLM residual files by the prefix EB before the corresponding variable names. (For a further discussion and illustration of OL and EB residuals see *Hierarchical Linear Models*, pp. 47-48; and 76-95).

1.7 Hypothesis testing

Corresponding to the three basic types of parameter estimates based on a hierarchical linear model (EB estimates of random level-1 coefficients, GLS estimates of the fixed level-2 coefficients, and the maximum-likelihood estimates of the variance and covariance components), are single-parameter and multi-parameter hypothesis-testing procedures. (See *Hierarchical Linear Models*, pp. 56-65). The current HLM programs execute a variety of hypothesis tests for the level-2 fixed effects and the variance-covariance components. These are summarized in Table 1.1.

1.8 Restricted versus full maximum likelihood

By default, two-level models are estimated by means of restricted maximum likelihood (REML). Using this approach, the variance-covariance components are estimated via maximum likelihood, averaging over all possible values of the fixed effects. The fixed effects are estimated via GLS given these variance-covariance estimates. Under full maximum likelihood (ML), variance-covariance parameters and fixed level-2 coefficients are estimated by maximizing their joint likelihood (see *Hierarchical Linear Models*, pp. 52-53). One practical consequence is that, under ML, any pair of nested models can be tested using a likelihood ratio test. In contrast, using REML, the likelihood ratio test is available only for testing the variance-covariance parameters, as indicated in Table 1.1.

Table 1.1 Hypothesis tests for the level-2 fixed effects and the variance-covariance components

Type of hypothesis	Test statistic	Program output
Fixed level-2 effects		
Single Parameter: $H_0 : \gamma_{qs} = 0$ $H_1 : \gamma_{qs} \neq 0$	t-ratio[1]	Standard feature of the Fixed Effects Table for all level-2 coefficients
Multi-parameter: $H_0 : C'\gamma = 0$ $H_1 : C'\gamma \neq 0$	general linear hypothesis test (Wald test), chi-square test[2]	Optional output specification (see Section 2.8)

Table 1.1 Hypothesis tests for the level-2 fixed effects and the variance-covariance components (continued)

Type of hypothesis	Test statistic	Program output
Variance-covariance components		
Single-Parameter: $H_0 : \tau_{qq} = 0$ $H_1 : \tau_{qq} > 0$	Chi-square test[3]	Standard feature of the Variance Components Table for all level-2 random effects
Multi-parameter: $H_0 : \mathbf{T} = \mathbf{T}_0$ $H_1 : \mathbf{T} \neq \mathbf{T}_0$	Difference in deviances, likelihood ratio test.[4]	Optional output specification (see Section 2.8)

[1]See Equation 3.83 in *Hierarchical Linear Models*.

[2]See Equation 3.91 in *Hierarchical Linear Models*.

[3]See Equation 3.103 in *Hierarchical Linear Models*.

[4]Here $\mathbf{T}_0$ is a reduced form of $\mathbf{T}_1$.

1.9 Generalized Estimating Equations

Statistical inferences about the fixed level-2 coefficients, γ_{qs}, using HLM are based on the assumption that random effects at each level are normally distributed; and on the assumed structure of variation and covariation of these random effects at each level. Given a reasonably large sample of level-2 units, it is possible to make sound statistical inferences about γ_{qs} that are not based on these assumptions by using the method of generalized estimating equations or "GEE" (Zeger & Liang, 1986). Comparing these GEE inferences to those based on HLM provides a way of assessing whether the HLM inferences about γ_{qs} are sensitive to the violations of these assumptions. The simplest GEE model assumes that the outcome Y_{ij} for case i in unit j is independent of the outcome $Y_{i'j}$ for some other case, i', in the same unit; and that these outcomes have constant variance. Under these simple assumptions, estimation of the γ coefficients by ordinary least squares (OLS) would be justified. If these OLS assumptions are incorrect, the OLS estimates of γ_{qs} will be consistent (accurate in large samples) but not efficient. However, the standard error estimates produced under OLS will generally be inconsistent (biased, often badly, even in large samples).

Version 7 of HLM produces the following tables, often useful for comparative purposes:

- A table of OLS estimates along with the OLS standard errors.
- A table including the OLS estimates, but accompanied by robust standard errors, that is, standard errors that are consistent even when the OLS assumptions are incorrect.
- A table of HLM estimates of γ_{qs}, based on GLS, and standard errors based on the assumptions underlying HLM.
- A table of the same HLM estimates, but now accompanied by robust standard errors, that is, standard errors that are consistent even when the HLM assumptions are mistaken.

By comparing these four tables, it is possible a) to discern how different the HLM estimates and standard errors are from those based on OLS; and b) to discern whether the HLM inferences are plausibly distorted by incorrect assumptions about the distribution of the random effects at each level. We illustrate the value of these comparisons in Chapter 2 (for further discussion, see *Hierarchical Linear Models*, pp. 276-280). The GEE approach is very useful for strengthening inferences about the fixed level-2 coefficients but does not provide a basis for inferences about the random, level-1 coefficients or the variance-covariance components. Cheong, Fotiu, and Raudenbush (2001) have intensively studied the properties of HLM and GEE estimators in the context of three-level models. GEE results are also available for three-level data.

2 Working with HLM2

Data analysis by means of the HLM2 program will typically involve three stages:

1. construction of the "MDM file" (the multivariate data matrix);
2. execution of analyses based on the MDM file; and
3. evaluation of fitted models based on a residual file.

We describe each stage below and then illustrate a number of special options. Data collected from a High School & Beyond (HS&B) survey on 7,185 students nested within 160 US high schools, as described in Chapter 4 of *Hierarchical Linear Models*, will be used for demonstrations.

2.1 Constructing the MDM file from raw data

We assume that a user has employed a standard computing package to clean the data, make necessary transformations, and conduct relevant exploratory and descriptive analyses. We also recommend exploratory graphical analyses within HLM prior to model building as described in detail in Section 18.1 of this manual.

The first task in using HLM2 is to construct the Multivariate Data Matrix (MDM) from raw data or from a statistical package. We generally work with two raw data files: a level-1 file and a level-2 file. Both files must be sorted by the level-2 ID (It is possible, however, to build the MDM file from the level-1 file above, though this option is not suggested when the level-1 file is very large. The level-1 file must be sorted by level-2 ID. The level-1 file name will be selected as both the level-1 and level-2 file).

For the HS&B example, the level-1 units are students and the level-2 units are schools. The two files are linked by a common level-2 unit ID, school id in our example, which must appear on every level-1 record. In constructing the MDM file, the HLM program will compute summary statistics based on the level-1 unit data and store these statistics together with level-2 data.

The procedure to create a MDM file consists of three major steps. The user needs to

- Inform HLM of the input and MDM file type.
- Supply HLM with the appropriate information for the data, the command and the MDM files.
- Check if the data have been properly read into HLM.

2.2 Executing analyses based on the MDM file

Once the MDM file is constructed, all subsequent analyses will be computed using the MDM file as input. It will therefore be unnecessary to read the larger student-level data file in computing these analyses. The efficient summary of data in the MDM file leads to faster computation. The MDM file is like a "system file" in a standard computing package in that it contains not only the summarized data but also the names of all of the variables.

Model specification has three steps:

- Specifying the level-1 model, which defines a set of level-1 coefficients to be computed for each level-2 unit.
- Specifying a level-2 structural model to predict each of the level-1 coefficients.
- Specifying the level-1 coefficients to be viewed as random or non-random.

The output produced from these analyses includes:

- Ordinary least squares and generalized least squares results for the fixed coefficients defined in the level-2 model.
- Estimates of variance and covariance components and approximate chi-square tests for the variance components.
- A variety of auxiliary diagnostic statistics.

Additional output options and hypothesis-testing procedures may be selected.

2.3 Model checking based on the residual file

After fitting a hierarchical model, it is wise to check the tenability of the assumptions underlying the model:

- Are the distributional assumptions realistic?
- Are results likely to be affected by outliers or influential observations?
- Have important variables been omitted or non-linear relationships been ignored?

These questions and others can be addressed by means of analyses of the HLM residual files. A level-1 residual file includes:

- The level-1 residuals (discrepancies between the observed and fitted values).
- Fitted values (FV) for each level-1 unit (that is, values predicted on the basis of the model).
- The observed values of all predictors included in the model.
- Selected level-2 predictors useful in exploring possible relationships between such predictors and level-1 residuals.

A level-2 residual file includes:

- Fitted values for each level-1 coefficient (that is, values predicted on the basis of the level-2 model).
- Ordinary least squares (OL) and empirical Bayes (EB) estimates of level-2 residuals (discrepancies between level-1 coefficients and fitted values).
- Empirical Bayes coefficients, which are the sum of the EB estimates and the fitted values.
- Dispersion estimates useful in exploring sources of variance heterogeneity at level 1.
- Expected and observed Mahalanobis distance measures useful in assessing the multivariate normality assumption for the level-2 residuals.
- Selected level-2 predictors useful in exploring possible relationships between such predictors and level-2 residuals.

- Posterior variances (PV).

For HLM2 FML analyses, there is an additional set of posterior variances. See Chapter 9 in *Hierarchical Linear Models* for a full discussion of these methods.

2.4 Windows, interactive, and batch execution

Formulation and testing of models using HLM programs can be achieved via Windows, interactive, or batch modes. Most PC users will find the Windows mode preferable. This draws on the visual features of Windows while preserving the speed of use associated with a command-oriented (batch) program. Non-PC users have the choice of interactive and batch modes only. Interactive execution guides the user through the steps of the analysis by posing questions and providing a menu of options. In this chapter, we employ the Windows mode for all the examples. Descriptions and examples on how to use HLM2 in interactive and batch modes are given in Appendix A.

2.5 An example using HLM2 in Window mode

Chapter 4 in *Hierarchical Linear Models* presents a series of analyses of data from the HS&B survey. A level-1 model specifies the relationship between student socioeconomic status (SES) and mathematics achievement in each of 160 schools; at level-2, each school's intercept and slope are predicted by school sector (Catholic versus public) and school mean social class. We reproduce one analysis here (see Table 4.5 in *Hierarchical Linear Models*, p. 82).

2.5.1 Constructing the MDM file from raw data

PC users may construct the MDM file directly from different types of input files including SPSS, ASCII, SAS, SYSTAT, and STATA, or indirectly from many additional types of data file formats through the third-party software module included in the HLM program.

Non-PC users may construct the MDM file with one of the following types of input files: ASCII data files, SYSTAT data files, or SAS V5 transport files.

In order for the program(s) to correctly read the data, the IDs need to conform to the following rules:

1. For ASCII data the ID variables must be read in as character (alphanumeric). These IDs are indicated by the A field(s) in the format statement. For all other types of data, the ID may be character or numeric.
2. The level-1 cases must be grouped together by their respective level-2 unit ID. To assure this, sort the level-1 file by the level-2 ID field prior to entering the data into HLM2.
3. If the ID is numeric, it must be in the range $-(10^{13}+1)$ to $+(10^{13}+1)$ (*i.e.* 12 digits). Although the ID may be a floating point number, only the integer part is used.
4. If the ID variable is character, the length must not exceed 12 characters. Furthermore, the IDs at a given level must all be the same length. *This is often a cause of problems*. For example, imagine your data has IDs ranging from “1” to “100”. You will need to recreate the IDs as “001” to “100”. In other words, all spaces (blank characters) should be coded as zeros.
5. For non-ASCII files, the program can only properly deal with numeric variables (with the exception of character ID variables). Other data types, such as a “Date format”, will not be processed properly.

6. For non-ASCII files with missing data, one should only use the "standard" missing value code. Some statistical packages (SAS, for example) allow for a number of missing value codes. The HLM modules are incapable of understanding these correctly, thus these additional missing codes need to be recoded to the more common "." (period) code.

2.5.1.1 SPSS file input

We first illustrate the use of SPSS file input and then consider input from ASCII data files. Data input requires a level-1 file and a level-2 file.

Level-1 file. For our HS&B example data, the level-1 file (HSB1.SAV) has 7,185 cases and four variables (not including the SCHOOL ID). The variables are:

- MINORITY, an indicator for student ethnicity (1 = minority, 0 = other)
- FEMALE, an indicator for student gender (1 = female, 0 = male)
- SES, a standardized scale constructed from variables measuring parental education, occupation, and income
- MATHACH, a measure of mathematics achievement

Data for the first ten cases in HSB1.SAV are shown in Fig. 2.1.

Note: level-1 cases must be grouped together by their respective level-2 unit ID. To assure this, sort the level-1 file by the level-2 unit ID field prior to entering the data into HLM2.

	id	minority	female	ses	mathach
1	1224	0	1	-1.528	5.876
2	1224	0	1	-.588	19.708
3	1224	0	0	-.528	20.349
4	1224	0	0	-.668	8.781
5	1224	0	0	-.158	17.898
6	1224	0	0	.022	4.583
7	1224	0	1	-.618	-2.832
8	1224	0	0	-.998	.523
9	1224	0	1	-.888	1.527
10	1224	0	0	-.458	21.521

Figure 2.1 First ten cases in HSB1.SAV

Level-2 file. At level 2, the illustrative data set HSB2.SAV consists of 160 schools with 6 variables per school. The variables are:

- SIZE (school enrollment)
- SECTOR (1 = Catholic, 0 = public)
- PRACAD (proportion of students in the academic track)
- DISCLIM (a scale measuring disciplinary climate)
- HIMNTY (1 = more than 40% minority enrollment, 0 = less than 40%)
- MEANSES (mean of the SES values for the students in this school who are included in the level-1 file)

The data for the first ten schools are displayed in Fig 2.2.

	id	size	sector	pracad	disclim	himinty	meanses
1	1224	842	0	.350	1.597	0	-.428
2	1288	1855	0	.270	.174	0	.128
3	1296	1719	0	.320	-.137	1	-.420
4	1308	716	1	.960	-.622	0	.534
5	1317	455	1	.950	-1.694	1	.351
6	1358	1430	0	.250	1.535	0	-.014
7	1374	2400	0	.500	2.016	0	-.007
8	1433	899	1	.960	-.321	0	.718
9	1436	185	1	1.000	-1.141	0	.569
10	1461	1672	0	.780	2.096	0	.683

Figure 2.2 First ten cases in HSB2.SAV

As mentioned earlier, the construction of an MDM file consists of three major steps. This will now be illustrated with the HS&B example.

To inform HLM of the input and MDM file type

1. At the **WHLM** window, open the **File** menu.
2. Choose **Make new MDM file…Stat package input** (see Figure 2.3). A **Select MDM type** dialog box opens (see Figure 2.4).
3. Select **HLM2** and click **OK**. A **Make MDM - HLM2** dialog box will open (see Figure 2.5).

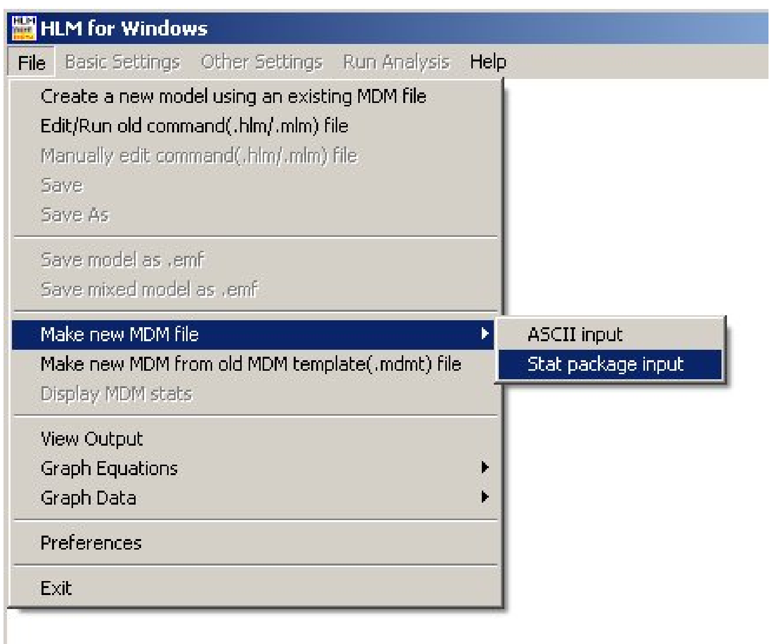


Figure 2.3 WHLM window

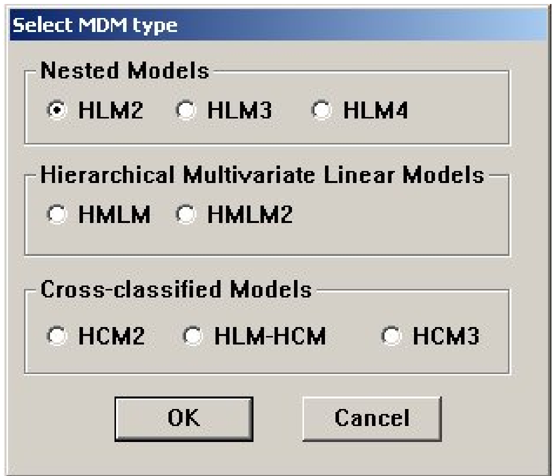


Figure 2.4 Select MDM type dialog box

To supply HLM with appropriate information for the data, the command, and the MDM files:

1. Select **SPSS/Windows** from the **Input File Type** pull-down menu (see Figure 2.5).
2. Specify the structure of data. The three choices are cross-sectional, longitudinal, and measures within groups. The data in HSB1.SAV are cross-sectional.
3. Click **Browse** in the **Level-1 Specification** section to open an **Open Data File** dialog box.
4. Open a level-1 SPSS system file in the HLM folder (HSB1.SAV in our example). The **Choose Variables** button will be activated.

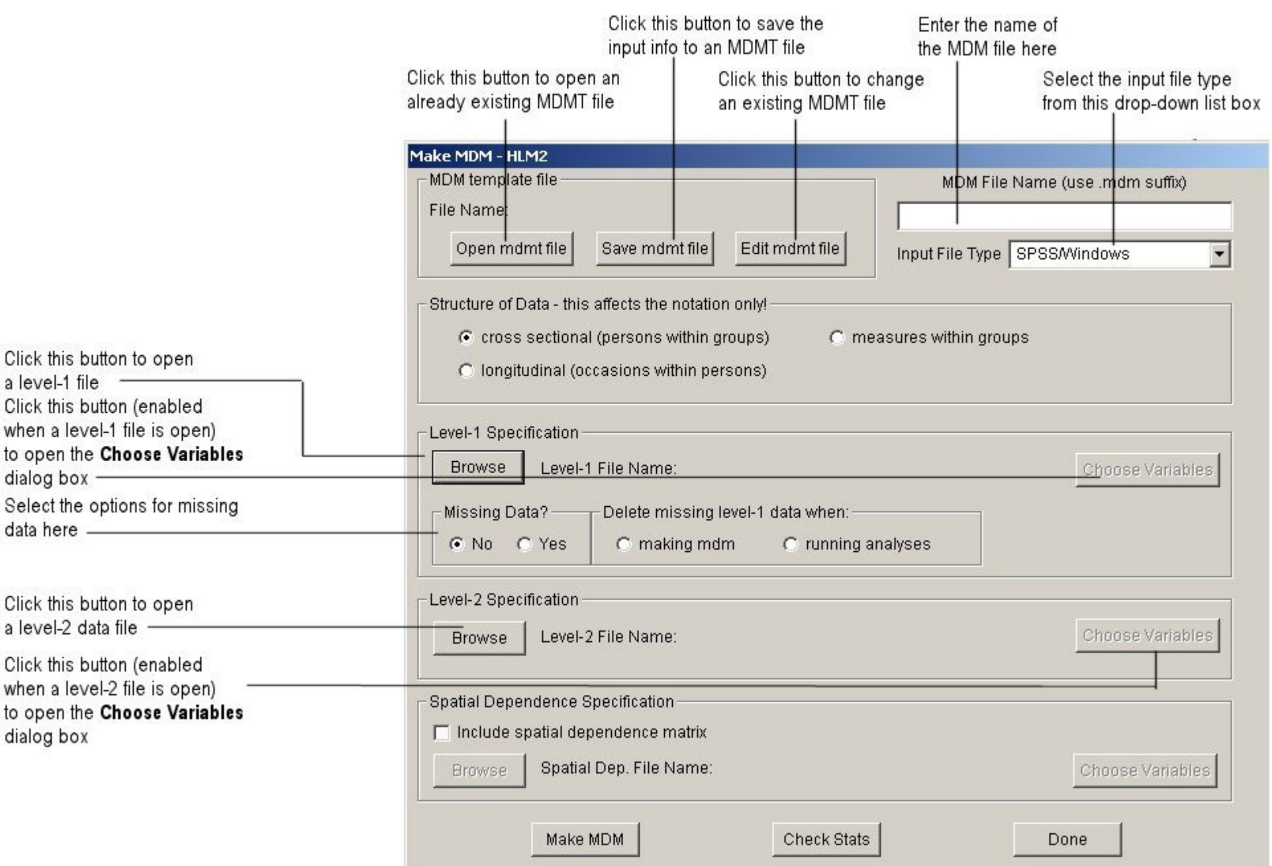


Figure 2.5 Make MDM - HLM2 dialog box

5. Click **Choose Variables** to open the **Choose Variables - HLM2** dialog box and choose the ID and variables by clicking the appropriate check boxes (See Figure 2.6). To deselect, click the box again.
6. Select the options for missing data in the level-1 file (there is no missing data in HSB1.SAV; see Section 2.6 for details).
7. Click the selection button for **measures within persons** for the **type of nesting of input data** if the level-1 data consist of repeated measures or item responses. With this selection, WHLM will use in its displays and output model notations that match those used in *Hierarchical Linear Models* for studies on individual change and latent variables (Chapters 6 and 11). The default type is **persons within groups**. It is generally used when the level-1 data are comprised of cross-sectional measures. With this option, WHLM will use model notations that correspond to those used for applications in organization research (Chapters 4 and 5).
8. Click **Browse** in the **Level-2 specification** section to open an **Open Data File** dialog box.
9. Open a level-2 SPSS system file in the HLM folder (HSB2.SAV in our example). The **Choose Variables** button below **Browse** will be activated.
10. Click **Choose Variables** to open the **Choose Variables - HLM2** dialog box and choose the ID and variables by clicking the appropriate check boxes (see Figure 2.7).
11. Check the box **include spatial dependence matrix** to specify spatial dependence, if applicable (see Section 11.4 for details). The **Spatial Dependence Specification** box should only be used if you have spatial dependence data and wish to run this kind of model.
12. Enter a name for the MDM file in the **MDM file name** box (for example, HSB.MDM).

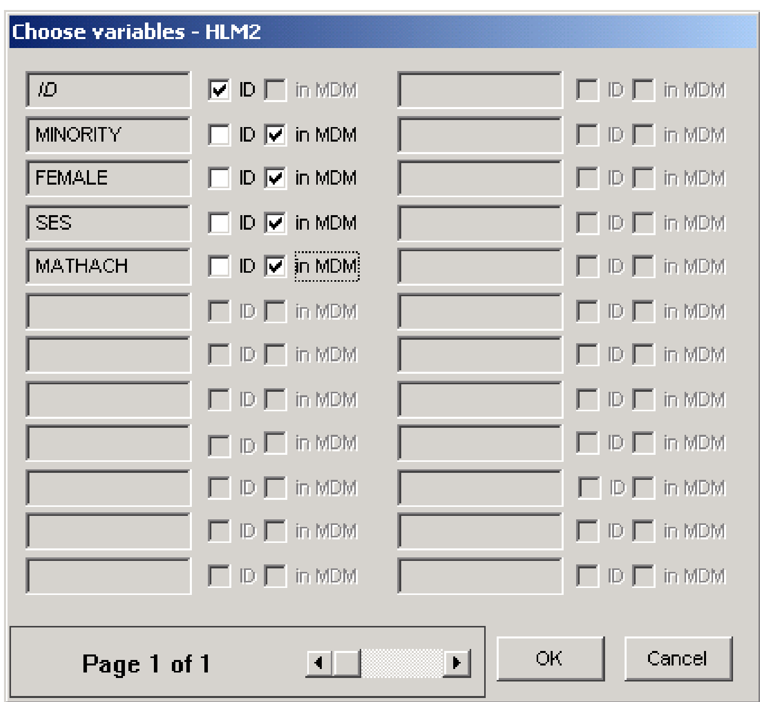


Figure 2.6 Choose Variables - HLM2 dialog box for the level-1 file, HSB1.SAV

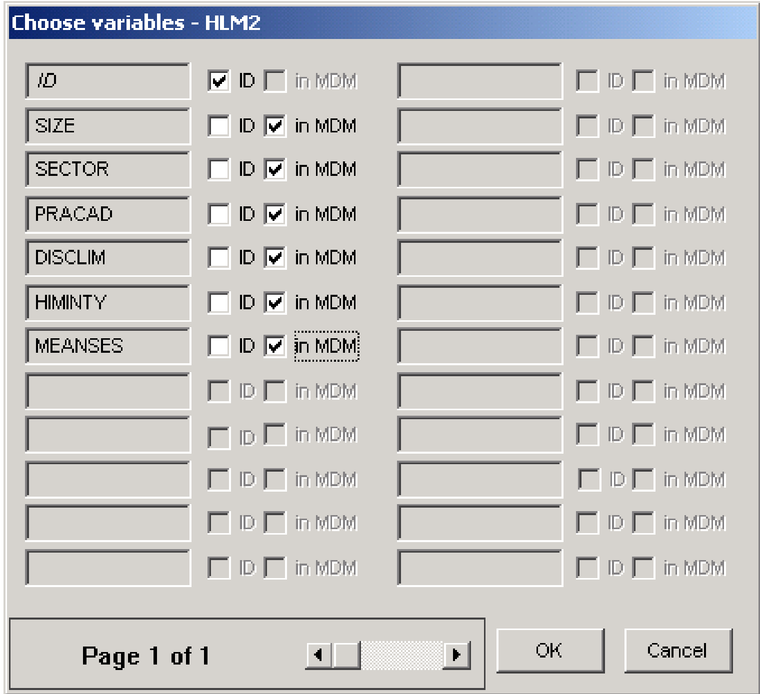
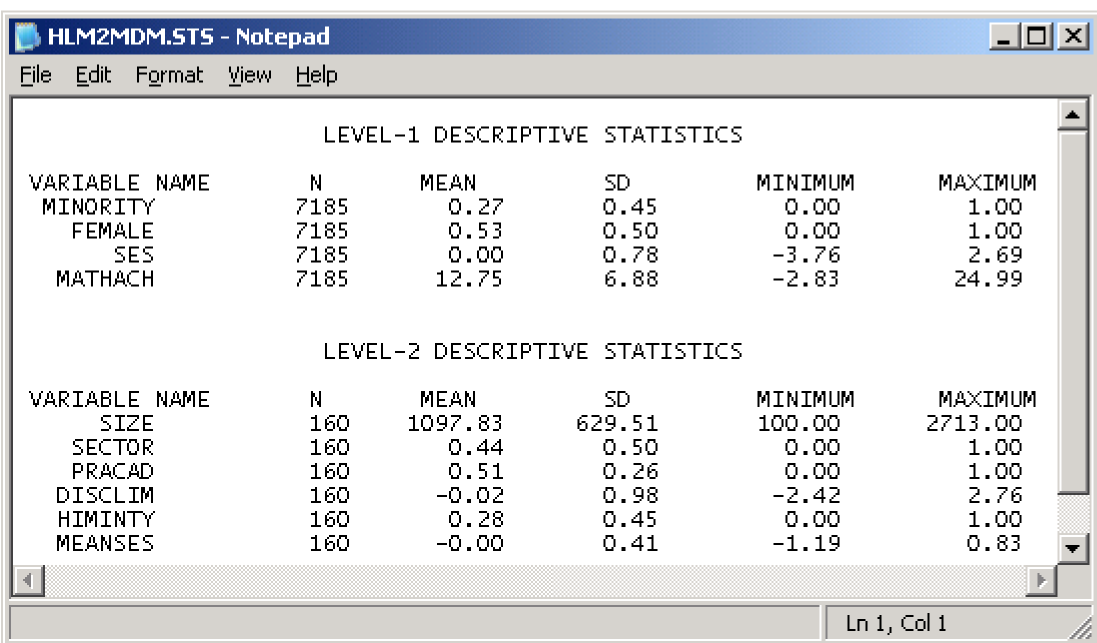


Figure 2.7 Choose variables - HLM2 dialog box for the level-2 file, HSB2.SAV

13. Click **Save mdmt file** in the **MDM template file** section to open a **Save MDM template file** dialog box. Enter a name for the MDMT file (for example, HSBSPSS.MDMT). Click **Save** to save the file. The command file saves all the input information entered by the user. It can be re-opened by clicking the **Open mdmt file** button (see Figure 2.5). To make changes to an existing MDMT file, click the **Edit mdmt file** button.
14. Note that HLM will also save the input information into another file called CREATMDM.MDMT when the MDM is created.
15. Click the **Make MDM** button. A screen displaying the prompts and responses for MDM creation will appear.




```
                  LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME        N       MEAN         SD       MINIMUM     MAXIMUM
 MINORITY         7185       0.27       0.45        0.00        1.00
   FEMALE         7185       0.53       0.50        0.00        1.00
      SES         7185       0.00       0.78       -3.76        2.69
  MATHACH         7185      12.75       6.88       -2.83       24.99


                  LEVEL-2 DESCRIPTIVE STATISTICS

VARIABLE NAME        N       MEAN         SD       MINIMUM     MAXIMUM
     SIZE          160    1097.83     629.51      100.00     2713.00
   SECTOR          160       0.44       0.50        0.00        1.00
   PRACAD          160       0.51       0.26        0.00        1.00
  DISCLIM          160      -0.02       0.98       -2.42        2.76
  HIMINTY          160       0.28       0.45        0.00        1.00
  MEANSES          160      -0.00       0.41       -1.19        0.83
```

Figure 2.8 Descriptive Statistics for the MDM file, HSB.MDM

To check whether the data have been properly read into HLM

3 When the screen disappears, the level-1 and level-2 descriptive statistics will automatically be displayed (See Figure 2.8). Pay particular attention to the N column. It is not an uncommon mistake to forget to sort by the ID variable, which can lead to a lot (or most) of the data not being processed. Close the Notepad window when done. Use the **Save As** option to give it a new name if later use of this file is anticipated. The file can also be opened by clicking on the **Display Stats** button.

4 Click **Done**. The **WHLM** window displays the type and name on its title bar (**hlm2** & **HSB.MDM**) and the level-1 variables on a drop-down menu (See Figure 2.9).

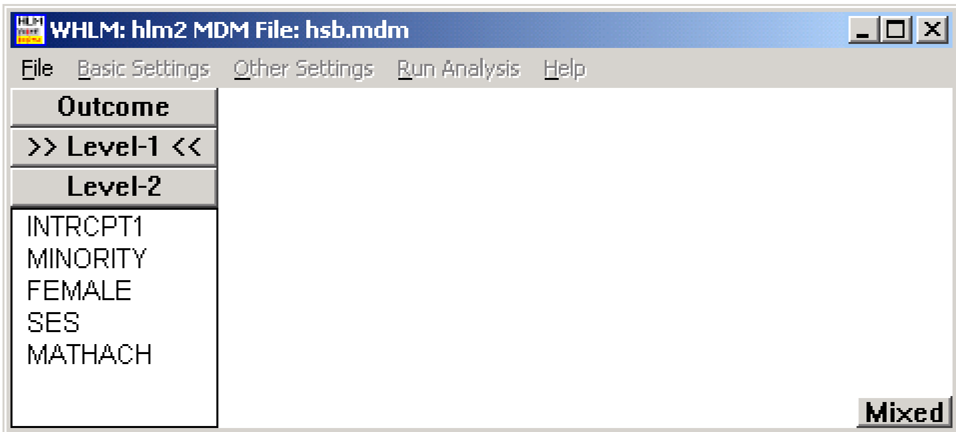


Figure 2.9 WHLM: hlm2 MDM File window for HSB.MDM

2.5.1.2 ASCII file input

Below is the procedure for creating a multivariate data matrix file with input from ASCII files.

To inform HLM of the input and MDM file type

1. At the **WHLM** window, open the **File** menu.
2. Choose **Make new MDM file…ASCII** input. A **Select MDM type** dialog box opens.
3. Select **HLM2** (see Figure 2.4) and click **OK**. A **Make MDM File – HLM2** will open (see Figure 2.10).

To supply HLM with appropriate information for the data, the command, and the MDM files

1. Click **Browse** in the **Level-1 specification** section to open an **Open Data File** dialog box. Open a level-1 ASCII data file in the HLM examples folder (HSB1.DAT in our example). The file name (HSB1.DAT) appears in the **Level-1 File Name** box.
2. Enter the number of variables into the **Number of Variables** box (4 in our example) and the data entry format in the **Data Format** box (A4,4F12.3 in our example).

Note that the ID is included in the format statement, but excluded in the **Number of Variables** box. Rules for input format statements are given in Section A.2 in Appendix A.

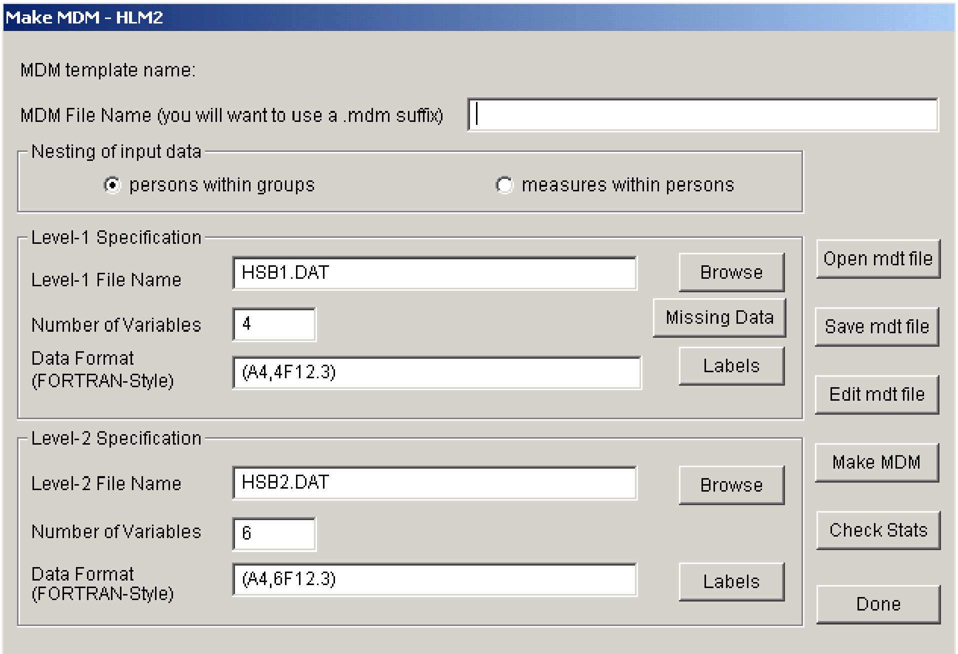


Figure 2.10 Make MDM – HLM2 dialog box

3. Click **Labels** to open the **Enter Variable Labels** dialog box.
4. Enter the variable names into the boxes (MINORITY, FEMALE, SES, MATHACH for our example, see Figure 2.11). Click **OK**.
5. Click the **Missing Data** button to enter level-1 missing data info (there is no missing data in HSB1.DAT; see Section 2.6 for details).
6. Click **Browse** in the **Level-2 specification** section to open an **Open Data File** dialog box. Open a level-2 ASCII data file in the HLM folder (HSB2.DAT in our example). The file name (HSB2.DAT in our example) will appear in the **Level-2 File Name** box.
7. Enter the number of variables into the **Number of Variables** box (6 in our example) and the data entry format in the **Data Format** box (A4,6F12.3 in our example).
8. Click **Labels** to open the **Enter Variable Labels** dialog box for the level-2 variables.
9. Enter the variable names into the **Variable** boxes (SIZE, SECTOR, PRACAD, DISCLIM, HIMINTY, MEANSES in our example, see Figure 2.12). Click **OK**.
10. Enter an MDM file name in the **MDM File Name** box (for example, HSB.MDM).
11. Click **Save mdmt file** in the **MDM template file** section to open a **Save MDM template file** dialog box. Enter a name for the MDMT file (for example, HSBASCII.MDMT). Click **Save** to save the file. The command file saves all the input information entered by the user. It can be re-opened or changed by clicking either the **Open mdmt file** or the **Edit mdmt file** button (see Figure 2.10).

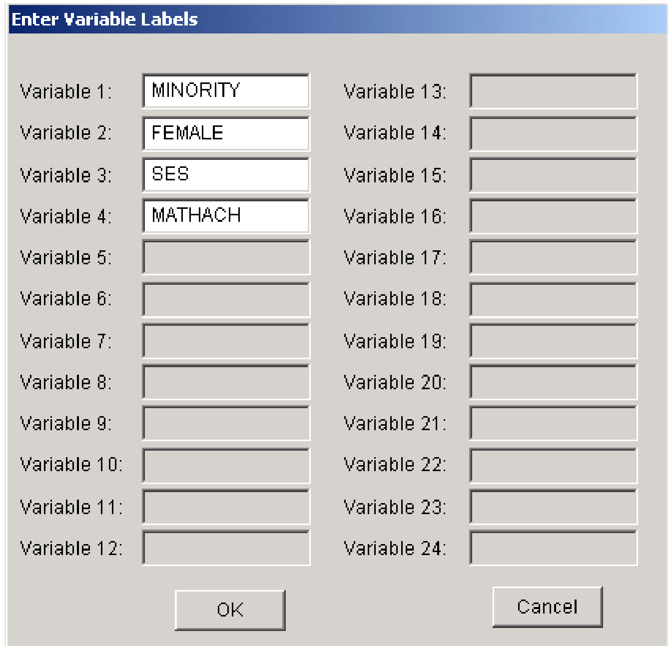


Figure 2.11 Enter Variable Labels dialog box for level-1 file, HSB1.DAT

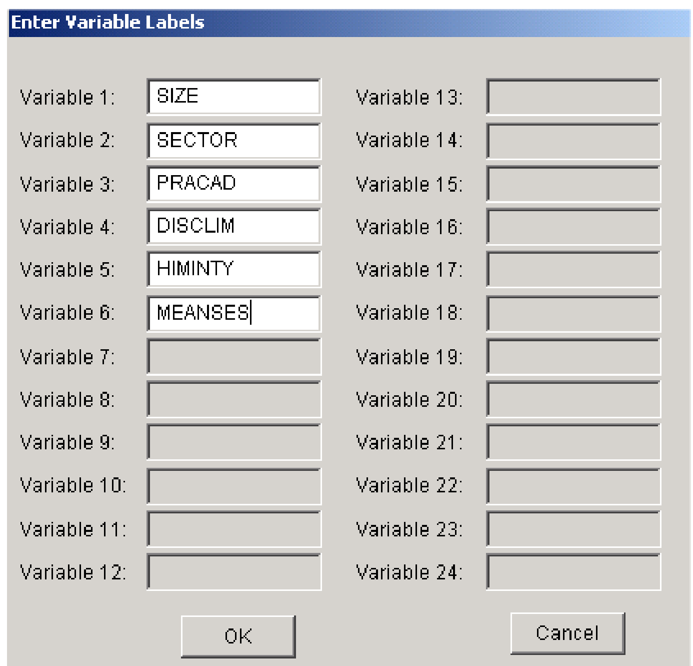


Figure 2.12 Enter Variable Labels dialog box for level-2 file, HSB2.DAT

To check whether the data have been properly read into HLM

The procedure is the same as for SPSS file input (see Section 2.5.1.1 for a complete description).

2.5.1.4 SAS transport, SYSTAT, STATA file input and other formats for raw data

For SAS transport, SYSTAT or STATA file input, a user selects either **SAS 5 transport**, **SYSTAT** or **STATA** from the **Input File Type** drop-down menu as appropriate to open the **Open Data File** dialog box. With the third-party software module included in the current version, HLM will read data from EXCEL, LOTUS and many other formats. Select **Anything else** from the **Input File Type** drop-down menu *before* clicking on the **Browse** button in the input file specifications sections. If the data type is set on the **File**, **Preferences** screen, the program will default to your selected type for both input data and residual files.

2.5.2 Executing analyses based on the MDM file

Once the MDM file is constructed, it can be used as input for the analysis. As mentioned earlier, model specification has three steps:

- Specification of the level-1 model. In our example, we shall model mathematics achievement (MATHACH) as the outcome, to be predicted by student SES. Hence, the level-1 model will have two coefficients: the intercept and the SES-MATHACH slope.
- Specification of the level-2 prediction model. We shall predict each school's intercept by school SECTOR and MEANSES in our example. Similarly, SECTOR and MEANSES will predict each school's SES-MATHACH slope.
- Specification of level-1 coefficients as random or non-random. We shall model both the intercept and the slope as having randomly varying residuals. That is, we are assuming that the intercept and slope vary not only as a function of the two predictors, SECTOR and MEANSES, but also as a function of a unique school effect. The two school residuals (*e.g.*, for the intercept and slope) are assumed sampled from a bivariate normal distribution.

The procedure for executing analyses based on the MDM file is described below.

Step 1: To specify the level-1 prediction model

1. From the **HLM** window, open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open an existing MDM file (HSB.MDM in our example). The name of the MDM file will be displayed on the title bar of the main window. A list box for level-1 variables (**>>Level-1<<**) will appear (see Figure 2.13).
3. Click on the name of the outcome variable (MATHACH in our example). Click **Outcome variable** (see Figure 2.13). The specified model will appear in equation format.

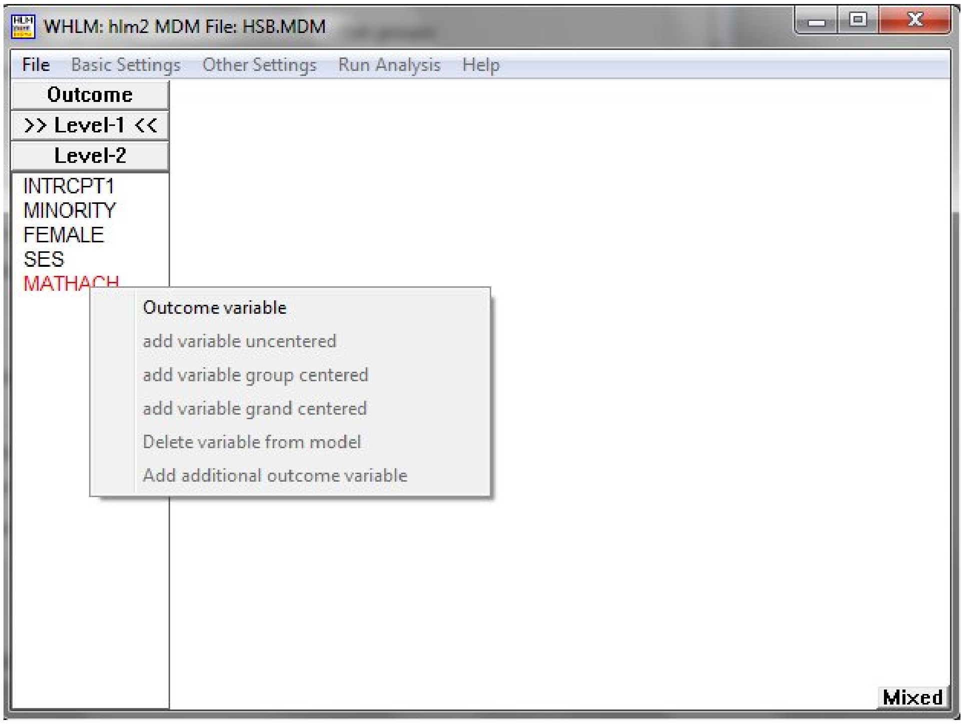


Figure 2.13 Model window for the HS&B example

4. Click on the name of a predictor variable and click the type of centering (SES and **add variable group centered**, see Figure 2.14). The predictor will appear on the equation screen and each regression coefficient associated with it will become an outcome in the Level-2 model (see Figure 2.15).

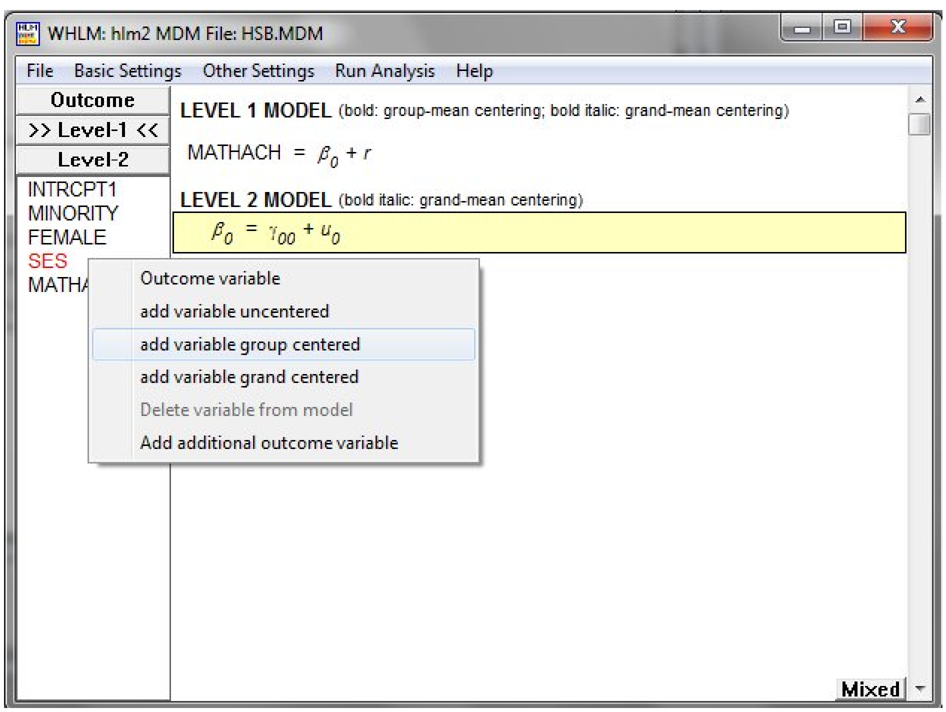


Figure 2.14 Specification of model predictor, SES, for the HS&B example

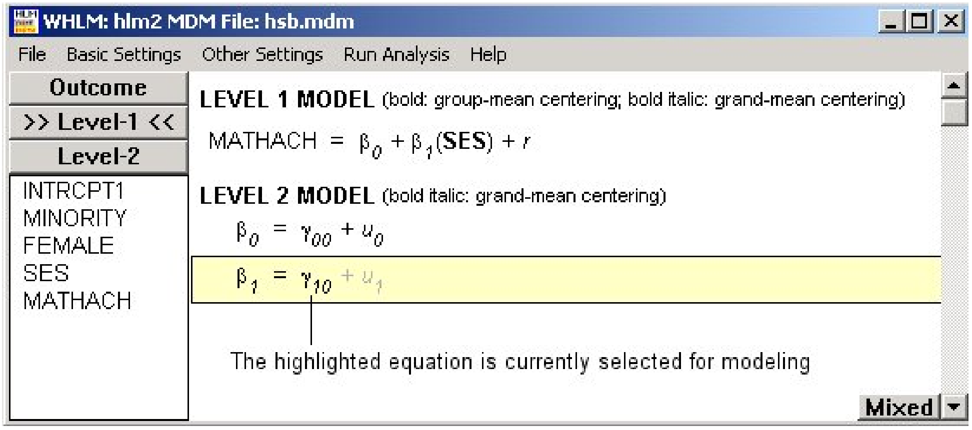


Figure 2.15 Model window for the HS&B example

Step 2: To specify the level-2 prediction model

1. Select the equation containing the regression coefficient(s) to be modeled by clicking on the equation (β_0 (intercept) and β_1 (SES slope) in our HS&B example). A list box for level-2 variables (**>>Level-2<<**) will appear (see Figure 2.16).
2. Click to select the variable(s) to be entered as predictor(s) and the type of centering. For our example, select SECTOR and **add variable uncentered**, and MEANSES and add variable **grand-mean centered** to model β_0 and β_1, see Figure 2.16.
3. HLM allows the model to be displayed in three alternative forms. Figure 2.17 displays the model specified in the default notation familiar to users of previous versions of HLM.

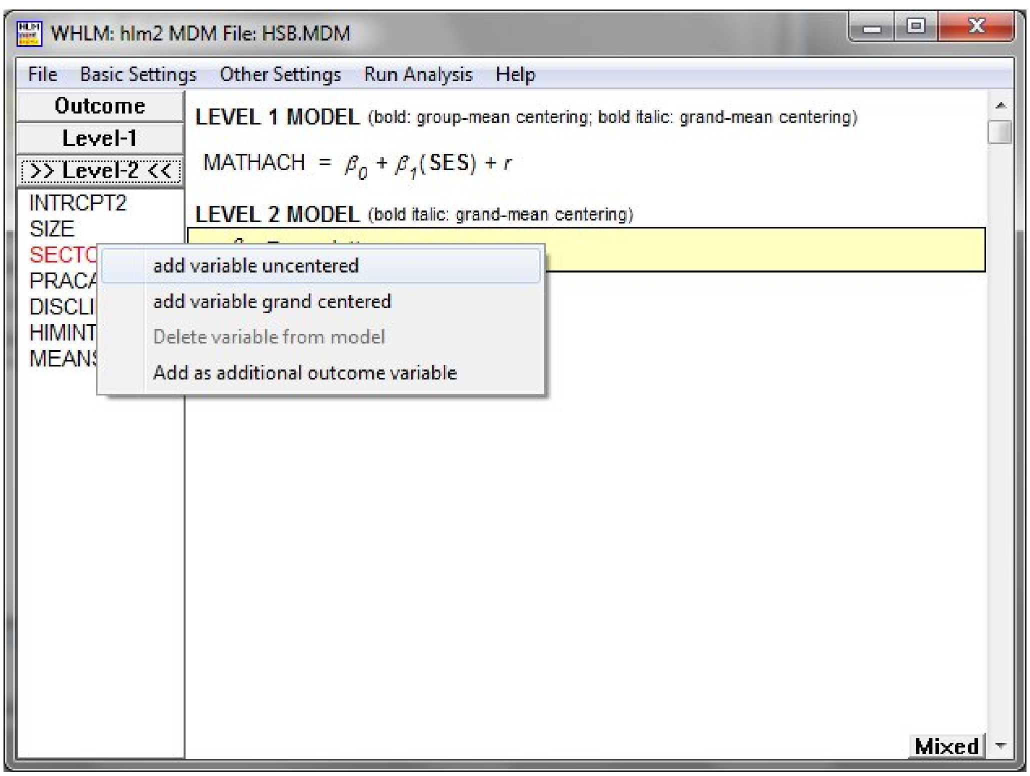


Figure 2.16 Specification of the level-2 model

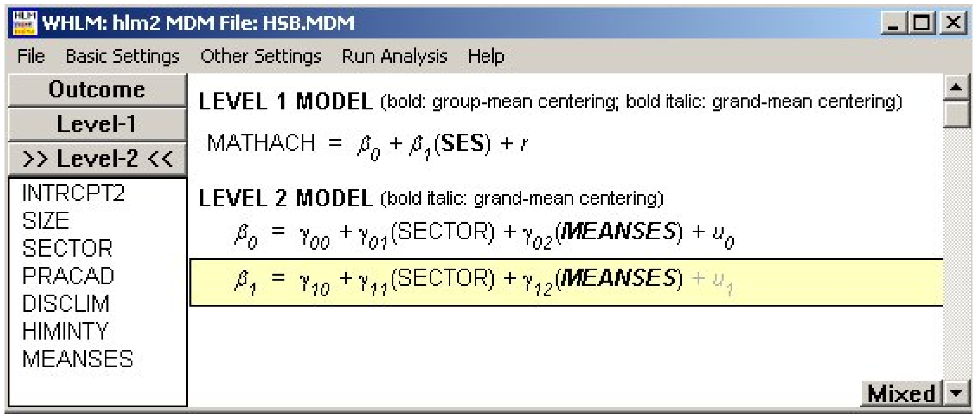


Figure 2.17 Model window for the HS&B example

4. In addition, the model can also be displayed in a mixed model formulation and with complete subscripts for all coefficients present in the model as illustrated in Figure 2.18. The mixed model is obtained by clicking the **Mixed** button at the bottom of the main window. The model is shown as a single equation, obtained by substituting the equations for β_0 and β_1 in the level-1 equation. This notation shows the model in a familiar linear regression format, and also draws attention to any cross-level interaction terms present in the combined model. By using the **Preferences** dialog box accessible via the **File** menu (see details in Section 2.8) both the mixed model formulation and the model with subscripts for all coefficients can be displayed automatically. The model can also be saved as an EMF file for later use in reports or papers.

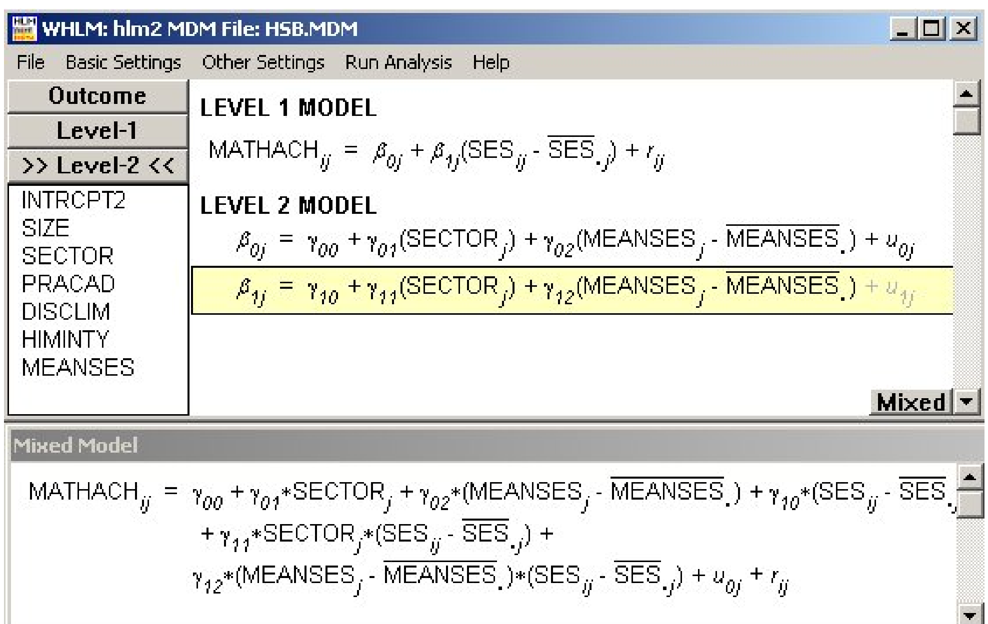


Figure 2.18 Alternative model window for the HS&B example

Step 3: To specify level-1 coefficients as random or non-random

The program begins by assuming that only the intercept (β_0) is specified as random. The u_1 at the end of the β_1 equation is grayed out and constrained to zero (See Figure 2.15), *i.e.* this level-1 coefficient is specified as "fixed". In the HS&B example, both level-1 coefficients, β_0 and β_1, are to be specified as random. To specify the SES slope as randomly varying, click on the equation for β_1 so that the error term u_1 is enabled. Note that one can toggle the error term in any of the three following ways:

- Click on the error term, u_1.
- Type *u*.
- Right-click on the yellow box, which will bring up a single-item menu **toggle error term**. Click on the button.

Steps 1 to 3 are the three major steps for executing analyses based on the MDM file. Other analytic options are described in Section 2.9. After specifying the model, a title can be given to the output and the output file can be named by the following procedure:

1. Select **Basic Settings** to open the **Basic Model Specifications – HLM2** dialog box. Enter a title in the **Title** field (for example, Intercept and slopes-as-Outcomes Model) and an output file name in **Output file name** field (see Figure 2.19). Click **OK**. See Section 2.8 for the definitions of entries and options in **Basic Model Specifications – HLM2** dialog box.
2. Open the **File** menu and choose **Save As** to open a **Save command file** dialog box.
3. Enter a command file name (for example, HSB1.MDM).
4. Click **Run Analysis**. A dialog box displaying the iterations will appear (see Figure 2.20).

Note: If you wish to terminate the computations early, press the Ctrl-C key combination once. This will stop the analysis after the current iteration and provide a full presentation of results based on that iteration. If you press Ctrl-C more than once, however, computation is terminated immediately and all output is lost.

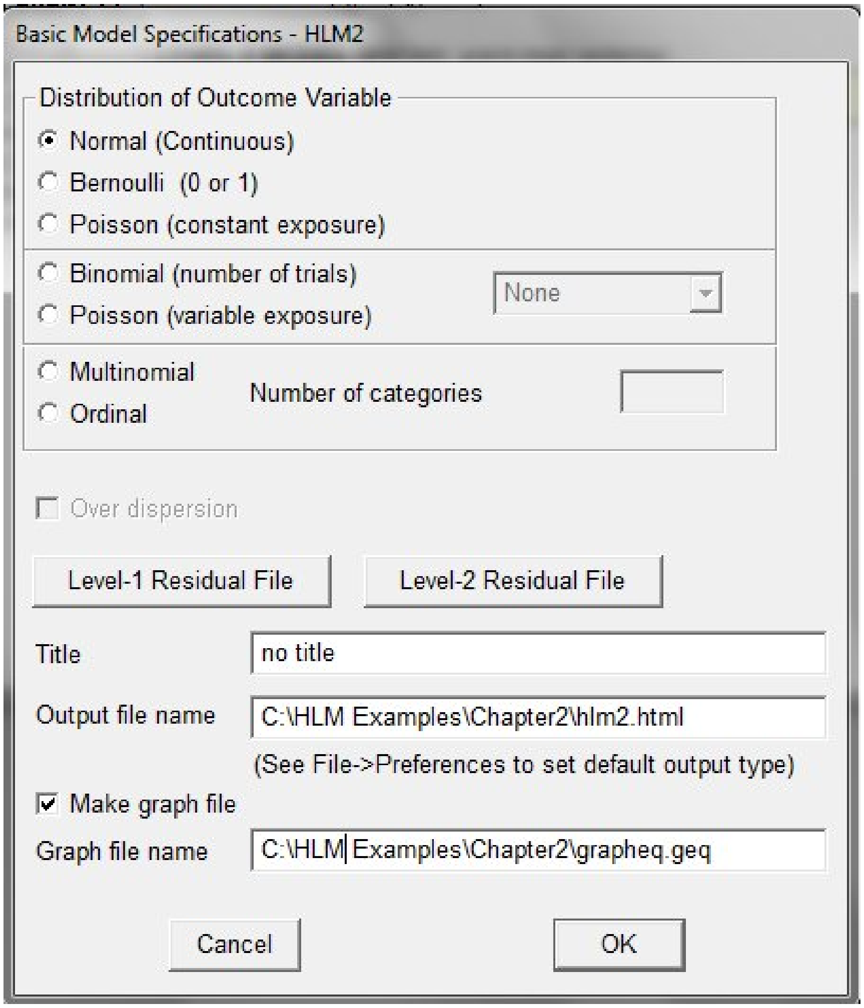


Figure 2.19 Basic Model Specifications – HLM2 dialog box for the HS&B example

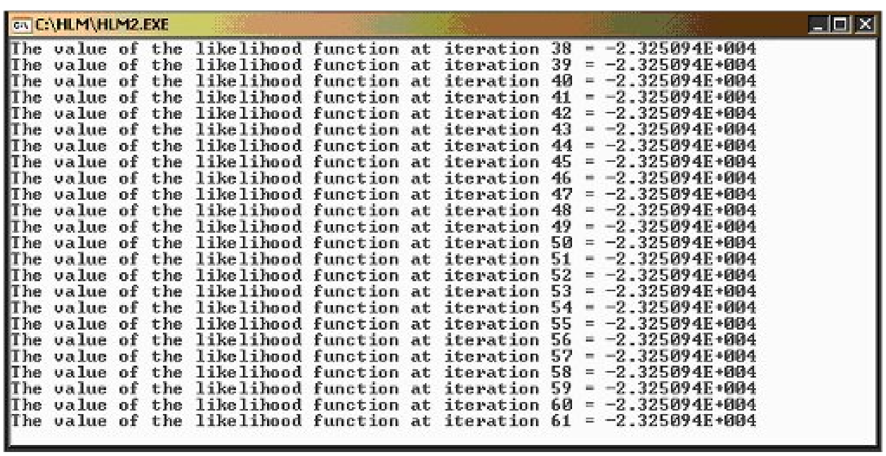


Figure 2.20 Iteration screen

2.5.3 Annotated HLM2 output

The output file will automatically be displayed in the format specified via the **Preference** menu. It can also be opened by selecting the **View Output** option from the **File** menu. Here is the output produced by the Windows session described above (see example HSB1.MDM).

Specifications for this HLM2 run

Problem Title: Intercepts and Slopes-as-outcomes Model

The data source for this run = HSB.MDM
The command file for this run = HSB1.MLM
Output file name = hlm2.html
The maximum number of level-1 units = 7185
The maximum number of level-2 units = 160
The maximum number of iterations = 100
Method of estimation: restricted maximum likelihood

The outcome variable is MATHACH

Summary of the model specified

Level-1 Model

$$MATHACH_{ij} = \beta_{0j} + \beta_{1j}*(SES_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(SECTOR_j) + \gamma_{02}*(MEANSES_j) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(SECTOR_j) + \gamma_{12}*(MEANSES_j) + u_{1j}$$

SES has been centered around the group mean.
MEANSES has been centered around the grand mean.

Mixed Model

$$MATHACH_{ij} = \gamma_{00} + \gamma_{01}*SECTOR_j + \gamma_{02}*MEANSES_j$$
$$+ \gamma_{10}*SES_{ij} + \gamma_{11}*SECTOR_j*SES_{ij} + \gamma_{12}*MEANSES_j*SES_{ij}$$
$$+ u_{0j} + u_{1j}*SES+ r_{ij}$$

The information presented on the first page or two of the HLM2 printout summarizes key details about the MDM file (*e.g.*, number of level-1 and level-2 units, whether weighting was specified), and about both the fixed and random effects models specified for this run. In this particular case, we are estimating the model specified by Equations 4.14 and 4.15 in *Hierarchical Linear Models*.

Level-1 OLS Regressions

Level-2 Unit	INTRCPT1	SES slope
1224	9.71545	2.50858
1288	13.51080	3.25545
1296	7.63596	1.07596
1308	16.25550	0.12602
1317	13.17769	1.27391
1358	11.20623	5.06801
1374	9.72846	3.85432
1433	19.71914	1.85429
1436	18.11161	1.60056
1461	16.84264	6.26650

When first analyzing a new data set, examining the OL equations for all of the units may be helpful in identifying possible outlying cases and bad data. By default, HLM2 does not print out the ordinary least squares (OL) regression equations, based on the level-1 model. The OLS regression equations for the first 10 units, as shown here, were obtained using optional settings on the **Other Settings** menu.

The average OLS level-1 coefficient for INTRCPT1 = 12.62075
The average OLS level-1 coefficient for SES = 2.20164

This is a simple average of the OLS coefficients across all units that had sufficient data to permit a separate OLS estimation.

Least Squares Estimates
σ^2 = 39.03409

Least-squares estimates of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.083837	0.106889	113.050	7179	<0.001
SECTOR, γ_{01}	1.280341	0.157845	8.111	7179	<0.001
MEANSES, γ_{02}	5.163791	0.190834	27.059	7179	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.935664	0.155268	18.907	7179	<0.001
SECTOR, γ_{11}	-1.642102	0.240178	-6.837	7179	<0.001
MEANSES, γ_{12}	1.044120	0.299885	3.482	7179	<0.001

Least-squares estimates of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.083837	0.169507	71.288	7179	<0.001
SECTOR, γ_{01}	1.280341	0.299077	4.281	7179	<0.001
MEANSES, γ_{02}	5.163791	0.334078	15.457	7179	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.935664	0.147576	19.893	7179	<0.001
SECTOR, γ_{11}	-1.642102	0.237223	-6.922	7179	<0.001
MEANSES, γ_{12}	1.044120	0.332897	3.136	7179	0.002

The first of the fixed effects tables are based on OLS estimation. The second table provides robust standard errors. Note that the standard errors associated with γ_{00}, γ_{01}, and γ_{12} are smaller than their robust counterparts.

The least-squares likelihood value = -2.336211E+004
Deviance = 46724.22267
Number of estimated parameters = 1

Starting Values

$\sigma^2_{(0)}$ = 36.72025

$\tau_{(0)}$

INTRCPT1, β_0	2.56964	0.28026
SES, β_1	0.28026	-0.01614

New $\tau_{(0)}$

INTRCPT1, β_0	2.56964	0.28026
SES, β_1	0.28026	-0.01614

The initial starting values failed to produce an appropriate variance-covariance matrix ($\tau_{(0)}$). An automatic fix-up was introduced to correct this problem (New $\tau_{(0)}$).

Estimation of fixed effects
(Based on starting values of covariance components)

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.094864	0.204326	59.194	157	<0.001
SECTOR, γ_{01}	1.226266	0.315204	3.890	157	<0.001
MEANSES, γ_{02}	5.335184	0.379879	14.044	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.935219	0.168674	17.402	157	<0.001
SECTOR, γ_{11}	-1.634083	0.260672	-6.269	157	<0.001
MEANSES, γ_{12}	1.015061	0.323523	3.138	157	0.002

Above are the initial estimates of the fixed effects. These are not to be used in drawing substantial conclusions.

The value of the log-likelihood function at iteration 1 = -2.325199E+004
The value of the log-likelihood function at iteration 2 = -2.325182E+004
The value of the log-likelihood function at iteration 3 = -2.325174E+004
The value of the log-likelihood function at iteration 4 = -2.325169E+004
The value of the log-likelihood function at iteration 5 = -2.325154E+004
. . .
The value of the log-likelihood function at iteration 57 = -2.325094E+004
The value of the log-likelihood function at iteration 58 = -2.325094E+004
The value of the log-likelihood function at iteration 59 = -2.325094E+004
The value of the log-likelihood function at iteration 60 = -2.325094E+004

Below are the estimates of the variance and covariance components from the final iteration and selected other statistics based on them.

```
******* ITERATION 61 *******
```

σ^2 = 36.70313 — *Level-1 variance components*

τ — *Level-2 variance-covariance components*

INTRCPT1, β_0	2.37996	0.19058
SES, β_1	0.19058	0.14892

τ (as correlations) — *Level-2 variance-covariance components expressed as*

INTRCPT1, β_0	1.000	0.320	*correlations*
SES, β_1	0.320	1.000	

Random level-1 coefficient	Reliability estimate
INTRCPT1, β_0	0.733
SES, β_1	0.073

These are average reliability estimates for the random level-1 coefficients

The value of the log-likelihood function at iteration 61 = -2.325094E+004

The next three tables present the final estimates for: the fixed effects with GLS and robust standard errors, variance components at level-1 and level-2, and related test statistics.

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.095006	0.198717	60.865	157	<0.001
SECTOR, γ_{01}	1.226384	0.306272	4.004	157	<0.001
MEANSES, γ_{02}	5.333056	0.369161	14.446	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.937787	0.157119	18.698	157	<0.001
SECTOR, γ_{11}	-1.640954	0.242905	-6.756	157	<0.001
MEANSES, γ_{12}	1.034427	0.302566	3.419	157	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.095006	0.173688	69.637	157	<0.001
SECTOR, γ_{01}	1.226384	0.308484	3.976	157	<0.001
MEANSES, γ_{02}	5.333056	0.334600	15.939	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.937787	0.147615	19.902	157	<0.001
SECTOR, γ_{11}	-1.640954	0.237401	-6.912	157	<0.001
MEANSES, γ_{12}	1.034427	0.332785	3.108	157	0.002

The first table provides model-based estimates of the standard errors while the second table provides robust estimates of the standard errors. Note that the two sets of standard errors are similar. If the robust and model-based standard errors are substantively different, it is recommended that the tenability of key assumptions should be investigated further (see Section 4.3 on examining residuals).

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, u_0	1.54271	2.37996	157	605.29503	<0.001
SES slope, u_1	0.38590	0.14892	157	162.30867	0.369
level-1, r	6.05831	36.70313			

Statistics for current covariance components model

Deviance = 46501.875643
Number of estimated parameters = 4

2.5.4 Model checking based on the residual file

HLM2 provides the data analyst with a means of checking the fit and distributional assumptions of the model by producing residual files for the level-1 and level-2 models. These files may be requested using the **Basic Model Specifications – HLM2** dialog box (see Fig. 2.19). The level-1 and level-2 residual files will be written as SPSS, SAS, STATA, SYSTAT or ASCII data files. In the case of SPSS and STATA, the residual files will be written out so that the respective packages may use them immediately. The other forms of raw data will require submitting them as command streams.

2.5.4.4 The level-1 residual file

2.5.4.1.1 Structure of the level-1 residual file

The level-1 residual file will contain level-1 residuals (the differences between the observed and fitted values), the fitted values, the square root of σ^2, the values of the level-1 and level-2 predictors entered in the model, and those of other level-1 and level-2 variables selected by the user. To illustrate, we show how to prepare SPSS residual files.

To create the SPSS level-1 residual file type

1. Select **Basic Settings** to open the **Basic Model Specifications – HLM2** dialog box.
2. Click **Level-1 Residual File** to open a **Create Level-1 Residual File** dialog box (see Figure 2.21).
3. For the level-1 and level-2 variables, the box displays two columns of variables. The predictor variables in the model are in the **Variables in residual file** column. Others are listed in the **Possible choices** column. To include any of them in the residual file for exploratory purposes, double-click on their labels.
4. Select **SPSS residual file type** (default).
5. Enter a name for the residual file in the **Residual File Name** box (for example, RESFIL1.SAV, see Figure 2.21). Click **OK**.

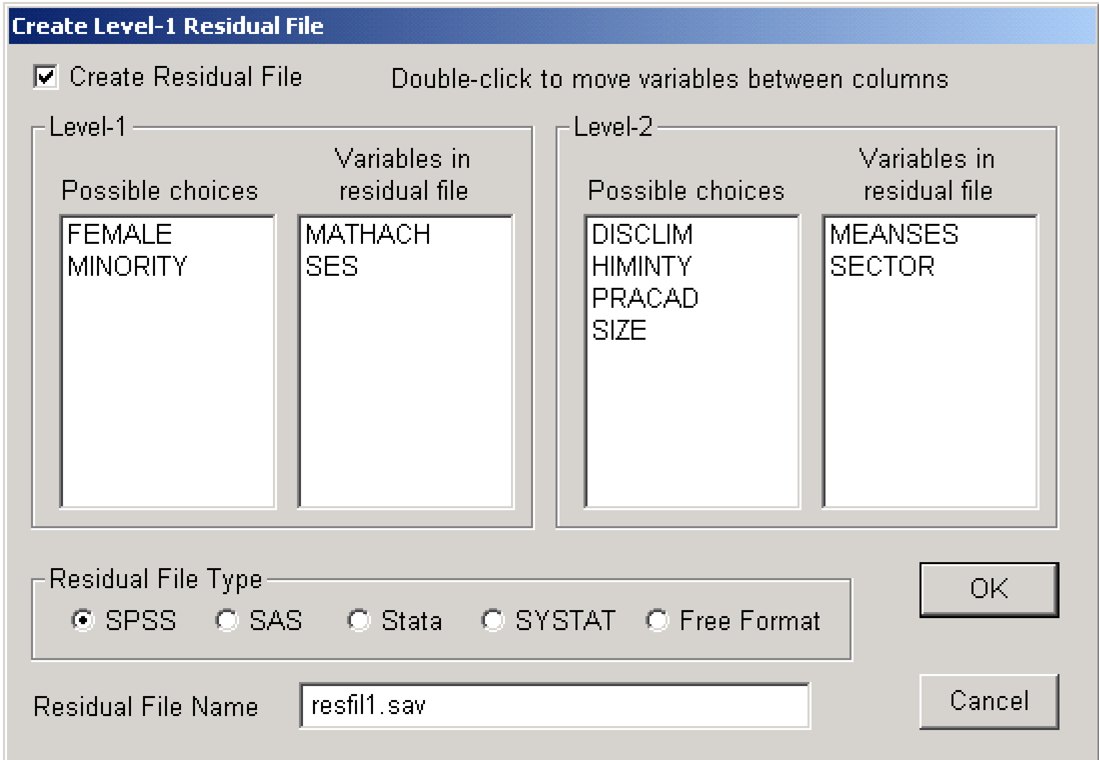
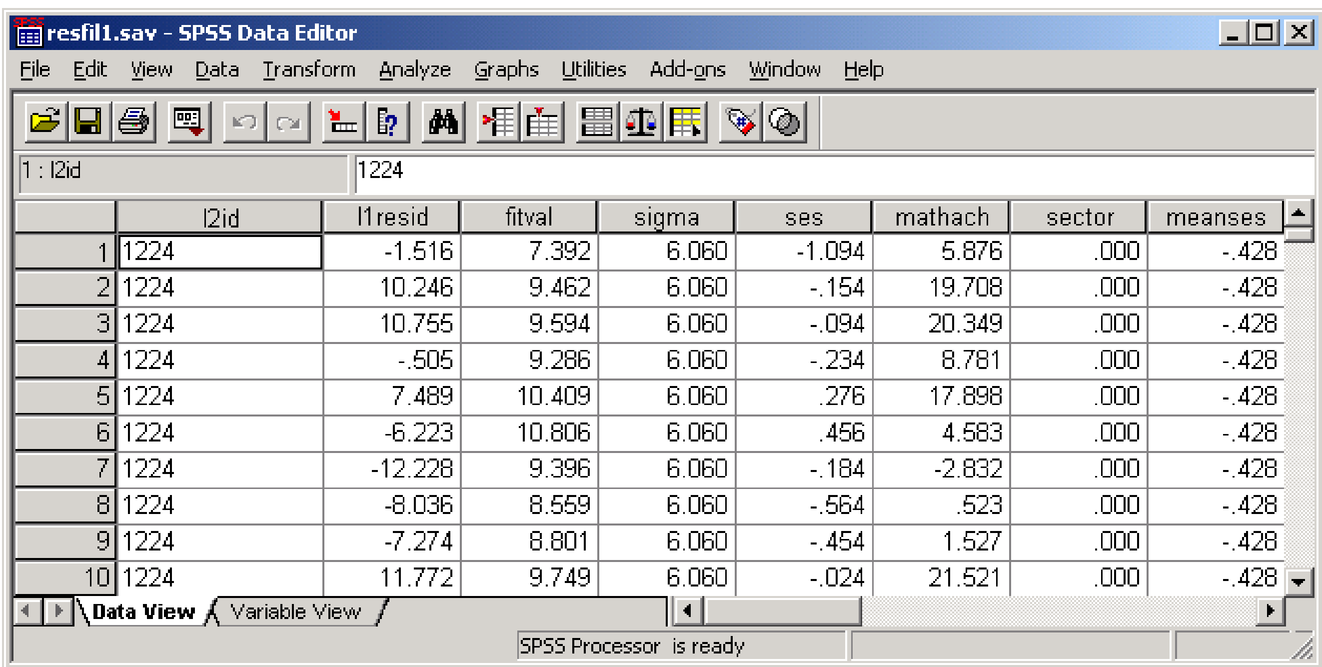


Figure 2.21 Create Level-1 Residual File dialog box



	l2id	l1resid	fitval	sigma	ses	mathach	sector	meanses
1	1224	-1.516	7.392	6.060	-1.094	5.876	.000	-.428
2	1224	10.246	9.462	6.060	-.154	19.708	.000	-.428
3	1224	10.755	9.594	6.060	-.094	20.349	.000	-.428
4	1224	-.505	9.286	6.060	-.234	8.781	.000	-.428
5	1224	7.489	10.409	6.060	.276	17.898	.000	-.428
6	1224	-6.223	10.806	6.060	.456	4.583	.000	-.428
7	1224	-12.228	9.396	6.060	-.184	-2.832	.000	-.428
8	1224	-8.036	8.559	6.060	-.564	.523	.000	-.428
9	1224	-7.274	8.801	6.060	-.454	1.527	.000	-.428
10	1224	11.772	9.749	6.060	-.024	21.521	.000	-.428

Figure 2.22 Level-1 Residual File

Data for the first ten cases in RESFIL1.SAV are shown in Figure 2.22. The file consists of the level-2 ID, L2ID, and the following variables:

- L1RESID: the difference between the fitted and observed value for each level-1 unit.
- FITVAL: the fitted value for each level-1 unit.
- SIGMA: the square root of σ^2.

The variables SES, MATHACH, SECTOR, and MEANSES are described in Section 2.5.1.1.

2.5.4.1.2 Some possible residual analyses

We illustrate a possible use of a residual file in examining the tenability of the assumption of normal distribution of level-1 errors, whose violations could adversely influence the estimated standard errors for the estimates of the fixed effects and inferential statistics (see *Hierarchical Linear Models* p. 266). Figure 2.23 displays a normal Q-Q plot of the level-1 residuals for the 7,185 students based on the final fitted model. The plot is approximately linear, suggesting there is not a serious departure from a normal distribution and that the assumption is tenable.

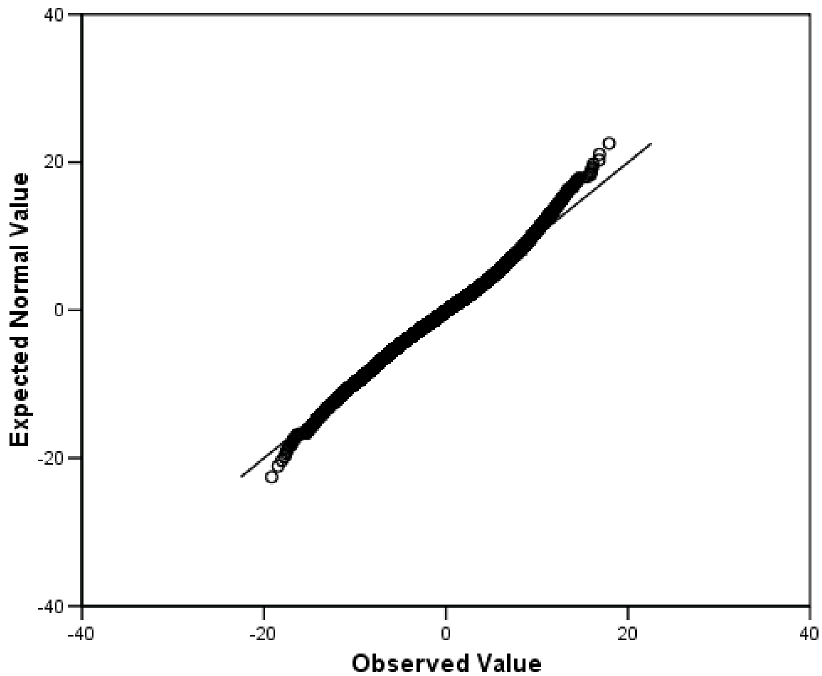


Figure 2.23 Q-Q plot of level-1 residuals

2.5.4.5 The level-2 residual file

This file will contain the EB residuals (see Equation 1.10 above), OL residuals (see Equation 1.9 above), and fitted values, *i.e.*,

$$\hat{\gamma}_{q0} + \sum \hat{\gamma}_{qs} W_{sj}$$

for each level-1 coefficient. By adding the OL residuals to the corresponding fitted values, the analyst can also obtain the OL estimate of the corresponding level-1 coefficient β_{qj}. The file also produces the EB estimate β^*_{qj} of each level-1 coefficient, β_{qj}.

In addition, the file will contain Mahalanobis distances (which are discussed below), estimates of the total and residual standard deviations (log metric) within each unit, the values of the predictors used in the level-2 model, and any other level-2 prediction variables selected by the user.

To create the SPSS level-2 residual file type

1. Select **Basic Settings** to open the **Basic Model Specifications – HLM2** dialog box.
2. Click **Level-2 Residual File** to open a **Create Level-2 Residual File** dialog box.
3. Double-click the variables to be entered into the residual file (for our example, select DISCLIM, PRACAD, HIMINTY and SIZE, see Figure 2.24).
4. Select **SPSS** as **Residual File Type**. Note that SYSTAT, STATA or SAS file type can be created as well, or the residuals written to file in free format. By default, a SYSTAT file will be created. To set the default file type created to one of the other formats, the **Preference** dialog box (see Section 2.8) can be used.

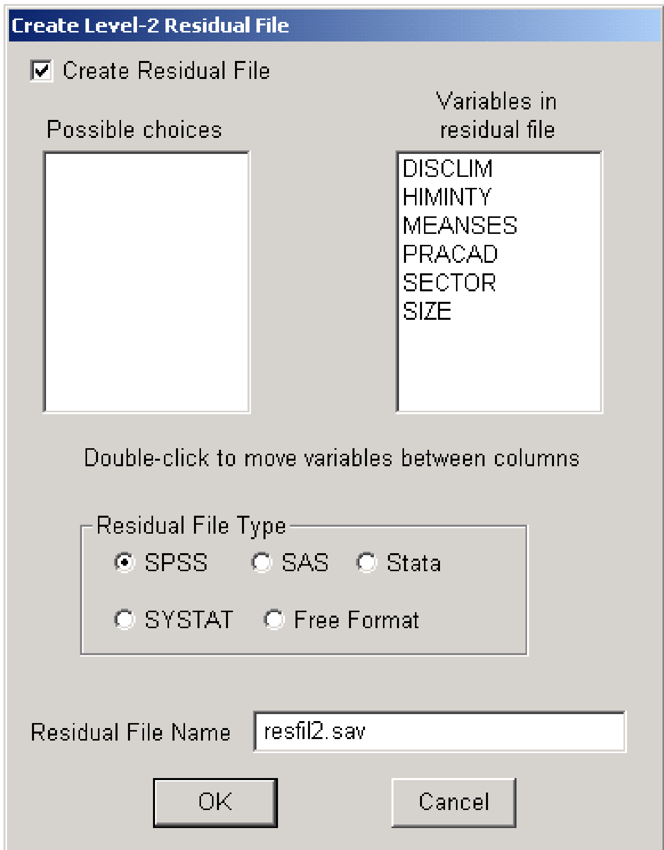


Figure 2.24 Create Residual File dialog box

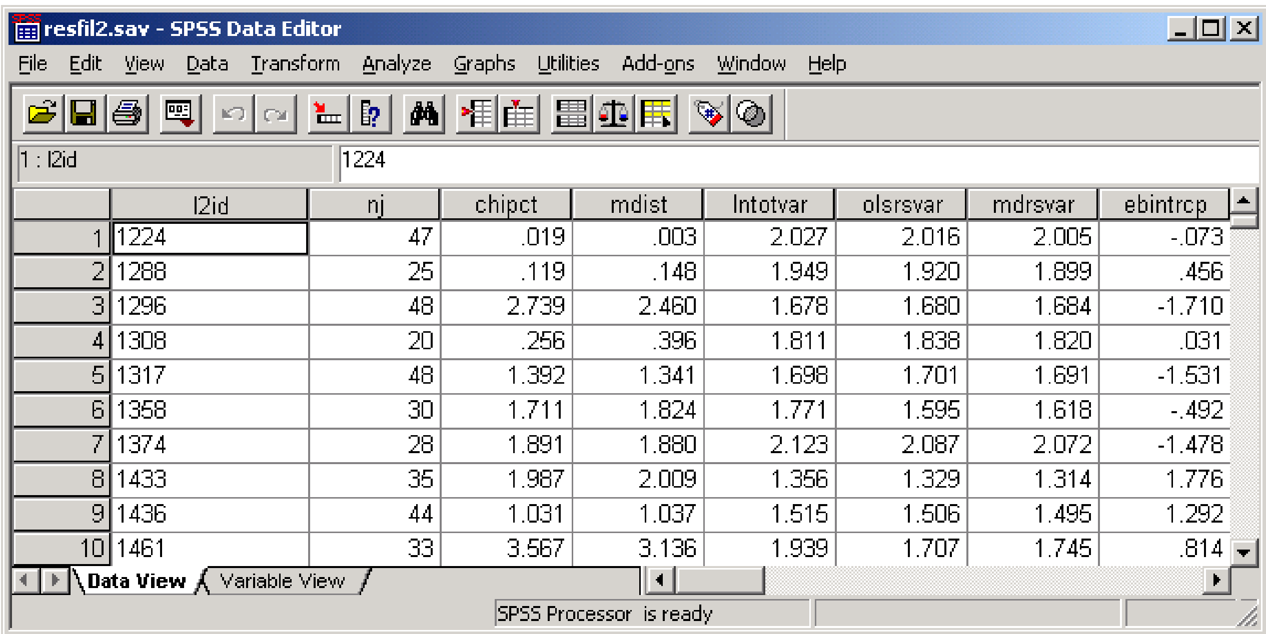
5. Enter a name for the residual file in the **Residual File Name** box (for example RESFIL2.SPS, see Figure 2.24). Click **OK**.

An example of an SPSS version of a level-2 residual file is shown in Figure 2.25. Only the data from the first ten units and the first 8 variables are reproduced here. This file can be used to construct various diagnostic plots.

2.5.4.2.1 Structure of the level-2 residual file

The residual file contains a single record per unit. The first variable in this file contains the unit ID, followed by the number of level-1 units within that level-2 unit (denoted by nj), and various summary statistics (chipct through mdrsvar). These are followed by the two EB residuals; the two OL residuals; and the fitted or predicted values of the level-1 coefficients based on the estimated level-2 models. Next are the EB coefficients ecintrcp and ecses, which are the sum of the fitted values plus the EB residuals. The posterior variances and covariances of the estimates of the intercept and the SES slopes are given next (pv00 to pvc10). Finally, the level-2 predictors used in the analysis plus those additional level-2 predictors requested by the user for inclusion in the file are given (not shown in Figure 2.24).

While most of this is straightforward, the information contained in the first set of variables for each unit merits elaboration. nj is the number of cases for level-2 unit j. It is followed by two variables, chipct and mdist. If we model q level-1 coefficients, mdist would be the Mahalanobis distance (*i.e.*, the standardized squared distance of a unit from the center of a v-dimensional distribution, where v is the number of random effects per unit). Essentially, mdist provides a single, summary measure of the distance of a unit's EB estimates, β^*_{qj}, from its “fitted value,” $\hat{\gamma}_{q0} + \sum \hat{\gamma}_{q0} W_{sj}$.



	l2id	nj	chipct	mdist	lntotvar	olsrsvar	mdrsvar	ebintrcp
1	1224	47	.019	.003	2.027	2.016	2.005	-.073
2	1288	25	.119	.148	1.949	1.920	1.899	.456
3	1296	48	2.739	2.460	1.678	1.680	1.684	-1.710
4	1308	20	.256	.396	1.811	1.838	1.820	.031
5	1317	48	1.392	1.341	1.698	1.701	1.691	-1.531
6	1358	30	1.711	1.824	1.771	1.595	1.618	-.492
7	1374	28	1.891	1.880	2.123	2.087	2.072	-1.478
8	1433	35	1.987	2.009	1.356	1.329	1.314	1.776
9	1436	44	1.031	1.037	1.515	1.506	1.495	1.292
10	1461	33	3.567	3.136	1.939	1.707	1.745	.814

Figure 2.25 SPSS version of residual file

If the normality assumption is true, then the Mahalanobis distances should be distributed approximately $\chi^2_{(v)}$. Analogous to univariate normal probability plotting, we can construct a Q-Q plot of mdist vs. chipct. chipct are the expected values of the order statistics for a sample of size J selected from a population that is distributed $\chi^2_{(v)}$. If the Q-Q plot resembles a 45 degree line, we have evidence that the random effects are distributed v-variate normal. In addition, the plot will help us detect outlying units (*i.e.*, units with large mdist values well above the 45 degree line). It should be noted that such plots are good diagnostic tools only when the level-1 sample sizes, nj, are at least moderately large. (For further discussion see *Hierarchical Linear Models*, pp. 274-280.)

After mdist, three estimates of the level-1 variability are given:

- The natural logarithm of the total standard deviation within each unit, lntotvar.

- The natural logarithm of the residual standard deviation within each unit based on its least squares regression, olsrsvar. Note, this estimate exists only for those units that have sufficient data to compute level-1 OLS estimates.
- The mdrsvar, the natural logarithm of the residual standard deviation from the final fitted fixed effects model.

The natural log of these three standard deviations (with the addition of a bias-correction factor for varying degrees of freedom) is reported (see *Hierarchical Linear Models*, p. 219). We note that these statistics can be used as input for the V-known option in HLM2 in research on group-level correlates of diversity (Raudenbush & Bryk, 1987; also see Sections 2.8.9 and 9.3).

2.5.4.2.2 Some possible residual analyses

We illustrate below some of the possible uses of a level-2 residual file in examining the adequacy of fitted models and in considering other possible level-2 predictor variables. (For a full discussion of this topic see Chapter 9 of *Hierarchical Linear Models*.) Here are the basic statistics for each of the variables created as part of the HLM2 residual file.

	N	Minimum	Maximum	Mean	Std. Deviation
nj	160	14	67	44.91	11.855
chipct	160	.006	11.537	1.99115	1.967047
mdist	160	.003	13.218	2.00727	2.144775
lntotvar	160	1.265	2.138	1.82057	.150434
olsrsvar	160	1.272	2.087	1.78983	.137449
mdrsvar	160	1.314	2.072	1.79039	.134968
ebintrcp	160	-3.718	4.162	.00000	1.312584
ebses	160	-.378	.438	.00000	.141577
olintrcp	160	-7.714	5.545	-.01079	1.847386
olses	160	-3.560	3.803	-.01823	1.460555
fvintrcp	160	5.760	17.754	12.63155	2.490807
fvses	160	.515	3.650	2.21987	.775690
ecintrcp	160	4.710	18.928	12.63155	2.815492
ecses	160	.288	3.845	2.21987	.788504
pv00	160	.486	1.255	.66785	.140621
pv10	160	.036	.098	.05033	.011378
pv11	160	.121	.138	.12900	.003583
pvc00	160	.449	1.257	.63936	.147741
pvc10	160	.030	.097	.04682	.011911
pvc11	160	.138	.247	.16255	.017345
size	160	100.000	2713.000	1097.82500	629.506431
sector	160	.000	1.000	.43750	.497636
pracad	160	.000	1.000	.51394	.255897
disclim	160	-2.416	2.756	-.01513	.976978
himinty	160	.000	1.000	.27500	.447916
meanses	160	-1.188	.831	.00000	.413973
Valid N (listwise)	160				

Examining OL and EB residuals. Figure 2.26 shows a plot of the OL vs EB residuals for the SES slopes. As expected, the EB residuals for the slope are much more compact than the OL residuals.

While the latter ranges between (–4.0, 4.0), the range for the EB residuals is only (–0.5, 0.5). (For a further discussion see *Hierarchical Linear Models*, pp. 87-92.)

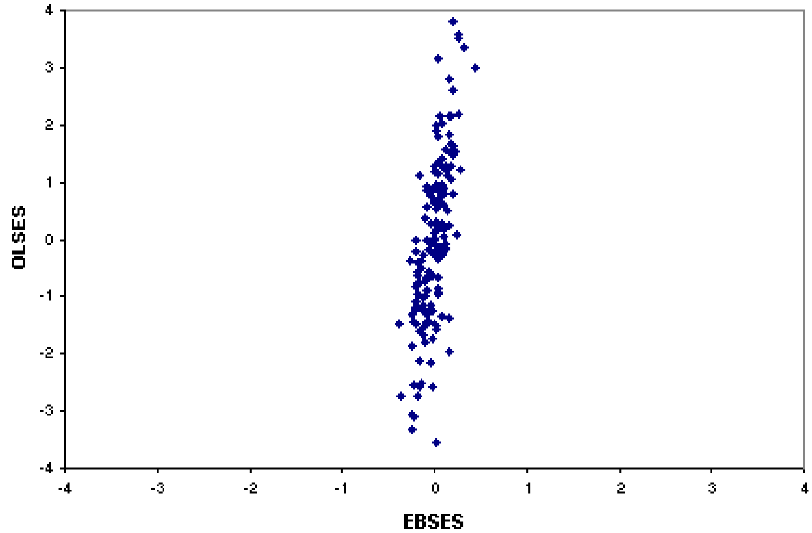


Figure 2.26 OL versus EB residuals for the SES slopes

Exploring the potential of other possible level-2 predictors. Figure 2.27 shows a plot of EB residuals against a possible additional level-2 predictor, PRACAD, for the intercept model. Although the relationship appears slight (a correlation of 0.15), PRACAD will enter this model as a significant predictor. (For a further discussion of the use of residual plots in identifying possible level-2 predictors see *Hierarchical Linear Models*, pp. 267-270.)

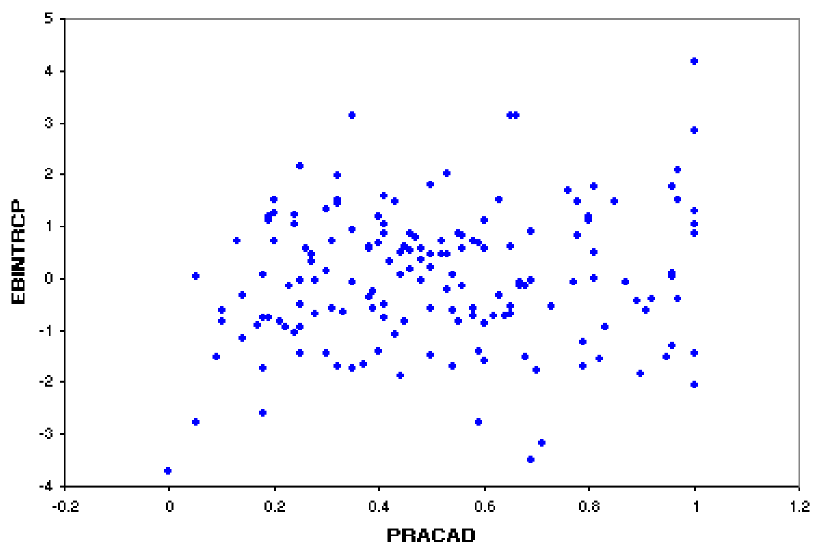


Figure 2.27 EB residuals against a possible additional level-2 predictor, PRACAD, for the intercept model

Next, in Figure 2.28, we see a plot of the OL vs EB residuals for the intercepts. Notice that while the EB intercepts are “shrunk” as compared to the OL estimates, the amount of shrinkage for the intercepts as shown in Figure 2.28 is far less than for the SES slopes as shown in Figure 2.26.

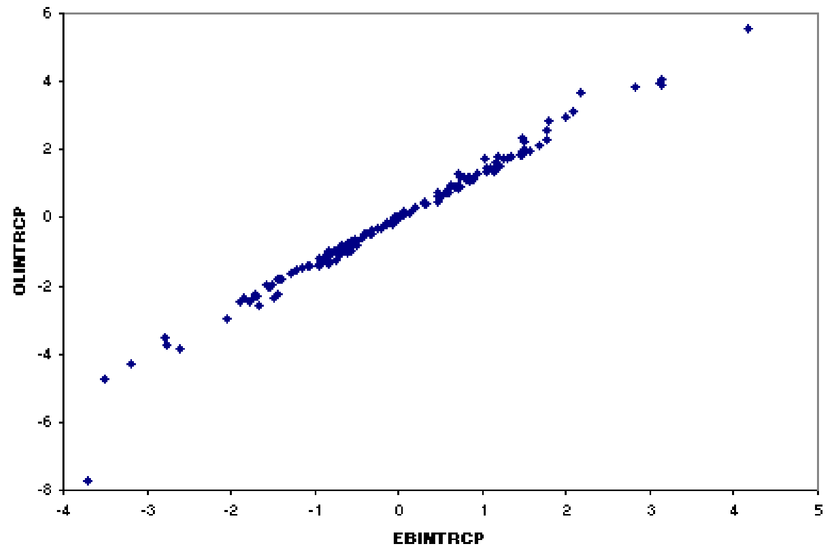


Figure 2.28 OL versus EB residuals for the intercepts

Examining possible nonlinearity of a level-2 predictor's relationship to an outcome. Next, in Fig. 2.29, is an example of a plot of EB residuals, in this case the SES slope, against a variable included in the model. This plot suggests that the assumption of a linear relationship between the SES slope and MEANSES is appropriate. (That is, the residuals appear randomly distributed around the zero line without regard to values of MEANSES.)

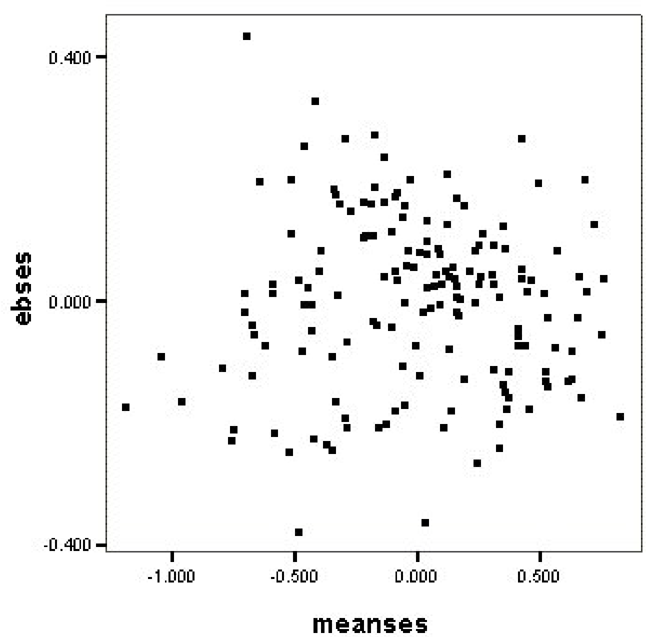


Figure 2.29 EB residuals for SES slope against MEANSES

2.6 Handling of missing data

HLM2 provides three options for handling missing data at level 1: listwise deletion of cases when the MDM file is made, listwise deletion of cases when running the analysis (See Figure 2.3), and analysis of multiply-imputed data (see Section 11.2). A set of level-1 variables to be used as basis for runtime deletion for a series of models based on the same MDM can also be selected via the

Other Settings, **Estimation Settings** menu by using the **Level-1 Deletion Variables** option. These follow the conventional routines used in standard statistical packages for regression analysis and the general linear model. Listwise deletion of cases when the MDM file is made is based on the variables selected for inclusion in the MDM file, while listwise deletion when running the analysis only takes the variables included in the model into account.

At level 2, HLM2 assumes complete data. If you have missing data at level 2, you should either impute a value for the missing information or delete the units in question, or preferably use methods described in Section 11.2. **Failure to do so will cause the automatic listwise deletion of level-2 units with missing data when the MDM file is created.**

For ASCII file input, click **Missing Data** in the **Make MDM – HLM2** dialog box. The dialog box displayed in Fig. 2.30 will open.

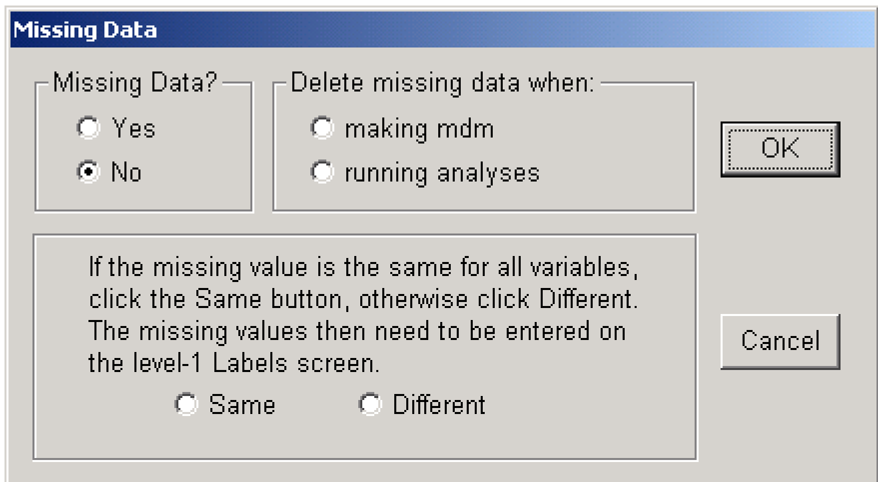


Figure 2.30 Missing Data dialog box

Assuming you have missing data, you should click **Yes** in the **Missing Data?** box, and select deletion when making the MDM file or when running analyses. Then, if you have coded all of your missing values for all of the variables to the same number, click the **Same** button. When you specify the variable names, enter this number in the box to right of the first variable in the **Enter Variable Labels** dialog box (see Fig. 2.31). If you have more than one missing value code, check the **Different** button, and enter these codes for each respective variable on the **Enter Variable Labels** screen.

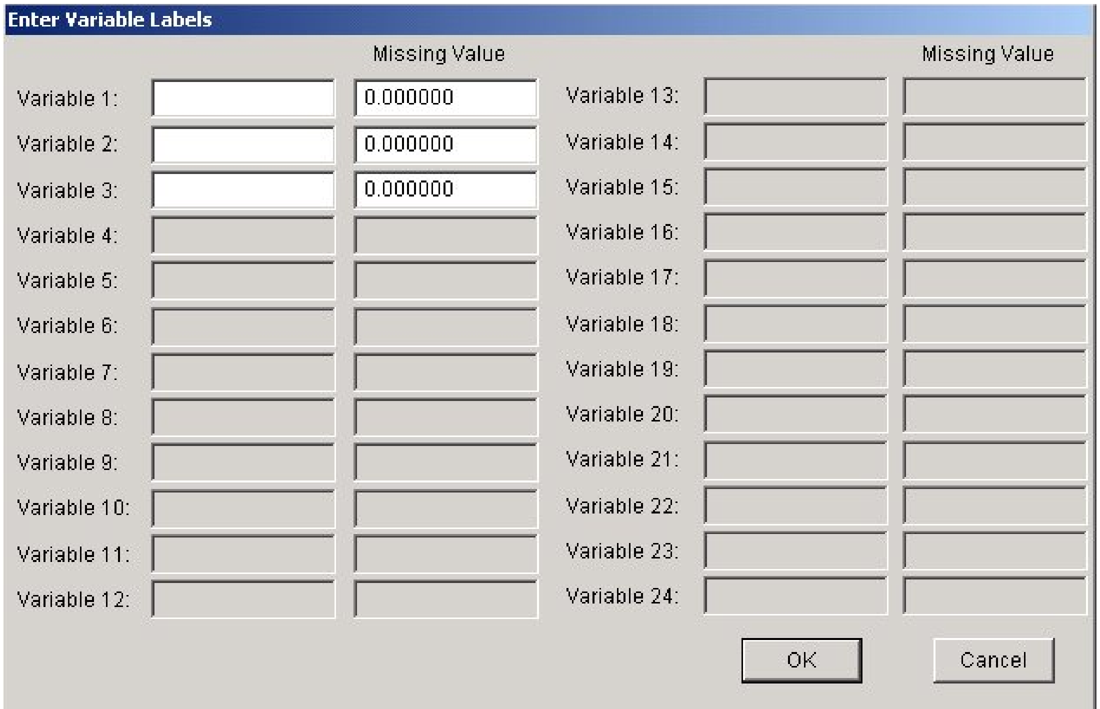


Figure 2.31 Enter Variable Labels dialog box for missing ASCII data

For non-ASCII data at level 1, you should click **Yes** in the **Missing Data?** field, and select when you want to implement the listwise deletion by selecting one of the two options in this group box. Then, when HLM2 encounters values coded as missing, it will recognize these properly. It is important to note that some statistics packages (*e.g.* SAS) allow for more than one kind of missing data code. HLM2 (and HLM3, etc.) will recognize only the standard, "system-missing" code.

How HLM2 handles missing data differs a bit in the ASCII and non-ASCII cases. *For ASCII data, it is very important that you don't have any missing data codes or blanks in the level 2 file*. HLM2 will read these as valid data; missing data codes as they are coded, and blanks will be read as zeros. For non-ASCII data, the program will skip over cases that have missing data in them, essentially performing listwise deletion on the level-2 data file. **Note:** For non-ASCII file input, the user has to either prepare system-missing values or missing value codes for the missing data.

2.7 The Basic Model Specifications - HLM2 dialog box

The **Basic Model Specifications – HLM2** dialog box (see Fig. 2.32) is used to indicate the distribution of the outcome variable, to request residual files and to provide a title and the locations and names of output files.

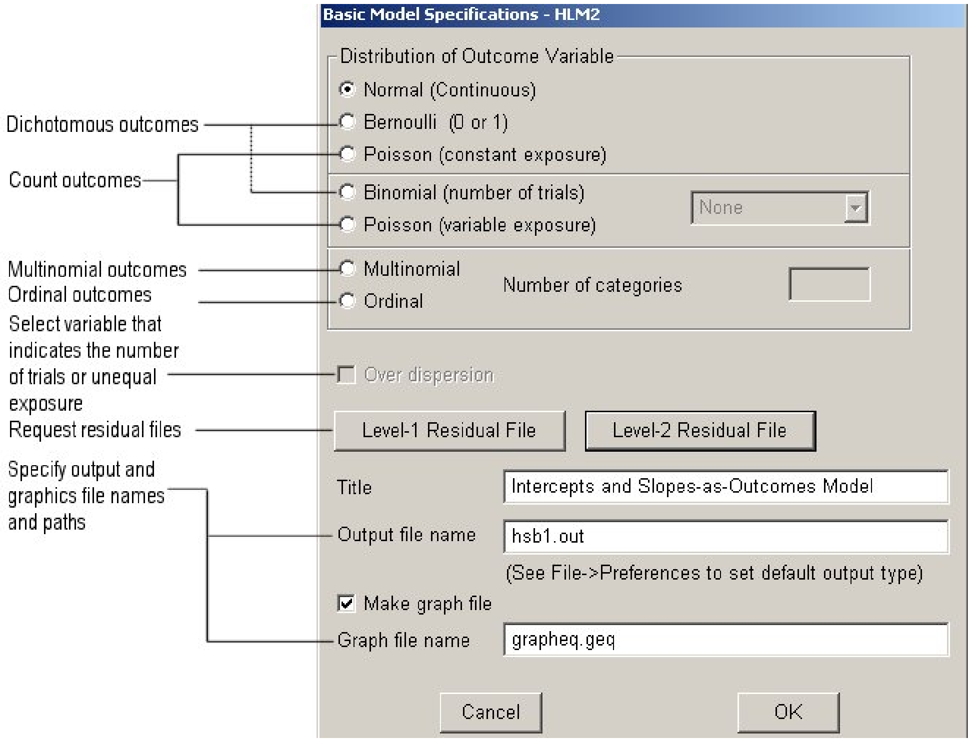


Figure 2.32 Basic Model Specifications - HLM2 dialog box

2.8 Other analytic options

2.8.1 Controlling the iterative procedure

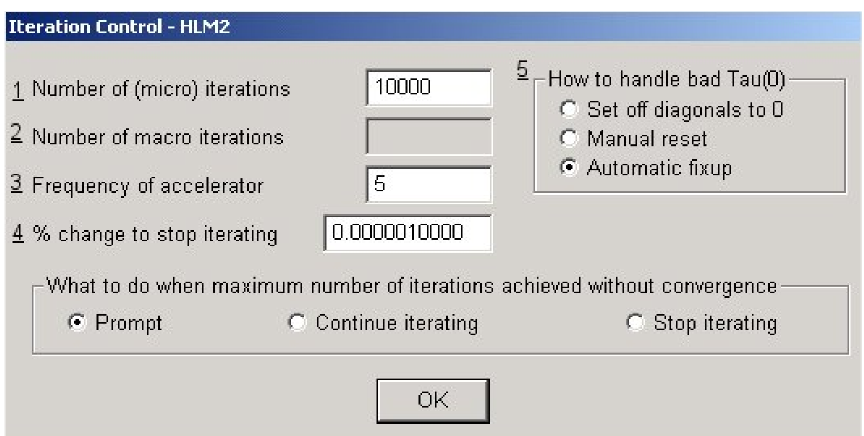


Figure 2.33 Iteration Control - HLM2 dialog box

The iterative procedure settings can be changed by opening the **Iteration Control – HLM2** dialog box. To do so, select the **Iteration Settings** option from the **Other Settings** menu. Table 2.1 lists the definitions and options in the **Iteration Control – HLM2** dialog box. See Fig. 2.33; note the linking numbers in figure and table.

Table 2.1 Table of definitions and options in Iteration Control - HLM2 dialog box

Key Terms	Function	Option	Definition
1, 2 Number of iterations	Maximum number of iterations	positive integer	
3 Frequency of accelerator	Controls frequency of use of acceleration	integer ^{3} 3	Selects how often the accelerator is used. Default is 10.
4 % change to stop iterating	Convergence criterion for maximum likelihood estimation	positive real number	Default: 0.000001. Can be specified to be more (or less) restrictive
5 How to handle bad Tau(0)	Method of correcting unacceptable starting values	3 choices	1. Set off-diagonal to 0 2. Manual reset (starting values) 3. Automatic fix-up (default)

2.8.2 Estimation control

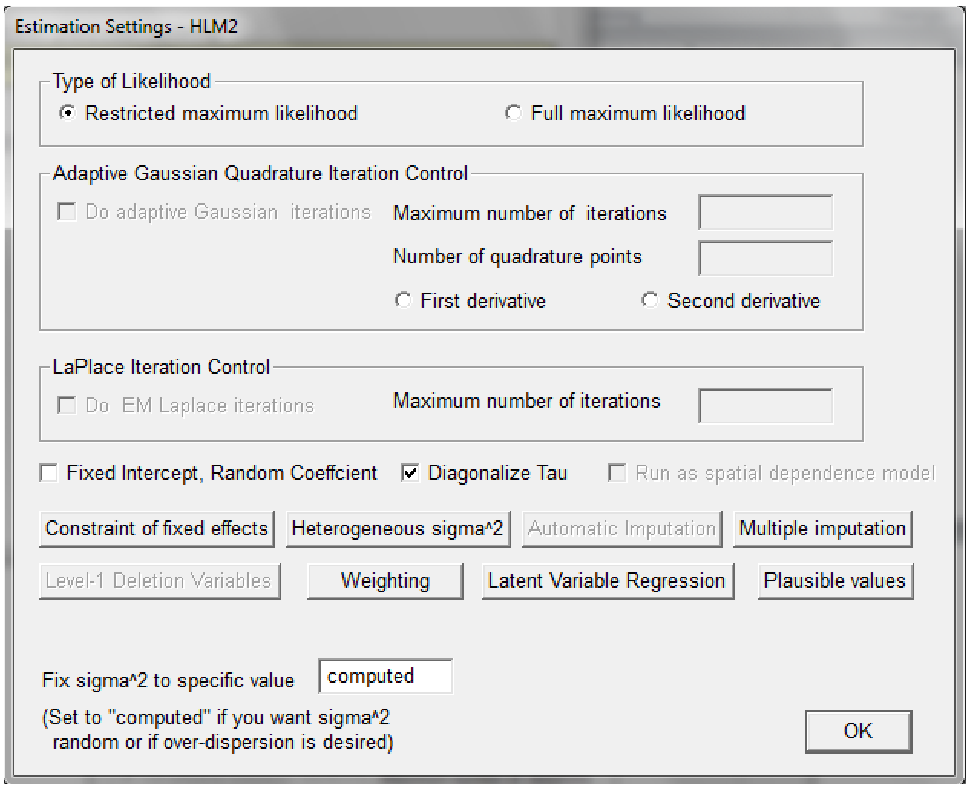


Figure 2.34 Estimation Settings – HLM2 dialog box

The **Estimation Settings – HLM2** dialog box, accessed via the **Estimation Settings** option on the **Other Settings** menu, offers additional control over the iterative procedure.

HLM2 will use restricted maximum likelihood estimation by default. The type of likelihood used is set in the **Type of Likelihood** group box (see Fig. 2.34), where full maximum likelihood estimation may alternatively be requested (see *Hierarchical Linear Models*, pp. 52-53.)

Full maximum Adaptive Gaussian Quadrature and LaPlace and EM LaPlace iterations may be requested when nonlinear (HGLM) models are fitted. The maximum number of iterations required, which has to be a positive integer, should be entered in the **LaPlace Iteration Control** or **EM LaPlace Iteration Control** group box (see Fig. 2.34).

The **Estimation Settings – HLM2** dialog box may also be used to access dialog boxes used in defining special analyses, *e.g.* latent variable regression, applying HLM to multiply-imputed data, and plausible value analysis. The **Fixed Intercept, Random Coefficient** option is used to invoke the fiting of fixed intercepts random coefficients in models as discussed in Chapter 19. The **Diagonalize Tau** option constrains the variance-covariance matrix to a diagonal matrix; in other words no covariation between random coefficients are assumed or estimated if this option is checked.

These special features, associated with the **Plausible values**, **Multiple imputation** and **Latent Variable Regression** buttons in the **Estimation Settings – HLM2** dialog box, are discussed in Chapter 11.

2.8.3 Constraints on the fixed effects

A user may wish to constrain two or more fixed effects to be equal. For example, Barnett, Marshall, Raudenbush, & Brennan (1993) applied this approach in studying correlates of psychological distress in married couples. Available for each person were two parallel measures of psychological distress. Hence, for each couple, there were four such measures (two per person). At level-1 these measures were modeled as the sum of a "true score" plus error:

$$Y_{ij} = \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + r_{ij},$$

where X_{1ij} is an indicator for females, X_{2ij} is an indicator for males, and r_{ij} is a measurement error. Hence β_{1j} is the "true score" for females and β_{2j} is the "true score" for males. At level 2, these true scores are modeled as a function of predictor variables, one of which was marital role quality, W_j, a measure of one's satisfaction with one's marriage. (Note that this is also a model without a level-1 intercept.) A simple level-2 model is then:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}W_j + u_{2j}.$$

The four coefficients to be considered are $\gamma_{10}, \gamma_{11}, \gamma_{20}, \gamma_{21}$. We may, for instance, wish to specify some constraints of fixed effects.

2.8.4 To put constraints on fixed effects

1. Open the **Other Settings**, **Estimation Settings** menu.
2. Click the **Constraint of fixed effects** button to open the **Constrain Gamma** dialog box. Enter **1** in the Sector boxes (see Figure 2.35 for an example). Click **OK**. The constraint imposed is $\gamma_{11} = \gamma_{21}$.

Coefficients with 0s are not constrained, and those with 1s are. A user is allowed to impose multiple constraints up to 5. Each set of the constrained coefficients will share the same value from 1 to 5.

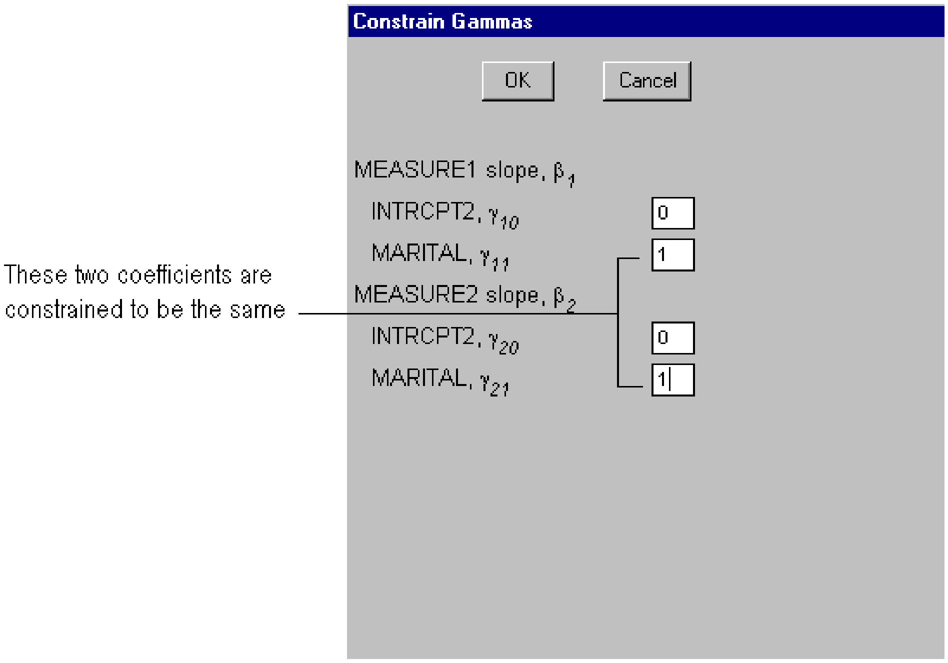


Figure 2.35 Constrain Gammas dialog box for the Barnett *et al.*'s (1993) example

2.8.5 Modeling heterogeneity of level-1 variances

Users may wish to estimate models that allow for heterogeneous level-1 variances. A simple example (see HSB3.HLM) using the HS&B data would be a model that postulates that the two genders have different means in and variances of math achievement scores. To specify a model that hypothesizes different central tendency and variability in math achievement for the two genders, the model displayed in Fig. 2.36 must first be set up.

To model heterogeneity of level-1 variances

1. Open the **Other Settings** menu and select the **Estimation Settings** option to open the **Estimation Settings – HLM2** dialog box.
2. Click the **Heterogeneous sigma^2** button to open the **Heterogeneous sigma^2 Predictors of level-1 variance** dialog box. Double-click FEMALE to enter as a variable in the **Predictors of level-1 variance** box (see Figure 2.37 for an example). Click **OK**.

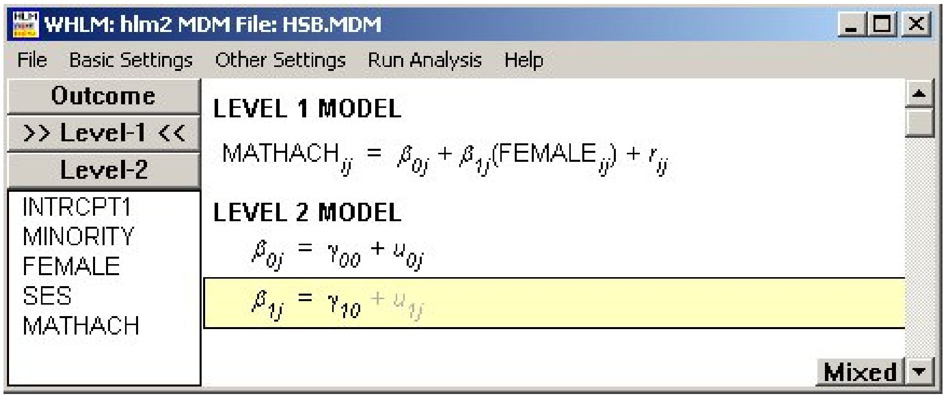


Figure 2.36 Model window for the modeling heterogeneity of level-1 variances example

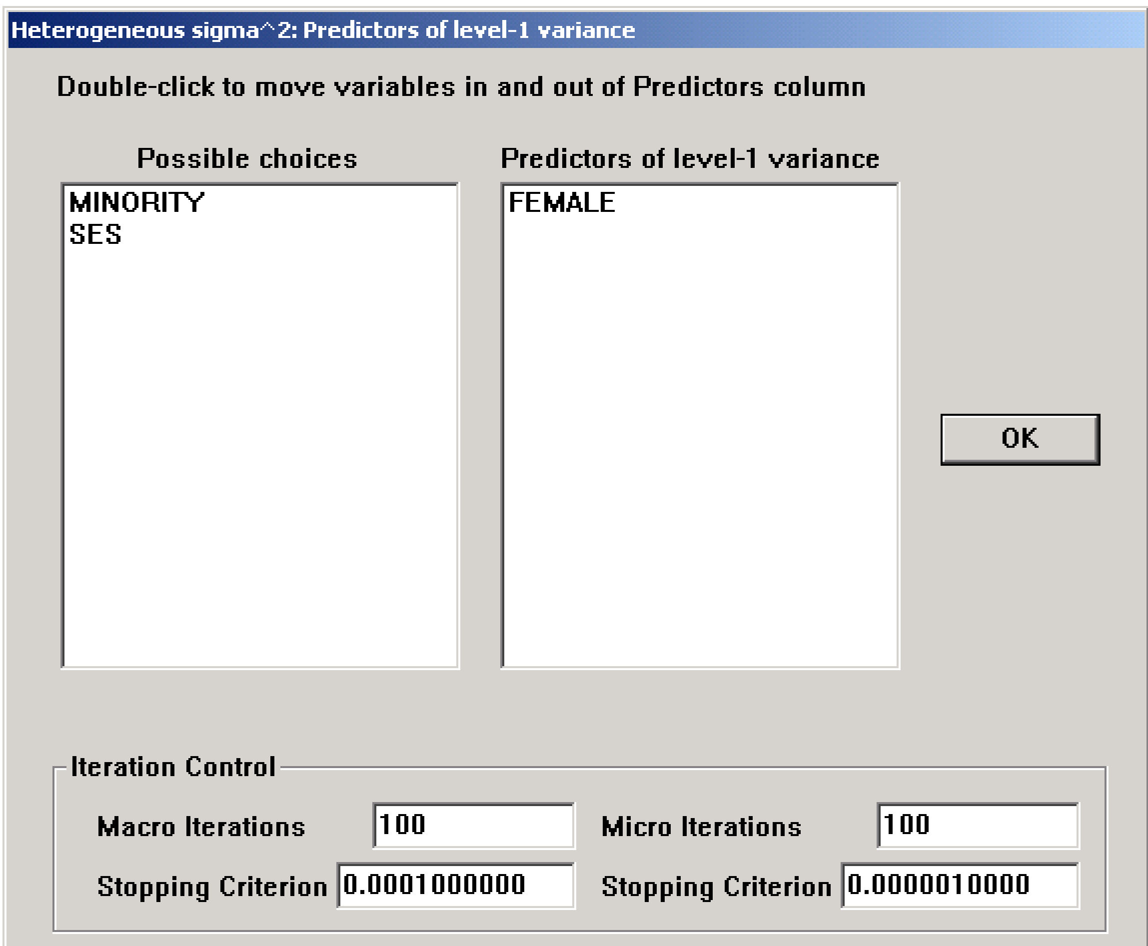


Figure 2.37 Heterogeneous sigma^2: Predictors of level-1 variance dialog box

The model estimated is a log linear-model for the level-1 variances, which can be generally stated as:

$$\sigma_{ij}^2 = \exp\left\{\alpha_0 + \alpha_1 \text{FEMALE}_{ij}\right\}$$

The following is a selected annotated output of the model run.

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	13.345271	0.253915	52.558	159	<0.001
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.359401	0.171411	-7.931	7024	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	13.345271	0.260426	51.244	159	<0.001
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.359401	0.185181	-7.341	7024	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, u_0	2.84757	8.10864	159	1601.08000	<0.001
level-1, r	6.23256	38.84483			

Statistics for the current model

Deviance = 47051.483085
Number of estimated parameters = 4

Results for Heterogeneous σ^2
(macro iteration 4)

Var(R) = σ^2 and $\log(\sigma^2) = \alpha_0 + \alpha_1$(FEMALE)

Model for level-1 variance

Parameter	Coefficient	Standard Error	*Z*-ratio	*p*-value
INTRCPT1 ,α_0	3.70771	0.024645	150.444	0.000
FEMALE ,α_1	-0.09307	0.034023	-2.736	0.007

Summary of Model Fit

Model	Number of Parameters	Deviance
1. Homogeneous σ^2	4	47051.48309
2. Heterogeneous σ^2	5	47044.02705

Model Comparison	χ^2	d.f.	*p*-value
Model 1 vs Model 2	7.45604	1	0.006

The Z-ratio for γ_0 (Z = -7.341) and Z-ratio for α_1 (Z =-2.736) for FEMALE indicate that the math achievement scores of males are on average higher as well as more variable than those for females. Furthermore, a comparison of the fits of the models suggests that the model with heterogeneous within-school variances appears appropriate (χ^2 = 7.45604, df = 1). See Chapter 10 in this manual for details on model comparisons.

2.8.6 Specifying level-1 deletion variables

If, when making the MDM file, "Delete missing data when running analyses" was specified, this feature may be used to alter the default behavior of the programs. By default, the programs will delete missing data on the basis of the level-1 variables actually in the model. While in many cases this is the desired behavior, in other situations it may not be. For instance, one might be running and comparing analyses that have different level-1 models. With many datasets, this can lead to comparing results that have a different number of level-1 records used. To solve this problem, check the option to delete missing data "when making the MDM file" (see Figure 2.30).

2.8.7 Using design weights

In many studies, data arise from sample surveys in which units have been selected with known but unequal probabilities. In these cases, it will often be desirable to weight observations in order to produce unbiased estimates of population parameters. According to standard practice in such cases, the information from each unit is weighted inversely proportional to its probability of selection.

Suppose, for instance, that in a pre-election poll, ethnic minority voters are over-sampled to insure that various ethnic groups are represented in the sample. Without weighting, the over-sampled groups would exert undue influence on estimates of the proportion of voters in the population favoring a specific candidate. Use of design weights can yield unbiased estimates of the population parameters.

Design weights are also commonly used to correct for differential non-response of sub-groups. Response rates are estimated for relevant sub-groups, and information from each respondent is weighted inversely proportional to the probability of response. That way, respondents who are over-represented in a sample as a function of non-response are appropriately weighted down.

2.8.4.1 Design weighting in the hierarchical context

Hierarchical data can be described as arising from a multi-stage sampling procedure. For example, schools might be sampled from a national frame of schools and then, within each school, students might then be sampled from a list of all students attending the school. Probabilities at each level might be known but unequal. For example, one might over-sample private schools and then over-sample minority students within each school. Weights might be constructed at each level to be inversely proportional to the probability of selection at that level. In some cases, weights might be available at only one level. For example, in a two-level design with students nested within schools, one might compute the marginal probability that a student is selected as the product of the probability that student's school is selected multiplied by the conditional probability that the student is selected given that his or her school is selected. In another context, suppose persons are selected with known probability and then followed longitudinally over time. In this case, we have occasions at level 1 nested within persons at level 2. The only weight may be a level-2 weight, inversely proportional to the probability of selection of that person. It is, of course, possible to include level-1 weights as well, but it is common to have weights only at level-2 in such longitudinal studies.

HLM 7 uses a method of computation devised by Pfefferman *et al*. (1998) for hierarchical data. This method, based on weighting the information of each case in the framework of maximum likelihood, is more appropriate than the method of weighting in earlier versions of HLM, which used a more conventional approach of weighting observations.

2.8.7.1 Weighting in two-level designs

In the two-level context, weights might be available at level 1, at level 2 or at both levels. If weights are available at level-1 only, the methodology used in HLM 7 assumes that these weights are inversely proportional to P_{ij}, the marginal probability of that student i in school j is selected into the sample. HLM 7 will then normalize the weight to have a mean of 1.0. Thus we have

$$w_{ij} = \frac{N / P_{ij}}{\sum_{j=1}^{J} \sum_{i=1}^{n_j} 1 / P_{ij}} \tag{0.0}$$

in which case

$$\sum_{J=1}^{J} \sum_{i=1}^{n_j} w_{ij} = N \tag{0.0}$$

where N is the total sample size of level-1 units. In contrast, if weights are available only at level 2, the methodology assumes that these weights are inversely proportional to P_j the probability of selection of the level-2 unit. In this case, HLM 7 will again normalize the weight to have a mean of 1.0, yielding

$$w_j = \frac{J/P_{ij}}{\sum_{j=1}^{J} 1 / P_{ij}}, \tag{0.0}$$

in which case

$$\sum_{j=1}^{J} w_j = J. \tag{0.0}$$

where J is the total number of level-2 units. If weights are available at both level-1 and level-2, the methodology assumes that the level-1 weight is $P_{i|j}$, the conditional probability of selection of unit i given that unit j was selected, so that $P_{i|j} = P_{ij} \mid P_j$. The level-2 weight is assumed to be inversely proportional to P_j. In this case, HLM will normalize the level-1 weight within level-2 units:

$$w_{i|j} = \frac{n_j / P_{i|j}}{\sum_{i=1}^{n_j} 1/P_{i|j}} \tag{0.0}$$

so that the sum of these weights within a level-2 unit will be

$$\sum_{i=1}^{n_j} w_{i|j} = n_j \tag{0.0}$$

where n_j is the sample size of level-1 units in level-2 unit j.

2.8.7.2 Weighting in three-level designs

In the three-level context, weights might be available at any one of the three levels, at any pair of them, or at all three levels. Normalization proceeds in a fashion completely analogous to that in the case of two levels. If weights are available only at level 1, we assume these are inversely proportional to P_{ijk}, the marginal probability of selection of unit *ijk*. Similarly, if weights are available only at level 2 or only at level 3, the corresponding probabilities are P_{jk} or P_k, respectively. If the weights are at levels 1 and 2 but not 3, the corresponding probabilities are $P_{i|jk}$ and P_{jk}; if at levels 2 and 3 (but not 1), the corresponding probabilities are $P_{j|k}$ and P_k; if the weights are at levels 1 and 3 (but not 2), the corresponding probabilities are $P_{i|k}$ and P_k. If weights are present at all three levels, the probabilities are $P_{i|jk}$, $P_{j|k}$ and P_k .

To apply weights for both levels

In HLM, weights are selected at the time of analysis, not when the MDM file is made:

1. Select the **Estimation Settings** option from the **Other Settings** menu.
2. Click the **Weighting** button to access the pull-down menus used to select the weighting variables at any level.

Note that the cover sheet of each HLM output reminds the user of the weighting specification chosen.

2.8.8 Hypothesis testing

2.8.8.1 Multivariate hypothesis tests for fixed effects

HLM allows multivariate hypothesis tests for the fixed effects. For instance, for the model displayed in Fig. 2.39, a user can test the following composite null hypothesis:

$$H_0 : \gamma_{01} = \gamma_{11} = 0,$$

where γ_{01} is the effect of SECTOR on the intercept and γ_{11} is the effect of sector on the SES slope.

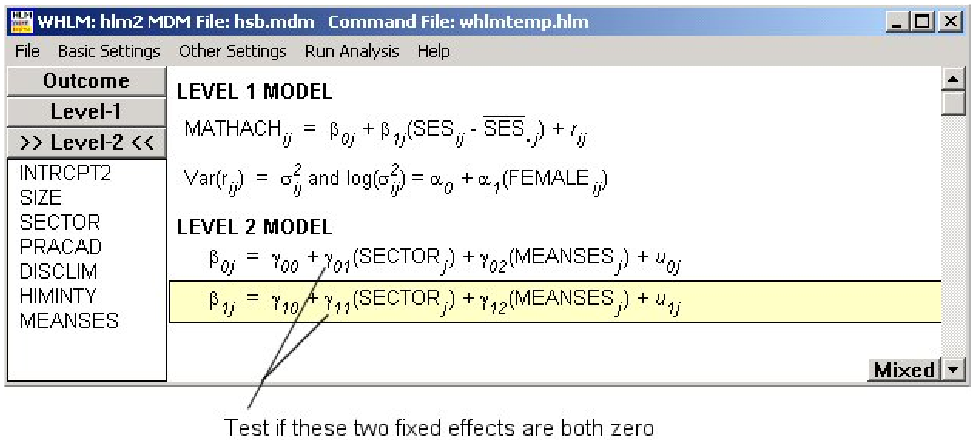


Figure 2.39 Model window

Below is a procedure that illustrates a Windows execution of the hypothesis test.

To pose a multivariate hypothesis test among the fixed effects

1. Open the **Other Settings** menu and select the **Hypothesis Settings** option to open the **Hypothesis Testing – HLM2** dialog box (See Figure 2.40).
2. Click "1" to open the **General Linear Hypothesis: Hypothesis 1** dialog box and to specify the first hypothesis (see Fig 2.41 for the contrasts for testing both of the effects of SECTOR on the intercept and on the SES slope as null, see *Hierarchical Linear Models,* p. 82). Then, click the "2" button for the second column and enter a 1 on the γ_{11} line in the second column. Click **OK**.

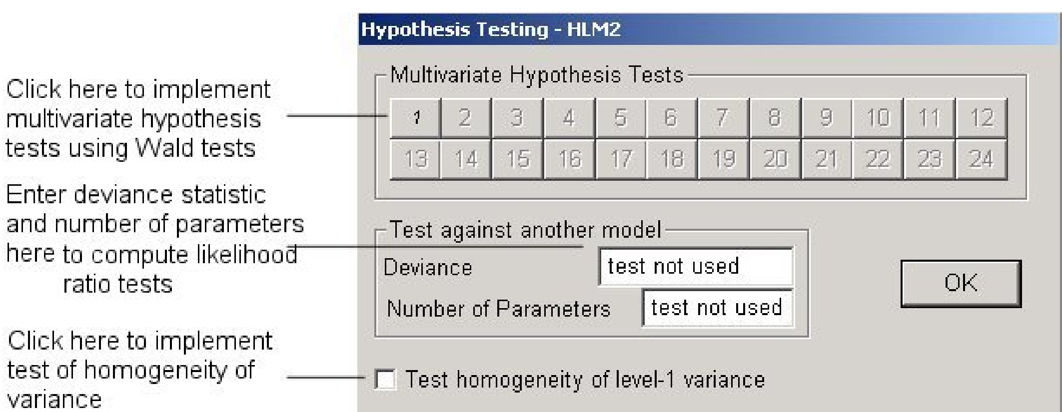


Figure 2.40 Optional Hypothesis Testing/Estimation dialog box

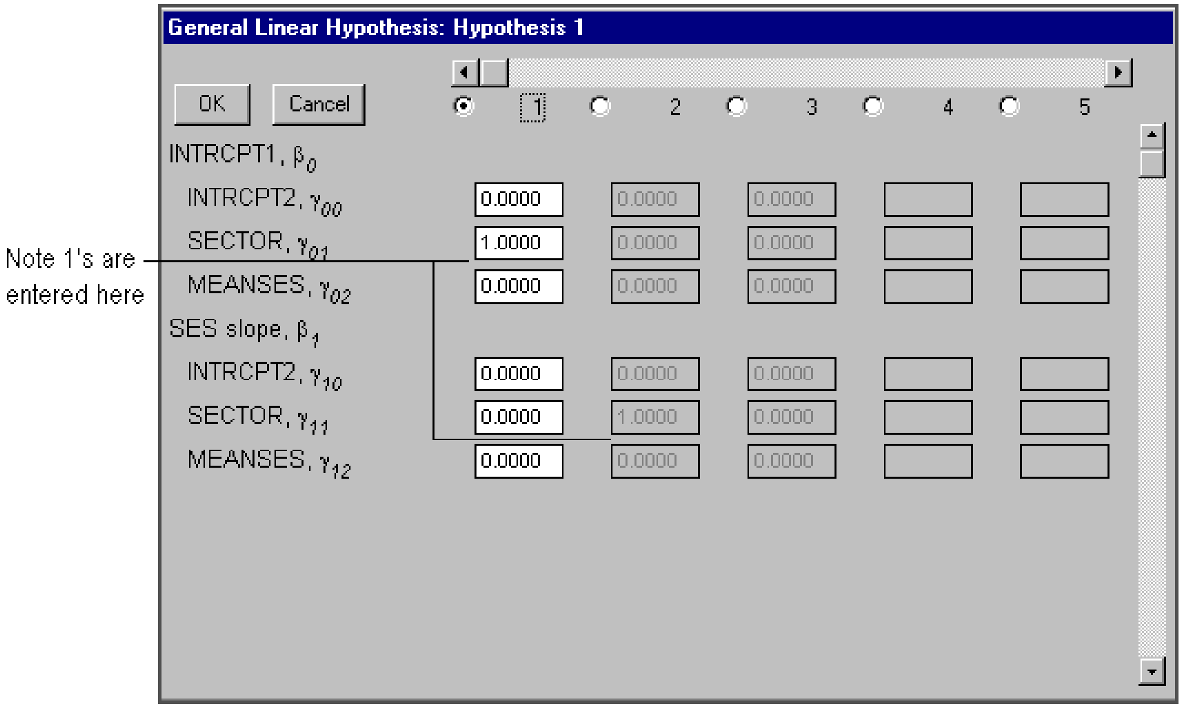


Figure 2.41 General Linear Hypothesis: Hypothesis 1 dialog box

The HLM2 output associated with this test appears in Section 2.8.8.3 below. (For a further discussion of this multivariate hypothesis test for fixed effects see *Hierarchical Linear Models*, pp. 58-61, 81-85).

2.8.8.2 Testing homogeneity of level-1 variances

By default, HLM2 assumes homogeneity of residual variance at level 1. That is, it specifies a common σ^2 within each of the J level-2 units. As an option, HLM2 tests the adequacy of this assumption.

To test homogeneity of level-1 variances

1. Click the **Test homogeneity of level-1 variance** box (Figure 2.40).
2. The HLM2 output associated with this test also appears in Section 2.8.8.3 below. (For a further discussion of this test see *Hierarchical Linear Models*, pp. 263-267. We advise that users review these pages carefully before using this procedure.)

2.8.8.3 Multivariate tests of variance-covariance components specification

HLM2 also provides, as an option, a multi-parameter test for the variance-covariance components. This likelihood-ratio test compares the deviance statistic of a restricted model with a more general alternative. The user must input the value of the deviance statistic and related degrees of freedom for the alternative specification. Below we compare the variance-covariance components of two Intercept-and-Slope-as-Outcome models. One treats β_1 as random and the other does not.

To specify a multivariate test of variance-covariance components

Enter the deviance and the number of parameters in the **Deviance Statistics** box and in the

Number of Parameters box (see Fig. 2.40) respectively (the two numbers for our example are 46512.978000 and 4, obtained in Section 2.5.3).

The HLM2 output associated with this test appears in the section below. (For a further discussion of this multi-parameter test see *Hierarchical Linear Models*, pp. 63-65, 83-85). Below is an example of a selected HLM2 output that illustrates optional hypothesis testing procedures.

The outcome variable is MATHACH

Summary of the model specified

Level-1 Model

$\text{MATHACH}_{ij} = \beta_{0j} + \beta_{1j}*(\text{SES}_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(\text{SECTOR}j) + \gamma_{02}*(\text{MEANSES}j) + u_{0j}$
$\beta_{1j} = \gamma_{10} + \gamma_{11}*(\text{SECTOR}j) + \gamma_{12}*(\text{MEANSES}j) + u_{1j}$

SES has been centered around the group mean.

MEANSES has been centered around the grand mean.

Mixed Model

$\text{MATHACH}_{ij} = \gamma_{00} + \gamma_{01}*\text{SECTOR}_j + \gamma_{02}*\text{MEANSES}_j$
$+ \gamma_{10}*\text{SES}_{ij} + \gamma_{11}*\text{SECTOR}_j*\text{SES}_{ij} + \gamma_{12}*\text{MEANSES}_j*\text{SES}_{ij}$
$+ u_{0j} + u_{1j}*\text{SES}$

Note, the middle section of output has been deleted. We proceed directly to the final results page.

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.095250	0.198627	60.894	157	<0.001
SECTOR, γ_{01}	1.224401	0.306117	4.000	157	<0.001
MEANSES, γ_{02}	5.336698	0.368978	14.463	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.935664	0.150690	19.482	7022	<0.001
SECTOR, γ_{11}	-1.642102	0.233097	-7.045	7022	<0.001
MEANSES, γ_{12}	1.044120	0.291042	3.588	7022	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.095250	0.173679	69.641	157	<0.001
SECTOR, γ_{01}	1.224401	0.308507	3.969	157	<0.001
MEANSES, γ_{02}	5.336698	0.334617	15.949	157	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.935664	0.147576	19.893	7022	<0.001
SECTOR, γ_{11}	-1.642102	0.237223	-6.922	7022	<0.001
MEANSES, γ_{12}	1.044120	0.332897	3.136	7022	0.002

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	1.54118	2.37524	157	604.29895	<0.001
level-1, r	6.06351	36.76611			

Statistics for current covariance components model

Deviance = 46502.952743
Number of estimated parameters = 2

For the likelihood ratio test, the deviance statistic reported above is compared with the value from the alternative model manually. The result of this test appears below.

Variance-Covariance components test

χ^2 statistic = 10.02526
Degrees of freedom = 2
p-value = 0.007

A model that constrains the residual variance for the SES slopes, β_1, to zero appears appropriate. (For a further discussion of this application see *Hierarchical Linear Models*, pp. 83-85.)

Test of homogeneity of level-1 variance

χ^2 statistic = 244.08638
degrees of freedom = 159
p-value = 0.000

These results indicate that there is variability among the (J = 160) level-2 units in terms of the residual within-school (*i.e.*, level-1) variance. (For a full discussion of these results see *Hierarchical Linear Models*, pp. 263-267.)

Results of General Linear Hypothesis Testing - Test 1

	Coefficients	Contrast	
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	12.095250	0.0000	0.0000
SECTOR, γ_{01}	1.224401	1.0000	0.0000
MEANSES, γ_{02}	5.336698	0.0000	0.0000
For SES slope, β_1			
INTRCPT2, γ_{10}	2.935664	0.0000	0.0000
SECTOR, γ_{11}	-1.642102	0.0000	1.0000
MEANSES, γ_{12}	1.044120	0.0000	0.0000
Estimate		1.2244	-1.6421
Standard error of estimate		0.3085	0.2372

χ^2 statistic = 60.527852
Degrees of freedom = 2
p-value = <0.001

The table above is a reminder of the multivariate contrast specified. The chi-square statistic and associated *p*-value indicate that it is highly unlikely that the observed estimates for γ_{01} and γ_{11} could have occurred under the specified null hypothesis.

2.9 Output options

There are a few options relating to the output that can be selected on the **Other Settings**, **Output Settings** menu:

- **# of OLS estimates shown** (HLM2 only) – this controls the number of OLS estimates printed in the output. See the output in Section 2.5.3.
- **Print variance-covariance matrices** – see Section A.5.
- **Print reduced output** – if this is checked, only the header page and the final results are printed.

Starting values, OLS estimates (if present), etc. will not be printed.

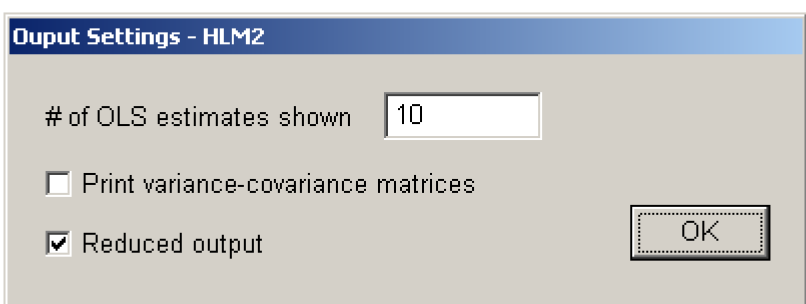


Figure 2.42 Output Settings – HLM2 dialog box

2.10 Models without a level-1 intercept

In some circumstances, users may wish to estimate models without a level-1 intercept. Consider, for example, a hypothetical study in which three alternative treatments are implemented within each of J hospitals. One might estimate the following level-1 (within-hospital) model:

$$Y_{ij} = \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \beta_{3j} X_{3ij} + r_{ij},$$

where X_{qij} (q = 1,2,3) are indicator variables taking on a value of 1 if patient i in hospital j has received treatment q, 0 otherwise; and β_{qj} is the mean outcome in hospital j of those receiving treatment q. At level-2, the treatment means β_{qj} are predicted by characteristics of the hospitals. Of course, the same data could alternatively be modeled by a level-1 intercept and two treatment contrasts per hospital, but users will sometimes find the no-intercept approach is more convenient.

An example of a no-intercept model appears on page 174 of *Hierarchical Linear Models*. The vocabulary growth of young children is of interest. Both common sense and the data indicated that children could be expected to have no vocabulary at 12 months of age. Hence, the level-1 model contained no intercept:

$$Y_{ti} = \pi_{1i}(\mathrm{AGE}_{ti} - 12) + \pi_{2i}(\mathrm{AGE}_{ti} - 12)^2 + e_{ti}$$

where AGE_{ti} is the age of child i at time t in months and Y_{ti} is the size of that child's vocabulary at that time.

To delete an intercept from a level-1 model

Click INTRCPT1 on the **>>Level-1<<** drop-down list. Click **delete variable from model**.

2.11 Coefficients having a random effect with no corresponding fixed effect

A user may find it useful at times to model a level-1 predictor as having a random effect but no fixed effect. For example, it might be that gender differences in educational achievement are, on average, null across a set of schools; yet, in some schools females outperform males while in other schools males outperform females. In this case, the fixed effect of gender could be set to zero while the variance of the gender effect across schools would be estimated.

The vocabulary analysis in *Hierarchical Linear Models* supplies an example of a level-1 predictor having a random effect without a corresponding fixed effect. For the age interval under study, it was found that, on average, the linear effect of age was zero. Yet this effect varied significantly across children. The level-1 model estimated was:

$$Y_{ti} = \pi_{1i}\left(\mathrm{AGE}_{ti} - 12\right) + \pi_{2i}\left(\mathrm{AGE}_{ti} - 12\right)^2 + e_{ti}$$

However, the level-2 model was:

$$\pi_{1i} = r_{1i}$$
$$\pi_{2i} = \beta_{20} + r_{2i}$$

Notice that AGE – 12 has a random effect but no fixed effect.

To delete the fixed effect from a level-2 model

1. Select the equation from which the fixed effect is to be removed.
2. Click INTRCPT2 on the **>>Level-2<<** drop-down list. Click **delete variable from model**.

2.12 Exploratory analysis of potential level-2 predictors

The user may be interested in computing "*t*-to-enter statistics" for potential level-2 predictors to guide specification of subsequent HLM2 models. The implementation procedure is as follows.

To implement exploratory analysis of potential level-2 predictors

1. Open the **Other Settings menu** and choose **Exploratory Analysis (level 2).** A **Select Variables For Exploratory Analysis** dialog box appears.
2. Click the equation associated with a regression coefficient to model the corresponding coefficient. Click to select variables for exploratory analysis. (Figure 2.43 displays the level-2 predictors chosen for our HS&B example).
3. Click **Return to Model Mode** to return to the model window.

The following contains a selected HLM2 output to illustrate exploratory analysis of potential level-2 predictors.

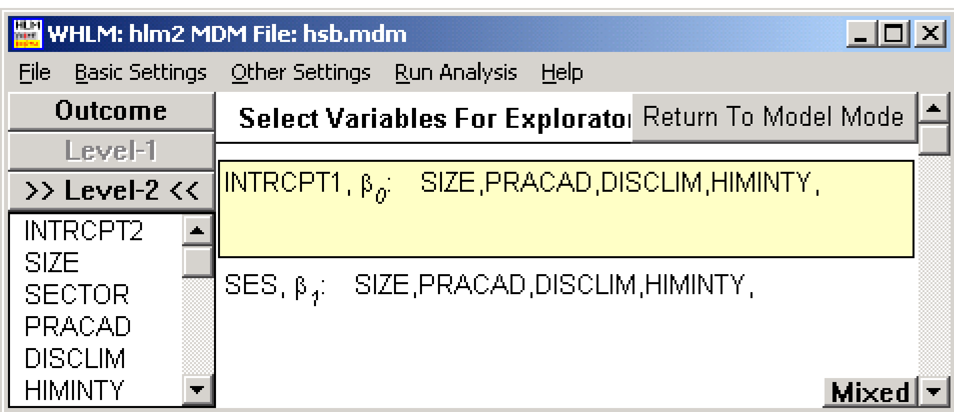


Figure 2.43 Select Variables For Exploratory Analysis dialog box for the HS&B example

Exploratory Analysis: estimated level-2 coefficients and their standard errors obtained by regressing EB residuals on level-2 predictors selected for possible inclusion in subsequent HLM runs

Level-1 Coefficient	Potential Level-2 Predictors			
INTRCPT1, β_0				
	SIZE	PRACAD	DISCLIM	HIMINTY
Coefficient	0.000	0.690	-0.161	-0.543
Standard Error	0.000	0.404	0.106	0.229
t-value	1.569	1.707	-1.515	-2.372
SES, β_1				
	SIZE	PRACAD	DISCLIM	HIMINTY
Coefficient	0.000	0.039	-0.005	-0.058
Standard Error	0.000	0.044	0.012	0.025
t-value	1.297	0.899	-0.425	-2.339

The results of this exploratory analysis suggest that HIMINTY might be a good candidate to include in the INTRCPT1 model. The t-values represent the approximate result that will be obtained when one additional predictor is added to any of the level-2 equations. This means that if HIMINTY is added to the model for the INTRCPT1, for example, the apparent relationship suggested above for HIMINTY in the SES slope model might disappear. (For a further discussion of the use of these statistics see discussion in *Hierarchical Linear Models*, p. 270 on "Approximate *t*-to-Enter Statistics.")

3. Conceptual and Statistical Background for Three-Level Models

The models estimated by HLM3 are applicable to a hierarchical data structure with three levels of random variation in which the errors of prediction at each level can be assumed to be approximately normally distributed. Consider, for example, a study in which achievement test scores are collected from a sample of children nested within classrooms that are in turn nested within schools. This data structure is hierarchical (each child belongs to one and only one classroom and each classroom belongs to one and only one school); and there are three levels of random variation: variation among children within classrooms, variation among classrooms within schools, and variation among schools. The outcome (achievement test scores) makes the normality assumption at level 1 reasonable, and the normality assumption at the classroom and school levels will often also be a sensible one.

Chapter 8 of *Hierarchical Linear Models* discusses several applications of a three-level model. The first is a three-level cross-sectional study as described above. A second case involves time-series data collected on each subject where the subjects are nested within organizations. This latter example is from the *Sustaining Effects Study*, where achievement data were collected at five time points for each child. Here the time-series data are nested within children and the children are nested within schools. A third example in Chapter 8 involves measures taken on each of the multiple classes taught by secondary school teachers. The classes are nested within teachers and the teachers within schools. A final example involves multiple items from a questionnaire administered to teachers. The items vary "within teachers" at level 1, the teachers vary within schools at level 2, and the schools vary at level 3. In effect, the level-1 model is a model for the measurement error associated with the questionnaire. Clearly, there are many interesting applications of a three-level model.

3.1 The general three-level model

The three-level model consists of three submodels, one for each level. For example, if the research problem consists of data on students nested within classrooms and classrooms within schools, the level-1 model will represent the relationships among the student-level variables, the level-2 model will capture the influence of class-level factors, and the level-3 model will incorporate school-level effects. Formally there are $i = 1, \ldots, n_{jk}$ level-1 units (*e.g.*, students), which are nested within each of $j = 1, \ldots, J_k$ level-2 units (*e.g.*, classrooms), which in turn are nested within each of $k = 1, \ldots, K$ level-3 units (*e.g.*, schools).

3.1.1 Level-1 model

In the level-1 model we represent the outcome for case i within level-2 unit j and level-3 unit k as:

$$
\begin{aligned}
Y_{ijk} &= \pi_{0jk} + \pi_{1jk}a_{1jk} + \pi_{2jk}a_{2jk} + \ldots + \pi_{pjk}a_{pjk} + e_{ijk} \\
&= \pi_{0jk} + \sum_{p=1}^{P} \pi_{pjk}a_{pjk} + e_{ijk}
\end{aligned} \qquad (0.0)
$$

where

π_{pjk} ($p = 0,1,..., P$) are *level-1 coefficients*,

a_{pjk} is a *level-1 predictor p* for case *i* in level-2 unit *j* and level-3 unit *k*,

e_{ijk} is the *level-1 random effect*, and

σ^2 is the variance of e_{ijk}, that is the *level-1 variance*.

Here we assume that the random term $e_{ijk} \sim N(0, \sigma^2)$.

3.1.2 Level-2 model

Each of the π_{pjk} coefficients in the level-1 model becomes an outcome variable in the level-2 model:

$$
\begin{aligned}
\pi_{pjk} &= \beta_{p0k} + \beta_{p1k}X_{1jk} + \beta_{p2k}X_{2jk} + \ldots + \beta_{pQ_pk}X_{Q_pjk} + r_{pjk} \\
&= \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk}X_{qjk} + r_{pjk},
\end{aligned} \qquad (0.0)
$$

where

β_{pqk} ($q = 0,1,...,Q_p$) are *level-2 coefficients*,

X_{qjk} is a *level-2 predictor*, and

r_{pjk} is a *level-2 random effect.*

We assume that, for each unit *j*, the vector $\left(r_{0jk},\ r_{1jk}, \ldots, r_{Pjk}\right)'$ is distributed as multivariate normal where each element has a mean of zero and the variance of r_{pjk} is:

$$\mathrm{Var}(r_{pjk}) = \tau_{\pi pp}. \qquad (0.0)$$

For any pair of random effects *p* and p',

$$\mathrm{Cov}(r_{pjk}, r_{p'jk}) = \tau_{\pi pp'}. \qquad (0.0)$$

These level-2 variance and covariance components can be collected into a dispersion matrix, $\mathbf{T}_\pi$, with a maximum dimension is $(P+1)\times(P+1)$.

We note that each level-1 coefficient can be modeled at level 2 as one of three general forms:

a level-1 coefficient that is fixed at the same value for all level-2 units; *e.g.*,

$$\pi_{pjk} = \beta_{p0k}\,, \tag{0.0}$$

a level-1 coefficient that varies non-randomly among level-2 units, *e.g.*,

$$\pi_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk}\,, \tag{0.0}$$

a level-1 coefficient that varies randomly among level-2 units, *e.g.*,

$$\pi_{pjk} = \beta_{p0k} + r_{pjk} \tag{0.0}$$

or

$$\pi_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk} + r_{pjk}. \tag{0.0}$$

The actual dimension of $\mathbf{T}_\pi$ in any application depends on the number of level-1 coefficients specified as randomly varying. We also note that a different set of level-2 predictors may be used in each of the $P+1$ equations that form the level-2 model.

3.1.3 Level-3 model

Each of the level-2 coefficients, β_{pqk}, defined in the level-2 model becomes an outcome variable in the level-3 model:

$$\begin{aligned} \beta_{pqk} &= \gamma_{pq0} + \gamma_{pq1} W_{1k} + \gamma_{pq2} W_{2k} + \ldots + \gamma_{pqS_{pq}} W_{S_{pq}k} + u_{pqk} \\ &= \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} W_{sk} + u_{pqk}, \end{aligned} \tag{0.0}$$

where

γ_{pqs} ($s = 0, 1, \ldots, S_{pq}$) are *level-3 coefficients*,

W_{sk} is a *level-3 predictor*, and

u_{pqk} is a *level-3 random effect*.

We assume that, for each level-3 unit, the vector of level-3 random effects (the u_{pqk} terms) is distributed as multivariate normal, with each having a mean of zero and with covariance matrix $\mathbf{T}_\beta$, whose maximum dimension is:

$$\sum_{p=0}^{p}(Q_p+1)\times\sum_{p=0}^{p}(Q_p+1), \tag{0.0}$$

We note that each level-2 coefficient can be modeled at level-3 as one of three general forms:

1. *as a fixed effect*, *e.g.*,

$$\beta_{pqk}=\gamma_{pq0}, \tag{0.0}$$

2. *as non-randomly varying*, *e.g.*

$$\beta_{pqk}=\gamma_{pq0}+\sum_{s=1}^{S_{pq}}\gamma_{pqs}W_{sk}, \tag{0.0}$$

3. *as randomly varying*, *e.g.*

$$\beta_{pqk}=\gamma_{pq0}+u_{pqk} \tag{0.0}$$

or

$$\beta_{pqk}=\gamma_{pq0}+\sum_{s=1}^{S_{pq}}\gamma_{pqs}W_{sk}+u_{pqk}. \tag{0.0}$$

The actual dimension of $\mathbf{T}_\beta$ in any application depends on the number of level-3 coefficients specified as randomly varying. We also note that a different set of level-3 predictors may be used in each equation of the level-3 model.

3.2 Parameter estimation

Three kinds of parameter estimates are available in a three-level model: empirical Bayes estimates of randomly varying level-1 and level-2 coefficients; maximum-likelihood estimates of the level-3 coefficients (note: these are also generalized least squares estimates); and maximum-likelihood estimates of the variance-covariance components. The maximum-likelihood estimate of the level-3 coefficients and the variance-covariance components are printed on the output for every run. The empirical Bayes estimates for the level-1 and level-2 coefficients may optionally be saved in the "residual files" at levels 2 and 3, respectively. Reliability estimates for each random level-1 and level-2 coefficient are always produced. The actual estimation procedure for the three-level model differs a bit from the default two-level model. By default, HLM2 uses a "restricted maximum likelihood" approach in which the variance-covariance components are estimated by means of maximum likelihood and then the fixed effects (level-2 coefficients) are estimated via generalized least squares given those variance-covariance estimates. In HLM3, not only the variance-covariance components, but also the fixed effects (level-3 coefficients) are estimated by means of maximum likelihood. This procedure is referred to as "full" as opposed to "restricted" maximum likelihood (For a further discussion of this see *Hierarchical Linear Models*, pp. 52-53). Note that full maximum likelihood is also available as an option for HLM2.

3.3 Hypothesis testing

As in the case of the two-level program, the three-level program routinely prints standard errors and *t*-tests for each of the level-3 coefficients ("the fixed effects") as well as a chi-square test of homogeneity for each random effect. In addition, optional "multivariate hypothesis tests" are available in the three-level program. Multivariate tests for the level-3 coefficients enable both omnibus tests and specific comparisons of the parameter estimates just as described in the section *Multivariate hypothesis tests for fixed effects* in this chapter. Multivariate tests regarding alternative variance-covariance structures at level 2 or level 3 proceed just as in the section *Multivariate tests of variance-covariance components specification* in this chapter.

The use of full maximum likelihood for parameter estimation in HLM3 has a consequence for hypothesis testing. For both restricted and full maximum likelihood, one can test alternative variance-covariance structures by means of the likelihood-ratio test as described in the section *Multivariate tests of variance-covariance components specification*. However, in the case of full maximum likelihood, it is also possible to test alternative specifications of the fixed coefficients by means of a likelihood-ratio test. In fact, any pair of nested models can be compared using the likelihood-ratio test under full maximum likelihood. By nested models, we refer to a pair of models in which the simpler model can be derived by imposing constraints on the parameters of the more complex model. Any pair of nested two-level models can be compared using a likelihood ratio test.

4 Working with HLM3

As in the case of the two-level program, data analysis by means of the HLM3 program will typically involve three stages:

- Construction of an MDM file (the multivariate data matrix)
- Execution of analyses based on the MDM file
- Evaluation of fitted models based on residual files

As in HLM2, HLM3 analyses can be executed in Windows, interactive, and batch modes. We describe a Windows execution below. We consider interactive and batch execution in Appendix B. A number of special options are presented at the end of the chapter.

4.1 An example using HLM3 in Windows mode

Chapter 8 in *Hierarchical Linear Models* presents a series of analyses of data from the US *Sustaining Effects Study*, a longitudinal study of children's growth in academic achievement during the primary years. A level-1 model specifies the relationship between age and academic achievement for each child. At level 2, the coefficients describing each child's growth vary across children within schools as a function of demographic variables. At level 3, the parameters that describe the distribution of growth curves within each school vary across schools as a function of school-level predictors.

To illustrate the operation of the HLM3 program, we analyze another data set having a similar structure. The level-1 data are time-series observations on 1721 students nested within 60 urban public primary schools and mathematics achievement is the outcome. These data are provided along with the HLM software so that a user may replicate our results in order to assure that the program is operating correctly.

4.1.1 Constructing the MDM file from raw data

In constructing the MDM file, the user has the same range of options for data input for HLM3 as for HLM2 (see Section 2.5.1). We first describe the use of SPSS file input and then consider ASCII, SYSTAT, SAS, and other data file formats.

4.1.1.1 SPSS input

Data input requires a level-1 file (in our illustration a time-series data file), a level-2 file (child-level file), and a level-3 (school-level) file.

Level-1 file.

The level-1 file, EG1.SAV, has 7242 observations collected on 1721 children beginning at the end of grade one and followed up annually thereafter until grade six. There are four level-1 variables (not including the schoolid and the childid). Time-series data for the first two children are shown in Figure 4.1.

There are eight records listed, three for the first child and five for the second. (Typically there are four or five observations per child with a maximum of six.) The first ID is the level-3 (*i.e.*, school)

ID and the second ID is the level-2 (*i.e.*, child) ID. We see that the first record comes from school 2020 and child 273026452 within that school. Notice that this child has three records, one for each of three measurement occasions. Following the two ID fields are that child's values on four variables:

- YEAR (year of the study minus 3.5)

  This variable can take on values of -2.5, -1.5, -0.5, 0.5, 1.5, and 2.5 for the six years of data collection.
- GRADE

  The grade level minus 1.0 of the child at each testing occasion. Therefore, it is 0 at grade 1, 1 at grade 2, etc.
- MATH

  A math test in an IRT scale score metric.
- RETAINED

  An indicator that a child is retained in grade for a particular year (1 = retained, 0 = not retained).

	schoolid	childid	year	grade	math	retained
1	2020	273026452	.50	2.00	1.15	.00
2	2020	273026452	1.50	3.00	1.13	.00
3	2020	273026452	2.50	4.00	2.30	.00
4	2020	273030991	.50	2.00	2.43	.00
5	2020	273030991	1.50	3.00	2.25	.00
6	2020	273030991	2.50	4.00	3.87	.00
7	2020	273030991	-.50	1.00	.44	.00
8	2020	273030991	-1.50	.00	-1.30	.00

Figure 4.1 First eight cases in EG1.SAV

We see that the first child, child 27306452 in school 2020, had values of 0.5, 1.5, and 2.5 on year. Clearly, that child had no data at the first three data collection waves (because we see no values of -2.5, -1.5, or -0.5 on year), but did have data at the last three waves. We see also that this child was not retained in grade during this period since the values for GRADE increase by 1 each year and since RETAINED takes on a value of 0 for each year. The three MATH scores of that child (1.15, 1.13, 2.30) show no growth in time period 1.5. Oddly enough, the time-series record for the second child (child 273030991 in school 2020) displays a similar pattern in the same testing.

Note: The level-1 and level-2 files must also be sorted in the same order of level-2 ID nested within level-3 ID, *e.g.*, children within schools. If this nested sorting is not performed, an incorrect multivariate data matrix file will result.

Level-2 file.

The level-2 units in the illustration are 1721 children. The data are stored in the file EG2.SAV. The level-2 data for the first eight children are listed below. The first field is the schoolid and the second is the childid. Note that each of the first ten children is in school 2020.

There are three variables:

- FEMALE (1 = female, 0 = male)
- BLACK (1 = African-American, 0 = other)
- HISPANIC (1= Hispanic, 0 = other)

We see, for example, that child 273026452 is a Hispanic male (FEMALE = 0, BLACK = 0, HISPANIC = 1).

	schoolid	childid	female	black	hispanic
1	2020	273026452	.00	.00	1.00
2	2020	273030991	.00	.00	.00
3	2020	273059461	.00	.00	1.00
4	2020	278058841	.00	.00	.00
5	2020	292017571	.00	.00	1.00
6	2020	292020281	.00	.00	.00
7	2020	292020361	.00	.00	.00
8	2020	292025081	.00	.00	.00

Figure 4.2 First eight children in EG2.SAV

Level-3 file.

The level-3 units in the illustration are 60 schools. Level-3 data for the first seven schools are printed below. The full data are in the file EG3.SAV. The first field on the left is the schoolid. There are three level-3 variables:

- SIZE, number of students enrolled in the school
- LOWINC, the percent of students from low income families
- MOBILE, the percent of students moving during the course of a single academic year

We see that the first school, school 2020, has 380 students, 40.3% of whom are low income. The school mobility rate is 12.5%.

	schoolid	size	lowinc	mobility
1	2020	380.00	40.30	12.50
2	2040	502.00	83.10	18.60
3	2180	777.00	96.60	44.40
4	2330	800.00	78.90	31.70
5	2340	1133.00	93.70	67.00
6	2380	439.00	36.90	39.30
7	2390	566.00	100.00	39.90

Figure 4.3 First seven schools in EG3.SAV

In sum, there are four variables at level 1, three at level 2 and three at level 3. Note that the ID variables do not count as variables. Once the user has identified the two sets of IDs, the number of variables in each file, the variable names, and the filenames, creation of the MDM file is exactly analogous to the three major steps described in the Section 2.5.1.1. The user first informs HLM that the input files are SPSS system files and the MDM is a three-level file. Then HLM is supplied with the appropriate information for the data. Note that the three files are linked by level-2 and level-3 IDs here.

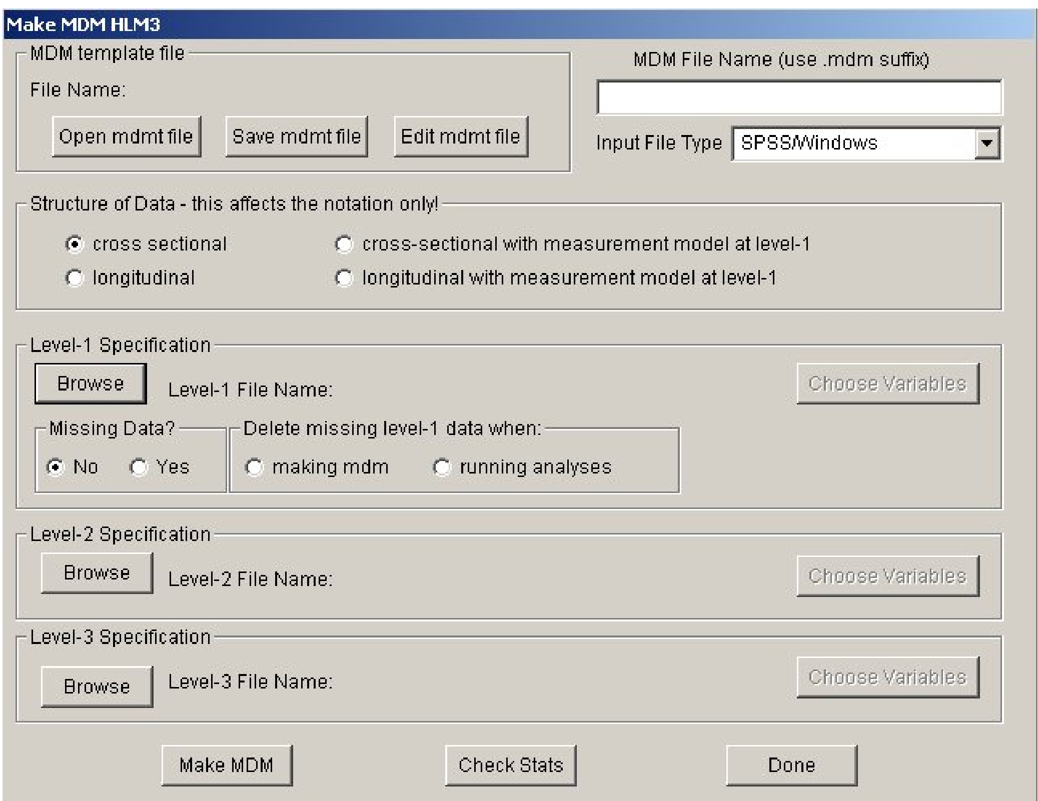


Figure 4.4 Make MDM – HLM3 dialog box for EG.MDM

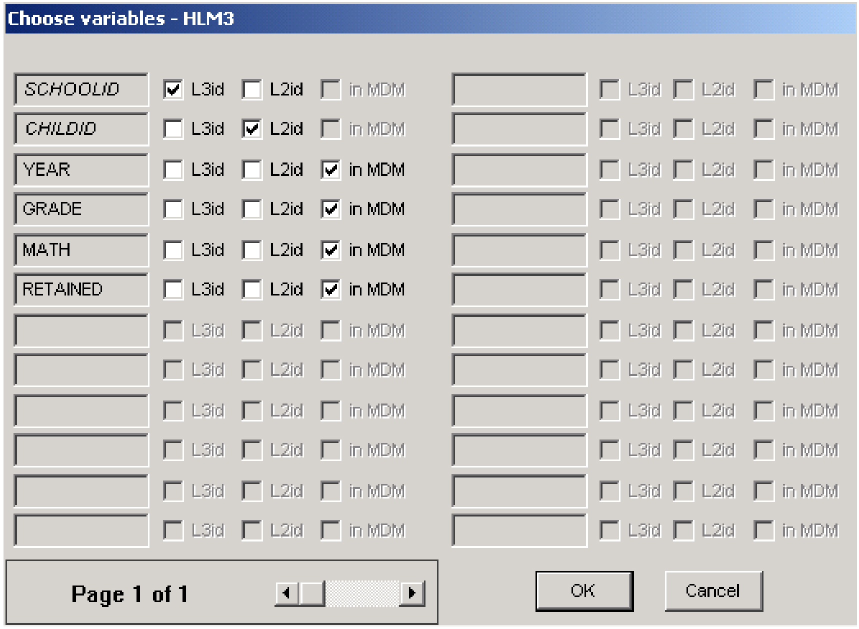


Figure 4.5 Choose Variables – HLM3 dialog box for level-1 file, EG1.SAV

Note: In addition, the program can handle **missing data at level-1 only**, with the same options available as discussed in HLM2. HLM3 will listwise delete cases with missing data at levels two and three. The three level program handles design weights at all three levels.

The response file, EGSPSS.MDMT, contains a log of the input responses used to create the MDM file, EG.MDM, using EG1.SAV, EG2.SAV, and EG3.SAV. Figure 4.4 displays the dialog box used to create the MDM file. Figure 4.5 shows the dialog box for the level-1 file, EG1.SAV.

Note: As in the case of HLM2, after constructing the MDM file, you should check whether the data have been properly read into HLM by examining the descriptive statistics of the MDM file.

4.1.1.2 ASCII input

The procedure for constructing an MDM file from ASCII data files is similar to that for SPSS file input. The major difference is that the format statements must be entered for the three data files, variable names, and missing value codes, if applicable. Rules about the format are included in the Appendix. An example is included in the response file, EGASCII.MDMT, which constructs the MDM file, EGASCII.MDM, using EG1.DAT, EG2.DAT, and EG3.DAT. Figure 4.6 shows the dialog box for creating the MDM file, displaying the input responses of EGASCII.MDMT.

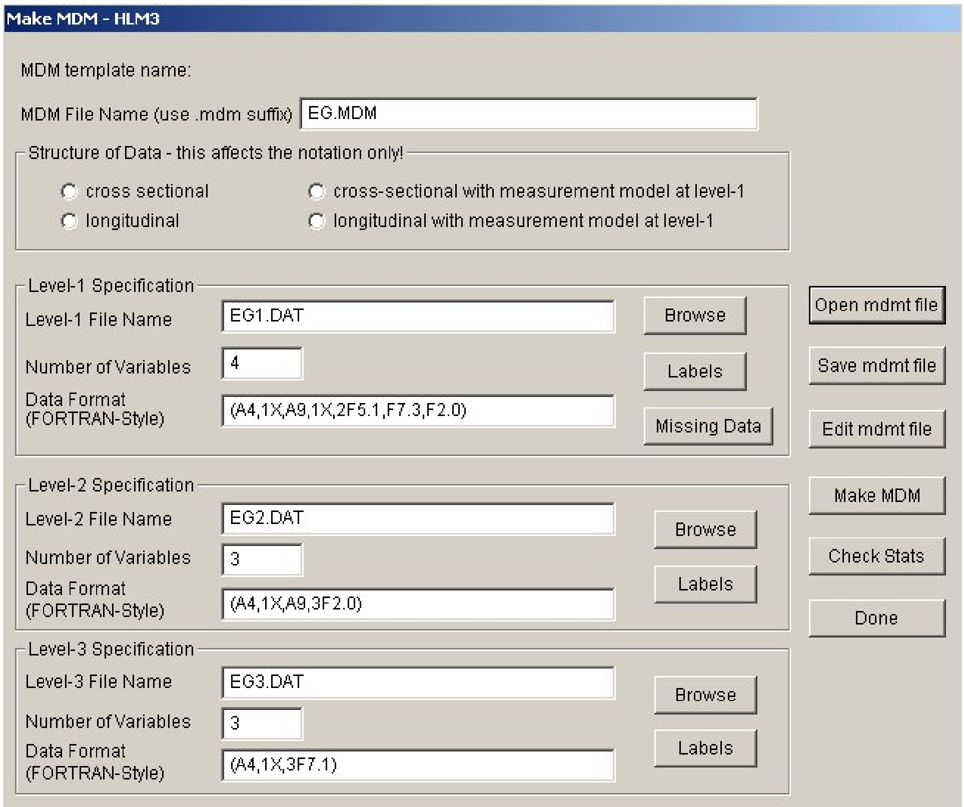


Figure 4.6 Make MDM – HLM3 dialog box for EGASCII.MDM

4.1.1.3 Other file input

For SAS and SYSTAT file input, a user selects either SAS5 transport or SYSTAT from the **Input File Type** drop-down list box as appropriate before clicking the **Browse** buttons in the file specification sections and follows the same steps for SPSS input type to create MDM files.

4.1.1.4 Other file type input

HLM3 has the same range of options for data input as HLM2. In addition to SYSTAT, SPSS, STATA, free format, and SAS, the Windows version (through a third-party module) allows numerous other data formats from, for example, EXCEL, and LOTUS input. See Section 2.5.1 for details.

4.2 Executing analyses based on the MDM file

Once the MDM file is constructed, it is used as input for the analysis. Model specification via the Windows mode has five steps:

1. Specification of the level-1 model. In our case we shall model mathematics achievement (MATH) as the outcome, to be predicted by YEAR in the study. Hence, the level-1 model will have two coefficients for each child: the intercept and the YEAR slope.
2. Specification of the level-2 prediction model. Here each level-1 coefficient – the intercept and the YEAR slope in our example – becomes an outcome variable. We may select certain child characteristics to predict each of these level-1 coefficients. In principle, the level-2 parameters then describe the distribution of growth curves within each school.

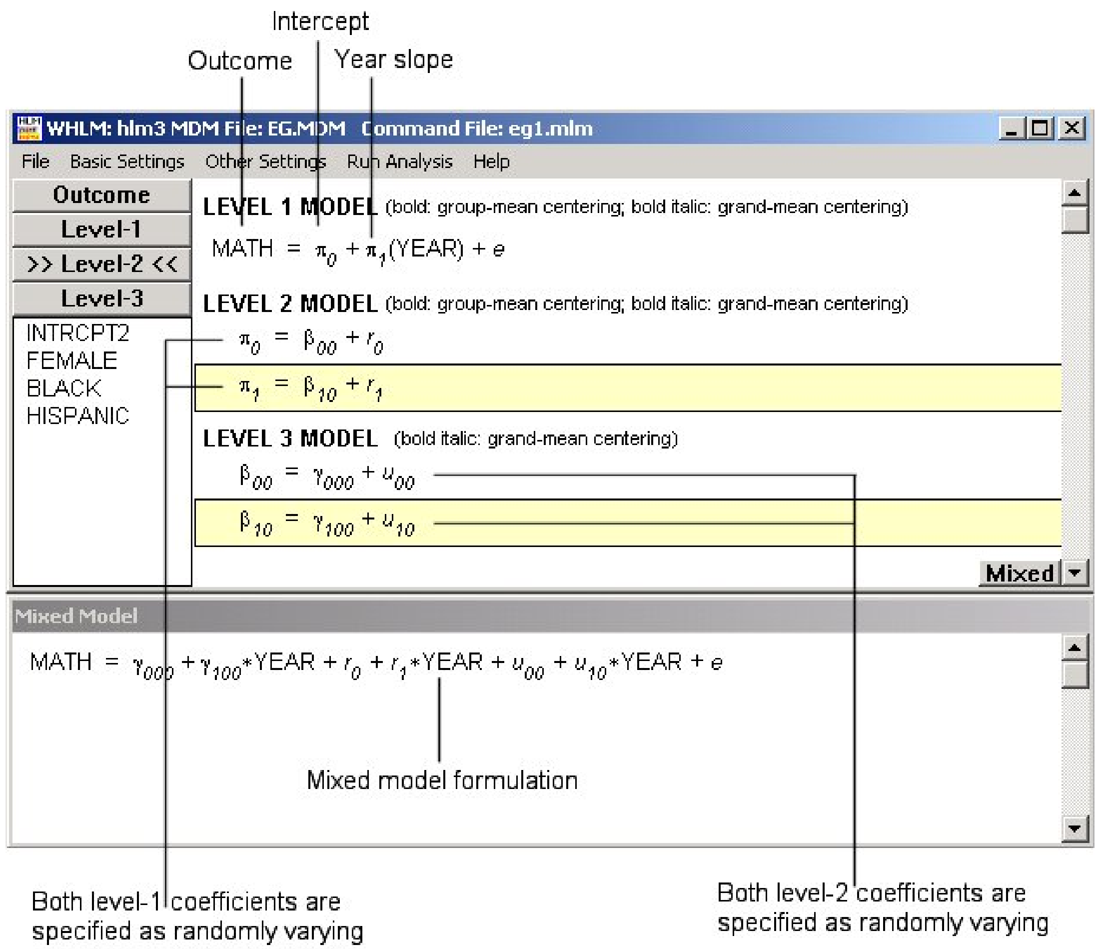


Figure 4.7 Model Window for the public school example

3. Specification of level-1 coefficients as random or non-random across level-two units. We shall model the intercept and the YEAR slope as varying randomly across the children within schools.
4. Specification of the level-3 prediction model. Here each level-2 coefficient becomes an outcome, and we can select level-3 variables to predict school-to-school variation in these level-2 coefficients. In principle, this model specifies how schools differ with respect to the distribution of growth curves within them.
5. Specification of the level-2 coefficients as random or non-random across level-3 units.

Following the five steps above, we first specify a model with no child- or school-level predictors. The Windows execution is very similar to the one for HLM2 as described in Section 2.5.2. The command file, EG1.HLM, contains the model specification input responses. To open the command file, open the **File** menu and choose **Edit/Run old command file**. Figure 4.7 displays the model specified in both standard and mixed model notation.

4.2.1 An annotated example of HLM3 output

Here is the output produced by the model described above. The first page of the output gives the specification of the model.

Problem Title: UNCONDITIONAL LINEAR GROWTH MODEL

The data source for this run = EG.MDM	*Name of the MDM file*
The command file for this run = eg1.mlm	*Name of the command file*
Output file name = hlm3.html	*Name of this output file*
The maximum number of level-1 units = 7230	*There are 7230 observations*
The maximum number of level-2 units = 1721	*There are 1721 children*
The maximum number of level-3 units = 60	*There are 60 schools*
The maximum number of iterations = 100	
Method of estimation: full maximum likelihood	

Level-1 Model

$MATH_{ijk} = \pi_{0jk} + \pi_{1jk}*(YEAR_{ijk}) + e_{ijk}$

Level-2 Model

$\pi_{0jk} = \beta_{00k} + r_{0jk}$
$\pi_{1jk} = \beta_{10k} + r_{1jk}$

Level-3 Model

$\beta_{00k} = \gamma_{000} + u_{00k}$
$\beta_{10k} = \gamma_{100} + u_{10k}$

Mixed Model

$MATH_{ijk} = \gamma_{000} + \gamma_{100}*YEAR_{ijk}$
$+ r_{0jk} + r_{1jk} *YEAR_{ijk}$
$+ u_{00k} + u_{10k} *YEAR_{ijk} + e_{ijk}$

Next come the initial parameter estimates or "starting values." Users should not base inferences on these values, the sole purpose of which is to get the iterations started.

Least Squares Estimates

$\sigma^2 = 1.21432$

Least-squares estimates of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	-0.827685	0.013431	-61.623	7228	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.765828	0.009293	82.410	7228	<0.001

Least-squares estimates of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	-0.827685	0.072631	-11.396	7228	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.765828	0.018892	40.537	7228	<0.001

The least-squares likelihood value = -1.096090E+004
Deviance = 21921.80879
Number of estimated parameters = 3

For starting values, data from 7230 level-1 and 1721 level-2 records were used

Starting Values

$\sigma^2_{(0)}$ = 0.29710

$\tau_{\pi(0)}$

INTRCPT1,π_0	0.71125	0.05143
YEAR,π_1	0.05143	0.01582

$\tau_{\beta(0)}$

INTRCPT1 INTRCPT2,β_{00}	YEAR INTRCPT2,β_{10}
0.14930	0.01473
0.01473	0.01196

The value of the log-likelihood function at iteration 1 = -8.169527E+003
The value of the log-likelihood function at iteration 2 = -8.165377E+003
. . .

Final Results - Iteration 9
Iterations stopped due to small change in likelihood function

******* ITERATION 9 *******

σ^2 = 0.30148

Standard error of σ2 = 0.00660

τ_π

INTRCPT1,π_0	0.64049	0.04676
YEAR,π_1	0.04676	0.01122

Standard errors of τ_π

INTRCPT1,π_0	0.02515	0.00499
YEAR,π_1	0.00499	0.00196

τ_π (as correlations)

INTRCPT1,π_0	1.000	0.551
YEAR,π_1	0.551	1.000

Note that the estimated correlation between true status at YEAR = 3.5 (halfway through third grade) and true rate of change is estimated to be 0.551 for children in the same school.

Random level-1 coefficient	Reliability estimate
INTRCPT1,π_0	0.839
YEAR,π_1	0.190

Reliabilities of child parameter estimates.

τ_β

INTRCPT1	YEAR
INTRCPT2,β_{00}	INTRCPT2,β_{10}
0.16531	0.01705
0.01705	0.01102

Standard errors of τ_β

INTRCPT1	YEAR
INTRCPT2,β_{00}	INTRCPT2,β_{10}
0.03641	0.00720
0.00720	0.00252

τ_β (as correlations)

INTRCPT1/INTRCPT2,β_{00}	1.000	0.399
YEAR/INTRCPT2,β_{10}	0.399	1.000

Notice that the estimated correlation between true school mean status at YEAR = 3.5 and true school-mean rate of change is 0.399.

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2,β_{00}	0.821
YEAR/INTRCPT2,β_{10}	0.786

Reliabilities of school-level parameter estimates. These indicate the reliability with which we can discriminate among level-2 units using their least-squares estimates of β_0 and β_1. Low reliabilities do not invalidate the HLM analysis. Very low reliabilities (*e.g.*, < 0.10), often indicate that a random coefficient might be considered fixed in subsequent analyses.

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	-0.779309	0.057829	-13.476	59	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.763029	0.015263	49.993	59	<0.001

The above table indicates that the average growth rate is significantly positive at 0.763 logits per year, $t = 49.997$.

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	-0.779309	0.057830	-13.476	59	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.763029	0.015260	50.000	59	<0.001

Note that the results with and without robust standard errors are nearly identical. If the robust and model-based standard errors are substantially different, further investigation of the tenability of key assumptions (see Section 4.3 on examining residuals) is recommended.

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, r_0	0.80030	0.64049	1661	13679.62589	<0.001
YEAR slope, r_1	0.10595	0.01122	1661	2132.50756	<0.001
level-1, *e*	0.54907	0.30148			

Final estimation of level-3 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/INTRCPT2, u_{00}	0.40658	0.16531	59	488.30922	<0.001
YEAR/INTRCPT2, u_{10}	0.10498	0.01102	59	377.43020	<0.001

The results above indicate significant variability among schools in terms of mean status at YEAR = 3.5 (χ^2 = 488.34499, df = 59) and in terms of school-mean rates of change (χ^2 of 377.40852, df = 59).

Statistics for the current model

Deviance = 16326.231407

Number of estimated parameters = 9

Exploratory Analysis: estimated level-2 coefficients and their standard errors obtained by regressing EB residuals on level-2 predictors selected for possible inclusion in subsequent HLM runs

Level-1 Coefficient	Potential Level-2 Predictors		
YEAR, π_1			
	FEMALE	BLACK	HISPANIC
Coefficient	0.001	-0.029	0.005
Standard Error	0.003	0.006	0.006
t-value	0.369	-4.835	0.761

Exploratory Analysis: estimated level-3 coefficients and their standard errors obtained by regressing EB residuals on level-3 predictors selected for possible inclusion in subsequent HLM runs

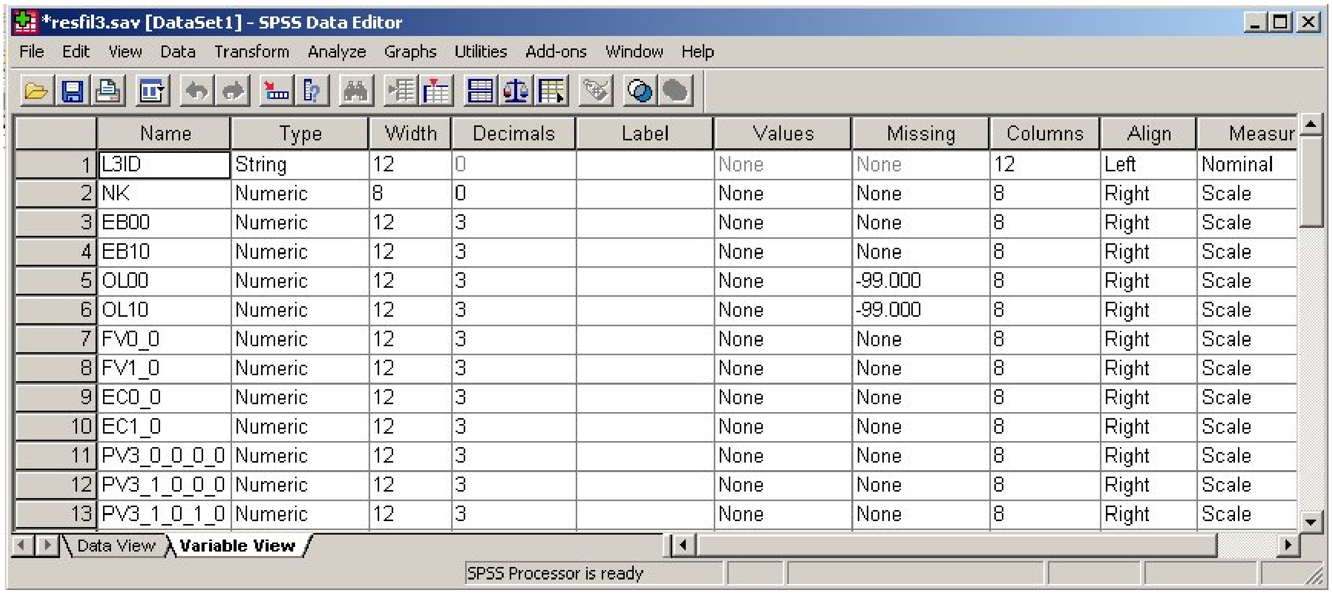
Level-1 Coefficient	Potential Level-3 Predictors		
YEAR/INTRCPT2, β_{10}			
	SIZE	LOWINC	MOBILITY
Coefficient	-0.000	-0.001	-0.002
Standard Error	0.000	0.000	0.001
t-value	-1.525	-2.871	-1.962

Just as in the case of the two-level program, the potential predictors not included in the model to be employed as significant predictors in subsequent models is indicated approximately by the "t-values" given above. **Note:** because of the metric of school size (100s and 1000s), the actual coefficients and standard errors are too small to be printed. The *t*-values are not, however.

4.3 Model checking based on the residual files

HLM3 produces three residual files, one each at levels 1 and 2 (see Chapter 2 for a discussion of these files) and one at level-3 (containing estimates of the βs). These files will contain the EB residuals defined at the various levels, fitted values, and OLS residuals, and EB coefficients. In addition, level-2 predictors can be included in the level-2 residual file and level-3 predictors in the level-3 residual file. However, other statistics provided in the residual file of HLM2, for example the Mahalanobis distance measures, are not available in the residual files produced by HLM3. The procedures for requesting level-3 residual files are similar to those for HLM2 as described in Section 2.5.4.

The files in this example are structured as SPSS data files and can be directly opened in SPSS. As with HLM2, the user can also specify STATA, SYSTAT or SAS command file format for the residual file. The result will be STATA, SYSTAT or SAS data files. (For more details see Section 2.5.4.) Alternatively, the data can be obtained in free form (*i.e.*, as a text file) by selecting the **Free Format** option on the **Create Level-3 Residual File** dialog box. These residual files can then be read into any other computing package. The list of variables in the level-3 residual file and their attributes are shown in Figure 4.8, while the first 10 records contained in this file are shown in Figure 4.10.


*resfil3.sav [DataSet1] - SPSS Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measur
1	L3ID	String	12	0		None	None	12	Left	Nominal
2	NK	Numeric	8	0		None	None	8	Right	Scale
3	EB00	Numeric	12	3		None	None	8	Right	Scale
4	EB10	Numeric	12	3		None	None	8	Right	Scale
5	OL00	Numeric	12	3		None	-99.000	8	Right	Scale
6	OL10	Numeric	12	3		None	-99.000	8	Right	Scale
7	FV0_0	Numeric	12	3		None	None	8	Right	Scale
8	FV1_0	Numeric	12	3		None	None	8	Right	Scale
9	EC0_0	Numeric	12	3		None	None	8	Right	Scale
10	EC1_0	Numeric	12	3		None	None	8	Right	Scale
11	PV3_0_0_0_0	Numeric	12	3		None	None	8	Right	Scale
12	PV3_1_0_0_0	Numeric	12	3		None	None	8	Right	Scale
13	PV3_1_0_1_0	Numeric	12	3		None	None	8	Right	Scale

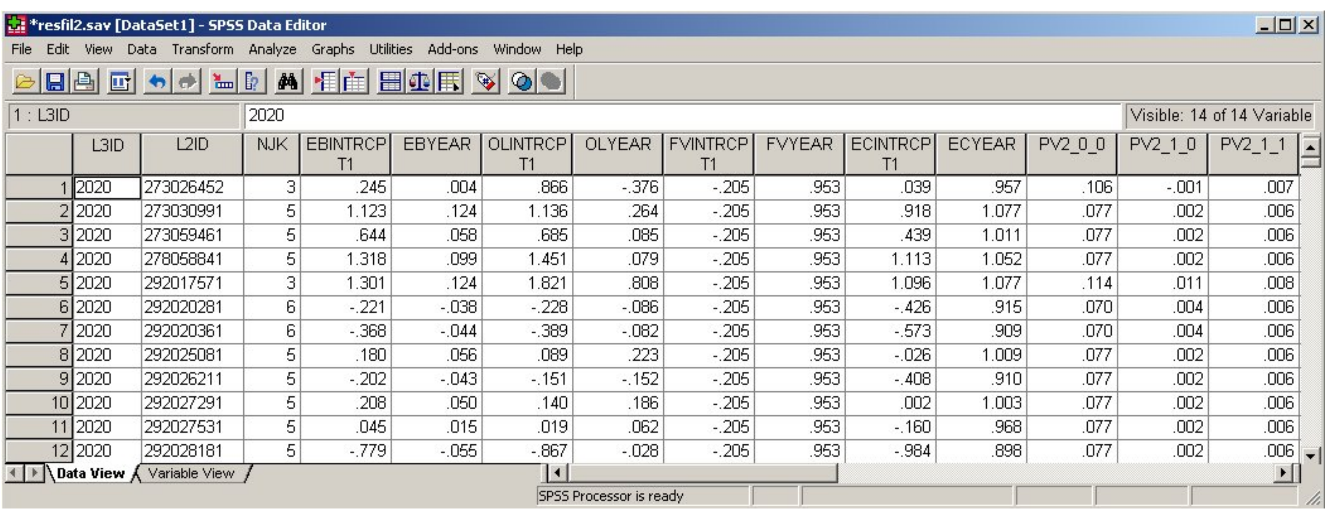
Data View | Variable View

SPSS Processor is ready

Figure 4.8 List of variables and attributes for level-3 residual file

An example of the level-2 residual file produced in the above analysis is shown in Figure 4.9. Only data from school 2020 are given.

We see that the level-3 ID (l3id) is the first variable and the level-2 ID (l2id) is the second. The third variable is njk, the number of observations associated with child j in school k. The empirical Bayes estimates of the residuals, r_{pjk}, are given next, including, respectively, the intercept (ebintrcpt1) and the year effect (ebyear). The ordinary least squares estimates of the same quantities (olintrcpt1 and olyear); and the fitted values, that is, the predicted values of the π_{pjk} s for a given child based on the fixed effects (fvintrcpt1 and fvyear) and random school effect, follow. These are followed by the EB coefficients. Finally, the posterior variances and covariances (pv2_0_0, pv2_1_0, and pv2_1_1) of the empirical Bayes estimates are given.


*resfil2.sav [DataSet1] - SPSS Data Editor

1 : L3ID 2020 Visible: 14 of 14 Variable

	L3ID	L2ID	NJK	EBINTRCPT1	EBYEAR	OLINTRCPT1	OLYEAR	FVINTRCPT1	FVYEAR	ECINTRCPT1	ECYEAR	PV2_0_0	PV2_1_0	PV2_1_1
1	2020	273026452	3	.245	.004	.866	-.376	-.205	.953	.039	.957	.106	-.001	.007
2	2020	273030991	5	1.123	.124	1.136	.264	-.205	.953	.918	1.077	.077	.002	.006
3	2020	273059461	5	.644	.058	.685	.085	-.205	.953	.439	1.011	.077	.002	.006
4	2020	278058841	5	1.318	.099	1.451	.079	-.205	.953	1.113	1.052	.077	.002	.006
5	2020	292017571	3	1.301	.124	1.821	.808	-.205	.953	1.096	1.077	.114	.011	.008
6	2020	292020281	6	-.221	-.038	-.228	-.086	-.205	.953	-.426	.915	.070	.004	.006
7	2020	292020361	6	-.368	-.044	-.389	-.082	-.205	.953	-.573	.909	.070	.004	.006
8	2020	292025081	5	.180	.056	.089	.223	-.205	.953	-.026	1.009	.077	.002	.006
9	2020	292026211	5	-.202	-.043	-.151	-.152	-.205	.953	-.408	.910	.077	.002	.006
10	2020	292027291	5	.208	.050	.140	.186	-.205	.953	.002	1.003	.077	.002	.006
11	2020	292027531	5	.045	.015	.019	.062	-.205	.953	-.160	.968	.077	.002	.006
12	2020	292028181	5	-.779	-.055	-.867	-.028	-.205	.953	-.984	.898	.077	.002	.006

Data View | Variable View

SPSS Processor is ready

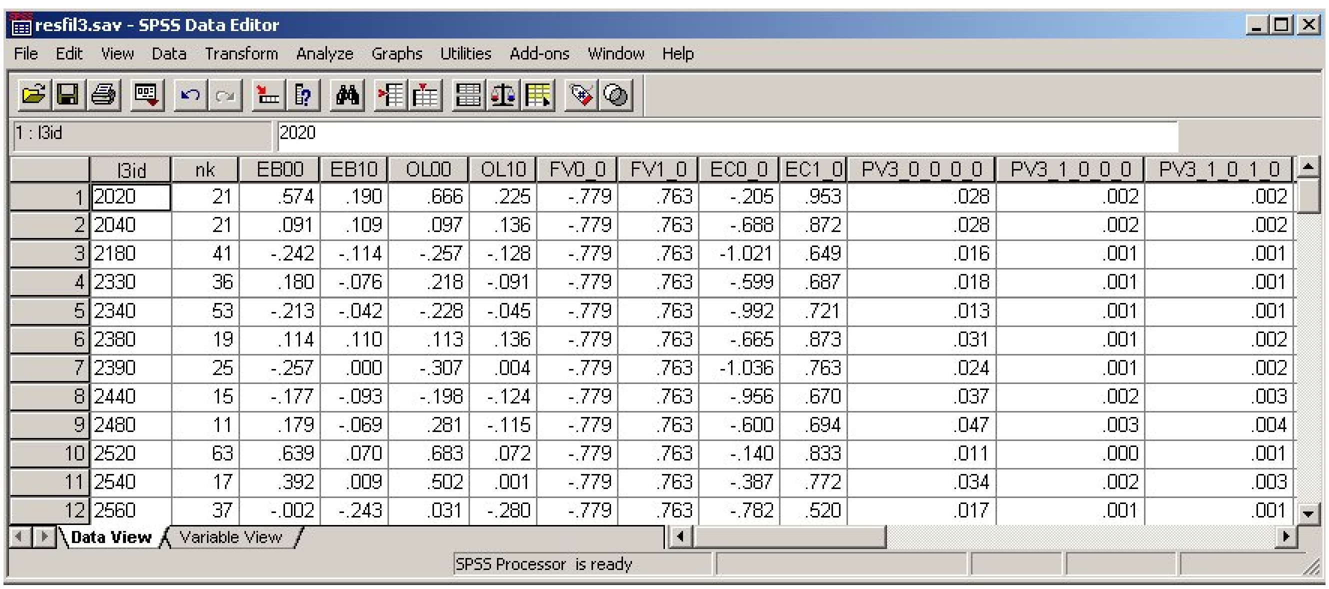
Figure 4.9 First 12 children in level-2 residual file

We see that the first child in the data set has schoolid 2020 and childid 273026452. That child has 3 time-series observations. The predicted growth rate for that child (the YEAR effect) is the fitted value .953. That child's empirical Bayes residual YEAR effect is .004. Thus, the EB coefficient ("ebyear") is computed as:

$$\begin{aligned} \pi_{1jk}^{*} &= \beta_{10k}^{*} + r_{1jk}^{*} \\ &= \text{FVYEAR} + \text{EBYEAR} \\ &= 0.953 + 0.004 \\ &= 0.957 \end{aligned} \tag{0.0}$$

The empirical Bayes estimate for the child's intercept, π_{0jk}^{*} ("ecintrcp"), is computed similarly.

The level-3 residual file, printed below, has a similar structure. Only the data for the first 10 schools are given. We see that the level-3 ID (l3id) is the first value given, and is followed by nk, the number of children in school k. This is followed by the empirical Bayes estimates of the βs, including, respectively, the intercept (eb00) and the *year* effect (eb10). The ordinary least squares estimates of the same quantities (ol00 and ol10); and the fitted values, that is, the predicted values of the βs for a given school based on that school's effect and the fixed effects (fv0_0 and fv1_0). The EB coefficients are given next. Finally, the posterior variances and covariances (pv3_0_0_0_0, pv3_1_0_0_0, and pv3_1_0_1_0) of the estimates are given.



	l3id	nk	EB00	EB10	OL00	OL10	FV0_0	FV1_0	EC0_0	EC1_0	PV3_0_0_0_0	PV3_1_0_0_0	PV3_1_0_1_0
1	2020	21	.574	.190	.666	.225	-.779	.763	-.205	.953	.028	.002	.002
2	2040	21	.091	.109	.097	.136	-.779	.763	-.688	.872	.028	.002	.002
3	2180	41	-.242	-.114	-.257	-.128	-.779	.763	-1.021	.649	.016	.001	.001
4	2330	36	.180	-.076	.218	-.091	-.779	.763	-.599	.687	.018	.001	.001
5	2340	53	-.213	-.042	-.228	-.045	-.779	.763	-.992	.721	.013	.001	.001
6	2380	19	.114	.110	.113	.136	-.779	.763	-.665	.873	.031	.001	.002
7	2390	25	-.257	.000	-.307	.004	-.779	.763	-1.036	.763	.024	.001	.002
8	2440	15	-.177	-.093	-.198	-.124	-.779	.763	-.956	.670	.037	.002	.003
9	2480	11	.179	-.069	.281	-.115	-.779	.763	-.600	.694	.047	.003	.004
10	2520	63	.639	.070	.683	.072	-.779	.763	-.140	.833	.011	.000	.001
11	2540	17	.392	.009	.502	.001	-.779	.763	-.387	.772	.034	.002	.003
12	2560	37	-.002	-.243	.031	-.280	-.779	.763	-.782	.520	.017	.001	.001

Figure 4.10 First 10 schools in level-3 residual file

We see that the first unit, school 2020, has nk = 21 children. The predicted YEAR effect for school 2020 is the fitted value .763, that is, the maximum-likelihood estimate of the school mean growth rate in the case of this unconditional model. That school's empirical Bayes residual YEAR effect is .190. Thus HLM3 constructs the empirical Bayes estimate of that school's YEAR effect (mean rate of growth, "ec_10") as

$$\beta_{10k}^{*} = \gamma_{100}^{*} + u_{1k}^{*} = \text{fv01} + \text{eb10} \tag{0.0}$$
$$= .763 + .190 = .953.$$

Similarly, HLM3 constructs the empirical Bayes estimate for the school's intercept, β_{00k}^{*} ("ec0_0"), using fv0_0 + eb00.

Note that the empirical Bayes estimate of the school YEAR effect, 0.953, is the fitted value for each child in that school (in the level-2 residual file). This will be true in any model that is unconditional at level 2, that is, any model with no child-level predictors such as race, ethnicity or female. When level-2 predictors are in the model, the level-2 fitted values will also depend on those predictors.

4.4 Specification of a conditional model

The above example involves a model that is "unconditional" at levels 2 and 3; that is, no predictors are specified at each of those levels. Such a model is useful for partitioning variation in intercepts and growth rates into components that lie within and between schools (see *Hierarchical Linear Models*, Chapter 8), but provides no information on how child or school characteristics relate to the growth curves. Figure 4.11 shows a model that incorporates information about a child's race and ethnicity and a school's percent low income. Moreover, we explore the possibility that several other predictors (gender, school enrollment, and percent mobility) might help account for variation in subsequent models.

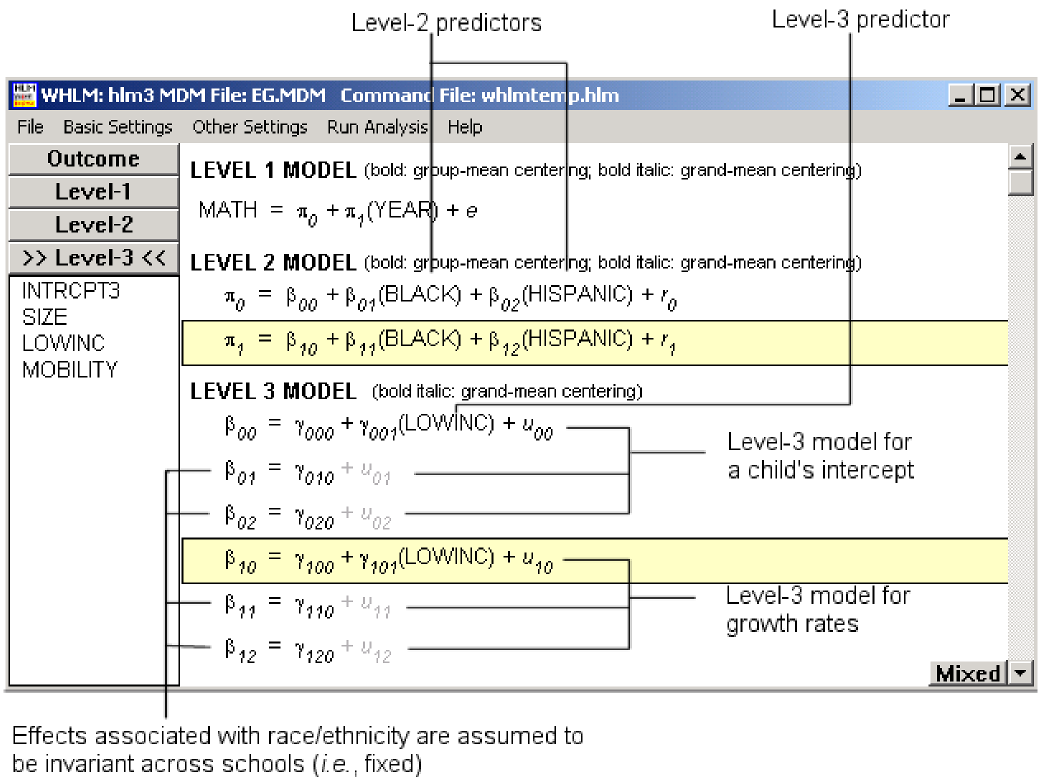


Figure 4.11 Model window for the public school example

The results of the analysis are given below.

Problem Title: LINEAR GROWTH OVER GRADE, MINORITY, LOW INCOME

The data source for this run = EG.MDM
The command file for this run = eg2.mlm
Output file name = hlm3.html
The maximum number of level-1 units = 7230
The maximum number of level-2 units = 1721
The maximum number of level-3 units = 60
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Level-1 Model

$$\mathrm{MATH}_{ijk} = \pi_{0jk} + \pi_{1jk}*(\mathrm{YEAR}_{ijk}) + e_{ijk}$$

Level-2 Model

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}*(\mathrm{BLACK}_{jk}) + \beta_{02k}*(\mathrm{HISPANIC}_{jk}) + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k} + \beta_{11k}*(\mathrm{BLACK}_{jk}) + \beta_{12k}*(\mathrm{HISPANIC}_{jk}) + r_{1jk}$$

Level-3 Model

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\mathrm{LOWINC}_k) + u_{00k}$$
$$\beta_{01k} = \gamma_{010}$$
$$\beta_{02k} = \gamma_{020}$$
$$\beta_{10k} = \gamma_{100} + \gamma_{101}(\mathrm{LOWINC}_k) + u_{10k}$$
$$\beta_{11k} = \gamma_{110}$$
$$\beta_{12k} = \gamma_{120}$$

Mixed Model

$$\mathrm{MATH}_{ijk} = \gamma_{000} + \gamma_{001}*\mathrm{LOWINC}_k + \gamma_{010}*\mathrm{BLACK}_{jk} + \gamma_{020}*\mathrm{HISPANIC}_{jk}$$
$$+ \gamma_{100}*\mathrm{YEAR}_{ijk} + \gamma_{101}*\mathrm{YEAR}_{ijk}*\mathrm{LOWINC}_k + \gamma_{110}*\mathrm{YEAR}_{ijk}*\mathrm{BLACK}_{jk} + \gamma_{120}*\mathrm{YEAR}_{ijk}*\mathrm{HISPANIC}_{jk}$$
$$+ r_{0jk} + r_{1jk}*\mathrm{YEAR}_{ijk} + u_{00k} + u_{10k}*\mathrm{YEAR}_{ijk} + e_{ijk}$$

Least Squares Estimates
$\sigma^2 = 1.07437$

Least-squares estimates of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	0.187343	0.040175	4.663	7222	<0.001
LOWINC, γ_{001}	-0.008941	0.000568	-15.733	7222	<0.001
For BLACK, β_{01}					
INTRCPT3, γ_{010}	-0.405550	0.041045	-9.881	7222	<0.001
For HISPANIC, β_{02}					
INTRCPT3, γ_{020}	-0.285918	0.049723	-5.750	7222	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.906001	0.027528	32.912	7222	<0.001
LOWINC, γ_{101}	-0.001768	0.000392	-4.512	7222	<0.001
For BLACK, β_{11}					
INTRCPT3, γ_{110}	-0.015548	0.028610	-0.543	7222	0.587
For HISPANIC, β_{12}					
INTRCPT3, γ_{120}	0.032732	0.034446	0.950	7222	0.342

Least-squares estimates of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	0.187343	0.106837	1.754	7222	0.080
LOWINC, γ_{001}	-0.008941	0.001287	-6.948	7222	<0.001
For BLACK, β_{01}					
INTRCPT3, γ_{010}	-0.405550	0.106437	-3.810	7222	<0.001
For HISPANIC, β_{02}					
INTRCPT3, γ_{020}	-0.285918	0.089893	-3.181	7222	0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.906001	0.031606	28.665	7222	<0.001
LOWINC, γ_{101}	-0.001768	0.000446	-3.968	7222	<0.001
For BLACK, β_{11}					
INTRCPT3, γ_{110}	-0.015548	0.030859	-0.504	7222	0.614
For HISPANIC, β_{12}					
INTRCPT3, γ_{120}	0.032732	0.037194	0.880	7222	0.379

The least-squares likelihood value = -1.051825E+004
Deviance = 21036.49127
Number of estimated parameters = 9

For starting values, data from 7230 level-1 and 1721 level-2 records were used

Starting Values

$\sigma^2_{(0)} = 0.29710$

$\tau_{\pi(0)}$

INTRCPT1, π_0	0.69259	0.04914
YEAR, π_1	0.04914	0.01481

$\tau_{\beta(0)}$

INTRCPT1	YEAR
INTRCPT2, β_{00}	INTRCPT2, β_{10}
0.05922	0.00290
0.00290	0.01057

The value of the log-likelihood function at iteration 1 = -8.127397E+003
The value of the log-likelihood function at iteration 2 = -8.121908E+003
The value of the log-likelihood function at iteration 3 = -8.121269E+003
The value of the log-likelihood function at iteration 4 = -8.121059E+003
The value of the log-likelihood function at iteration 5 = -8.120942E+003
. . .

Final Results - Iteration 9
Iterations stopped due to small change in likelihood function

$\sigma^2 = 0.30162$
Standard error of $\sigma^2 = 0.00660$

τ_π

INTRCPT1,π_0	0.62231	0.04657
YEAR,π_1	0.04657	0.01106

Standard errors of τ_π

INTRCPT1,π_0	0.02451	0.00491
YEAR,π_1	0.00491	0.00196

τ_π (as correlations)

INTRCPT1,π_0	1.000	0.561
YEAR,π_1	0.561	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1,π_0	0.835
YEAR,π_1	0.188

τ_β

INTRCPT1	YEAR
INTRCPT2,β_{00}	INTRCPT2,β_{10}
0.07808	0.00082
0.00082	0.00798

Standard errors of τ_β

INTRCPT1	YEAR
INTRCPT2,β_{00}	INTRCPT2,β_{10}
0.01991	0.00441
0.00441	0.00194

τ_β (as correlations)

INTRCPT1/INTRCPT2,β_{00}	1.000	0.033
YEAR/INTRCPT2,β_{10}	0.033	1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/INTRCPT2,β_{00}	0.702
YEAR/INTRCPT2,β_{10}	0.735

The value of the log-likelihood function at iteration 9 = -8.119604E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	0.140628	0.127486	1.103	58	0.275
LOWINC, γ_{001}	-0.007578	0.001691	-4.482	58	<0.001
For BLACK, β_{01}					
INTRCPT3, γ_{010}	-0.502091	0.077879	-6.447	1597	<0.001
For HISPANIC, β_{02}					
INTRCPT3, γ_{020}	-0.319381	0.086099	-3.709	1597	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.874501	0.039144	22.340	58	<0.001
LOWINC, γ_{101}	-0.001369	0.000523	-2.619	58	0.011
For BLACK, β_{11}					
INTRCPT3, γ_{110}	-0.030918	0.022453	-1.377	1597	0.169
For HISPANIC, β_{12}					
INTRCPT3, γ_{120}	0.043085	0.024652	1.748	1597	0.081

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	0.140628	0.113814	1.236	58	0.222
LOWINC, γ_{001}	-0.007578	0.001396	-5.428	58	<0.001
For BLACK, β_{01}					
INTRCPT3, γ_{010}	-0.502091	0.076842	-6.534	1597	<0.001
For HISPANIC, β_{02}					
INTRCPT3, γ_{020}	-0.319381	0.081918	-3.899	1597	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.874501	0.037287	23.453	58	<0.001
LOWINC, γ_{101}	-0.001369	0.000499	-2.744	58	0.008
For BLACK, β_{11}					
INTRCPT3, γ_{110}	-0.030918	0.022274	-1.388	1597	0.165
For HISPANIC, β_{12}					
INTRCPT3, γ_{120}	0.043085	0.024368	1.768	1597	0.077

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, r_0	0.78886	0.62231	1659	13364.57298	<0.001
YEAR slope, r_1	0.10518	0.01106	1659	2126.73092	<0.001
level-1, *e*	0.54920	0.30162			

Final estimation of level-3 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/INTRCPT2, u_{00}	0.27943	0.07808	58	254.96395	<0.001
YEAR/INTRCPT2, u_{10}	0.08935	0.00798	58	277.26967	<0.001

Statistics for the current model

Deviance = 16239.207347
Number of estimated parameters = 15

Exploratory Analysis: estimated level-2 coefficients and their standard errors obtained by regressing EB residuals on level-2 predictors selected for possible inclusion in subsequent HLM runs

Level-1 Coefficient	Potential Level-2 Predictors
YEAR,π_1	
	FEMALE
Coefficient	0.001
Standard Error	0.003
t-value	0.454

Exploratory Analysis: estimated level-3 coefficients and their standard errors obtained by regressing EB residuals on level-3 predictors selected for possible inclusion in subsequent HLM runs

Level-1 Coefficient	Potential Level-3 Predictors	
YEAR/INTRCPT2,β_{10}		
	SIZE	MOBILITY
Coefficient	-0.000	-0.000
Standard Error	0.000	0.001
t-value	-1.155	-0.540

4.5 Other program features

The options available for HLM3 are similar to those available with HLM2. The differences are outlined below.

4.5.1 Basic specifications

The level-3 residual files may also be specified. They are specified similarly to the level-2 residuals.

4.5.2 Iteration control

The **Mode of iteration acceleration** section of this screen is primarily intended for people who have data large enough to cause the accelerator (and final) iterations to take a prohibitive amount of time. While for most data the 2nd derivative option is recommended, users with large amounts of data (particularly with large ratios of level-1 to level-2 data) may find the 1st derivative Fisher useful, although this will make the standard errors of σ^2 and the τ matrices more crude. If the third option, **No accelerator**, is selected, there will be *no* Fisher iterations will be performed. This will make large MDMs run faster, but will have the side effect of not producing standard errors of σ^2 and the tau matrices. If you want to suppress any Fisher iterations, but do want to have the above mentioned standard errors, choose 1st or 2nd derivative Fisher, and set the value in the **Frequency of accelerator** box to the number of iterations + 1.

4.5.3 Estimation settings

HLM3 has the same options as HLM2.

4.5.4 Hypothesis testing

HLM3 does not have the test of level-1 homogeneity.

4.5.5 Output settings

HLM3 output does not include OLS estimates.

5 Conceptual and Statistical Background for Four-Level Models

HLM4 handles models with data that have a four-level nested structure. A four-level hierarchy would arise in the HLM3 illustrative example described in the last chapter, for example, if the students who were repeatedly observed while attending a given school were also nested within classrooms. With an additional clustering unit of classrooms, the achievement data would be triply nested. The time-series data are nested within students, the students nested within classrooms, and the classrooms nested within schools. In a different scenario, with the incorporation of a measurement model for the repeated measures on mathematics achievement for the same example, one would implement four-level analyses. Hough, Bryk, Pinnell, Kerbow, Fountas, and Scharer (2008), for example, used this approach with four-level models to study the effect of school-based coaching on the growth in teacher expertise in literary practices. The level-1 model in their study was a measurement error model associated with repeated measures on teacher expertise, the level-2 model studied the growth trajectories of the "true scores" on the expertise, and the level-3 and level-4 models investigated the associations of the growth trajectory parameters with teacher- and school-level correlates, respectively. For examples of similar level-1 measurement error models (in three-level analyses), see pp. 248-249 in Chapter 8 and Chapter 11 of *Hierarchical Linear Models*.

5.1 The general four-level model

The four-level model consists of four submodels, one for each level. For example, if the research problem consists of data on students nested within classrooms, classrooms within schools, and classrooms within school districts, the level-1 model will represent the relationships among the student-level variables, the level-2 model will capture the influences of class-level correlates, the level-3 model will incorporate school-level effects, the level-4 model will handle district-level factors.

Formally there are $i = 1, ..., n_{jkl}$ level-1 units (*e.g.*, students), which are nested within each of $j = 1,..., J_{kl}$ level-2 units (*e.g.*, classrooms) nested within each of $k = 1,..., K_l$ level-3 units (*e.g.*, schools) nested within each of $l = 1,…, L$ level-4 units (*e.g.*, school districts).

5.1.1 Level-1 model

In the level-1 model, the user can select notation according to the type of application (*e.g.*, a cross-sectional model versus a model with longitudinal data). In the case of a cross-sectional model, we represent the outcome for case *i* within level-2 unit *j*, level-3 unit *k* and level-4 unit *l* as:

$$\begin{aligned} Y_{ijkl} &= \pi_{0jkl} + \pi_{1jkl}a_{1ijkl} + \pi_{2jkl}a_{2ijkl} + \dots + \pi_{pjkl}a_{pijkl} + e_{ijkl} \\ &= \pi_{0jkl} + \sum_{p=1}^{P} \pi_{pjkl}a_{pijkl} + e_{ijkl}, \end{aligned} \tag{5.1}$$

where

π_{pjkl} $(p = 0,1,..., P)$ are *level-1 coefficients*;

a_{pijkl} is a *level-1 predictor p* for case i in level-2 unit j, level-3 unit k, and level-4 unit l;

e_{ijkl} is the *level-1 random effect*; and

σ^2 is the variance of e_{ijkl}, that is the *level-1 variance*.

Here we assume that the random term $e_{ijkl} \sim N(0,\ \sigma^2)$.

5.1.2 Level-2 model

Each of the π_{pjkl} coefficients in the level-1 model becomes an outcome variable in the level-2 model:

$$\begin{aligned}\pi_{pjkl} &= \beta_{p0kl} + \beta_{p1kl}X_{1jkl} + \beta_{p2kl}X_{2jkl} + \cdots + \beta_{pQ_p kl}X_{Q_p jkl} + r_{pjkl} \\ &= \beta_{p0kl} + \sum_{q=1}^{Q_p} \beta_{pqkl}X_{qjkl} + r_{pjkl},\end{aligned} \tag{5.2}$$

where

β_{pqkl} $(q = 0,1,...,Q_p)$ are *level-2 coefficients*;

X_{qjkl} is a *level-2 predictor*; and

r_{pjkl} is a *level-2 random effect.*

We assume that, for each level-2 unit, the vector of level-1 random effects (the r_{pqkl} terms) is distributed as multivariate normal, with each having a mean of zero and with covariance matrix $\mathbf{T}_\pi$, with a maximum dimension $(P+1)\times(P+1)$.

5.1.3 Level-3 model

Each of the level-2 coefficients, β_{pqkl}, defined in the level-2 model, becomes an outcome variable in the level-3 model:

$$\begin{aligned}\beta_{pqkl} &= \gamma_{pq0l} + \gamma_{pq1l}W_{1kl} + \gamma_{pq2l}W_{2kl} + \cdots + \gamma_{pqS_{pq}l}W_{S_{pq}kl} + u_{pqkl} \\ &= \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqsl}W_{skl} + u_{pqkl},\end{aligned} \tag{5.3}$$

where

γ_{pqsl} $(s = 0, 1, \ldots, S_{pq})$ are *level-3 coefficients*,

W_{skl} is a *level-3 predictor*, and

u_{pqkl} is a *level-3 random effect*.

We assume that, for each level-3 unit, the vector of level-3 random effects (the u_{pqkl} terms) is distributed as multivariate normal, with each having a mean of zero and with covariance matrix $\mathbf{T}_\beta$, whose maximum dimension is:

$$\sum_{p=0}^{p}(Q_p + 1) \times \sum_{p=0}^{p}(Q_p + 1), \tag{5.4}$$

5.1.4 Level-4 model

Each of the level-3 coefficients, γ_{pqsl}, defined in the level-3 model, becomes an outcome variable in the level-4 model:

$$\begin{aligned} \gamma_{pqsl} &= \delta_{pqs0} + \delta_{pqs1}Z_{1l} + \delta_{pqs2}Z_{2l} + \ldots + \delta_{pqG_{pqs}}Z_{G_{pqs}l} + \upsilon_{pqsl} \\ &= \delta_{pqs0} + \sum_{g=1}^{G_{pqs}} \delta_{pqsg}Z_{gl} + \upsilon_{pqsl}, \end{aligned} \tag{5.5}$$

where

δ_{pqsg} $(g = 0, 1, \ldots, G_{pqs})$ are *level-4 coefficients*,

Z_{gl} is a *level-4 predictor*, and

υ_{pqsl} is a *level-4 random effect*.

We assume that, for each level-4 unit, the vector of level-4 random effects (the υ_{pqsl} terms) is distributed as multivariate normal, with each having a mean of zero and with covariance matrix $\mathbf{T}_\gamma$, whose maximum dimension is:

$$\sum_{pq=0}^{pq}(S_{pq} + 1) \times \sum_{pq=0}^{pq}(S_{pq} + 1), \tag{5.6}$$

5.2 Parameter estimation

Three kinds of parameter estimates are available in a four-level model: empirical Bayes estimates of randomly varying level-1, level-2, and level-3 coefficients; maximum-likelihood estimates of the level-4 coefficients (note: these are also generalized least squares estimates); and maximum-likelihood estimates of the variance-covariance components. Both HLM3 and HLM4 estimate the variance-covariance components and the fixed effects (level-4 coefficients) by means of full maximum likelihood. In nonlinear models, the coefficients are estimated via penalized quasi-likelihood. Unlike HGLM, however, only unit-specific and not population-averaged results are available.

5.3 Hypothesis testing

As in the case of the three-level program, the three-level program routinely prints standard errors and *t*-tests for each of the level-3 coefficients ("the fixed effects") as well as a chi-square test of homogeneity for each random effect. In addition, optional "multivariate hypothesis tests" and residual files are available in the four-level program.

6 Working with HLM4

Data analysis by means of the HLM4 program involves similar stages regarding MDM creation, analyses, and fit evaluation as in the case of the two- and three-level programs. HLM4 analyses can be executed in Windows, interactive, and batch modes. We describe a Windows execution below. We consider interactive and batch execution in Appendix D.

6.1 An example using HLM4 in Windows mode

To illustrate the operation of the HLM4 program, we reanalyze a subset of data from Hough, Bryk, Pinnell, Kerbow, Fountas, and Scharer (2008). Hough *et al.* used a four-level model to examine the association between school-based coaching and the development of teachers' expertise in literary instruction. The level-1 model in their study was a measurement error model associated with 1317 repeated observations on a measure of classroom instruction, which they called teaching *expertise*. (This measurement model relates the observed data to a "true" or latent score plus some error of measurement. See below.) The level-2 model represented a growth model for each teacher's "true scores" on teaching expertise, and the level-3 and level-4 models investigated the associations of the growth trajectory parameters with teacher- and school-level correlates with data from 219 teachers from 17 schools, respectively.

The example illustrates the use of a level-1 in HLM as a measurement model. In brief,

$$Y_{mtij} = \psi_{0tij} + \varepsilon_{mtij}, \quad \varepsilon_{mtij} \sim N\left(0, \sigma^2_{mtij}\right)$$

where

Y_{mtij} is the observed measure on occasion t for teacher i in school j,
ψ_{tij} is the true or latent value for teacher expertise, and
ε_{mtij} is the error of measurement associated with the observed rating m on occasion t for teacher i in school j.

(Note, in this data set there is only one observed rating per occasion. As a result the number of level-1 and level-2 units are identical.)

In most applications, ε_{mtij} is unknown and assumed normally distributed with constant variance. In contrast in this application, the Rasch measurement model for the observed outcomes, Y_{mtij}, also provides a standard error estimate for each observed measure, s_{mtij}. We explicitly represent this by multiplying both sides of the level-1 model by the inverse of the standard error, $a_{mtij} = s^{-1}_{mtij}$, yielding

$$Y^*_{mtij} = a_{mtij}\psi_{0tij} + e^*_{mtij}, \quad e^*_{mtij} \sim N\left(0,1\right).$$

The variance at level-1 is now assumed known and fixed at a value of 1.0.

6.1.1 Constructing the MDM file from raw data

The user has the same range of options for data input for HLM4 as for HLM3. We will use SPSS file input for the illustrative example.

6.1.1.1 SPSS input

Data input requires a level-1 file (in our illustration a measurement data file), a level-2 file ("true scores" file), a level-3 (teacher level), and a level-4 (school level) file.

Level-1 file. The level-1 file, MEASURE.SAV, has 1317 observations collected on 219 teachers on up to 9 different occasions. Data for the first three teachers are shown in Fig. 6.1. Each of these teachers was observed on three occasions. (Some teachers in the study were observed on as many as nine occasions over three years.)

The first column contains the level-4 (*i.e.*, school) ID, next is the level-3 (*i.e.*, teacher) ID, and this is followed by the level-2 (*i.e.*, occasion) ID. We see that the first record comes from school 1100, teacher 1100002, and occasion 11000026. Following the teacher ID fields are that teacher's values on two variables:

- expertis
  A composite Rasch measure of teachers' classroom literacy practice rated on some particular occasion (weighted by the inverse of its standard error of measurement.)
- invstder
  The inverse of the standard error of measurement associated with that individual rating (the standard errors are generated as part of the Rasch rating scale model.)

	schid	tchrid	occasid	expertis	invstderr
1	1100	1100002	11000026	-2.862	4.472
2	1100	1100002	11000027	-1.850	5.000
3	1100	1100002	11000028	-2.182	4.642
4	1100	1100011	11000116	5.750	5.000
5	1100	1100011	11000117	4.105	5.263
6	1100	1100011	11000118	7.150	5.000
7	1100	1100012	11000123	2.227	4.545
8	1100	1100012	11000124	.591	4.545
9	1100	1100012	11000125	2.913	4.348
10	1100	1100013	11000136	.400	5.000

Figure 6.1 First nine cases in MEASURE.SAV

Level-2 file. The level-2 units consisted of the 1317 occasions when measurements on classroom literary practice were made. The data are stored in the file OCCAS.SAV. The level-2 data for the first nine records are listed below. It has the same three ID's as the level-1 file. The two occasion-level variables are included in the file:

- occasion
  This variable identifies the specific data collection time point, counted up from the first study occasion in the fall of year1 (a value of 0) through the end of the study in the spring of year 3 (a value of 8).

* artifact
  A dummy variable introduced into the analysis to adjust for a measurement artifact that occurred with the first-year spring scores (at occasion = 2).

	schid	tchrid	occasid	occasion	artifact
1	1100	1100002	11000026	3.000	.000
2	1100	1100002	11000027	4.000	.000
3	1100	1100002	11000028	5.000	.000
4	1100	1100011	11000116	3.000	.000
5	1100	1100011	11000117	4.000	.000
6	1100	1100011	11000118	5.000	.000
7	1100	1100012	11000123	3.000	.000
8	1100	1100012	11000124	4.000	.000
9	1100	1100012	11000125	5.000	.000
10	1100	1100013	11000136	3.000	.000

Figure 6.2 First nine cases in OCCAS.SAV

The first teacher in this data file, Teacher 1100002 in school 1100, was observed on three occasions during the second year of the study (*i.e.* occasions 3 through 5). The same was true for the next two teachers. In general, the data collection patterns vary among teachers in this study depending upon their employment history at the school and when they first became eligible for classroom coaching.

Level-3 file. The level-3 units are the 219 teachers. The data are stored in the TCHR.SAV file. The first field is the school ID and the second is the teacher ID. Note that each of the first ten teachers is in school 1100. There are six variables in this file:

* coach
  The average number of one-on-one coaching sessions per month that each teacher received over the course of the study
* newwtch
  A dummy variable indicating that the teacher had three or fewer years of classroom teaching experience at onset of study participation
* pdpart
  A composite measure of teachers' exposure to literacy professional development prior to the onset of the study
* scmt
  A scale score on the teacher's commitment to the school measured at study onset
* y2ent
  A dummy variable indicating the teacher began work at the school during the second year of the study

- y3ent
  A dummy variable indicating the teacher began work at the school during the third year of the study

	schlid	tchrid	coach	newtchr	pdpart	scmt	y2ent	y3ent
1	1100.00	1100002	.571	.000	.842	-.292	1.000	.000
2	1100.00	1100011	.571	1.000	-.361	-.813	1.000	.000
3	1100.00	1100012	.755	.000	1.653	.267	.000	.000
4	1100.00	1100013	.571	1.000	1.115	.774	1.000	.000
5	1100.00	1100020	.496	.000	.856	1.150	.000	.000
6	1100.00	1100023	.878	.000	-.248	-1.379	.000	.000
7	1100.00	1100025	.731	.000	-.631	1.150	.000	.000
8	1100.00	1100026	.831	.000	-.248	-.813	.000	.000
9	1100.00	1100027	.736	.000	.307	.164	.000	.000
10	1100.00	1100029	.695	1.000	-.292	-.340	.000	.000

Figure 6.3 First ten teachers in TCHR.SAV

Level-4 file. The school level data from 17 schools appear in SCH.SAV. The first field is the school ID. This is followed by:

- chgcoach
  A dummy variable indicating that a coaching change occurred during the course of the study. This happened with only one school in the sample.

	schid	chgcoach
1	1100.00	0.000
2	1200.00	0.000
3	1300.00	0.000
4	1400.00	0.000
5	1600.00	0.000
6	1700.00	0.000
7	1800.00	0.000
8	1900.00	0.000
9	2000.00	0.000
10	2100.00	0.000

Figure 6.4 First ten schools in SCH.SAV

The response file, LITERACY.MDMT, contains a log of the input responses used to create the MDM file, LITERACY.MDM, using MEASURE.SAV, OCCAS.SAV, TCHR.SAV, and SCH.SAV. Figure 6.5 shows the dialog box used to create the MDM file. Note that the model notation selected is **longitudinal with measurement model data**. Choosing this option affects the notation used for subscripts and model parameters in the Windows interface and program output.

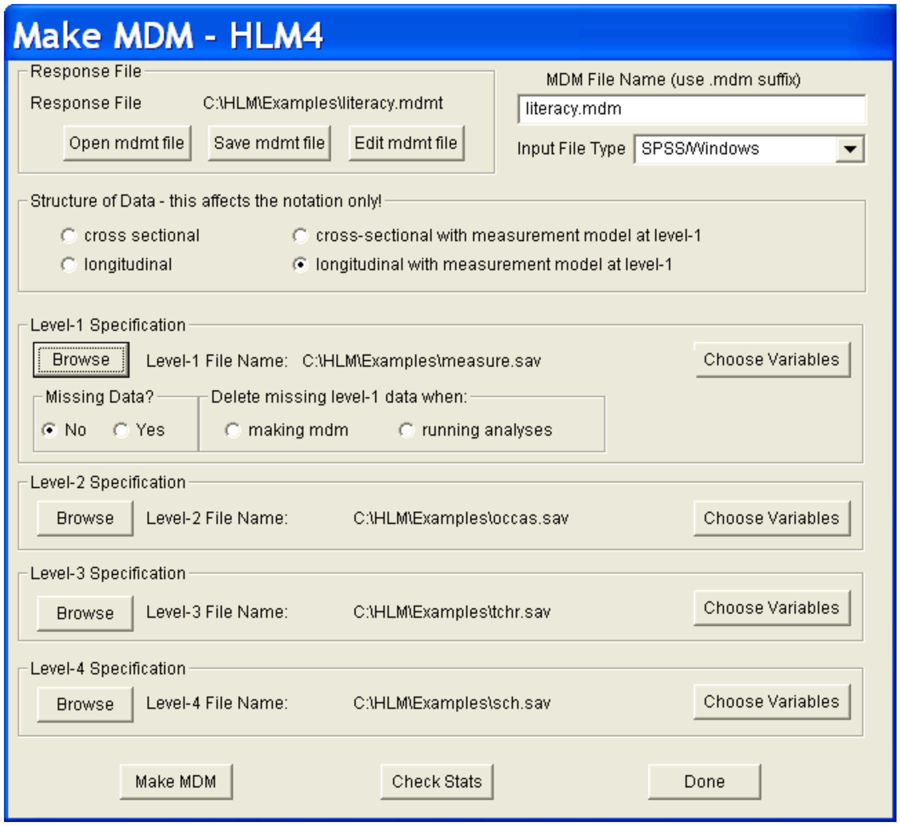


Figure 6.5 Make MDM – HLM4 dialog box for LITERACY.MDM

6.2 Executing analyses based on the MDM file

The MDM file can now be used as input for analysis. Model specification via the Windows mode has seven steps:

1. Specification of the level-1 model. In our example data set, EXPERTIS is the outcome and we use INVSTDER as a level-1 predictor. We also delete the standard intercept from the level-1 model. At a subsequent step (see step 8 below) we will specify the level-1 random effect as having a known variance of 1.0.
2. Specification of the level-2 prediction model. In this measurement model application, the level-1 coefficient associated with INVSTDER becomes the outcome variable. (As noted above, this coefficient now represents the true or latent score on a particular occasion.) At level 2, we model this outcome as a function of OCCAS. That is, we specify a linear growth model for teacher's expertise development over the course of the study. This allows us to represent for every teacher both their initial status and growth rate on the expertise measure over time. We also include as a fixed effect in the level-2 model for the measurement artifact that occurred at the third time point, ARTIFACT.
3. The "true score" level-2 outcomes are specified as randomly varying between teachers.
4. Specification of the level-3 prediction model. In general, one may select different level-3 predictors for each level-3 equation. In the example below, we illustrate this with four of the variables included in the MDM file.
5. Specification of level-3 equations as fixed, random or non-randomly varying. The intercept and the OCCASION slope, which capture the initial status and growth rate of expertise in literary practice, are specified as randomly varying within schools. The effect for ARTIFACT is fixed to the same value for all teachers within a given school.

6. Specification of the level-4 prediction model. In general, each level-3 coefficient becomes an outcome, and we can select level-4 variables to predict school-to-school variation in these level-3 coefficients. Given the relatively small number of school in the data set ($J = 17$) no level-3 predictors are used in the example.
7. Specification of the level-4 equations as fixed, random or non-randomly varying. In the example, mean school initial status on expertise, mean growth rates for teacher expertise and the size of the measurement artifact are all allowed to vary randomly across schools.

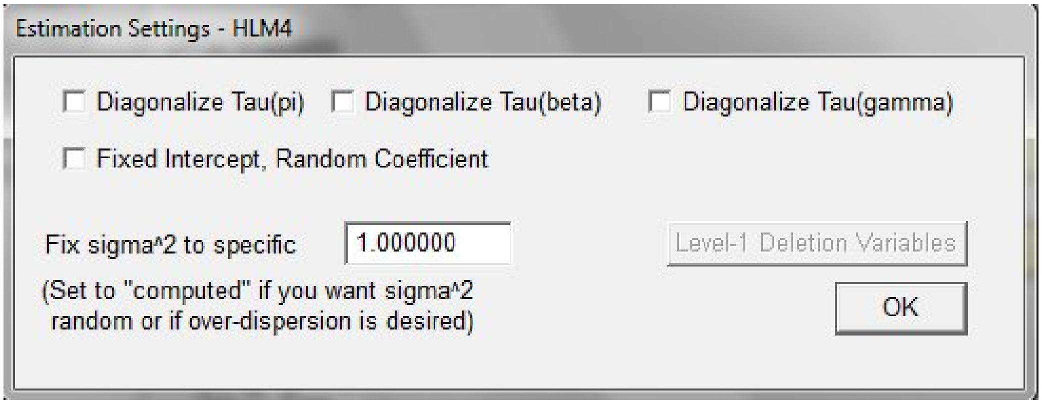


Figure 6.6 Estimation Settings dialog box

8. Finally, to specify the level-1 variance as fixed at a value of 1.0, per the measurement model described above, open the **Other Settings menu**, select **Estimation Settings**, enter 1.0 in the text box for **Fix Sigma^2 to specific value**.

6.2.1 A 4-level measurement model example

To illustrate the use of HLM4 we posed the following model for teacher expertise development:

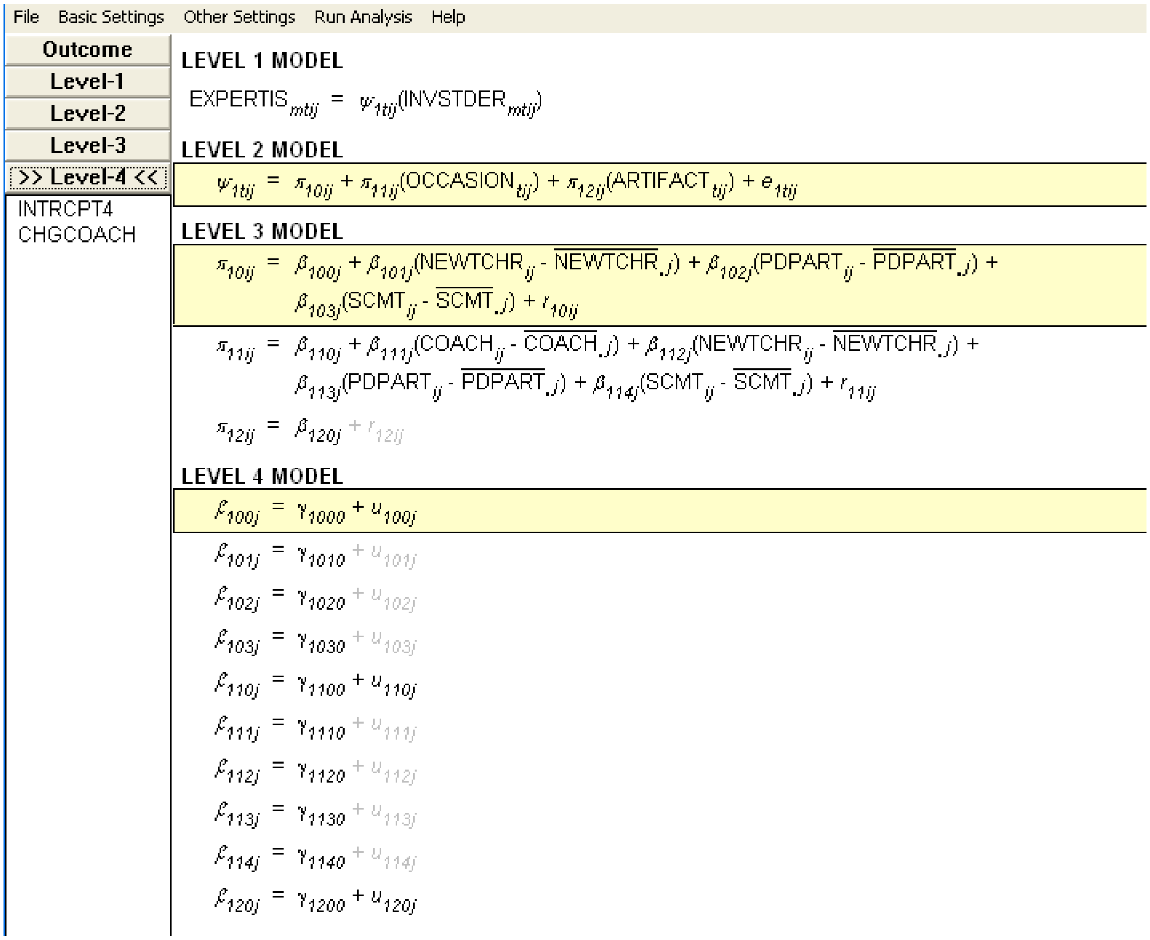


Figure 6.7 Model window for the conditional model for the literacy program example

6.3 An annotated example of HLM4 output

Problem Title: HLM4 example, measurement model
The data source for this run = literacy.mdm
The command file for this run = C:\whlmtemp.hlm
Output file name = C:\hlm4measurement model example.html
The maximum number of level-1 units = 1317
The maximum number of level-2 units = 1317
The maximum number of level-3 units = 219
The maximum number of level-4 units = 17
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is EXPERTIS

Summary of the model specified

Level-1 Model

$$EXPERTIS_{mtij} = \psi_{1tij}*(INVSTDER_{mtij})$$

Level-2 Model

$$\psi_{1tij} = \pi_{10ij} + \pi_{11ij}*(OCCASION_{tij}) + \pi_{12ij}*(ARTIFACT_{tij}) + e_{1tij}$$

Level-3 Model

$\pi_{10ij} = \beta_{100j} + \beta_{101j}$*(NEWTCHR$_{ij}$) + β_{102j}*(PDPART$_{ij}$) + β_{103j}*(SCMT$_{ij}$) + r_{10ij}
$\pi_{11ij} = \beta_{110j} + \beta_{111j}$*(COACH$_{ij}$) + β_{112j}*(NEWTCHR$_{ij}$) + β_{113j}*(PDPART$_{ij}$) + β_{114j}*(SCMT$_{ij}$) + r_{11ij}
$\pi_{12ij} = \beta_{120j}$

Level-4 Model

$\beta_{100j} = \gamma_{1000} + u_{100j}$
$\beta_{101j} = \gamma_{1010}$
$\beta_{102j} = \gamma_{1020}$
$\beta_{103j} = \gamma_{1030}$
$\beta_{110j} = \gamma_{1100} + u_{110j}$
$\beta_{111j} = \gamma_{1110}$
$\beta_{112j} = \gamma_{1120}$
$\beta_{113j} = \gamma_{1130}$
$\beta_{114j} = \gamma_{1140}$
$\beta_{120j} = \gamma_{1200} + u_{120j}$

COACH NEWTCHR PDPART SCMT have been centered around the level-4 mean.

For starting values, data from 1317 level-1, 1312 level-2, 214 level-3 and 17 level-4 records were used

Final Results - Iteration 61
Iterations stopped due to small change in likelihood function

σ^2_e
INVSTDER,ψ_1 0.31788

σ^2_e (as correlations)
INVSTDER,ψ_1 1.000

Random level-1 coefficient	Reliability estimate
INVSTDER	0.821

τ_π

INVSTDER INTRCPT2,π_{10}	INVSTDER OCCASION,π_{11}
0.93753	0.01861
0.01861	0.00113

τ_π (as correlations)

INVSTDER/INTRCPT2,π_{10}	1.000	0.571
INVSTDER/OCCASION,π_{11}	0.571	1.000

Random level-2 coefficient	Reliability estimate
INVSTDER/INTRCPT2	0.740
INVSTDER/OCCASION	0.077

Note: The reliability estimates reported above are based on only 214 of 219 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Note, among teachers within schools, there is a positive correlation of 0.571 between their initial status and expertise development.

τ_β

INVSTDER	INVSTDER	INVSTDER
INTRCPT2	OCCASION	ARTIFACT
INTRCPT3,β_{100}	INTRCPT3,β_{110}	INTRCPT3,β_{120}
0.28840	-0.03214	0.16341
-0.03214	0.03798	-0.05972
0.16341	-0.05972	0.22678

τ_β (as correlations)

INVSTDER/INTRCPT2/INTRCPT3,β_{100}	1.000	-0.307	0.639
INVSTDER/OCCASION/INTRCPT3,β_{110}	-0.307	1.000	-0.643
INVSTDER/ARTIFACT/INTRCPT3,β_{120}	0.639	-0.643	1.000

In contrast, at the school level a negative correlation, -.307, exists between school mean initial status on teachers' expertise and school-level growth rates.

Random level-3 coefficient	Reliability estimate
INVSTDER/INTRCPT2/INTRCPT3	0.727
INVSTDER/OCCASION/INTRCPT3	0.965
INVSTDER/ARTIFACT/INTRCPT3	0.747

The value of the log-likelihood function at iteration 61 = -3.447675E+003

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INVSTDER, ψ_1					
For INTRCPT2, π_{10}					
For INTRCPT3, β_{100}					
INTRCPT4, γ_{1000}	-0.042320	0.152308	-0.278	32	0.783
For NEWTCHR, β_{101}					
INTRCPT4, γ_{1010}	-0.520219	0.226444	-2.297	178	0.022
For PDPART, β_{102}					
INTRCPT4, γ_{1020}	0.167179	0.092189	1.813	178	0.069
For SCMT, β_{103}					
INTRCPT4, γ_{1030}	0.137797	0.085591	1.610	178	0.107
For OCCASION, π_{11}					
For INTRCPT3, β_{110}					
INTRCPT4, γ_{1100}	0.208296	0.048144	4.327	32	<0.001
For COACH, β_{111}					
INTRCPT4, γ_{1110}	0.261937	0.078204	3.349	178	0.001
For NEWTCHR, β_{112}					
INTRCPT4, γ_{1120}	0.009542	0.027833	0.343	178	0.731
For PDPART, β_{113}					
INTRCPT4, γ_{1130}	0.004064	0.009894	0.411	178	0.681
For SCMT, β_{114}					
INTRCPT4, γ_{1140}	0.014517	0.010328	1.406	178	0.160
For ARTIFACT, π_{12}					
For INTRCPT3, β_{120}					
INTRCPT4, γ_{1200}	0.569328	0.133191	4.275	16	<0.001

New teachers scored considerably lower on initial status than more experienced teachers (γ_{1010} = –0.520, t = –2.297, *p*-value = 0.022.) As hypothesized by the study, both prior professional development experience PDPART and commitment to school improvement SCMT were positively related to differences among schools in initial expertise ratings (*p-values* of 0.069 and 0.107 respectively.)

In terms of teachers' growth in expertise over the course of the study, OCCASION, the study hypothesized that this would be related to differential exposure to coaching, COACH. A highly significant relationship was found, (γ_{1110} = 0.262, with associated *t*-value of 3.349 and a *p*-value = 0.001). A significant measurement artifact also occurred, see results for γ_{1200}.

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INVSTDER, e_1	0.56381	0.31788	1078	4729.76970	<0.001

Note: The chi-square statistics reported above are based on only 1312 of 1317 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of level-3 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INVSTDER/INTRCPT2,r_{10}	0.96826	0.93753	193	734.15590	<0.001
INVSTDER/OCCASION,r_{11}	0.03365	0.00113	192	267.53588	<0.001

Note: The chi-square statistics reported above are based on only 214 of 219 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

The variation on among teachers within schools on expertise ratings at the study onset, var(r_{10}), is 0.937 and the variation within schools on teachers' rate of growth in expertise, var (r_{11}), is 0.001. Both variance components are statistically significant.

Final estimation of level-4 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INVSTDER/ INTRCPT2/INTRCPT3, u_{100}	0.53703	0.28840	16	65.90635	<0.001
INVSTDER/ OCCASION/INTRCPT3, u_{110}	0.19489	0.03798	16	599.59968	<0.001
INVSTDER/ ARTIFACT/INTRCPT3, u_{120}	0.47622	0.22678	16	71.51494	<0.001

We see evidence of considerable variability among schools in teachers' initial expertise ratings, u_{110}, $\left(\chi^2 = 65.906,\ p-value < 0.001\right)$. Significant variation was also found in school growth rates, u_{110}, and in the magnitude of the measurement artifact at each school, u_{120}.

Statistics for the current model

Deviance = 6895.349602
Number of estimated parameters = 20

6.4 Other program features

Multivariate hypothesis testing and residual files at all four levels are available in HLM4. Other options found in HLM2 and HLM3 are not currently operational. For a list of all options currently available in HLM4, please see the table in Appendix J.

7 Conceptual and Statistical Background for Hierarchical Generalized Linear Models (HGLM)

The hierarchical linear model (HLM) as described in the previous six chapters is appropriate for two- and three-level data where the random effects at each level are normally distributed. The assumption of normality at level-1 is quite widely applicable when the outcome variable is continuous. Even when a continuous outcome is highly skewed, a transformation can often be found that will make the distribution of level-1 random effects (residuals) at least roughly normal. Methods for assessing the normality of random effects at higher levels are discussed on page 38 and on page 274 of *Hierarchical Linear Models.*

There are important cases, however, where the assumption of normality at level-1 is clearly not realistic and no transformation can make it so. Examples of a binary outcome, Y, are: the presence of a disease ($Y = 1$ if the disease is present, $Y = 0$ if the disease is absent), graduation from high school ($Y = 1$ if a student graduates on time, $Y = 0$ if not), or the commission of a crime ($Y = 1$ if a person commits a crime during a given time interval, $Y = 0$ if not). The use of the standard level-1 model in this case would be inappropriate for three reasons:

- Given the predicted value of the outcome, the level-1 random effect can take on only one of two values, and therefore cannot be normally distributed.
- The level-1 random effect cannot have homogeneous variance. Instead, the variance of this random effect depends on the predicted value as specified below.
- Finally, there are no restrictions on the predicted values of the level-1 outcome in the standard model: they can legitimately take on any real value. In contrast, the predicted value of a binary outcome Y, if viewed as the predicted probability that $Y = 1$, cannot meaningfully be less than zero or greater than unity. Thus, an appropriate model for predicting Y ought to constrain the predicted values to lie in the interval (0, 1). Without this constraint the effect sizes estimated by the model are, in general, uninterpretable.

Another example involves count data, where Y is the number of crimes a person commits during a year or Y is the number of questions a child asks during the course of a one-hour class period. In these cases, the possible values of Y are non-negative integers 0, 1, 2, ... Such data will typically be positively skewed. If there are very few zeros in the data, a transformation, *e.g.*, $Y^* = \log(1+Y)$, may solve this problem and allow sensible use of the standard HLM. However, in the cases mentioned above, there will typically be many zeros (many persons will not commit a crime during a given year and many children will not raise a question during a one-hour class). When there are many zeros, the normality assumption cannot be approximated by a transformation. Also, as in the case of the binary outcome, the variance of the level-1 random effects will depend on the predicted value (higher predicted values will have larger variance). Similarly, the predicted values ought to be constrained to be positive.

Another example involves multi-category (³ 2) data, where the outcome consists of responses tapping teachers' commitment to their career choice. Teachers are asked if they would choose the teaching profession if they could go back to college and start over again. The three response categories are:

1. yes, I would choose teaching again
2. not sure
3. no, I would not choose teaching again.

Such outcomes can be studied using a multinomial model. Thus, as discussed previously for models with binary outcomes, the use of the standard level-1 model would be inappropriate. Another model one may use is an ordinal model, which treats the categories as ordered.

Within HLM, the user can specify a non-linear analysis appropriate for counts and binary, multinomial, or ordinal data. The approach is a direct extension of the generalized linear model of McCullagh & Nelder (1989) to the case of hierarchical data. We therefore refer to this approach as a “hierarchical generalized linear model” (HGLM). The execution of these analyses is in many ways similar to that in HLM, but there are also important differences.

7.1 The two-level HLM as a special case of HGLM

The level-1 model in the HGLM may be viewed as consisting of three parts: a sampling model, a link function, and a structural model. In fact, the standard HLM can be viewed as a special case of the HGLM where the sampling model is normal and the link function is the identity link.

7.1.1 Level-1 sampling model

The sampling model for a two-level HLM might be written as

$$Y_{ij} \mid \mu_{ij} \sim NID(\mu_{ij}, \sigma^2) \tag{0.0}$$

meaning that the level-one outcome Y_{ij}, given the predicted value, μ_{ij}, is normally and independently distributed with an expected value of μ_{ij} and a constant variance, σ^2. The level-1 expected value and variance may alternatively be written as

$$E(Y_{ij} \mid \mu_{ij}) = \mu_{ij} \qquad Var(Y_{ij} \mid \mu_{ij}) = \sigma^2 . \tag{0.0}$$

7.1.2 Level-1 link function

In general it is possible to transform the level-1 predicted value, μ_{ij}, to η_{ij} to insure that the predictions are constrained to lie within a given interval. Such a transformation is called a link function. In the normal case, no transformation is necessary. However, this decision not to transform may be made explicit by writing

$$\eta_{ij} = \mu_{ij} . \tag{0.0}$$

The *link function* in this case is viewed as the "identity link function."

7.1.3 Level-1 structural model

The transformed predicted value is now related to the predictors through the linear model or "structural model"

$$\eta_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \cdots + \beta_{Qj} X_{Qij}. \qquad (0.0)$$

It is clear that combining the level-1 sampling model (7.1), the level-1 link function (7.3), and the level-1 structural model (7.4) reproduces the level-1 model of HLM (1.1). In the context of a standard HLM, it seems silly to write three equations where only one is needed, but the value of the extra equations becomes apparent in the case of binary, count, and multi-categorical data.

7.2 Two-, three-, and four- level models for binary outcomes

While the standard HLM uses a normal sampling model and an identity link function, the binary outcome model uses a binomial sampling model and a logit link. Only the level-1 models differ from the linear case.

7.2.1 Level-1 sampling model

Let Y_{ij} be the number of "successes" in m_{ij} trials. Then we write that

$$Y_{ij} \mid \phi_{ij} \sim B\left(m_{ij}, \phi_{ij}\right), \qquad (0.0)$$

to denote that Y_{ij} has a binomial distribution with m_{ij} trials and probability of success ϕ_{ij}. According to the binomial distribution, the expected value and variance of Y_{ij} are then

$$E(Y_{ij} \mid \phi_{ij}) = m_{ij}\phi_{ij} \quad Var(Y_{ij} \mid \phi_{ij}) = m_{ij}\phi_{ij}\left(1-\phi_{ij}\right). \qquad (0.0)$$

When m_{ij} = 1, Y_{ij} may take on values of either zero or unity. This is a special case of the binomial distribution known as the Bernoulli distribution. HGLM allows estimation of models in which m_{ij} = 1 (Bernoulli case) or m_{ij} > 1 (other binomial cases). The case with m_{ij} >1 will be treated later.

For the Bernoulli case, the predicted value of the binary Y_{ij} is equal to the probability of a success, ϕ_{ij}.

7.2.2 Level-1 link function

When the level-1 sampling model is binomial, HGLM uses the logit link function

$$\eta_{ij} = \log\left(\frac{\phi_{ij}}{1-\phi_{ij}}\right). \tag{0.0}$$

In words, η_{ij} is the log of the odds of success. Thus if the probability of success, ϕ_{ij}, is 0.5, the odds of success is 1.0 and the log-odds or "logit" is zero. When the probability of success is less than 0.5, the odds are less than one and the logit is negative; when the probability is greater than 0.5, the odds are greater than unity and the logit is positive. Thus, while ϕ_{ij} is constrained to be in the interval $(0,1)$, η_{ij} can take on any real value.

7.2.3 Level-1 structural model

This will have exactly the same form as (7.4). Note that estimates of the βs in (7.4) make it possible to generate a predicted log-odds (η_{ij}) for any case. Such a predicted log-odds can be converted to an odds by computing odds = exponential (η_{ij}). Similarly, predicted log-odds can be converted to a *predicted probability* by computing

$$\phi_{ij} = \frac{1}{1+\exp\left(-\eta_{ij}\right)}. \tag{0.0}$$

Clearly, whatever the value of η_{ij}, applying (7.8) will produce a ϕ_{ij} between zero and unity.

7.2.4 Level-2 and Level-3 and Level-4 models

In the case of a two-level analysis, the level-2 model has the same form as used in a standard 2-level HLM (equations 1.2, 1.3, and 1.4). In the case of a three-level analysis, the level-2 and level-3 models are also the same as in a standard 3-level HLM. The same applies for 4-level HLM.

7.3 The model for count data

For count data, we use a Poisson sampling model and a log link function.

7.3.1 Level-1 sampling model

Let Y_{ij} be the number of events occurring during an interval of time having length m_{ij}. For example, Y_{ij} could be the number of crimes a person *i* from group *j* commits during five years, so that m_{ij} = 5. The time-interval of m_{ij} units may be termed the "exposure." Then we write that

$$Y_{ij} \mid \lambda_{ij} \sim P(m_{ij}, \lambda_{ij}) \tag{0.0}$$

to denote that Y_{ij} has a Poisson distribution with exposure m_{ij} and event rate λ_{ij}. According to the Poisson distribution, the expected value and variance of Y_{ij} are then

$$E(Y_{ij} \mid \lambda_{ij}) = m_{ij}\lambda_{ij} \qquad Var(Y_{ij} \mid \lambda_{ij}) = m_{ij}\lambda_{ij}. \tag{0.0}$$

The exposure m_{ij} need not be a measure of time. For example, if Y_{ij} is the number of bombs dropping on neighborhood *i* of city *j* during a war, m_{ij} could be the area of that neighborhood. A common case arises when, for each *i* and *j,* the exposure is the same (*e.g.*, Y_{ij} is the number of crimes committed during one year for each person *i* within each neighborhood *j*). In this case, we set m_{ij} = 1 for simplicity. HGLM allows estimation of models in which m_{ij} = 1 or m_{ij} ³ 1. (The case with m_{ij} ³ 1 will be treated later.)

According to our level-1 model, the predicted value of Y_{ij} when m_{ij} = 1 will be the event rate λ_{ij} .

7.3.2 Level-1 link function

HGLM uses the log link function when the level-1 sampling model is Poisson, that is

$$\eta_{ij} = \log(\lambda_{ij}). \tag{0.0}$$

In words, η_{ij} is the log of the event rate. Thus, if the event rate, λ_{ij}, is one, the log is zero. When the event rate is less than one, the log is negative; when the event rate is greater than one, the log is positive. Thus, while λ_{ij} is constrained to be non-negative, η_{ij} can take on any real value.

7.3.3 Level-1 structural model

This will have exactly the same form as (7.4). Note that estimates of the βs in (7.4) make it possible to generate a predicted log-event rate (η_{ij}) for any case. Such a predicted log-event rate can be converted to an event rate by computing

$$\lambda_{ij} = \text{event rate} = \exp(\eta_{ij})$$

Clearly, whatever the value of η_{ij}, λ_{ij} will be non-negative.

7.3.4 Level-2 model

The level-2 model has the same form as the level-2 model for HLM2 (equations 1.2, 1.3, and 1.4), and the level-2 and level-3 models have the same form in the three- and four-level case as in HLM3 and HLM4, respectively.

7.4 The model for multinomial data

For multi-category nominal data, we use a multinomial model and a logit link function. This is an extension of the Bernoulli model with more than two possible outcomes. This feature is not available in HLM4.

7.4.1 Level-1 sampling model

Let

$$\text{Prob}\left(R_{ij}=m\right)=\phi_{ij},$$

that is, the probability that person *i* in group *j* lands in category *m* is ϕ_{ij}, for categories *m* = 1, ..., *M*, there being *M* possible categories.

For example, R_{ij} = 1 if high school student *i* in school *j* goes on to college; R_{ij} = 2 if that student goes on to a job; R_{ij} = 3 if that student becomes unemployed. Here *M* = 3. The analysis is facilitated by constructing dummy variables $Y_1, Y_2, \ldots, Y_M$, where Y_{mij} = 1 if R_{ij} = *m*, 0 otherwise. For example, if student *ij* goes to college, R_{ij} = 1, so Y_{1ij} = 1, Y_{2ij} = 0, Y_{3ij} = 0; if student *ij* goes to work, R_{ij} = 2, so Y_{1ij} = 0, Y_{2ij} = 1, Y_{3ij} = 0; if that student becomes unemployed, R_{ij} = 3, so Y_{1ij} = 0, Y_{2ij} = 0, Y_{3ij} = 1. This leads to a definition of the probabilities as $\text{Prob}\left(Y_{mij}=1\right)=\phi_{mij}$. For example, for *M* = 3,

$$\begin{aligned}\text{Prob}\left(Y_{1ij}=1\right)&=\phi_{1ij}\\ \text{Prob}\left(Y_{2ij}=1\right)&=\phi_{2ij}\\ \text{Prob}\left(Y_{3ij}=1\right)&=\phi_{3ij}=1-\phi_{1ij}-\phi_{2ij}\end{aligned} \tag{0.0}$$

Note that because $Y_{3ij}=1-Y_{1ij}-Y_{2ij}$, Y_{3ij} is redundant.

According to the multinomial distribution, the expected value and variance of Y_{mij} given ϕ_{mij}, are then

$$E\left(Y_{mij}\mid\phi_{mij}\right)=\phi_{mij}\qquad Var\left(Y_{mij}\mid\phi_{mij}\right)=\phi_{mij}\left(1-\phi_{mij}\right). \tag{0.0}$$

The covariance between outcomes Y_{mij} and $Y_{m'ij}$ is

$$Cov\left(Y_{mij},Y_{m'ij}\right)=-\phi_{mij}\phi_{m'ij}. \tag{0.0}$$

7.4.2 Level-1 link function

HGLM uses the logit link function when the level-1 sampling model is multinomial. Define η_{mij} as the log-odds of falling into category *m* relative to that of falling into category *M*. Specifically

$$\eta_{mij} = \log\left(\frac{\phi_{mij}}{\phi_{Mij}}\right) \tag{0.0}$$

where

$$\phi_{Mij} = 1 - \sum_{m=1}^{M-1} \phi_{mij}. \tag{0.0}$$

In words, η_{mij} is the log odds of being in *m*-th category relative to the *M*-th category, which is known as the "reference category."

7.4.3 Level-1 structural model

At level-1, we have

$$\eta_{mij} = \beta_{0j(m)} + \sum_{q=1}^{Q} \beta_{qj(m)} X_{qij}, \tag{0.0}$$

for $m = 1, ..., (M-1)$. For example, with $M = 3$, there would be two level-1 equations, for η_{1ij} and η_{2ij}.

7.4.4 Level-2 model

The level-2 model has a parallel form

$$\beta_{qj(m)} = \gamma_{q0(m)} + \sum_{s=1}^{S_q} \gamma_{qs(m)} W_{sj} + u_{qj(m)}. \tag{0.0}$$

Thus, for $M = 3$, there would be two sets of level-2 equations.

7.5 The model for ordinal data

7.5.1 Level-1 sampling model

Again a person falls into category *m* and there are *M* possible categories, so $m = 1, ..., M$. But now the categories are ordered. Given the ordered nature of the data, we derive the *M* dummy variables $Y_{mij}, \ldots, Y_{(M-1)ij}$ for case *i* in unit *j* as

$$Y_{mij} = 1 \quad \text{if} \quad R_{ij} \le m, \; 0 \; \text{otherwise}. \tag{0.0}$$

For example, with $M = 3$, we have

$$\begin{aligned} Y_{1ij} &= 1 \quad \text{if} \quad R_{ij} = 1 \\ Y_{2ij} &= 1 \quad \text{if} \quad R_{ij} \le 2 \end{aligned} \tag{0.0}$$

The probabilities $\text{Prob}\left(Y_{mij} = 1\right)$ are thus cumulative probabilities. For example, with $M = 3$,

$$\begin{aligned} \text{Prob}\left(Y_{1ij} = 1\right) &= \text{Prob}\left(R_{ij} = 1\right) = \phi_{1ij} \\ \text{Prob}\left(Y_{2ij} = 1\right) &= \text{Prob}\left(R_{ij} = 1\right) + \text{Prob}\left(R_{ij} = 2\right) = \phi_{2ij} \\ \text{Prob}\left(Y_{3ij} = 1\right) &= \text{Prob}\left(R_{ij} = 1\right) + \text{Prob}\left(R_{ij} = 2\right) + \text{Prob}\left(R_{ij} = 3\right) = 1 \end{aligned} \tag{0.0}$$

Since $Y_{3ij} = 1 - Y_{2ij}$, Y_{3ij} is redundant. We actually need only $M - 1$ dummy variables.

Associated with the cumulative probabilities are the cumulative logits,

$$\eta_{mij} = \log\left(\frac{\text{Prob}\left(R_{ij} \le m\right)}{\text{Prob}\left(R_{ij} > m\right)}\right) = \log\left(\frac{\phi_{mij}}{1 = \phi_{mij}}\right). \tag{0.0}$$

7.5.2 Level-1 structural model

The level-1 structural model assumes "proportional odds",

$$\eta_{mij} = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \sum_{m=2}^{M} \delta_m. \tag{0.0}$$

Under the proportional odds assumption, the relative odds that $R_{ij} \le m$, associated with a unit increase in the predictor, does not depend on m.

Here δ_m is a "threshold" that separates categories $m - 1$ and m. For example, when $M = 4$,

$$\begin{aligned} \eta_{1ij} &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} \\ \eta_{2ij} &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \delta_2 \\ \eta_{3ij} &= \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj} X_{qij} + \delta_2 + \delta_3 \end{aligned} \tag{0.0}$$

7.6 Parameter estimation

HLM2 and HLM3 use three approaches to estimation for HGLM. The first method bases inference on the joint posterior modes of the level-1 and level-2 (and level-3) regression coefficients given the variance-covariance estimates. The variance-covariance estimates are based on a normal approximation to the restricted likelihood. Stiratelli, Laird, & Ware (1984) and Wong & Mason (1985) developed this approach for the binary case. Schall (1991) discusses the extension of this approach to the wider class of generalized linear models. Breslow & Clayton (1993) refer to this estimation approach as "penalized quasi-likelihood" or PQL. Extending HLM to HGLM requires a doubly iterative algorithm, significantly increasing computational time. Related approaches are described by Goldstein (1991), Longford (1993), and Hedeker & Gibbons (1994).

The second and third methods of estimation ("Laplace and "adaptive Gaussian quadrature") involve somewhat more computationally intensive algorithms but provide accurate approximation to maximum likelihood (ML). These two approaches are currently available for two-level and three-level Bernoulli models and for two-level Poisson models with $m_{ij} = 1$. We consider PQL below in some detail followed by a brief discussion of Laplace and adaptive Gaussian quadrature.

7.6.1 Estimation via PQL

The approach can be presented heuristically by computing a "linearized dependent variable" as in the generalized linear model of McCullagh and Nelder (1989). Basically, the analysis involves use of a standard HLM model with the introduction of special weighting at level-1. However, after this standard HLM analysis has converged, the linearized dependent variable and the weights must be recomputed. Then, the standard HLM analysis is re-computed. This iterative process of analyses and recomputing weights and linearized dependent variable continues until estimates converge.

We term the standard HLM iterations "micro-iterations." The recomputation of the linearized dependent variable and the weights constitute a "macro iteration." The approach is outlined below for four cases: Bernoulli (binomial with $m_{ij} = 1$), Poisson with $m_{ij} = 1$, binomial with $m_{ij} > 1$, and Poisson with $m_{ij} > 1$.

7.6.1.1 Bernoulli (binomial with $m_{ij} = 1$)

Consider the model

$$Y_{ij} = \phi_{ij} + \varepsilon_{ij} \qquad (0.0)$$

with ϕ_{ij} defined as in Equation 7.8 and

$$E(\varepsilon_{ij}) = 0 \qquad Var(\varepsilon_{ij}) = w_{ij} = \phi_{ij}(1 - \phi_{ij}). \qquad (0.0)$$

We now substitute for ϕ_{ij} its linear approximation with

$$\phi_{ij} \approx \phi_{ij}^{(0)} + \frac{\partial \phi_{ij}^{(i)}}{\partial \eta_{ij}^{(i)}} \left(\eta_{ij} - \eta_{ij}^{(0)} \right) \qquad (0.0)$$

$$\eta_{ij}^{(0)} = \log\left(\frac{\phi_{ij}^{(0)}}{1-\phi_{ij}^{(0)}}\right), \tag{0.0}$$

where $\phi_{ij}^{(0)}$ is an initial estimate and

$$\frac{\partial \phi_{ij}}{\partial \eta_{ij}} = w_{ij} = \phi_{ij}\,(1-\phi_{ij})\,. \tag{0.0}$$

If we evaluate w_{ij} at its initial estimates

$$w_{ij}^{(0)} = \phi_{ij}^{(0)}\,(1 - \phi_{ij}^{(0)})\,. \tag{0.0}$$

(7.25) can be written as

$$Y_{ij} = \phi_{ij}^{(0)} + w_{ij}^{(0)}\left(\eta_{ij} - \eta_{ij}^{(0)}\right) + \varepsilon_{ij}. \tag{0.0}$$

Algebraically rearranging the equation so that all observables are on the left-hand side yields

$$\begin{aligned} Z_{ij}^{(0)} &= \eta_{ij} + \frac{\varepsilon_{ij}}{w_{ij}^{(0)}} \\ &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{Qj}X_{Qij} + e_{ij}, \end{aligned} \tag{0.0}$$

where

$$Z_{ij}^{(0)} = \frac{Y_{ij} - \phi_{ij}^{(0)}}{w_{ij}^{(0)}} + \eta_{ij}^{(0)} \tag{0.0}$$

is the linearized dependent variable and

$$Var(e_{ij}) = Var\left(\frac{\varepsilon_{ij}}{w_{ij}^{(0)}}\right) \approx \frac{1}{w_{ij}^{(0)}}. \tag{0.0}$$

Thus, (7.32) is a standard HLM level-1 model with outcome $Z_{ij}^{(0)}$ and level-1 weighting variable $w_{ij}^{(0)}$.

The algorithm works as follows.

1. Given initial estimates of the predicted value, ϕ_{ij}, and therefore of the linearized dependent variable, Z_{ij}, and the weight, w_{ij}, compute a weighted HLM analysis with (7.32) as the level-1 model.
2. The HLM analysis from step 1 will produce new predicted values and thus new linearized dependent variables and weights. HLM will now compute a new, re-weighted MDM file with the appropriate linearized dependent variable and weights.
3. Based on the new linearized dependent variable and weights, re-compute step 1.

This process goes on until the linearized dependent variable, the weights, and therefore, the parameter estimates, converge to a pre-specified tolerance. The program then stops.

7.6.1.2 Poisson with $m_{ij} = 1$

The procedure is exactly the same as in the binomial case with $m_{ij} = 1$ except that

$$Var(\varepsilon_{ij}) = w_{ij} = \frac{\partial \lambda_{ij}}{\partial \eta_{ij}} = \lambda_{ij}. \qquad (0.0)$$

7.6.1.3 Binomial with $m_{ij} > 1$

In the previous example, Y_{ij} was formally the number of successes in one trial and therefore could take on a value of 0 or 1. We now consider the case where Y_{ij} is the number of successes in m_{ij} trials, where Y_{ij} and m_{ij} are non-negative integers, $Y_{ij} \le m_{ij}$.

Suppose that a researcher is interested in examining the relationship between pre-school experience (yes or no) and grade retention and wonders whether this relationship is similar for males and females. The design involves students at level 1 nested within schools at level 2. In this case, each school would have four “cell counts” (boys with and without pre-school and girls with and without pre-school). Thus, the data could be organized so that every school had four observations (except possibly schools without variation on pre-school or sex), where each observation was a cell having a cell size m_{ij} and a cell count Y_{ij} of students in that cell who were, in fact, retained. One could then re-conceptualize the study as having up to four level-1 units (cells); the outcome Y_{ij}, given the cell probability ϕ_{ij}, would be distributed as $B\left(m_{ij}, \phi_{ij}\right)$. There would be three level-1 predictors (a contrast for pre-school experience, a contrast for sex, and an interaction contrast). This problem then has the structure of a $2 \times 2 \times J$ contingency table (pre-school experience by sex by school) with the last factor viewed as random.

The structure of a level-1 file for group 2 might appear as follows.

Group	ID	n_{ij}	Y_{ij}	X_{1ij}	X_{2ij}	X_{3ij}
Girls with pre-school	2	n_{12}	Y_{12}	0.50	0.50	0.25
Girls without pre-school	2	n_{22}	Y_{22}	0.50	-0.50	-0.25
Boys with pre-school	2	n_{32}	Y_{32}	-0.50	0.50	-0.25
Boys without pre-school	2	n_{42}	Y_{42}	-0.50	-0.50	0.25

For example, n_{12} is the number of girls in school 2 with pre-school and Y_{12} is the number of those girls who were retained. The predictor X_{1ij} is a contrast coefficient to assess the effect of sex (0.5 if female, –0.5 if male); X_{2ij} is a contrast for pre-school experience (0.5 if yes, –0.5 if no), and $X_{3ij} = X_{1ij} \times X_{2ij}$ is the interaction contrast.

Estimation works the same in this case as in the binomial case except that

$$Z_{ij} = \frac{Y_{ij} - m_{ij}\phi_{ij}}{w_{ij}} + \eta_{ij} \tag{0.0}$$

with

$$w_{ij} = m_{ij}\phi_{ij}(1-\phi_{ij}). \tag{0.0}$$

7.6.1.4 Poisson with $m_{ij} > 1$

Consider now a study of the number of homicides committed within each of j neighborhoods in a large city. Many neighborhoods will have no homicides. The expected number of homicides in a neighborhood will depend not only on the homicide rate for that neighborhood, but also on the size of that neighborhood as indexed by its number of residents, m_{ij}. Level-1 variables might include characteristics of the homicide (*e.g.*, whether the homicide involved a domestic dispute, whether it involved use of a gun). Each cell (*e.g.*, the four types of homicide as defined by the cross-classification of domestic – yes or no – and use of a gun – yes or no) would be a level-1 unit.

Estimation in this case is the same as in the Poisson case with $m_{ij} = 1$ except that

$$Z_{ij} = \frac{Y_{ij} - m_{ij}\lambda_{ij}}{w_{ij}} + \eta_{ij} \tag{0.0}$$

and

$$w_{ij} = m_{ij}\lambda_{ij}. \quad (0.0)$$

7.6.2 Properties of the estimators

Using PQL, HGLM produces approximate empirical Bayes estimates of the randomly varying level-1 coefficients, generalized least squares estimators of the level-2 (and level-3 or level-4) coefficients, and approximate maximum-likelihood estimators of the variance and covariance parameters. Yang (1995) has conducted a simulation study of these estimators in comparison with an alternative approach used by some programs that sets the level-2 random coefficients to zero in computing the linearized dependent variables. Breslow & Clayton (1993) refer to this alternative approach as “marginalized quasi-likelihood” or MQL. Rodriquez & Goldman (1995) had found that MQL produced biased estimates of the level-2 variance and the level-2 regression coefficients. Yang's results showed a substantial improvement (reduction in bias and mean squared error) in using the approach of HGLM. In particular, the bias in estimation of the level-2 coefficients was never more than 10 percent for HGLM, while the MQL approach commonly produced a bias between 10 and 20 percent. HGLM performed better than the alternative approach in estimating a level-2 variance component as well. However, a negative bias was found in estimating this variance component, ranging between two percent and 21 percent. The bias was most severe when the true variance was very large and the typical “probability of success” was very small (or, equivalently, very large). Initial simulation results under the Poisson model appear somewhat more favorable than this. Breslow & Clayton (1993) suggest that the estimation will be more efficient as the level-1 sample size increases.

7.6.3 Parameter estimation: A high-order Laplace and adaptive Gaussian Quadrature approximation of maximum likelihood

For two- and three-level models with binary and count outcomes, HGLM provides two alternatives to estimation via PQL: a high-order Laplace and an adaptive Gaussian Quadrature approximation. Figure 7.1 displays the dialog box for the estimation settings for two-level models.

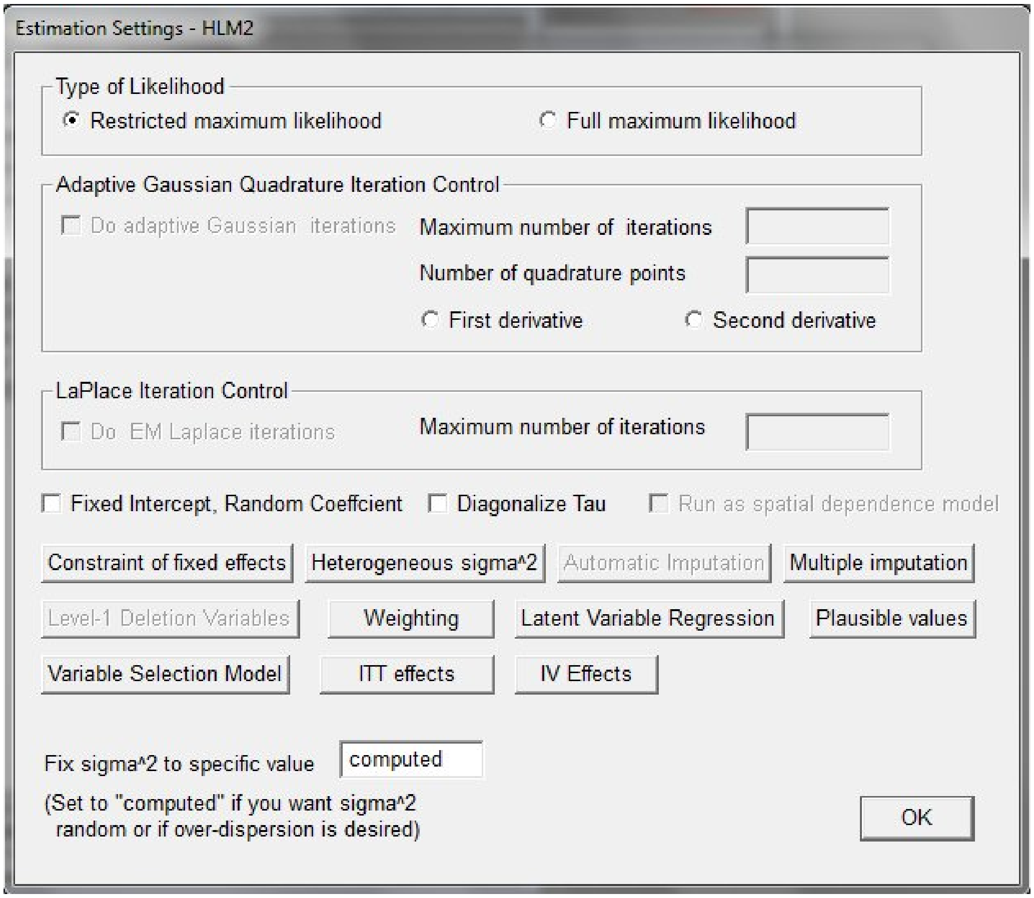


Figure 7.1 Estimation settings for two-level hierarchical generalized linear models

One alternative for two- and three-level Bernoulli and Poisson models with constant and variable exposure uses a high-order approximation to the likelihood based on a Laplace transform. The adaptive Gauss-Hermite quadrature (AGQ) technique (Pinheiro & Bates, 1995) is another approximation option available for two- and three-level binomial and Poisson models with constant and variable exposure. For AGQ, users have the options to specify the number of quadrature points and to choose the use of a first or a second derivative approximation. Both accuracy in approximation and computational demands increase as the number of nodes specified increases and when the second derivative option is used.

For two-level Bernoulli models, Yang (1998), Raudenbush, Yang, and Yosef (2000) and Yosef (2001) found that both the Laplace and AGQ techniques yielded accurate estimates. Results of Yosef (2001) suggested AGQ performed better for models with small cluster size ($n_{ij} = 2$) in terms of smaller means-squared errors and biases. The Laplace method, on the other hand, gave more accurate approximation in models with bivariate random effects. Johnson (2006) showed in his simulation study that for two-level Poisson models with equal exposure, the Laplace and AGQ estimates in general displayed less bias than those of PQL. However, AGQ gave more accurate approximation when the event rate was low and the level-2 variance was large ($\tau_{00} = 1$). Based on his results, he recommended AGQ be used with small event rate and small cluster size ($n_{ij} = 2$).

7.7 Unit-specific and population-average models

The models described above have been termed "unit-specific" models. They model the expected outcome for a level-2 unit conditional on a given set of random effects. For example, in the Bernoulli case ($m_{ij} = 1$), we might have a level-1 (within-school) model

$$\eta_{ij} = \beta_{0j} + \beta_{qj} X_{ij}, \tag{0.0}$$

and a level-2 (between-school) model

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01} W_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned} \tag{0.0}$$

leading to the combined model

$$\eta_{ij} = \gamma_{00} + \gamma_{01} W_j + \gamma_{10} X_{ij} + u_{0j}. \tag{0.0}$$

Under this model, the predicted probability for case *ij*, given u_{0j}, would be

$$E\left(Y_{ij} \mid u_{0j}\right) = \frac{1}{1 + \exp\left\{-\left(\gamma_{00} + \gamma_{01} W_j + \gamma_{10} X_{ij} + u_{0j}\right)\right\}}. \tag{0.0}$$

In this model γ_{10} is the expected difference in the log-odds of "success" between two students who attend the same school but differ by one unit on X (holding u_{0j} constant); γ_{01} is the expected difference in the log-odds of success between two students who have the same value on W but attend schools differing by one unit on W (holding u_{0j} constant). These definitions parallel definitions used in a standard HLM for continuous outcomes.

However, one might also want to know the average difference between log-odds of success of students having the same X but attending schools differing by one unit on W, that is, the difference of interest *averaging over all possible values of* u_{0j}. In this case, the unit-specific model would not be appropriate. The model that would be appropriate would be a "population-average" model (Zeger, Liang, & Albert, 1988). The distinction is tricky in part because it does not arise in the standard HLM (with an identity link function). It arises only in the case of a non-linear link function.

Using the same example as above, the population average model would be

$$E\left(Y_{ij}\right) = \frac{1}{1 + \exp\left\{-\left(\gamma_{00}^{*} + \gamma_{01}^{*} W_j + \gamma_{10}^{*} X_{ij}\right)\right\}}. \tag{0.0}$$

Notice that (7.41) does not condition on (or "hold constant") the random effect u_{0j}. Thus, γ_{01}^{*} gives the expected difference in log-odds of success between two students with the same X who attend schools differing by one unit on W – without respect to the random effect, u_{0j}. If one had a nationally representative sample and could validly assign a causal inference to W, γ_{01}^{*} would be the change in the log-odds of success in the whole society associated with boosting W by one unit while γ_{01} would be the change in log-odds associated with boosting W one unit for those schools sharing the same value of u_{0j}.

HGLM produces estimates for both the unit-specific and population-average models. The population-average results are based on generalized least squares given the variance-covariance estimates from the unit-specific model. Moreover, HGLM produces robust standard error estimates for the population-average model (Zeger, *et al.*, 1988). These standard errors are relatively insensitive to misspecification of the variances and covariances at the two levels and to the distributional assumptions at each level. The method of estimation used in HGLM for the population-average model is equivalent to the "generalized estimating equation" (GEE) approach popularized by Zeger, *et al.* (1988).

The following differences between unit-specific and population-average results are to be expected:

- If all predictors are held constant at their means, and if their means are zero, the population-average intercept can be used to estimate the average probability of success across the entire population, that is

$$\hat{\varphi}_{ij} = \frac{1}{1+\exp(-\gamma_{00}^{*})}. \tag{0.0}$$

  This will not be true of unit-specific intercepts unless the average probability of success is very close to .5.
- Coefficient estimates (other than the intercept) based on the population-average model will often tend to be similar to those based on the unit-specific model but will tend to be smaller in absolute value.

Users will need to take care in choosing unit-specific versus population-average results for their research. The choice will depend on the specific research questions that are of interest. In the previous example, if one were primarily interested in how a change in W can be expected to affect a particular individual school's mean, one would use the unit-specific model. If one were interested in how a change in W can be expected to affect the overall population mean, one would use the population-average model.

7.8 Over-dispersion and under-dispersion

As mentioned earlier, if the data follow the assumed level-1 sampling model, the level-1 variance of the Y_{ij} will be w_{ij} where

$$
\begin{aligned}
w_{ij} &= m_{ij}\phi_{ij}(1-\phi_{ij}), && \text{Binomial case, or} \\
w_{ij} &= m_{ij}\lambda_{ij}, && \text{Poisson case.}
\end{aligned} \tag{0.0}
$$

However, if the level-1 data do not follow this model, the actual level-1 variance may be larger than that assumed (over-dispersion) or smaller than that assumed (under-dispersion). For example, if undetected clustering exists within level-1 units or if the level-1 model is under-specified, extra-binomial or extra-Poisson dispersion may arise. This problem can be handled in a variety of ways; HGLM allows estimation of a scalar variance so that the level-1 variance will be $\sigma^2 w_{ij}$.

7.9 Restricted versus full PQL versus full ML

The default method of estimation for HGLM is restricted PQL, while full PQL is an option. For the three-and four-level HGLM, PQL estimation is by means of full PQL only. All estimates based on Laplace and adaptive Gauss-Hermite Quadratures are based on full ML.

7.10 Hypothesis testing

The logic of hypothesis testing with HGLM is quite similar to that used in the case of HLM. Thus, for the fixed effects (the γ s), a table of approximate *t*-values is routinely printed for univariate tests; multivariate tests for the fixed effects are available using the approach described earlier in Chapter 2. Similarly, univariate tests for variance components (approximate chi-squares) are also routinely printed out. The one exception is that multivariate tests based on comparing model deviances (-2 log likelihood at convergence) are not available using PQL, because PQL is based on quasi-likelihood rather than maximum-likelihood estimation. These are available using Laplace or adaptive Gauss-Hermite quadrature.

8 Fitting HGLMs (Nonlinear Models)

There is no difference between HGLM ("nonlinear analysis") and HLM ("linear analysis") in the construction of the MDM file. Thus, the same MDM file can be used for nonlinear and linear analysis.

8.1 Executing nonlinear analyses based on the MDM file

Model specification for nonlinear analyses, as in the case of linear analyses, can be achieved via Windows (PC implementation only), interactive execution, or batch execution. The mechanics of model specification are generally the same as in linear analyses with the following differences:

- Six types of nonlinear analysis are available. With Windows execution, these options are displayed in the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1). This dialog box is accessed by clicking the **Outcome** button at the top of the variable list box to the left of the main HLM window. There are two choices for dichotomous outcomes, two for count outcomes, one for multinomial outcomes, and one for ordinal outcomes.
- Highly accurate approximations to maximum likelihood based on either the Laplace approximation or adaptive Gauss-Hermite Quadrature are available for 2- and 3-level Bernoulli models and for 2-level Poisson models through the **Estimation Settings – HLM2** dialog box shown in Figure 8.3.
- If desired, an over-dispersion option is available for binomial and Poisson models. This option is not available with Laplace (see Figure 8.3). To specify over-dispersion, set the σ^2 value to **computed** in the **Estimation Settings – HLM2** dialog box (see Figure 8.3).
- As mentioned, the nonlinear analysis is doubly iterative so the maximum number of macro iterations can be specified as well as the maximum number of micro iterations. Similarly, convergence criteria can be reset for macro iterations as well as micro iterations.[2] The number of iterations and method of estimation is set through the **Iteration Control – HLM2** dialog box shown in Figure 8.2.

[2]The overall accuracy of the parameter estimates is determined by the convergence criterion for macro iterations. The convergence criterion for micro iterations will influence the number of micro iterations per macro iteration. The default specifications stop macro iterations when the largest parameter estimate change is less than 10^{-4}; micro iterations within macro iterations stop when the conditional log likelihood (conditional on the current weights and values of the linearized dependent variable) changes by less than 10^{-6}.

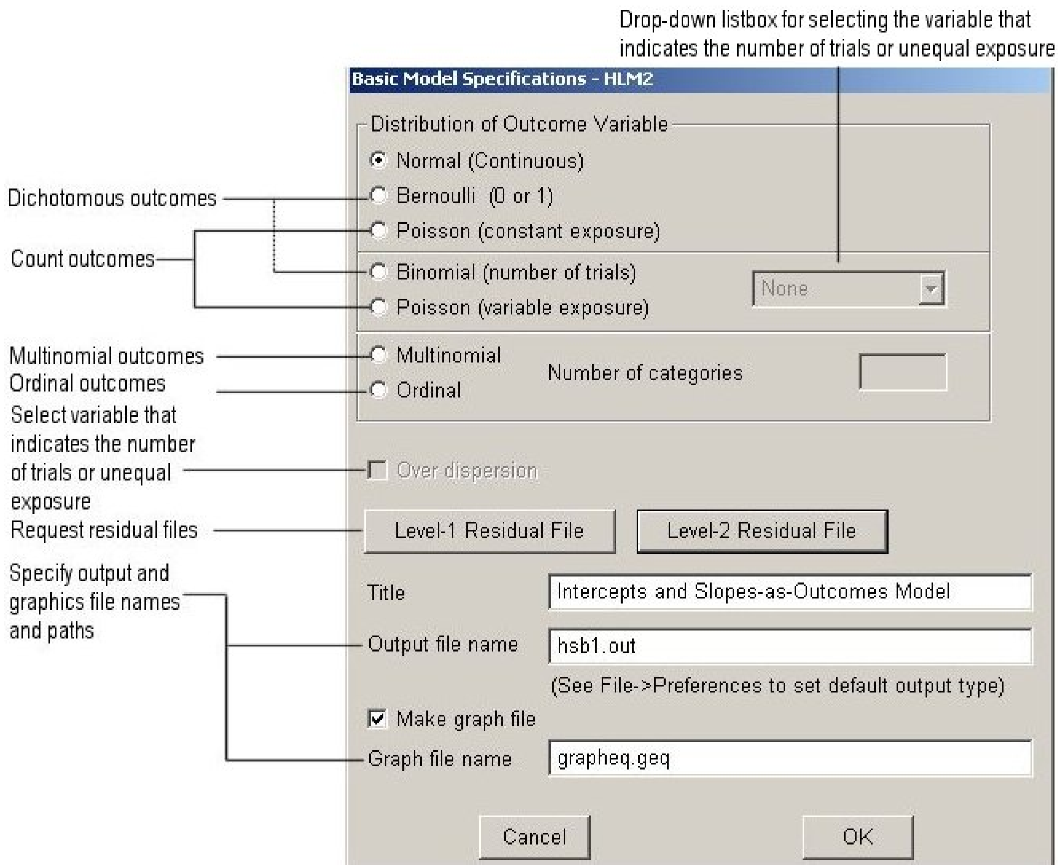


Figure 8.1 Basic Model Specifications – HLM2 dialog box

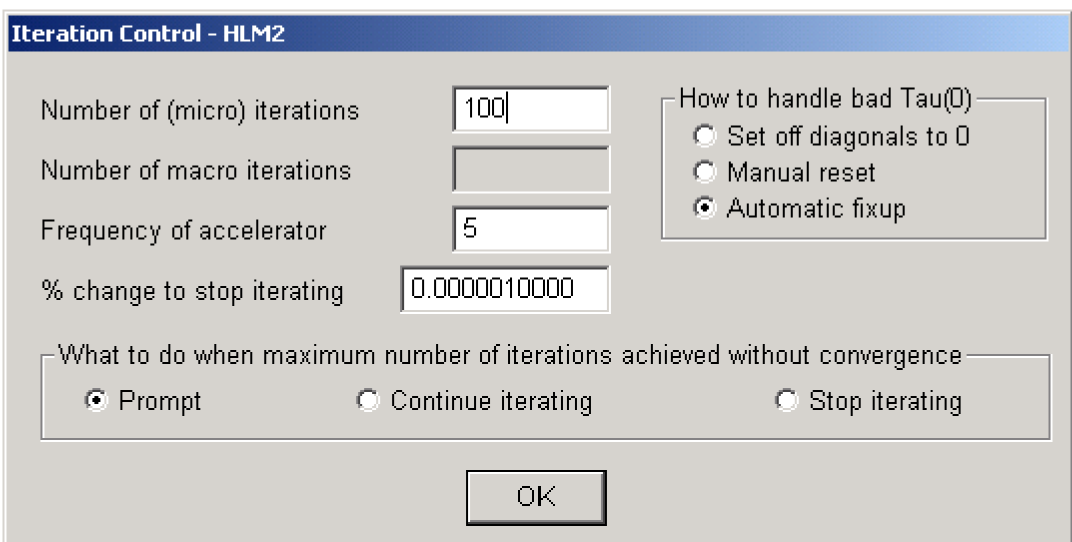


Figure 8.2 Iteration Control – HLM2 dialog box

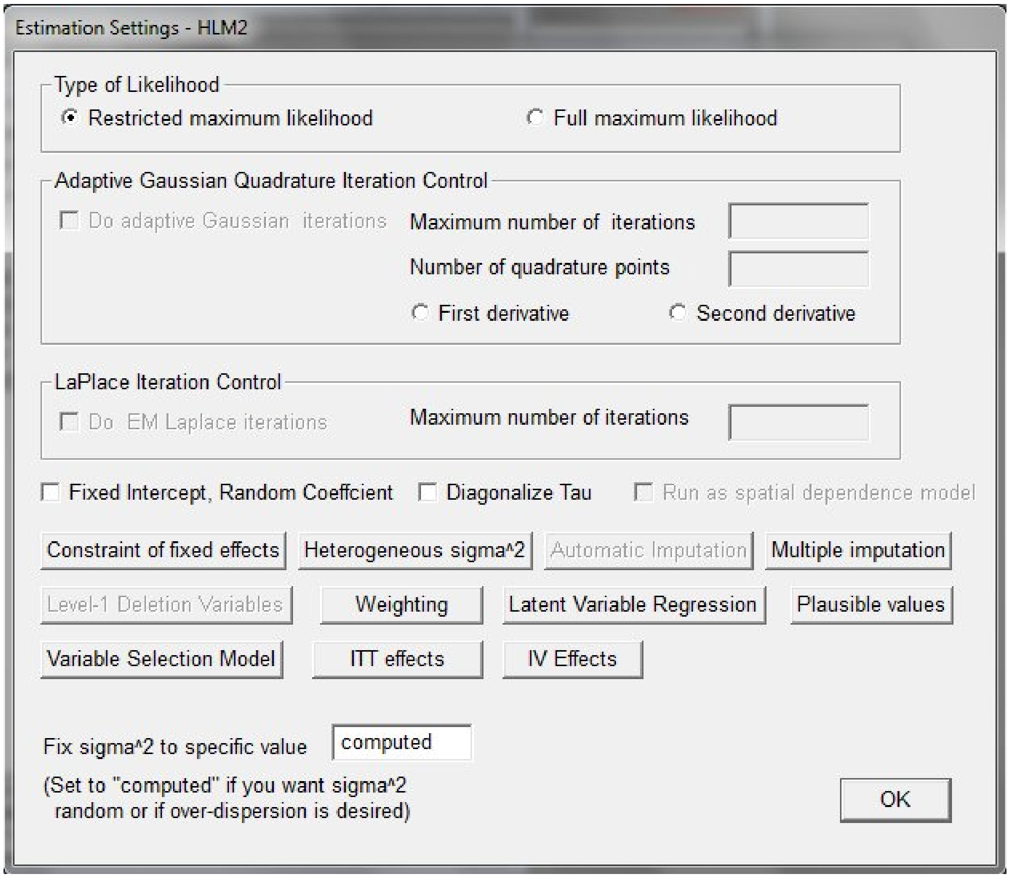


Figure 8.3 Estimation Settings – HLM2 dialog box

Below we provide two detailed examples of nonlinear analyses: the first uses the Bernoulli model, that is, a binomial model with the number of trials, m_{ij}, equal to one. The second example uses a binomial model with $m_{ij} > 1$. The analogs of these two analyses for count data are, respectively, the Poisson model with equal exposure and the Poisson case with variable exposure (some brief notes about these two applications are also included). Finally, we furnish two examples for multi-category outcomes, one for multinomial data and one for ordinal data. Windows mode specification is illustrated. See Appendix D for interactive and batch specification.

8.2 Case 1: a Bernoulli model

Data are from a national survey of primary education in Thailand (see Raudenbush & Bhumirat, 1992, for details), conducted in 1988, and yielding, for our analysis, complete data on 7516 sixth graders nested within 356 primary schools. Of interest is the probability that a child will repeat a grade during the primary years (REP1 = 1 if yes, 0 if no). It is hypothesized that the sex of the child (MALE = 1 if male, 0 of female), the child's pre-primary experience (PPED = 1 if yes, 0 if no), and the school mean SES (MSESC) will be associated with the probability of repetition. Every level-1 record corresponds to a student, with a single binary outcome per student, so the model type is Bernoulli. These data (level-1 and level-2) data files are UTHAIL1.SAV and THAI2.SAV.

Below are the Windows commands for specifying a Bernoulli model.

To specify a Bernoulli model

1. After specifying the outcome in the model specification window (REP1 in our example), click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1).
2. Select **Bernoulli (0 or 1)** as there is one binary outcome per level-1 unit.
3. (Optional) Specify the maximum number of macro and micro iterations by selecting the **Iteration Settings** option from the **Other Settings** menu.
4. (Optional) Select **Laplace approximation** or **Adaptive Gaussian iteration control** from the options on the **Estimation Settings – HLM2** dialog box, which is accessed by selecting the **Estimation Settings** options from the **Other Settings** menu (See sections 8.8 and 7.6.3).

The model described above is displayed in Figure 8.4 in both standard and mixed model notation. The command file for the model is THAIU1.HLM.

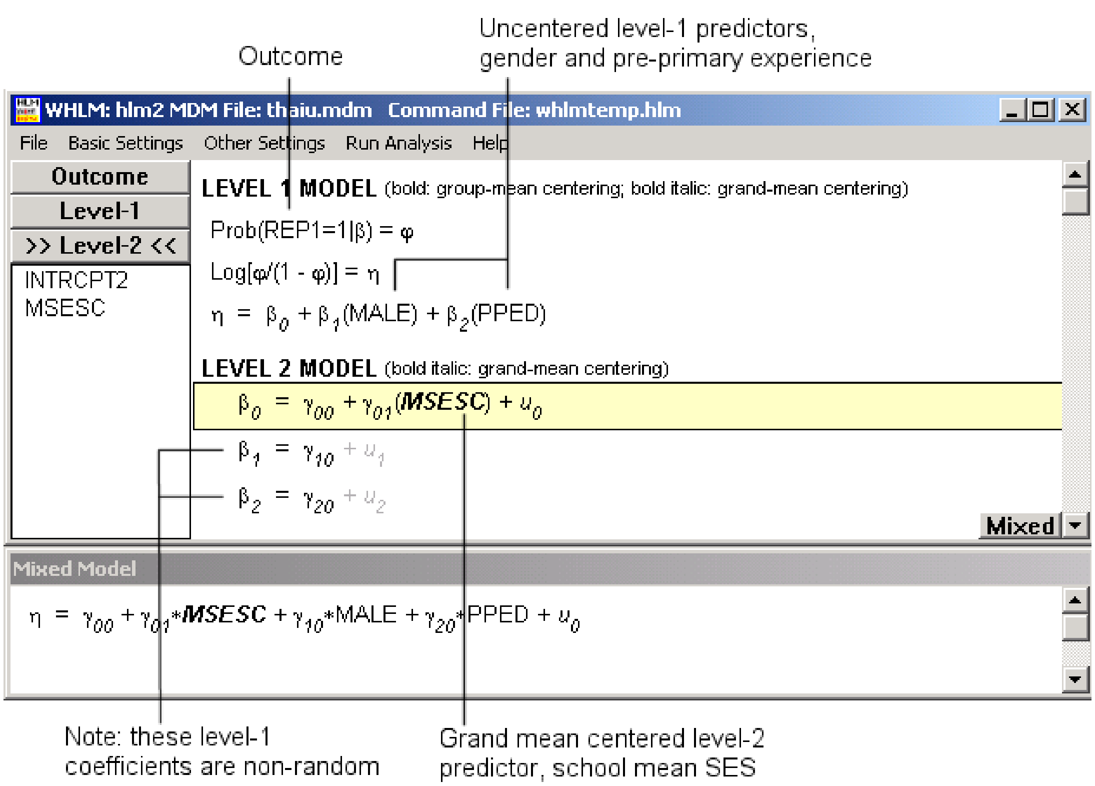


Figure 8.4 Model specification window for the Bernoulli model

Below we provide a transcript of the messages that HLM2 sent to the iteration window during computation of the results.

MACRO ITERATION 1

```
Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -2.400265E+003
The value of the likelihood function at iteration 2 = -2.399651E+003
The value of the likelihood function at iteration 3 = -2.399620E+003
The value of the likelihood function at iteration 4 = -2.399614E+003
The value of the likelihood function at iteration 5 = -2.399612E+003
The value of the likelihood function at iteration 6 = -2.399612E+003
The value of the likelihood function at iteration 7 = -2.399612E+003
```

Macro iteration number 1 has converged after seven micro iterations. This macro iteration actually computes the linear-model estimates (using the identity link function as if the level-1 errors were assumed normal). These results are then transformed and input to start macro iteration 2, which is, in fact, the first nonlinear iteration.

MACRO ITERATION 2

```
Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.067218E+004
The value of the likelihood function at iteration 2 = -1.013726E+004
The value of the likelihood function at iteration 3 = -1.011008E+004
The value of the likelihood function at iteration 4 = -1.010428E+004
The value of the likelihood function at iteration 5 = -1.010265E+004
The value of the likelihood function at iteration 6 = -1.010193E+004
The value of the likelihood function at iteration 7 = -1.010188E+004
The value of the likelihood function at iteration 8 = -1.010188E+004
The value of the likelihood function at iteration 9 = -1.010187E+004
The value of the likelihood function at iteration 10 = -1.010187E+004
The value of the likelihood function at iteration 11 = -1.010187E+004
The value of the likelihood function at iteration 12 = -1.010187E+004
```

Macro iteration 2, the first nonlinear macro iteration, converged after twelve micro iterations.

.

.

MACRO ITERATION 8

```
Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000374E+004
The value of the likelihood function at iteration 2 = -1.000374E+004
```

Note that macro iteration 8 converged with just 2 micro iterations. Macro iteration 8 was the final “unit-specific“ macro iteration. One final “population-average“ iteration is computed. Its output is given below.

MACRO ITERATION 9

```
Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.011638E+004
The value of the likelihood function at iteration 2 = -1.010710E+004
The value of the likelihood function at iteration 3 = -1.010710E+004
```

Next, we examine the output file THAIBERN.OUT.

SPECIFICATIONS FOR THIS NONLINEAR HLM RUN

Problem Title: Bernoulli output, Thailand data

The data source for this run = THAIUGRP.MDM
The command file for this run = THAIBERN.HLM
Output file name = THAIBERN.HTML
The maximum number of level-1 units = 7516
The maximum number of level-2 units = 356
The maximum number of micro iterations = 20
Method of estimation: restricted PQL
Maximum number of macro iterations = 25

Distribution at Level-1: Bernoulli

The outcome variable is REP1

Summary of the model specified

Level-1 Model

$Prob(\text{REP1}_{ij}=1|\beta_j) = \phi_{ij}$

$log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij}$

$\eta_{ij} = \beta_{0j} + \beta_{1j}*(\text{MALE}_{ij}) + \beta_{2j}*(\text{PPED}_{ij})$

Thus, the level-1 structural model is

$$\eta_{ij} = \log\left[\frac{\phi_{ij}}{1-\phi_{ij}}\right] = \beta_{0j} + \beta_{1j}\left(\text{MALE}\right)_{ij} + \beta_{2j}\left(\text{PPED}\right)_{ij}$$

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(\text{MSESC}_j) + u_{0j}$
$\beta_{1j} = \gamma_{10}$
$\beta_{2j} = \gamma_{20}$

MSESC has been centered around the grand mean**.**

And the level-2 structural model is

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}\left(\text{MSESC}\right)_{ij} + u_{0j}\\ \beta_{1j} &= \gamma_{10}\\ \beta_{2j} &= \gamma_{20}.\end{aligned}$$

Level-1 variance = $1/[\phi_{ij}\ (1-\phi_{ij})]$

In the metric of the linearized dependent variable, the level-1 variance is the reciprocal of the Bernoulli variance, $\phi_{ij}(1-\phi_{ij})$.

Mixed Model

$\eta_{ij} = \gamma_{00} + \gamma_{01}$*MSESC$_j$ + γ_{10}*MALE$_{ij}$ + γ_{20}*PPED$_{ij}$ + u_{0j}

Three sets of output results appear below: those for the normal linear model with identity link function, those for the unit-specific model with logit link function, and those for the population-average model with logit link. Typically, only the latter 2 sets of results will be relevant for drawing conclusions. The linear model with identity link is estimated simply to obtain starting values for the estimation of the models with logit link.

Final Results for Linear Model with the Identity Link Function

σ^2 **=** 0.12181

τ

INTRCPT1,β_0 0.01897

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.749

The value of the log-likelihood function at iteration 6 = -2.413825E+003

Estimation of fixed effects: (linear model with identity link function)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	0.153756	0.010812	14.221	354	<0.001
MSESC, γ_{01}	-0.033414	0.022465	-1.487	354	0.138
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.054131	0.008330	6.498	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.064613	0.010926	-5.914	7158	<0.001

Results for Non-linear Model with the Logit Link Function
Unit-Specific Model, PQL Estimation - (macro iteration 8)

τ

INTRCPT1,β_0 1.29571

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.682

The value of the log-likelihood function at iteration 2 = -1.001031E+004

Final estimation of fixed effects: (Unit-specific model)

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-2.046961	0.093985	-21.780	354	<0.001
MSESC, γ_{01}	-0.254412	0.193319	-1.316	354	0.189
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.508561	0.073935	6.879	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.594375	0.095962	-6.194	7158	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	-2.046961	0.129127	(0.107,0.155)
MSESC, γ_{01}	-0.254412	0.775372	(0.530,1.134)
For MALE slope, β_1			
INTRCPT2, γ_{10}	0.508561	1.662897	(1.439,1.922)
For PPED slope, β_2			
INTRCPT2, γ_{20}	-0.594375	0.551908	(0.457,0.666)

Final estimation of fixed effects
(Unit-specific model with robust standard errors)

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-2.046961	0.094872	-21.576	354	<0.001
MSESC, γ_{01}	-0.254412	0.204048	-1.247	354	0.213
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.508561	0.075994	6.692	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.594375	0.094840	-6.267	7158	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	-2.046961	0.129127	(0.107,0.156)
MSESC, γ_{01}	-0.254412	0.775372	(0.519,1.158)
For MALE slope, β_1			
INTRCPT2, γ_{10}	0.508561	1.662897	(1.433,1.930)
For PPED slope, β_2			
INTRCPT2, γ_{20}	-0.594375	0.551908	(0.458,0.665)

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
INTRCPT1, u_0	1.13829	1.29571	354	1431.43082	<0.001

Results for Population-Average Model

The value of the log-likelihood function at iteration 2 = -1.010987E+004

Final estimation of fixed effects: (Population-average model)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-1.748402	0.087969	-19.875	354	<0.001
MSESC, γ_{01}	-0.283620	0.185179	-1.532	354	0.127
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.446546	0.066993	6.666	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.536378	0.088479	-6.062	7158	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	-1.748402	0.174052	(0.146,0.207)
MSESC, γ_{01}	-0.283620	0.753053	(0.523,1.084)
For MALE slope, β_1			
INTRCPT2, γ_{10}	0.446546	1.562905	(1.371,1.782)
For PPED slope, β_2			
INTRCPT2, γ_{20}	-0.536378	0.584863	(0.492,0.696)

Notice that the results for the population-average model are quite similar to the results for the unit-specific model except in the case of the intercept. The intercept in the population-average model in this case is the expected log-odds of repetition for a person with values of zero on the predictors (and therefore, for a female without pre-primary experience attending a school of average SES). In this case, this expected log-odds corresponds to a probability of $1/(1 + \exp\{1.748402\}) = .148$, which is the "population-average" repetition rate for this group. In contrast, the unit-specific intercept is the expected log-odds of repetition rate for the same kind of student, but one who attends a school that not only has a mean SES of 0, but also has a random effect of zero (that is, a school with a "typical" repetition rate for the school of its type). This conditional expected log-odds is -2.046961, corresponding to a probability of $1/(1 + \exp\{2.046961\}) = .114$. Thus the probability of repetition is lower in a school with a random effect of zero than the average in the population of schools having mean SES of zero taken as a whole. This is a typical result. Population-average probabilities will be closer to .50 (than will the corresponding unit-specific probabilities).

One final set of results is printed out: population-average results with robust standard errors (below). Note that the robust standard errors in this case are very similar to the model-based standard errors, with a slight increase for the level-2 predictor and slight decreases for level-1 predictors. Results for other data may not follow this pattern.

Final estimation of fixed effects
(Population-average model with robust standard errors)

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-1.748402	0.082158	-21.281	354	<0.001
MSESC, γ_{01}	-0.283620	0.196005	-1.447	354	0.149
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.446546	0.062788	7.112	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.536378	0.082221	-6.524	7158	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	-1.748402	0.174052	(0.148,0.205)
MSESC, γ_{01}	-0.283620	0.753053	(0.512,1.107)
For MALE slope, β_1			
INTRCPT2, γ_{10}	0.446546	1.562905	(1.382,1.768)
For PPED slope, β_2			
INTRCPT2, γ_{20}	-0.536378	0.584863	(0.498,0.687)

8.3 Case 2: a binomial model (number of trials, m_{ij} □1)

A familiar example of two-level binomial data is the number of hits, Y_{ij}, in game i for baseball player j based on m_{ij} at bats. In an experimental setting, a subject j under condition i might produce Y_{ij} successes in m_{ij} trials.

A common use of a binomial model is when analysts do not have access to the raw data at level 1. For example, one might know the proportion of children passing a criterion-referenced test within each of many schools. This proportion might be broken down within schools by sex and grade. A binomial model could be used to analyze such data. The cases would be sex-by-age "cells" within each school where Y_{ij} is the number passing within cell i of school j and m_{ij} is the number of "trials," that is, the number of children in that cell. Sex and grade would be level-1 predictors.

Indeed, in the previous example, although raw level-1 data were available, the two level-1 predictors, MALE and pre-primary experience, were categorical. For illustration, we reorganized these data so that each school had, potentially, four cells defined by the cross-classification of sex and pre-primary experience:

- females without pre-primary experience
- females with pre-primary experience
- males without pre-primary experience
- males with pre-primary experience

Level-1 predictors were the same as before, with MALE = 1 if male, 0 if female; PPED = 1 if pre-primary experience, 0 if not. The outcome is the number of children in a particular cell who repeated a grade, and we created a variable TRIAL, which is the number of children in each cell. In some schools there were no children of a certain type (*e.g.*, no females with pre-primary experience). Such schools would have fewer than four cells. The necessary steps for executing the analysis via the Windows interface are given below.

To specify a Binomial model

1. After specifying the outcome in the model specification window (REP1 in our example), click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1).
2. Select **Binomial (number of trials)**.
3. Select the variable from the pull down menu in the dialog box, which indicates number of trials (TRIAL in our example) (See Figure 8.1).
4. (Optional) Specify the maximum number of macro and micro iterations by selecting the **Iteration Settings** option from the **Other Settings** menu.
5. (Optional) Select the **Over-dispersion** option if appropriate (See section on *Additional Features* at the end of the chapter).

The model described above uses the same predictors at level-1 and level-2 as those in the Bernoulli example (see Figure 8.5). The command file for the example is THAIBNML.HLM.

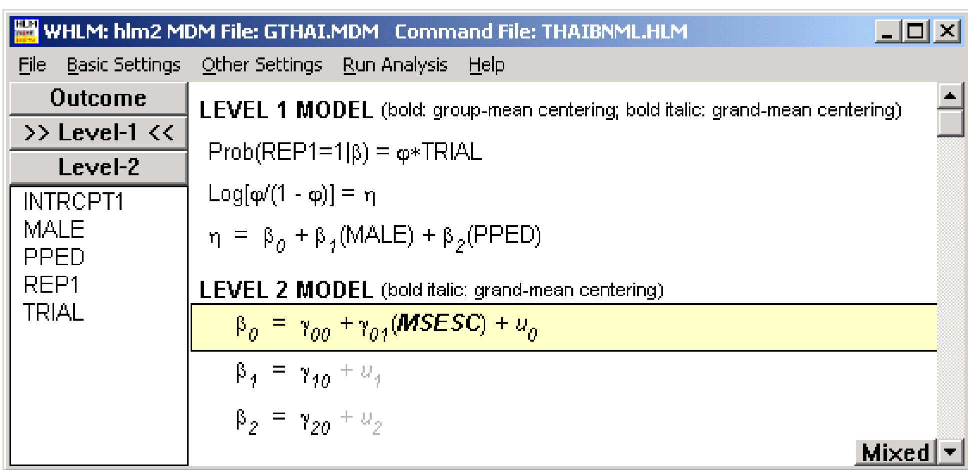


Figure 8.5 Model specification window for the Binomial model

Problem Title: BINOMIAL ANALYSIS, THAILAND DATA

The data source for this run = THAIGRP.MDM
The command file for this run = thaibnml.hlm
Output file name = thaibnml.out
The maximum number of level-1 units = 1097
The maximum number of level-2 units = 356
The maximum number of micro iterations = 50
Method of estimation: restricted PQL
Maximum number of macro iterations = 50
Distribution at Level-1: Binomial

Summary of the model specified

Level-1 Model

$E(\text{REP1}_{ij}=1|\beta_j) = \phi_{ij}$ * TRIAL $_{ij}$

$\log[\phi_{ij}/(1 - \phi_{ij})] = \eta_{ij}$

$\eta_{ij} = \beta_{0j} + \beta_{1j}*(\text{MALE}_{ij}) + \beta_{2j}*(\text{PPED}_{ij})$

This is the program's way of saying that the level-1 sampling model is binomial with “TRIAL” indicating the number of trials, so that the above equation, written with subscripts and Greek letters, is

$$E(Y_{ij} \mid \beta_j) = m_{ij}\phi_{ij}$$
$$Var(Y_{ij} \mid \beta_j) = m_{ij}\phi_{ij}(1-\phi_{ij}),$$

where m_{ij} = TRIAL.

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(\text{MSESC}_j) + u_{0j}$
$\beta_{1j} = \gamma_{10}$
$\beta_{2j} = \gamma_{20}$

MSESC has been centered around the grand mean**.**

Notice that the level-1 and level-2 structural models are identical to those in Case 1.

Level-1 variance = 1/[TRIAL* ϕ_{ij} (1-ϕ_{ij})]

In the metric of the linearized dependent variable, the level-1 variance is the reciprocal of the binomial variance,

$$m_{ij}\phi_{ij}(1-\phi_{ij}).$$

Results for the unit-specific model, population-average model, and population-average model with robust standard errors, are not printed below. They are essentially identical to the results using the Bernoulli model.

8.4 Case 3: Poisson model with equal exposure

Suppose that the outcome variable in Case 1 had been the number of days absent during the previous year rather than grade repetition. This outcome would be a non-negative integer, that is, a count rather than a dichotomy. Thus, the Poisson model with a log link would be a reasonable choice for the model. Notice that the time interval during which the absences could accumulate, that is, one year, would be the same for each student. We call this a case of “equal exposure,” meaning that each level-1 case had an “equal opportunity” to accumulate absences. (Case 4 describes an example where exposure varies across level-1 cases.)

This model has exactly the same logic as in Case 1 except that the type of model and therefore the corresponding link function will be different.

To specify a Poisson model with equal exposure

1. After specifying the outcome in the model specification window (REP1 in our example), click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1).
2. Select **Poisson (constant exposure)** to tell HLM that the level-1 sampling model is Poisson with equal exposure per level-1 case.
3. (Optional) Specify the maximum number of macro and micro iterations by selecting the **Iteration Settings** option from the **Other Settings** menu.
4. (Optional) Select the **Over-dispersion** option if appropriate (See section on *Additional Features* at the end of the chapter).

The HLM output would describe the model as follows

Level-1 Model

$E(\text{REP1}_{ij}|\beta_j) = \lambda_{ij}$
$log[\lambda_{ij}] = \eta_{ij}$

The above equation, written with subscripts and Greek letters, is

$$E(Y_{ij} \mid \beta_j) = \lambda_{ij}$$
$$Var(Y_{ij} \mid \beta_j) = \lambda_{ij}$$

where λ_{ij} is the "true" rate of absence for child *ij*.

$\eta_{ij} = \beta_{0j} + \beta_{1j}*(\text{MALE}_{ij}) + \beta_{2j}*(\text{PPED}_{ij})$

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(\text{MSESC}_j) + u_{0j}$
$\beta_{1j} = \gamma_{10}$
$\beta_{2j} = \gamma_{20}$

MSESC has been centered around the grand mean.

Notice that the log link replaces the logit link when we have count data. In the example above, β_2 is the expected difference in log-absenteeism between two children of the same sex attending the same school. To translate back to the rate of absenteeism, we would expect a child with pre-primary experience to have exp $\{\beta_2\}$ times the absenteeism rate of a child attending the same school who did not have pre-primary experience (holding sex constant). In this particular case, the estimated effect for β_2 is most plausibly negative; exp $\{\beta_2\}$ is less than 1.0 so that pre-primary experience would reduce the rate of absenteeism. Notice that the level-2 structural models are identical to those in Case 1.

Notice that the level-1 and level-2 structural models are identical to those in Case 1.

Level-1 variance = $1/\lambda_{ij}$

In the metric of the linearized dependent variable, the level-1 variance is the reciprocal of the Poisson variance, λ_{ij}.

8.5 Case 4: Poisson model with variable exposure

Suppose that the frequency of a given kind of cancer were tabulated for each of many counties. For example, with five age-groups, the data could be organized so that each county had five counts, with Y_{ij} being the number of cancers in age-group *i* of county *j* and m_{ij} being the population size of that age group in that county. A Poisson model with variable exposure would be appropriate, with m_{ij} the variable measuring exposure.

To specify a Poisson model with variable exposure

1. After specifying the outcome in the model specification window (REP1 in our example), click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1).
2. Select **Poisson (variable exposure)** to tell HLM that the level-1 sampling model is Poisson with variable exposure per level-1 case.
3. Select the variable that indicates variable exposure from the drop-down list box (See Figure 8.1). (In the illustration below, we use TRIAL as the variable to indicate variable exposure).
4. (Optional) Specify the maximum number of macro and micro iterations by selecting the **Iteration Settings** option from the **Other Settings** menu.
5. (Optional) Select the **Over-dispersion** option if appropriate (See section on *Additional Features* at the end of the chapter).

The HLM output would describe the model as follows:

Level-1 Model

$E(\text{REP1}_{ij}|\beta_j) = \lambda_{ij}{*}\ \text{TRIAL}_{ij}$
$log[\lambda_{ij}] = \eta_{ij}$
$\eta_{ij} = \beta_{0j} + \beta_{1j}{*}(\text{MALE}_{ij}) + \beta_{2j}{*}(\text{PPED}_{ij})$

This is the program's way of saying that the level-1 sampling model is Poisson with variable exposure per level-1 case, so that the above equation, written with subscripts and Greek letters, is

$$E(Y_{ij} \mid \beta_j) = m_{ij}\lambda_{ij}$$
$$Var(Y_{ij} \mid \beta_j) = m_{ij}\lambda_{ij},$$

Notice that the log link replaces the logit link when we have count data.

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}$*(MSESC$_j$) + u_{0j}
$\beta_{1j} = \gamma_{10}$
$\beta_{2j} = \gamma_{20}$

Notice that the level-1 and level-2 structural models are identical to those in Case 1.

Level-1 variance = 1/(TRIAL*λ_{ij})

In the metric of the linearized dependent variable, the level-1 variance is the reciprocal of the Poisson variance, $m_{ij}\lambda_{ij}$.

8.6 Case 5: Multinomial model

Data are from a 1990 survey of teachers in 16 high schools in California and Michigan. In the MDM file, not included with the software, there are a total of 650 teachers. The level-1 SPSS input file is TCHR1.SAV, and the level-2 file is TCHR2.SAV.

An outcome with three response categories tapping teachers' commitment to their career choice is derived from teachers' responses to the hypothetical question of whether they would become a teacher if they could go back to college and start over again. The possible responses are:

- yes, I would choose teaching again
- not sure
- no, I would not choose teaching again.

At the teacher level, it is hypothesized that teachers' perception of task variety is positively associated with greater odds of a teacher choosing the first category relative to the third category, and with greater odds of a teacher choosing the second category relative to the third category. The perception is measured by a task variety scale that assessed the extent to which teachers followed the same teaching routines each day, performed the same tasks each day, had something new happening in their job each day, and liked the variety present in their work (Rowan, Raudenbush & Cheong, 1993).

At the school level, it is postulated that the extent of teacher control has the same relationship to the two log odds as perception of task variety does. The teacher control scale is constructed by aggregating nine-item scale scores of teachers within a school. This scale indicates teacher control over school policy issues such as student behavior codes, content of in-service programs, student grouping, school curriculum, and text selection; and control over classroom issues such as teaching content and techniques, and amount of homework assigned (Rowan, Raudenbush & Kang, 1991).

As a previous analysis showed that there is little between-teacher variability in their log-odds of choosing the second category relative to the third category, the level-1 coefficient associated with it is fixed. Furthermore, the effects associated with perception of task variety are constrained to be the same across teachers for the sake of parsimony.

The general procedure to specify a multinomial logit model is given below. Note that the multinomial and ordinal analyses provide unit-specific estimates only. They do not currently produce population-average estimates.

To specify a multinomial model

1. After specifying the outcome in the model specification window, click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1).
2. Select **Multinomial** to tell HLM that the level-1 sampling model is multinomial.
3. Enter the number of categories into the **Number of Categories** box.
4. (Optional) Specify the maximum number of macro and micro iterations by selecting the **Iteration Settings** option from the **Other Settings** menu.

Figure 8.6 displays the model discussed above.

The output obtained for this model follows.

```
Specifications for this multinomial HLM run

Problem Title: Multinomial Output, High School Context Data

The data source for this run  = tchr.MDM
The command file for this run = tchr1.hlm
Output file name              = tchr1.html
The maximum number of level-1 units = 650
The maximum number of level-2 units = 16
The maximum number of micro iterations = 14
Number of categories = 3
Distribution at Level-1: Multinomial

The outcome variable is TCOMMIT
```

Summary of the model specified

Level-1 Model

$Prob[\text{TCOMMIT}(1) = 1|\beta_j] = \phi_{1_{ij}}$

$Prob[\text{TCOMMIT}(2) = 1|\beta_j] = \phi_{1_{ij}}$

$Prob[\text{TCOMMIT}(3) = 1|\beta_j] = \phi_{1_{ij}} = 1 - \phi_{1_{ij}} - \phi_{1_{ij}}$

$log[\phi_{1_{ij}} / \phi_{1_{ij}}] = \beta_{0j(1)} + \beta_{1j(1)}$*(TASKVAR$_{ij}$)

$log[\phi_{1_{ij}} / \phi_{1_{ij}}] = \beta_{0j(2)} + \beta_{1j(2)}$*(TASKVAR$_{ij}$)

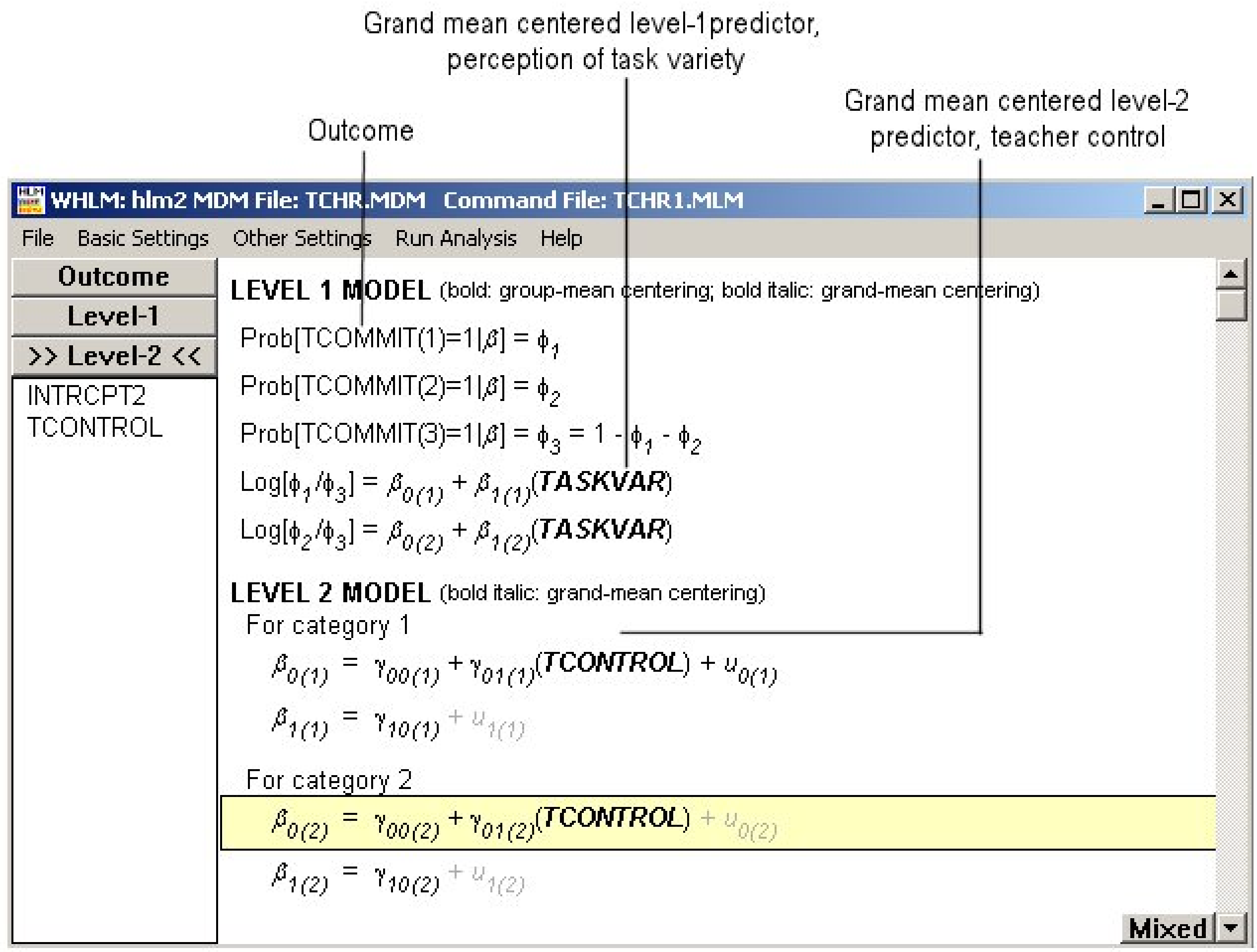


Figure 8.6 Model specification window for the multinomial example

Thus, the level-1 structural models are

$$\eta_{ij(1)} = \log\left[\frac{\phi_{ij(1)}}{\phi_{ij(3)}}\right] = \beta_{0j(1)} + \beta_{1j(1)}(\text{TASKVAR})_{ij}$$

$$\eta_{ij(2)} = \log\left[\frac{\phi_{ij(2)}}{\phi_{ij(3)}}\right] = \beta_{0j(2)} + \beta_{1j(2)}(\text{TASKVAR})_{ij}$$

Level-2 Model

$\beta_{0(1)} = \gamma_{00(1)} + \gamma_{01(1)}$*(TCONTROL$_j$) + $u_{0j(1)}$
$\beta_{1(1)} = \gamma_{10(1)}$
$\beta_{0(2)} = \gamma_{00(2)} + \gamma_{01(2)}$*(TCONTROL$_j$)
$\beta_{1(2)} = \gamma_{10(2)}$

TASKVAR has been centered around the grand mean.
TCONTROL has been centered around the grand mean.

The level-2 structural models are

$$\beta_{0j(1)} = \gamma_{00(1)} + \gamma_{01(1)}(\text{TCONTROL})_{ij} + u_{0j(1)}$$
$$\beta_{1j(1)} = \gamma_{10(1)}$$
$$\beta_{0j(2)} = \gamma_{00(2)} + \gamma_{01(2)}(\text{TCONTROL})_{ij}$$
$$\beta_{1j(2)} = \gamma_{10(2)}$$

τ

INTRCPT1(1) 0.00986

Random level-1 coefficient	Reliability estimate
INTRCPT1(1), $\beta_{0(1)}$	0.083

The value of the log-likelihood function at iteration 2 = -1.246191E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For Category 1					
For INTRCPT1, $\beta_{0(1)}$					
INTRCPT2, $\gamma_{00(1)}$	1.079269	0.123439	8.743	14	<0.001
TCONTROL, $\gamma_{01(1)}$	2.090207	0.508369	4.112	14	0.001
For TASKVAR slope, $\beta_{1(1)}$					
INTRCPT2, $\gamma_{10(1)}$	0.398355	0.113650	3.505	630	<0.001
For Category 2					
For INTRCPT1, $\beta_{0(2)}$					
INTRCPT2, $\gamma_{00(2)}$	0.091930	0.141643	0.649	630	0.517
TCONTROL, $\gamma_{01(2)}$	1.057285	0.577673	1.830	630	0.068
For TASKVAR slope, $\beta_{1(2)}$					
INTRCPT2, $\gamma_{10(2)}$	0.030693	0.130029	0.236	630	0.813

$\gamma_{00(1)}$, the unit-specific intercept, is the expected log-odds of an affirmative response relative to a negative response for a teacher with mean perception of task variety and working in a school with average teacher control and a random effect of zero. It is adjusted for the between-school heterogeneity in the likelihood of an affirmative response relative to a negative response, which is independent of the effect of task variety and teacher control. The estimated conditional expected log-odds is 1.079269.

The predicted probability that the same teacher responds affirmatively (Category 1) is exp{1.079269}/(1 + exp{1.079269} + exp{0.091930}) = .584. The predicted probability of responding "not sure" (category 2) is exp{0.091930}/(1 + exp{1.079269} + exp{0.091930}) = 1 - .584 - .218 = .198.

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For Category 1			
For INTRCPT1, $\beta_{0(1)}$			
INTRCPT2, $\gamma_{00(1)}$	1.079269	2.942528	(2.258,3.835)
TCONTROL, $\gamma_{01(1)}$	2.090207	8.086586	(2.718,24.063)
For TASKVAR slope, $\beta_{1(1)}$			
INTRCPT2, $\gamma_{10(1)}$	0.398355	1.489373	(1.191,1.862)
For Category 2			
For INTRCPT1, $\beta_{0(2)}$			
INTRCPT2, $\gamma_{00(2)}$	0.091930	1.096288	(0.830,1.448)
TCONTROL, $\gamma_{01(2)}$	1.057285	2.878545	(0.926,8.952)
For TASKVAR slope, $\beta_{1(2)}$			
INTRCPT2, $\gamma_{10(2)}$	0.030693	1.031169	(0.799,1.331)

The sets of γ_{01} and γ_{10} give the estimates of the change in the respective log-odds given one-unit change in the predictors, holding all other variables constant. For instance, all else being equal, a standard deviation increase in TCONTROL (.32) will nearly double the odds of an affirmative response to a negative response ($\exp\{2.090207 * .32\} = 1.952$). Note that the partial effect associated with perception of task variety is statistically significant for the logit of affirmative versus negative responses but not for the logit of undecided versus negative responses.

Below is a table for the results for the fixed effects with robust standard errors.

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For Category 1					
For INTRCPT1, $\beta_{0(1)}$					
INTRCPT2, $\gamma_{00(1)}$	1.079269	0.128263	8.415	14	<0.001
TCONTROL, $\gamma_{01(1)}$	2.090207	0.409607	5.103	14	<0.001
For TASKVAR slope, $\beta_{1(1)}$					
INTRCPT2, $\gamma_{10(1)}$	0.398355	0.127511	3.124	630	0.002
For Category 2					
For INTRCPT1, $\beta_{0(2)}$					
INTRCPT2, $\gamma_{00(2)}$	0.091930	0.139637	0.658	630	0.511
TCONTROL, $\gamma_{01(2)}$	1.057285	0.529606	1.996	630	0.046
For TASKVAR slope, $\beta_{1(2)}$					
INTRCPT2, $\gamma_{10(2)}$	0.030693	0.126446	0.243	630	0.808

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For Category 1			
For INTRCPT1, $\beta_{0(1)}$			
INTRCPT2, $\gamma_{00(1)}$	1.079269	2.942528	(2.235,3.874)
TCONTROL, $\gamma_{01(1)}$	2.090207	8.086586	(3.359,19.469)
For TASKVAR slope, $\beta_{1(1)}$			
INTRCPT2, $\gamma_{10(1)}$	0.398355	1.489373	(1.159,1.913)
For Category 2			
For INTRCPT1, $\beta_{0(2)}$			
INTRCPT2, $\gamma_{00(2)}$	0.091930	1.096288	(0.833,1.442)
TCONTROL, $\gamma_{01(2)}$	1.057285	2.878545	(1.017,8.145)
For TASKVAR slope, $\beta_{1(2)}$			
INTRCPT2, $\gamma_{10(2)}$	0.030693	1.031169	(0.804,1.322)

The robust standard errors are appropriate for datasets having a moderate to large number of level 2 units. These data do not meet this criterion.

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1(1), $u_{0(1)}$	0.09931	0.00986	14	16.16473	0.303

Note that the residual variance of $\beta_{00(1)}$ is not statistically different from zero. The model may be re-run with the coefficient set to be non-random.

8.7 Case 6: Ordinal model

The same data set, the multi-category outcome, and the same predictors in Case 5 are used here. The procedure for specifying an ordinal model is very similar to that of a multinomial model. Select the **Ordinal** instead of **Multinomial** option in the **Basic Model Specifications – HLM2** dialog box (See Figure 8.1). Figure 8.7 displays the model specified for the example (TCHR2.HLM).

Note: The multinomial and ordinal analyses currently produce unit-specific results only. They do not provide population-average results.

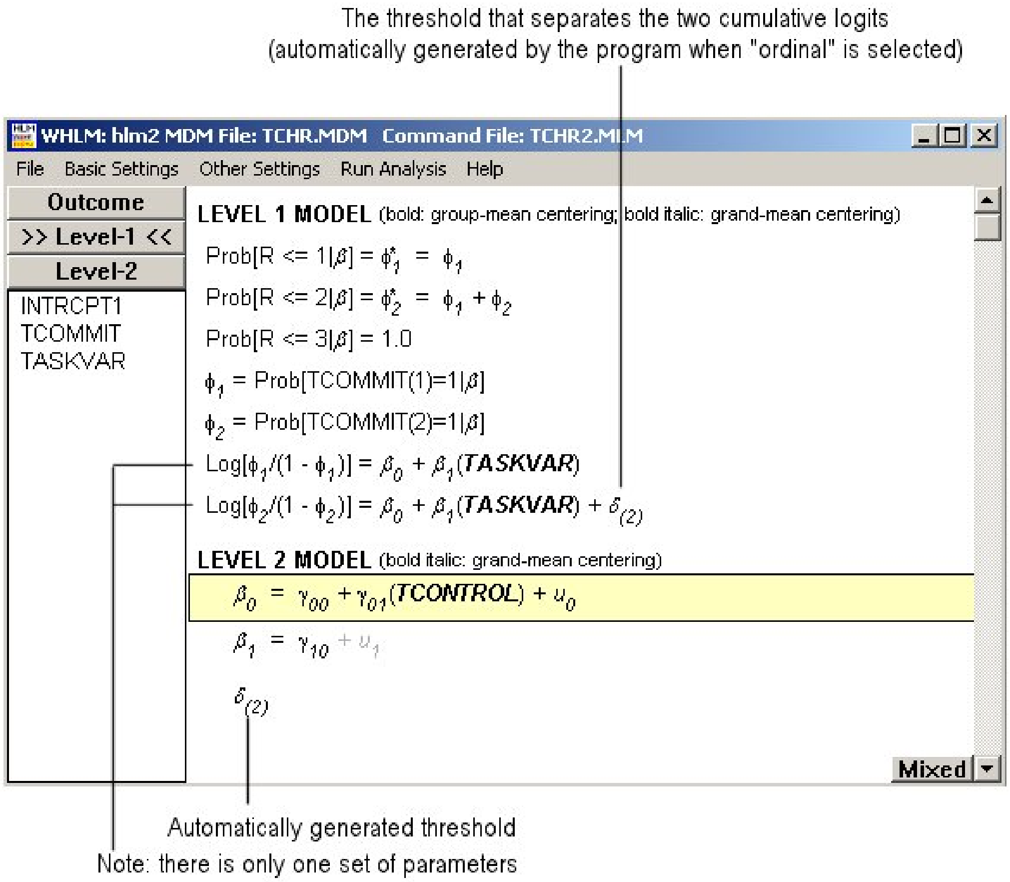


Figure 8.7 Model specification window for the ordinal model

The output obtained for this model follows.

Specifications for this ordinal HLM run

Problem Title: Ordinal Output, HIGH SCHOOL CONTEXT DATA

The data source for this run = TCHR.MDM
The command file for this run = TCHR2.HLM
Output file name = TCHR2.HTML
The maximum number of level-1 units = 650
The maximum number of level-2 units = 16
The maximum number of micro iterations = 14
Number of categories = 3
Method of estimation: restricted PQL

Distribution at Level-1: Ordinal

The outcome variable is TCOMMIT

Summary of the model specified

Level-1 Model

$Prob[R_{ij} <= 1|\beta_j] = \phi^*_{1ij} = \phi_{1ij}$

$Prob[R_{ij} <= 2|\beta_j] = \phi^*_{2ij} = \phi_{1ij} + \phi_{2ij}$

$Prob[R_{ij} <= 3|\beta_j] = 1.0$

ϕ_{1ij} = *Prob*[TCOMMIT(*1*) = *1*|β_j]

ϕ_{2ij} = *Prob*[TCOMMIT(2) = 1|β_j]

log[ϕ_{1ij} /(1 - ϕ_{1ij})] = β_{0j} + β_{1j}*(TASKVAR$_{ij}$)

log[ϕ_{2ij} /(1 - ϕ_{2ij})] = β_{0j} + β_{1j}*(TASKVAR$_{ij}$) + δ_2

Thus, the level-1 structural models are

$$\eta'_{ij(1)} = \left[\frac{\phi'_{ij(1)}}{1-\phi'_{ij(1)}}\right] = \beta_{0j} + \beta_{1j}(\text{TASKVAR})_{ij}$$

$$\eta'_{ij(2)} = \left[\frac{\phi'_{ij(2)}}{1-\phi'_{ij(2)}}\right] = \beta_{0j} + \beta_{1j}(\text{TASKVAR})_{ij} + \delta_{(2)}.$$

Level-2 Model

β_{0j} = γ_{00} + γ_{01}*(TCONTROL$_j$) + u_{0j}
β_{1j} = γ_{10}
δ_2

TASKVAR has been centered around the grand mean.
TCONTROL has been centered around the grand mean.

The level-2 structural model is

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{TCONTROL})_{ij} + u_{0j}$$
$$\beta_{1j} = \gamma_{10}$$

Final Results for Ordinal Iteration 9173

The extremely large number of iterations reflects the fact that the final estimate of the between-school variance, τ_{00}, is near zero, after adjusting for TCONTROL.

τ

INTRCPT1,β_0 0.00010

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.001

The value of the log-likelihood function at iteration 2 = -1.249070E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
For INTRCPT1 slope, β_0					
INTRCPT2, γ_{00}	0.333918	0.089735	3.721	14	0.002
TCONTROL, γ_{01}	1.541051	0.365624	4.215	14	<0.001
For TASKVAR slope, β_1					
INTRCPT2, γ_{10}	0.348801	0.087280	3.996	633	<0.001
For THOLD2,					
δ_2	1.054888	0.080868	13.045	633	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1 slope, β_0			
INTRCPT2, γ_{00}	0.333918	1.396429	(1.152,1.693)
TCONTROL, γ_{01}	1.541051	4.669496	(2.131,10.230)
For TASKVAR slope, β_1			
INTRCPT2, γ_{10}	0.348801	1.417367	(1.194,1.682)
For THOLD2,			
δ_2	1.054888	2.871653	(2.450,3.366)

γ_{00}, the unit-specific intercept, is the expected log-odds of an affirmative response relative to an undecided or negative response for a teacher with mean perception of task variety and working in a school with average teacher control and a random effect of zero. It is adjusted for the between-school heterogeneity in the likelihood of an affirmative response relative to a negative response, which is independent of the effect of task variety and teacher control. This conditional expected log-odds, is 0.333918. The expected log-odds for a teacher to give an affirmative or undecided response relative to a negative response is 0.333918 + 1.054888 = 1.388806. γ_{01} and γ_{10} give the estimates of the change in the respective cumulative logits, holding all other variables constant. For instance, all else being equal, a standard deviation increase in TCONTROL (.32) will increase the odds of an affirmative response to an undecided or negative response as well as the odds of an affirmative or undecided response to a negative response by a factor of 1.637 ($\exp\{1.541051 * .32\} = 1.637$).

Below is a table for the results for the fixed effects with robust standard errors.

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
For INTRCPT1 slope, β_0					
INTRCPT2, γ_{00}	0.333918	0.092707	3.602	14	0.003
TCONTROL, γ_{01}	1.541051	0.340944	4.520	14	<0.001
For TASKVAR slope, β_1					
INTRCPT2, γ_{10}	0.348801	0.092285	3.780	633	<0.001
For THOLD2,					
δ_2	1.054888	0.080353	13.128	633	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1 slope, β_0			
INTRCPT2, γ_{00}	0.333918	1.396429	(1.145,1.704)
TCONTROL, γ_{01}	1.541051	4.669496	(2.247,9.702)
For TASKVAR slope, β_1			
INTRCPT2, γ_{10}	0.348801	1.417367	(1.182,1.699)
For THOLD2,			
δ_2	1.054888	2.871653	(2.452,3.363)

The robust standard errors are appropriate for datasets having a moderate to large number of level 2 units. These data do not meet this criterion.

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, u_0	0.01016	0.00010	14	14.57034	0.408

Note that the residual variance of $\beta_{00(1)}$ is not statistically different from zero. In fact, it is very close to zero, which accounts for the large number of iterations required to achieve convergence. The model may be re-run with the coefficient set to be non-random.

8.8 Additional features

8.8.1 Over-dispersion

For binomial models with $m_{ij} > 1$ and for all Poisson models, there is an option to estimate a level-1 dispersion parameter σ^2 (See Figure 8.1). If the assumption of no dispersion holds, $\sigma^2 = 1.0$. If the data are over-dispersed, $\sigma^2 > 1.0$; if the data are under-dispersed, $\sigma^2 < 1.0$.

8.8.2 Adaptive Gauss-Hermite Quadrature and Laplace approximations for binary models

For two- and three-level binary outcome models, the highly accurate approximations to maximum likelihood based on adaptive Gauss-Hermite Quadrature and Laplace approximation (See Figure 8.1) can be selected. When estimating the model parameters, the program will send messages, similar to the following, to the iteration window during computation of the results.

The following is an example of an output for Laplace6 iterations.

Results for Unit-Specific Model, EM Laplace-2 Estimation
Iteration 33

τ

INTRCPT1,β_0 1.61733

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.724

The log-likelihood at EM Laplace-2 iteration 10 is -9.627480E+003

Final estimation of fixed effects (Unit-specific model)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-2.239223	0.100384	-22.307	354	<0.001
MSESC, γ_{01}	-0.297322	0.200573	-1.482	354	0.139
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.533635	0.072623	7.348	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.626218	0.099789	-6.275	7158	<0.001

Fixed Effect	Coefficient	Odds Ratio	Confidence Interval
For INTRCPT1, β_0			
INTRCPT2, γ_{00}	-2.239223	0.106541	(0.087,0.130)
MSESC, γ_{01}	-0.297322	0.742805	(0.501,1.102)
For MALE slope, β_1			
INTRCPT2, γ_{10}	0.533635	1.705119	(1.479,1.966)
For PPED slope, β_2			
INTRCPT2, γ_{20}	-0.626218	0.534610	(0.440,0.650)

Statistics for the current model

Deviance = 19254.960974
Number of estimated parameters = 5

Results for Unit-Specific Model, Adaptive Gaussian Quadrature Iteration 3

τ
INTRCPT1,β_0 1.68320

Standard error of τ
INTRCPT1,β_0 0.20904

Final estimation of fixed effects (Unit-specific model)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	-2.242961	0.106249	-21.110	354	<0.001
MSESC, γ_{01}	-0.295119	0.215888	-1.367	354	0.172
For MALE slope, β_1					
INTRCPT2, γ_{10}	0.535156	0.075975	7.044	7158	<0.001
For PPED slope, β_2					
INTRCPT2, γ_{20}	-0.626872	0.100135	-6.260	7158	<0.001

Statistics for the current model

Deviance = 19255.057516
Number of estimated parameters = 5

8.8.3 Printing variance-covariance matrices for fixed effects

Files containing variance-covariances for the fixed effects for the unit-specific, population-averaged and Laplace and adaptive Gaussian quadrature estimates can be requested. See Appendix A for more details, or Appendix J for a complete list of options available in each of the modules.

8.9 Fitting HGLMs with three and four levels

For simplicity of exposition, all of the examples above have used the two-level HGLM. These procedures generalize directly to three-and four-level applications. Again the type of *nonlinear* model desired at level-1 must be specified. There are now, however, structural models at both levels 2 and 3 as in the case of HLM3. The same idea applies to HLM4.

9 Conceptual and Statistical Background for Hierarchical Multivariate Linear Models (HMLM)

One of the most frequent applications of hierarchical models involves repeated observations (level 1) nested within persons (level 2). These are described in Chapter 6 of *Hierarchical Linear Models*. In these models, the outcome Y_{ij} for occasion i within person j is conceived as a univariate outcome, observed under different conditions or at different times. An advantage of viewing the repeated observations as nested within the person is that it allows each person to have a different repeated measures design. For example, in a longitudinal study, the number of time points may vary across persons, and the spacing between time points may be different for different persons. Such unbalanced designs would pose problems for standard methods of analysis such as the analysis of variance.

Suppose, however, that the aim of the study is to observe every participant according to a fixed design with, say, T observations per person. The design might involve T observation times or T different outcome variables or even T different experimental conditions. Given the fixed design, the analysis can be reconceived as a multivariate repeated measures analysis. The multivariate model is flexible in allowing a wide variety of assumptions about the variation and covariation of the T repeated measures (Bock, 1985). In the standard application of multivariate repeated measures, there can be no missing outcomes: every participant must have a full complement of T repeated observations.

Advances in statistical computation, beginning with the EM algorithm (Dempster, Laird, & Rubin, 1977; see also Jennrich & Schluchter, 1986), allow the estimation of multivariate normal models from incomplete data. In this case, the aim of the study was to collect T observations per person, but only n_j observations were collected ($n_j \leq T$). These n_j observations are indeed collected according to a fixed design, but $T - n_j$ data points are missing at random.

HMLM allows estimation of multivariate normal models from incomplete data; HMLM2 allows for study of multivariate outcomes for persons who are, in turn, nested within higher-level units. Within the framework of HMLM, it is possible to estimate models having

0. An unrestricted covariance structure, that is a full $T \times T$ covariance matrix.
1. A model with homogenous level-1 variance and random intercepts and/or slopes at level 2.
2. A model with heterogeneous variances at level 1 (a different variance for each occasion) and random intercepts and/or slopes at level 2.
3. A model that includes a log-linear structure for the level-1 variance and random intercepts and/or slopes at level 2.
4. A model with first-order auto-regressive level-1 random errors and random intercepts and/or slopes at level 2.

We note that applications 2 - 4 are available within the standard HLM2. However, within HMLM, models 2 - 4 can be compared to the unrestricted model (model 1), using a likelihood ratio test. No “unrestricted model“ can be meaningfully defined within the standard HLM2; such a model is definable only within the confines of a fixed design with T measurements.

HMLM2 allows the five models listed above to be embedded within a nested structure, *e.g.*, the persons who are repeatedly observed may be nested within schools.

9.1 Unrestricted model

This model is appropriate when the aim of the study is to collect T observations per participant according to a fixed design. However, one or more observations may be missing at random. We assume a constant but otherwise arbitrary $T \times T$ covariance matrix for each person's "complete data."

9.1.1 Level-1 model

The level-1 model relates the observed data, Y, to the complete data, Y^*:

$$Y_{hi} = \sum_{t=1}^{T} m_{thi} Y_{ti}^* \tag{0.0}$$

where Y_{hi} is the r-th outcome for person i associated with time h. Here Y_{ti}^* is the value that person i would have displayed if that person had been observed at time t, and m_{thi} is an indicator variable taking on a value of 1 if the h-th measurement for person i did occur at time t, 0 if not. Thus, Y_{ti}^*, $t = 1, ..., T$, represent the complete data for person i while Y_{hi}, $h = 1, ..., T_i$ are the observed data, and the indicators m_{thi} tell us the pattern of missing data for person i.

To make this clear, consider $T = 5$ and a person who has data at occasions 1,2, and 4, but not at occasions 3 and 5. Then Equation 9.1 expands to

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} Y_{1i}^* \\ Y_{2i}^* \\ Y_{3i}^* \\ Y_{4i}^* \\ Y_{5i}^* \end{pmatrix} \tag{0.0}$$

or, in matrix notation,

$$Y_i = M_i Y_i^* \tag{0.0}$$

This model says simply that the three observed data points for person i were observed at times 1, 2, and 4, so that data were missing at times 3 and 5. Although these data were missing, they do exist, in principle. Thus, every participant has a full 5×1 vector of "complete data" even though the $T_i \times 1$ vector of observed data will vary in length across persons.

We now pose a structural model for the within-person variation in Y^*:

$$Y_{ti}^{*} = \pi_{0i} + \sum_{p=1}^{P} \pi_{pi} a_{pt} + \varepsilon_{ti} \tag{0.0}$$

or, in matrix notation

$$Y_i^{*} = A\pi_i + \varepsilon_i, \tag{0.0}$$

where we assume that ε_i is multivariate normal in distribution with a mean vector of 0 and an arbitrary $T \times T$ covariance matrix Δ. In fact, Δ is not a "within-person" covariance. Rather, it captures all variation and covariation among the T repeated observations.

9.1.2 Level-2 model

The level-2 model includes covariates, X_i, that vary between persons:

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q} \beta_{pq} X_{qi} \tag{0.0}$$

or in matrix notation

$$\pi_i = X_i \beta \tag{0.0}$$

Note there is no random variation between persons in the regression coefficients π_{pi} because all random variation has been absorbed into Δ (see the text below Equation 9.5).

9.1.3 Combined model

Substituting the level-2 model into the level-1 model gives the combined model for the complete data, in matrix form:

$$Y_i^{*} = AX_i\beta + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \Delta) \tag{0.0}$$

Here the design matrix captures main effects of within-person covariates (the *a*s), main effects of person-level covariates (*X*s), and two-way interaction effects between them ($a \times X$ terms).

In sum, our reformulation poses a "multiple measures" model (Equation 9.3) that relates the observed data Y_i to the "complete data" Y_i^{*}, that is, the data that would have been observed if the researcher had been successful in obtaining outcome data at every time point. Our combined model is a standard multivariate normal regression model for the complete data.

Algebraically substituting the combined model expression for Y_i^{*} into the model for the observed data (Equation 9.3) yields the combined model

$$Y_i = M_i A X_i \beta + M_i \varepsilon_i . \tag{0.0}$$

Under the unrestricted model, the number of parameters estimated is $f + T(T+1)/2$, where f is the number of fixed effects and T is the number of observations intended for each person. The models below impose constraints on the unrestricted model, and therefore include fewer parameters. The fit of these simpler models to the data can be compared to the fit of the unrestricted model using a likelihood ratio test.

9.2 HLM with homogenous level-1 variance

Under the special case in which the within-person design is fixed[1], with T observations per person and randomly missing time points, the two-level HLM can be derived from the unrestricted model by imposing restrictions on the covariance matrix, Δ. (**Note:** regressors A_i having varying designs may be included in the level-1 model, but coefficients associated with such A_i values must not have random effects at level 2). The most frequently used assumption in the standard HLM is that the within-person residuals are independent with a constant variance, σ^2.

9.2.1 Level-1 model

The level-1 model has a similar form to that in the case of the unrestricted model

$$Y_i^* = A\pi_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \Sigma) \tag{0.0}$$

with $\Sigma = \sigma^2 I_T$.

9.2.2 Level-2 model

The level-2 model includes covariates, X_i, that vary between persons. Degrees of freedom are now available to estimate randomly varying intercepts and slopes across people:

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q_p} X_{qi}\beta_{pq} + r_{qi} \tag{0.0}$$

or in matrix notation

$$\pi_i = X_i \beta + r_i \tag{0.0}$$

All of the usual forms are now available for the intercepts and slopes (fixed, randomly varying, non-randomly varying), provided T is large enough.

9.2.3 Combined model

Substituting the level-2 model into the level-1 model gives the combined model for the complete data, in matrix form:

1 That is, $A_i = A$ for all i.

$$
\begin{aligned}
Y_i^* &= AX_i\beta + A_i r_i + e_i \\
&= AX_i\beta + \varepsilon_i,
\end{aligned} \tag{0.0}
$$

where $\varepsilon_i = Ar_i + e_i$ has variance-covariance matrix

$$
\begin{aligned}
Var(\varepsilon_i) &= Var(Ar_i + e_i) \\
&= A\tau A' + \sigma^2 I_T = \Delta.
\end{aligned} \tag{0.0}
$$

Under the HLM with homogenous level-1 variance, the number of parameters estimated is $f + r(r+1)/2 + 1$, where r is the dimension of τ. Thus, r must be less than T.

9.2 HLM with varying level-1 variance

One can model heterogeneity of level-1 variance as a function of the occasion of measurement. Such a model is suitable when we suspect that the level-1 residual variance varies across occasions. The models that can be estimated are a subset of the models that can be estimated within the standard HLM2 (see Section 2.8.8.2 on the option for heterogeneity of level-1 variance).The level-1 model is the same as in the case of homogenous variances (equations 9.11 and 9.12) except that now

$$
Var(e_i) = \Sigma = diag\left\{\sigma_t^2\right\}, \tag{0.0}
$$

that is, Σ is now diagonal with elements σ_t^2, the variance associated with occasion t, $t = 1, \ldots, T$.

The number of parameters estimated is $f + r(r+1)/2 + T$. Now r must be no larger than $T-1$. When $r = T-1$, the results will duplicate those based on the unrestricted model.

9.3 HLM with a log-linear model for the level-1 variance

The model with varying level-1 variance, described above, assumes a unique level-1 variance for every occasion. A more parsimonious model would specify a functional relationship between aspects of the occasion (*e.g.* time or age) and the variance. We would again have $\Sigma = diag\left\{\sigma_t^2\right\}$, but now

$$
\log\left(\sigma_t^2\right) = \alpha_0 + \sum_{I=1}^{L} \alpha_I c_{It}. \tag{0.0}
$$

Thus, the natural log of the level-1 variance may be a linear or quadratic function of age. If the explanatory variables c_I are $T-1$ dummy variables, each indicating the occasion of measurement, the results will duplicate those of the previous section.

The number of parameters estimated is now $f + r(r+1)/2 + L + 1$. Again, r must be no larger than $T-1$ and L must be no larger than $T-1$.

9.4 First-order auto-regressive model for the level-1 residuals

This model allows the level-1 residuals to be correlated under Markov assumptions (a level-1 residual depends on previous level-1 residuals only through the immediately preceding level-1 residuals). This leads to the level-1 covariance structure

$$Cov(e_{ti}, e_{t'i}) = \sigma^2 \rho^{|t-t'|}. \tag{0.0}$$

Thus, the variance at each time point is σ^2 and each correlation diminishes with the distance between time points, so that the correlations are $\rho, \rho^2, \rho^3, ...$ as the distance between occasions is 1, 2, 3, ... The number of parameters estimated is now $f + r(r+1)/2 + 2$. Again, r must be no larger than $T-1$.

Note that level-1 predictors are assumed to have the same values for all level-2 units of the complete data. This assumption can be relaxed. However, if the design for a_{pti} varies over *i*, its coefficient cannot vary randomly at level 2. In this regard, the standard 2-level model (See Chapters 2, 3) is more flexible than HMLM.

9.5 HMLM2: A multilevel, multivariate model

Suppose now that the persons yielding multiple outcomes are nested within higher-level units such as schools. We can embed the multivariate model for incomplete data within this multilevel structure.

9.5.1 Level-1 model

The level-1 model again relates the observed data, *Y*, to the complete data, Y^*. We simply add a subscript to the HMLM model to create the HMLM equation for the observed data:

$$Y_{hij} = \sum_{t=1}^{T} m_{thij} Y^*_{tij}. \tag{0.0}$$

Here individual *i* is nested within group *j* (*j* = 1, …, *J*) and we have Y_{hij}, the *h*-th outcome observed for person *i* in group *j*. Here Y^*_{tij} is the value that person *i* would have displayed if that person had been observed at time *t*, and m_{thij} is an indicator variable taking on a value of 1 if the *h*-th measurement for that person did occur at time *t*, 0 if not. Thus Y^*_{tij}, *t* = 1, …, *T* represent the complete data for person *i* in group *j* while Y_{hij}, *h* = 1, …, T_i are the observed data, and the indicators m_{thij} tell us the pattern of the missing data. Again, we pose a structural model for the within-person variation in Y^*:

$$Y^*_{tij} = \pi_{0ij} + \sum_{p=1}^{P} \pi_{pij} a_{pt} + e_{tij}, \tag{0.0}$$

or, in matrix notation

$$Y_{tj}^{*} = A\pi_{ij} + e_{ij}, \tag{0.0}$$

where we assume that e_{ij} is multivariate normal in distribution with a mean vector of 0 and an arbitrary $T \times T$ covariance matrix Σ.

9.5.2 The combined model

The level-2 model includes covariates, X_{ij}, that vary between persons within groups:

$$\pi_{pij} = \beta_{p0j} + \sum_{q=1}^{Q_{pq}} \beta_{pqj} X_{qij} \tag{0.0}$$

or, in matrix notation

$$\pi_{ij} = X_{ij}\beta_j \tag{0.0}$$

9.5.3 Level-3 model

Now the coefficients defined on persons (in the level-2 model) are specified as possibly varying at level-3 over groups:

$$\beta_{pqj} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} W_{sqj} + u_{pqj}. \tag{0.0}$$

Here the vector u_j, composed of elements u_{pqj} is multivariate normal in distribution with a zero mean vector and covariance matrix τ_β.

9.5.4 Level-2 model

The combined model can then be written in matrix notation as

$$Y_{ij}^{*} = AX_{ij}W_j\gamma + AX_{ij}u_j + \varepsilon_{ij}, \tag{0.0}$$

where

$$\varepsilon_{ij} = Ar_{ij} + e_{ij} \tag{0.0}$$

where ε_{ij} has a variance-covariance matrix

$$Var(\varepsilon_{ij}) = \Sigma \tag{0.0}$$

and Σ is modeled just as in the case of HMLM, depending on which submodel is of interest. The next chapter provides an illustration.

Note that level-1 predictors a_{pt} are assumed to have the same values for all level-2 units of the complete data. This assumption can be relaxed. However, if the design for a_{ptij} varies over i and j, the coefficient for a_{ptij}, that is π_{ptij}, must have no random effect at level 2. In this regard, the standard three-level model (see Chapters 3 and 4) is more flexible than is HMLM2.

10 Working with HMLM/HMLM2

Like the other programs, HMLM and HMLM2 execute analyses using MDM (multivariate data matrix) files, which consist of the combined level-1 and level-2 data files.

The procedures for constructing the MDM file are similar to the ones for HLM2 and HLM3 with one major difference: the user has to create and input indicator variables for the outcome(s) while constructing the MDM file. Model specification for HMLM and HMLM2 involves the same mechanics as in HLM2 and HLM3 with an extra step of model covariance structure selection.

Below we provide two examples using data sets from the first cohort of the National Youth Survey (Elliot, Huizinga, & Menard, 1989, Raudenbush, 1999) and the time-series observations on 1,721 students nested within 60 public primary schools as described in Chapter 8. Windows mode execution is illustrated. See Appendix E for interactive and batch mode execution.

10.1 An analysis using HMLM via Windows mode

10.1.1 Constructing the MDM from raw data

The range of options for data input are the same as for HLM2 and HLM3. We will use SPSS file input in our example.

10.1.1.1 Level-1 file

The level-1 file, NYS1.SAV, has 1,079 observations collected from interviewing annually 239 eleven-year-old youths beginning at 1976 for five consecutive years. Therefore, T = 5. The variables and the T indicator variables are:

ATTIT — a 9-item scale assessing attitudes favorable to deviant behavior.

Subjects were asked how wrong (very wrong, wrong, a little bit wrong, not wrong at all) they believe it is for someone their age to, for example, damage and destroy property, use marijuana, use alcohol, sell hard drugs, or steal.

The measure was positively skewed, so a logarithmic transformation was performed to reduce the skewness.

EXPO — Exposure to deviant peers.

Subjects were asked how wrong their best friends thought the nine deviant behaviors surveyed in the ATTIT scale were.

AGE — age of the participant
AGE11 — age of participant at a specific time minus 11
AGE13 — age of participant at a specific time minus 13
AGE11s — AGE11* AGE11

AGE13s	AGE13* AGE13
IND1	indicator for measure at time 1
IND2	indicator for measure at time 2
IND3	indicator for measure at time 3
IND4	indicator for measure at time 4
IND5	indicator for measure at time 5

The five indicators were created to facilitate use of HMLM. Data for the first two children are shown in Fig. 10.1.

Child 15 had data at all five years. Child 33, however, did not have data for the fourth year.

Indicators for the repeated measures

	id	attit	age	age11	age13	age11s	age13s	ind1	ind2	ind3	ind4	ind5
16	15	.44	11.00	.00	-2.00	.00	4.00	1.0	.00	.00	.00	.00
17	15	.44	12.00	1.00	-1.00	1.00	1.00	.00	1.0	.00	.00	.00
18	15	.89	13.00	2.00	.00	4.00	.00	.00	.00	1.0	.00	.00
19	15	.75	14.00	3.00	1.00	9.00	1.00	.00	.00	.00	1.0	.00
20	15	.80	15.00	4.00	2.00	16.00	4.00	.00	.00	.00	.00	1.0
21	33	.20	11.00	.00	-2.00	.00	4.00	1.0	.00	.00	.00	.00
22	33	.64	12.00	1.00	-1.00	1.00	1.00	.00	1.0	.00	.00	.00
23	33	.69	13.00	2.00	.00	4.00	.00	.00	.00	1.0	.00	.00
24	33	.11	15.00	4.00	2.00	16.00	4.00	.00	.00	.00	.00	1.0

Figure 10.1 Two children in the NYS1.SAV data set

10.1.1.2 Level-2 file

The level-2 data file, NYSB.SAV, consists of three variables on 239 youths. The file has the same structure as that for HLM2. The variables are:

FEMALE	an indicator for gender (1 = female, 0 = male)
MINORITY	an indicator for ethnicity (1 = minority, 0 = other)
INCOME	income

The construction of the MDM involves three major steps:

1. Select type of input data.
2. Supply the program with the appropriate data-defining information.
3. Check whether the data have been properly read into the program.

The steps are very similar to the ones described in Section 2.5.1. Select **HMLM** as the **MDM** type at the **Select MDM type** dialog box (see Figure 2.4) and inform WHLM the type of data input.

While the structure of HMLM input files is almost the same as in HLM2, there is one important difference: the indicator variables. In order to create these, one first needs to know *the maximum number of level-1 records per level-2 group; this determines the number of indicators*. We shall call them the number of "occasions." (This is the number of time points in a repeated measures study or the number of outcome variables in a cross-sectional multivariate study. Also note that each person does not need to have this number of occasions.) Then create the indicator variables so that a given variable takes on the value of 1.0 if the given occasion is at this time point, 0.0 otherwise. Looking at Figure 10.1, we see that IND1 is 1 if AGE11 is 0, IND2 is 1 if AGE11 is 1, IND3 is 1 if AGE11 is 2, and so on. Fig 10.2 shows the **Choose variables – HMLM** dialog box where the indicator variables are checked before the MDM file is created. This dialog box can be opened from the **Level-1 specification** section in the **Make MDM – HMLM** dialog box.

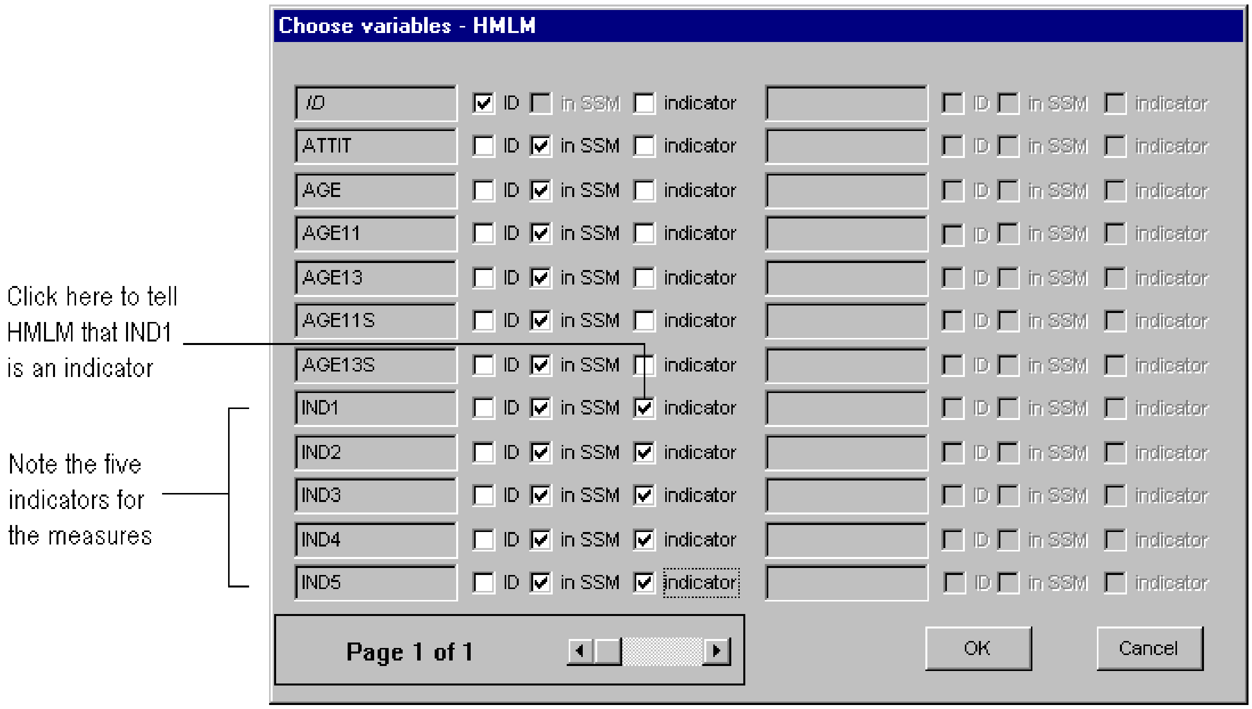


Figure 10.2 Choose variables – HMLM dialog box

10.2 Executing analyses based on the MDM file

The steps involved are similar to the ones for HLM2 as described on Section 2.5.2. It is necessary to specify

1. the level-1 model,
2. the level-2 structural model, and
3. the level-1 coefficients as random or non-random.

Under HMLM, level-1 predictors having random effects must have the same value for all participants at a given occasion. If the user specifies a predictor not fulfilling this condition to have a random effect, such coefficients will be automatically set as non-random by the program. Furthermore, an extra step for selecting the covariance structure for the models to be estimated is needed. Figure 10.3 displays the model specified for our example. Figure 10.4 shows the dialog box where the covariance structure is selected.

10.3 An annotated example of HMLM

In the example below (see NYS1.MLM) we specify AGE13 and AGE13S as predictors at level 1. At level 2, the model is unconditional. This is displayed in Fig. 10.3. We shall compare three alternative covariance structures:

- an unrestricted model,
- the homogeneous model, $\sigma_t^2 = \sigma^2$ for all t, and
- the heterogeneous model, which allows σ_t^2 to vary over time.

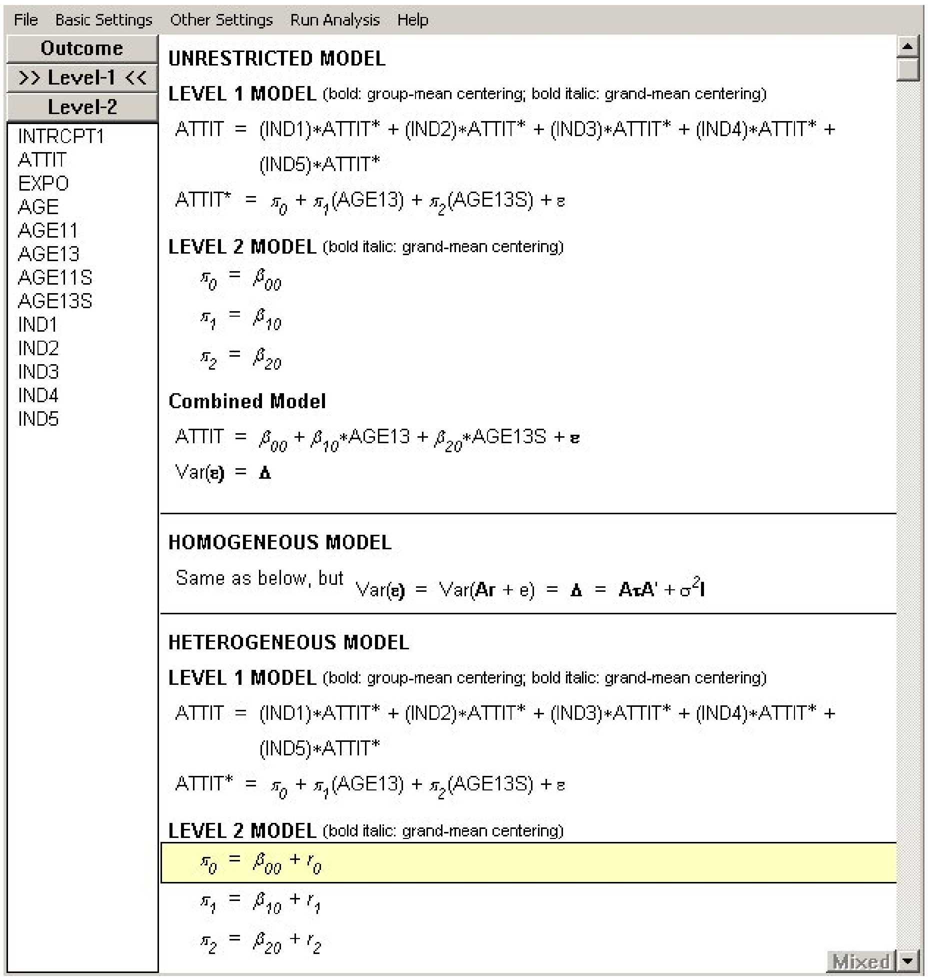


Figure 10.3 Model specification window for the NYS example

These three models are requested simply by checking the **Heterogeneous** option in the **Basic Model Specifications – HMLM** dialog box, as shown in Fig. 10.4.

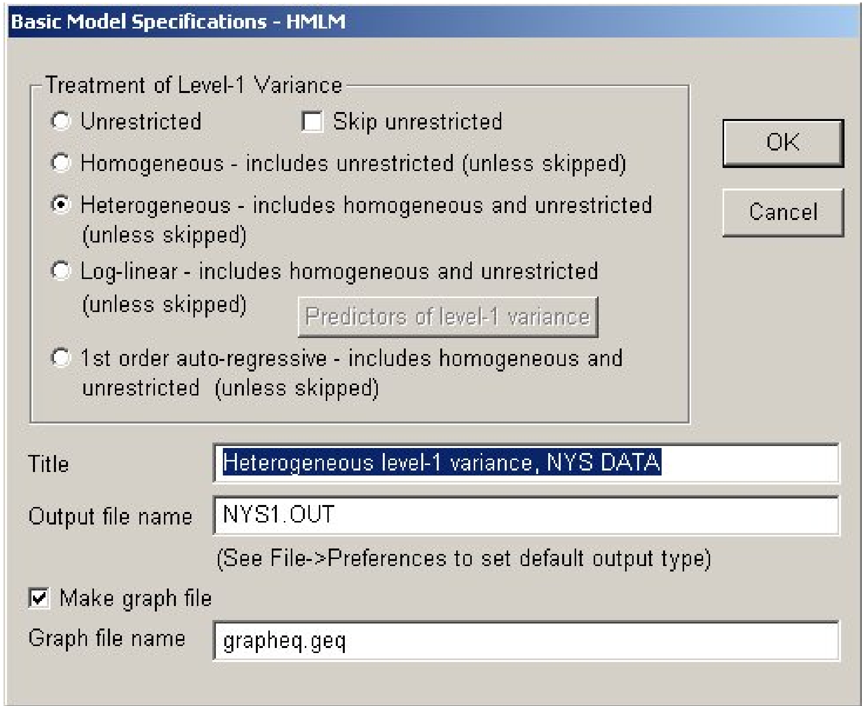


Figure 10.4 Basic Model Specifications - HMLM dialog box

Similarly, checking the **Log-linear** button will produce output on:

- the unrestricted model,
- the homogeneous model, and
- the log-linear model for σ_t^2.

In this case a modified model will be displayed, as shown in Fig. 10.5. To obtain this model, the **Predictors of level-1 variance** dialog box was used to select the variable EXPO.

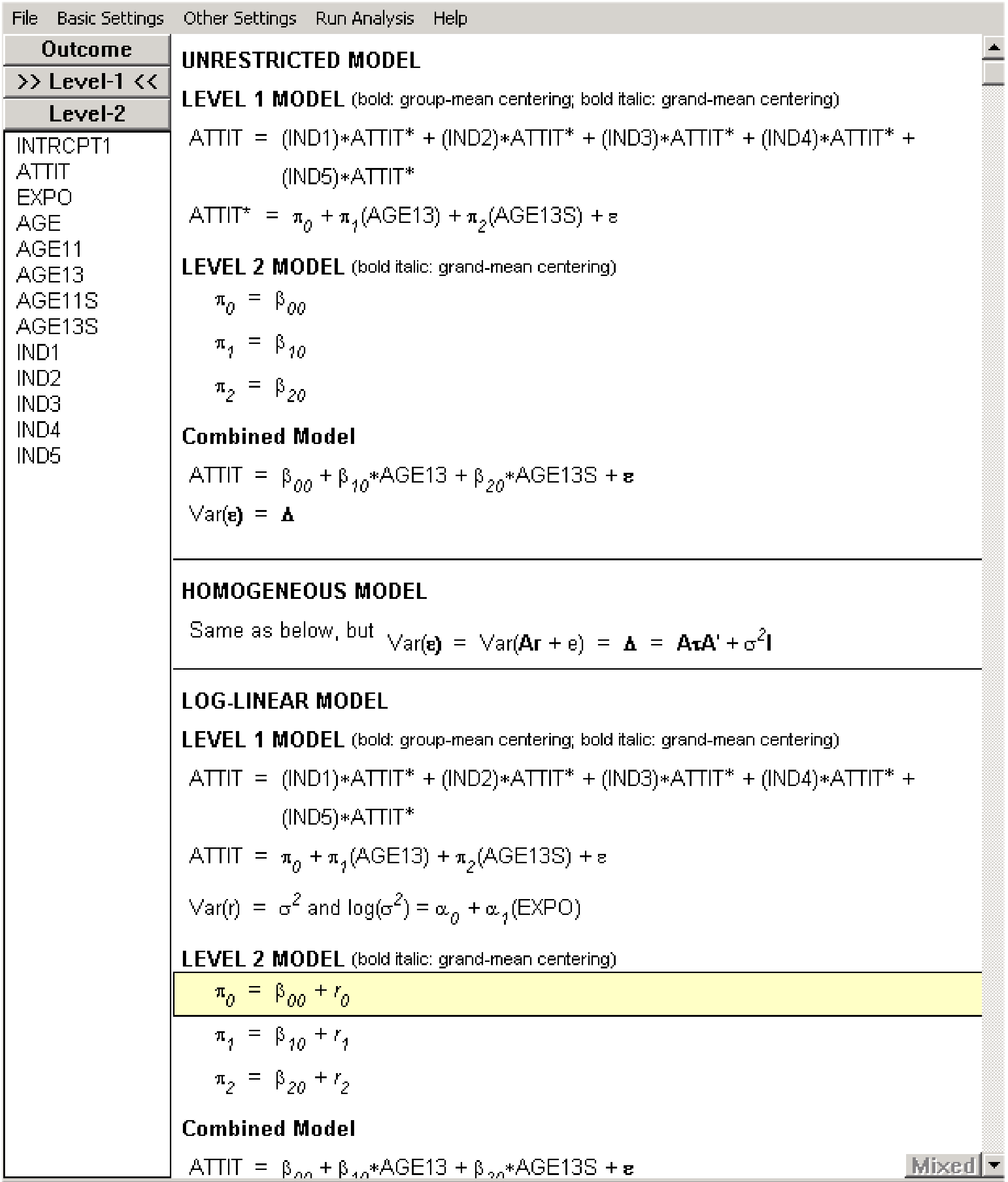


Figure 10.5 Model specification window for the NYS example: loglinear model selection

And, again similarly, choosing the **1st order auto-regressive** option will produce unrestricted and homogeneous results in addition to first-order auto-regressive results.

```
The data source for this run = NYS.MDM
The command file for this run = nys1.hlm
Output file name = nys1.html
The maximum number of level-1 units = 1079
The maximum number of level-2 units = 239
The maximum number of iterations = 100

The outcome variable is ATTIT

The model specified for the fixed effects was:
```

Level-1 Coefficients	Level-2 Predictors
INTRCPT1, π_0	INTRCPT2, β_{00}
# AGE13 slope, π_1	INTRCPT2, β_{10}
# AGE13S slope, π_2	INTRCPT2, β_{20}

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

Output for the Unrestricted Model

Summary of the model specified

Level-1 Model

$$\text{ATTIT}_{mi} = (\text{IND1}_{mi})\text{*ATTIT}_{1i}^{*} + (\text{IND2}_{mi})\text{*ATTIT}_{2i}^{*} + (\text{IND3}_{mi})\text{*ATTIT}_{3i}^{*} + (\text{IND4}_{mi})\text{*ATTIT}_{4i}^{*} + (\text{IND5}_{mi})\text{*ATTIT}_{5i}^{*}$$

The level-1 model relates the observed data, Y, to the complete data, Y^*.

$$\text{ATTIT}_{ti}^{*} = \pi_{0i} + \pi_{1i}\text{*(AGE13}_{ti}) + \pi_{2i}\text{*(AGE13S}_{ti}) + \varepsilon_{ti}$$

Level-2 Model

$\pi_{0i} = \beta_{00}$
$\pi_{1i} = \beta_{10}$
$\pi_{2i} = \beta_{20}$

For the restricted model, there is no random variation between persons in regression coefficient β_0, β_1, and β_2 because all random variation has been absorbed into $\mathbf{\Delta}$.

$\text{Var}(\varepsilon_i) = \mathbf{\Delta}$

$\mathbf{\Delta}_{(0)}$

IND1	0.03507	0.01671	0.01889	0.02149	0.02486
IND2	0.01671	0.04458	0.02779	0.02468	0.02714
IND3	0.01889	0.02779	0.07272	0.05303	0.04801
IND4	0.02149	0.02468	0.05303	0.08574	0.06636
IND5	0.02486	0.02714	0.04801	0.06636	0.08985

The 5×5 matrix $\mathbf{\Delta}$ contains the maximum likelihood estimates of the five variances (one for each time point) and ten covariances (one for each pair of time points). The associated correlation matrix is printed below.

Standard errors of $\mathbf{\Delta}$

IND1	0.00347	0.00304	0.00375	0.00413	0.00429
IND2	0.00304	0.00434	0.00430	0.00457	0.00473
IND3	0.00375	0.00430	0.00678	0.00631	0.00625
IND4	0.00413	0.00457	0.00631	0.00811	0.00736
IND5	0.00429	0.00473	0.00625	0.00736	0.00853

$\mathbf{\Delta}$ (as correlations)

IND1	1.000	0.423	0.374	0.392	0.443
IND2	0.423	1.000	0.488	0.399	0.429
IND3	0.374	0.488	1.000	0.672	0.594
IND4	0.392	0.399	0.672	1.000	0.756
IND5	0.443	0.429	0.594	0.756	1.000

The 5×5 matrix above contains estimated standard errors for each element of $\mathbf{\Delta}$.

The value of the log-likelihood function at iteration 8 = 1.891335E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.320244	0.014981	21.377	238	<0.001
For AGE13 slope, π_1					
INTRCPT2, β_{10}	0.059335	0.004710	12.598	238	<0.001
For AGE13S slope, π_2					
INTRCPT2, β_{20}	0.000330	0.003146	0.105	238	0.917

The expected log attitude at age 13 is 0.320244. The mean linear growth rate of increase is estimated to be 0.059335, t = 12.598, indicating a highly significantly positive average rate of increase in deviant attitude at age 13. The quadratic rate is not statistically significant.

Statistics for the current model

Deviance = -378.266936
Number of estimated parameters = 18

There are 3 fixed effects (f = 3) and five observations in the "complete data" for each person (T = 5). Thus, there are a total of $f + T(T+1)/2 = 3 + 5(5+1)/2 = 18$ parameters. This is the end of the unrestricted model output.

Next follows the results for the homogeneous level-1 variance.

Output for Random Effects Model with Homogeneous Level-1 Variance

Summary of the model specified

Level-1 Model

$\text{ATTIT}_{mi} = (\text{IND1}_{mi})*\text{ATTIT}_{1i}^{*} + (\text{IND2}_{mi})*\text{ATTIT}_{2i}^{*} + (\text{IND3}_{mi})*\text{ATTIT}_{3i}^{*} + (\text{IND4}_{mi})*\text{ATTIT}_{4i}^{*} + (\text{IND5}_{mi})*\text{ATTIT}_{5i}^{*}$

$\text{ATTIT}_{ti}^{*} = \pi_{0i} + \pi_{1i}*(\text{AGE13}_{ti}) + \pi_{2i}*(\text{AGE13S}_{ti}) + \varepsilon_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00} + r_{0i}$
$\pi_{1i} = \beta_{10} + r_{1i}$
$\pi_{2i} = \beta_{20} + r_{2i}$

$Var(\varepsilon_i) = Var(\boldsymbol{A}\boldsymbol{r}_i + e_i) = \boldsymbol{\Delta} = \boldsymbol{A\tau A'} + \sigma^2\boldsymbol{I}$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^*) = \Delta = ATA' + \Sigma$$

where $\Sigma = \sigma^2 I_T$.

A

IND1	1.00000	-2.00000	4.00000
IND2	1.00000	-1.00000	1.00000
IND3	1.00000	0.00000	0.00000
IND4	1.00000	1.00000	1.00000
IND5	1.00000	2.00000	4.00000

The above matrix describes the design matrix on occasions one through five.

Iterations stopped due to small change in likelihood function

Note: The results below duplicate exactly the results produced by a standard HLM2 run using homogeneous level-1 variance.

Final Results - Iteration 5

	Parameter	Standard Error
σ^2	0.02421	0.001672

τ

INTRCPT1,r_0	0.04200	0.00808	-0.00242
AGE13,r_1	0.00808	0.00277	-0.00012
AGE13S,r_2	-0.00242	-0.00012	0.00049

Standard errors of τ

INTRCPT1,r_0	0.00513	0.00127	0.00089
AGE13,r_1	0.00127	0.00054	0.00024
AGE13S,r_2	0.00089	0.00024	0.00025

τ (as correlations)

INTRCPT1,r_0	1.000	0.749	-0.532
AGE13,r_1	0.749	1.000	-0.101
AGE13S,r_2	-0.532	-0.101	1.000

$\mathbf{\Delta}$

IND1	0.03536	0.01388	0.01616	0.01801	0.01943
IND2	0.01388	0.04870	0.03150	0.03488	0.03464
IND3	0.01616	0.03150	0.06620	0.04766	0.04849
IND4	0.01801	0.03488	0.04766	0.08056	0.06095
IND5	0.01943	0.03464	0.04849	0.06095	0.09625

The 5×5 matrix above contains the five variance and ten covariance estimates implied by the "homogeneous level-1 variance" model.

$\mathbf{\Delta}$ (as correlations)

IND1	1.000	0.334	0.334	0.338	0.333
IND2	0.334	1.000	0.555	0.557	0.506
IND3	0.334	0.555	1.000	0.653	0.607
IND4	0.338	0.557	0.653	1.000	0.692
IND5	0.333	0.506	0.607	0.692	1.000

The value of the log-likelihood function at iteration 5 = 1.741132E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.327231	0.015306	21.379	238	<0.001
For AGE13 slope, π_1					
INTRCPT2, β_{10}	0.064704	0.004926	13.135	238	<0.001
For AGE13S slope, π_2					
INTRCPT2, β_{20}	0.000171	0.003218	0.053	238	0.958

Statistics for the current model

Deviance = -348.226421
Number of estimated parameters = 10

There are 3 fixed effects (f = 3); the dimension of τ is 3, and a common σ^2 is estimated at level-1. Thus, there are a total of $f + r(r+1)/2 + 1 = 3 + 3(3+1)/2 + 1 = 10$ parameters.

This is the end of the output for the "homogeneous level-1 variance" model. Finally, the heterogeneous level-1 variance solution is listed.

Output for Random Effects Model with Heterogeneous Level-1 Variance

Summary of the model specified

Level-1 Model

$$\text{ATTIT}_{mi} = (\text{IND1}_{mi})^*\text{ATTIT}_{1i}^* + (\text{IND2}_{mi})^*\text{ATTIT}_{2i}^* + (\text{IND3}_{mi})^*\text{ATTIT}_{3i}^* + (\text{IND4}_{mi})^*\text{ATTIT}_{4i}^* + (\text{IND5}_{mi})^*\text{ATTIT}_{5i}^*$$

$$\text{ATTIT}_{ti}^* = \pi_{0i} + \pi_{1i}^*(\text{AGE13}_{ti}) + \pi_{2i}^*(\text{AGE13S}_{ti}) + \varepsilon_{ti}$$

Level-2 Model

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$\pi_{1i} = \beta_{10} + r_{1i}$
$\pi_{2i} = \beta_{20} + r_{2i}$

$Var(\varepsilon_i) = Var(\mathbf{A}\mathbf{r}_i + e_i) = \boldsymbol{\Delta} = \mathbf{A}^*\boldsymbol{\tau}^*\mathbf{A}' + diag(\sigma^2{}_1,\ldots,\sigma^2{}_5)$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^*) = ATA' + \Sigma$$

where $\Sigma = diag\left\{\sigma_t^2\right\}$, i.e. that is, Σ is now a diagonal matrix with diagonal elements σ_t^2, the variance associated with occasion t, $t = 1, 2, \ldots, T$.

A

IND1	1.00000	-2.00000	4.00000
IND2	1.00000	-1.00000	1.00000
IND3	1.00000	0.00000	0.00000
IND4	1.00000	1.00000	1.00000
IND5	1.00000	2.00000	4.00000

Iterations stopped due to small change in likelihood function

Final Results - Iteration 8

	σ^2	Standard Error
IND1	0.01373	0.005672
IND2	0.02600	0.003296
IND3	0.02685	0.003658
IND4	0.02602	0.003633
IND5	0.00275	0.007377

The five estimates above are the estimates of the level-1 variance for each time point.

τ

INTRCPT1,r_0	0.04079	0.00736	-0.00241
AGE13,r_1	0.00736	0.00382	0.00025
AGE13S,r_2	-0.00241	0.00025	0.00106

Standard errors of τ

INTRCPT1,r_0	0.00512	0.00124	0.00088
AGE13,r_1	0.00124	0.00066	0.00042
AGE13S,r_2	0.00088	0.00042	0.00030

τ (as correlations)

INTRCPT1,r_0	1.000	0.590	-0.366
AGE13,r_1	0.590	1.000	0.124
AGE13S,r_2	-0.366	0.124	1.000

$\mathbf{\Delta}$

IND1	0.03410	0.01707	0.01646	0.01851	0.02325
IND2	0.01707	0.05165	0.03103	0.03322	0.03223
IND3	0.01646	0.03103	0.06764	0.04574	0.04588
IND4	0.01851	0.03322	0.04574	0.08208	0.06421
IND5	0.02325	0.03223	0.04588	0.06421	0.08996

The 5×5 matrix above contains the estimates of five variances and ten covariances implied by the "heterogeneous level-1 variance" model.

$\mathbf{\Delta}$ (as correlations)

IND1	1.000	0.407	0.343	0.350	0.420
IND2	0.407	1.000	0.525	0.510	0.473
IND3	0.343	0.525	1.000	0.614	0.588
IND4	0.350	0.510	0.614	1.000	0.747
IND5	0.420	0.473	0.588	0.747	1.000

The value of the log-likelihood function at iteration 8 = 1.816074E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.327646	0.015252	21.482	238	<0.001
For AGE13 slope, π_1					
INTRCPT2, β_{10}	0.060864	0.004737	12.849	238	<0.001
For AGE13S slope, π_2					
INTRCPT2, β_{20}	-0.000541	0.003178	-0.170	238	0.865

Statistics for the current model

Deviance = -363.214879
Number of estimated parameters = 14

There are 3 fixed effects (f = 3), the dimension of τ is 3, and there are five observations intended for each person, each associated with a unique level-1 variance. Thus, there are a total of $f + r(r+1)/2 + T = 3 + 3(4)/2 + 5 = 14$ parameters.

Summary of Model Fit

Model	Number of Parameters	Deviance
1. Unrestricted	18	-378.26694
2. Homogeneous σ^2	10	-348.22642
3. Heterogeneous σ^2	14	-363.21488

Model Comparison	χ^2	*d.f.*	*p*-value
Model 1 vs Model 2	30.04052	8	<0.001
Model 1 vs Model 3	15.05206	4	0.005
Model 2 vs Model 3	14.98846	4	0.005

The model deviances are employed to evaluate the fits of the three models (unrestricted, homogeneous σ^2, and heterogeneous σ^2). Differences between deviances are distributed asymptotically as chi-square variates under the null hypothesis that the simpler model fits the data as well as the more complex model does. The results show that Model 1 fits better than does the homogeneous sigma squared model χ^2 = 30.04052, df = 8; it also fits better than does the heterogeneous sigma squared model χ^2 = 15.05206, df = 4.

In addition to the evaluation of models based on their fit to the data, the above results can be used to check the sensitivity of key inferences to alternative specifications of the variance-covariance structure. For instance, one could compare the mean and variance in the rate of change at age 13 obtained in Model 2 and Model 3 to assess how robust the results are to alternative plausible covariance specifications. The mean rate, γ_{10}, for Model 2 is 0.064704 (s.e. = 0.004926), and the variance, τ_{22}, is 0.00277 (s.e. = 0.00054). The mean rate, G10, for Model 3 is 0.060864 (s.e. = 0.004737), and the variance, τ_{22}, is 0.00382 (s.e. = 0.00066). The results are basically similar. See Raudenbush (2001) for a more detailed analysis of alternative covariance structures for polynomial models of individual growth and change using the same NYS data sets employed here for the illustrations.

Below are partial outputs for two random effect models.

Output for Random Effects Model for Log-linear model for Level-1 Variance

Summary of the model specified

Level-1 Model

$\text{ATTIT}_{mi} = (\text{IND1}_{mi})*\text{ATTIT}_{1i}^{*} + (\text{IND2}_{mi})*\text{ATTIT}_{2i}^{*} + (\text{IND3}_{mi})*\text{ATTIT}_{3i}^{*} + (\text{IND4}_{mi})*\text{ATTIT}_{4i}^{*} + (\text{IND5}_{mi})*\text{ATTIT}_{5i}^{*}$

$\text{ATTIT}_{ti}^{*} = \pi_{0i} + \pi_{1i}*(\text{AGE13}_{ti}) + \pi_{2i}*(\text{AGE13S}_{ti}) + \varepsilon_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00} + r_{0i}$
$\pi_{1i} = \beta_{10} + r_{1i}$
$\pi_{2i} = \beta_{20} + r_{2i}$

$Var(\varepsilon_i) = Var(\boldsymbol{A}\boldsymbol{r}_i + e_i) = \boldsymbol{\Delta} = \boldsymbol{A\tau A'} + diag(\sigma^2_1,\ldots,\sigma^2_5)$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^*) = ATA' + \Sigma$$

where $\Sigma = diag(\sigma_t^2)$, and

$$\log(\sigma_t^2) = \alpha_0 + \alpha_1\,(EXPO)_t.$$

A

IND1	1.00000	-2.00000	4.00000
IND2	1.00000	-1.00000	1.00000
IND3	1.00000	0.00000	0.00000
IND4	1.00000	1.00000	1.00000
IND5	1.00000	2.00000	4.00000

Iterations stopped due to small change in likelihood function

Final results – Iteration 7

	Parameter	Standard Error
α_0	-3.72883	0.069238
α_1	-1.43639	1.053241

	σ^2
IND1	0.02690
IND2	0.02677
IND3	0.02419
IND4	0.02188
IND5	0.02136

τ

INTRCPT1,r_0	0.04255	0.00831	-0.00257
AGE13,r_1	0.00831	0.00277	-0.00005
AGE13S,r_2	-0.00257	-0.00005	0.00051

Standard errors of τ

INTRCPT1,r_0	0.00517	0.00128	0.00089
AGE13,r_1	0.00128	0.00054	0.00025
AGE13S,r_2	0.00089	0.00025	0.00025

τ (as correlations)

INTRCPT1,r_0	1.000	0.766	-0.549
AGE13,r_1	0.766	1.000	-0.042
AGE13S,r_2	-0.549	-0.042	1.000

Δ

IND1	0.03576	0.01267	0.01566	0.01782	0.01917
IND2	0.01267	0.05095	0.03168	0.03516	0.03464
IND3	0.01566	0.03168	0.06674	0.04829	0.04889
IND4	0.01782	0.03516	0.04829	0.07909	0.06192
IND5	0.01917	0.03464	0.04889	0.06192	0.09510

The 5×5 matrix above contains the variance and covariance estimates implied by the "log-linear" model for the level-1 variance.

$\boldsymbol{\Delta}$ (as correlations)

IND1	1.000	0.297	0.320	0.335	0.329
IND2	0.297	1.000	0.543	0.554	0.498
IND3	0.320	0.543	1.000	0.665	0.614
IND4	0.335	0.554	0.665	1.000	0.714
IND5	0.329	0.498	0.614	0.714	1.000

The value of the log-likelihood function at iteration 7 = 1.749582E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.328946	0.015379	21.390	238	<0.001
For AGE13 slope, π_1					
INTRCPT2, β_{10}	0.064661	0.004923	13.135	238	<0.001
For AGE13S slope, π_2					
INTRCPT2, β_{20}	-0.000535	0.003222	-0.166	238	0.869

Statistics for the current model

Deviance = -349.916489
Number of estimated parameters = 11

There are 3 fixed effects ($f = 3$), the dimension of τ is 3 ($r = 3$), and there is 1 intercept and 1 explanatory ($H = 1$) variable. Thus, there are a total of $f + r(r+1)/2 + 1 + H = 3 + 3(3+1)/2 + 1 + 1 = 11$ parameters.

Next are the results for the first-order auto-regressive model (Example: NYS4.MLM)

Output for Random Effects Model First-order Autoregressive Model for Level-1 Variance

Summary of the model specified

Level-1 Model

$$\text{ATTIT}_{mi} = (\text{IND1}_{mi})*\text{ATTIT}_{1i}^{*} + (\text{IND2}_{mi})*\text{ATTIT}_{2i}^{*} + (\text{IND3}_{mi})*\text{ATTIT}_{3i}^{*} + (\text{IND4}_{mi})*\text{ATTIT}_{4i}^{*} + (\text{IND5}_{mi})*\text{ATTIT}_{5i}^{*}$$

$$\text{ATTIT}_{ti}^{*} = \pi_{0i} + \pi_{1i}*(\text{AGE13}_{ti}) + \pi_{2i}*(\text{AGE13S}_{ti}) + \varepsilon_{ti}$$

Level-2 Model

$$\pi_{0i} = \beta_{00} + r_{0i}$$
$$\pi_{1i} = \beta_{10}$$
$$\pi_{2i} = \beta_{20}$$

Note that β_1 and β_2 are specified as non-random due to the fact that the time-series is relatively short and therefore the data do not allow the estimation of both random slopes and an autocorrelation parameter.

$$Var(\varepsilon_i) = Var(\boldsymbol{A}\boldsymbol{r}_i + e_i) = \boldsymbol{\Delta} = \boldsymbol{A\tau A'} + \sigma^2\rho^{|t-t'|}$$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^*) = ATA' + \Sigma$$

where

$$\Sigma = \sigma^2 \rho^{|t-t'|}.$$

A

IND1	1.00000
IND2	1.00000
IND3	1.00000
IND4	1.00000
IND5	1.00000

Iterations stopped due to small change in likelihood function

Final Results - Iteration 6

	Parameter	Standard Error
ρ	0.39675	0.053849
σ^2	0.04158	0.003582

Note that the maximum-likelihood estimate of $\hat{\rho}$ = 0.397 is much larger than its standard error (0.054), suggesting a significantly positive autocorrelation.

τ

INTRCPT1,r_0	0.02427

Standard error of τ

INTRCPT1,r_0	0.00450

Δ

IND1	0.06585	0.04077	0.03081	0.02686	0.02530
IND2	0.04077	0.06585	0.04077	0.03081	0.02686
IND3	0.03081	0.04077	0.06585	0.04077	0.03081
IND4	0.02686	0.03081	0.04077	0.06585	0.04077
IND5	0.02530	0.02686	0.03081	0.04077	0.06585

The 5×5 matrix above contains the variance and covariance estimates implied by the "auto-correlation" model for the level-1 variance.

Δ (as correlations)

IND1	1.000	0.619	0.468	0.408	0.384
IND2	0.619	1.000	0.619	0.468	0.408
IND3	0.468	0.619	1.000	0.619	0.468
IND4	0.408	0.468	0.619	1.000	0.619
IND5	0.384	0.408	0.468	0.619	1.000

The value of the log-likelihood function at iteration 6 = 1.471600E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.327579	0.015265	21.459	238	<0.001
For AGE13 slope, π_1					
INTRCPT2, β_{10}	0.061428	0.004836	12.703	1076	<0.001
For AGE13S slope, π_2					
INTRCPT2, β_{20}	0.000211	0.003373	0.062	1076	0.951

Statistics for the current model

Deviance = -294.319916
Number of estimated parameters = 6

Summary of Model Fit

Model	Number of Parameters	Deviance
1. Unrestricted	18	-378.26694
2. Homogeneous σ^2	5	-229.01630
3. First order Autoregressive	6	-294.31992

10.4 An analysis using HMLM2 via Windows mode

To illustrate how to use HMLM2, we use the data files from the public school example described in Section 4.1.1.1. We prepared six indicators for the measures of mathematics proficiency collected over the six years and put them in the level-1 file, EG1.SAV. The new level-1 file is called EG1HMLM2.SAV. The same level-2 and level-3 files, EG2.SAV and EG3.SAV are used. The MDM file created is EGHMLM2.MDM. Like in the case in HMLM, users need to tell the program what the indicator variables are while creating the MDM file (see Fig. 10.2).

10.5 Executing analyses based on the MDM file

The steps involved are similar to the ones for HMLM outlined previously and for HLM3 as described in Section 4.2. The user specifies

1. the level-1 model,
2. the level-2 structural model, and
3. the level-1 coefficients as random or non-random.

In addition, the user selects the covariance structure for the models to be estimated. Below is the output for the linear growth model specified in Section 4.2. As in the case for HMLM, the results allow us to compare model fit and assess sensitivity of inferences with alternative specification of variance-covariance structures.

10.5.1 Specifications for this HMLM2 run

Problem Title: no title

The data source for this run = EGHMLM2.MDM
The command file for this run = EG.HLM
Output file name = hmlm2.html
The maximum number of level-1 units = 7230

The maximum number of level-2 units = 1721
The maximum number of level-3 units = 60
The maximum number of iterations = 100

The outcome variable is MATH

The model specified for the fixed effects was:

The model specified for the fixed effects

Level-1	Level-2	Level-3
INTRCPT1, π_0	INTRCPT2, β_{00}	INTRCPT3 ,γ_{000}
YEAR slope, π_1	INTRCPT2, β_{10}	INTRCPT3 ,γ_{100}

Output for the Unrestricted Model

Summary of the model specified

Level-1 Model

$MATH_{mij}$ = $(IND1_{mij})*MATH_{1ij}^*$ + $(IND2_{mij})*MATH_{2ij}^*$ + $(IND3_{mij})*MATH_{3ij}^*$ + $(IND4_{mij})*MATH_{4ij}^*$ + $(IND5_{mij})*MATH_{5ij}^*$ + $(IND6_{mij})*MATH_{6ij}^*$

$MATH_{tij}^* = \pi_{0ij} + \pi_{1ij}*(YEAR_{tij}) + \varepsilon_{tij}$

Level-2 Model

$\pi_{0ij} = \beta_{00j}$
$\pi_{1ij} = \beta_{10j}$

Level-3 Model

$\beta_{00j} = \gamma_{000} + u_{00j}$
$\beta_{10j} = \gamma_{100} + u_{10j}$

$Var(\varepsilon_{ij}) = \boldsymbol{\Delta}$

$\boldsymbol{\Delta}_{(0)}$

IND1	0.04268	0.01233	0.01919	0.01968	0.01506	0.00898
IND2	0.01233	0.60634	0.35457	0.42101	0.31132	0.24927
IND3	0.01919	0.35457	0.76957	0.62363	0.42394	0.35205
IND4	0.01968	0.42101	0.62363	1.15453	0.67302	0.52773
IND5	0.01506	0.31132	0.42394	0.67302	0.81870	0.55086
IND6	0.00898	0.24927	0.35205	0.52773	0.55086	0.65701

$\tau_{\beta(0)}$

INTRCPT1	YEAR
INTRCPT2 ,β_{00}	INTRCPT2 ,β_{10}
0.20128	0.01542
0.01542	0.01608

The value of the log-likelihood function at iteration 1 = -8.445655E+003
The value of the log-likelihood function at iteration 2 = -8.228973E+003
The value of the log-likelihood function at iteration 3 = -8.166659E+003
The value of the log-likelihood function at iteration 4 = -8.126574E+003
The value of the log-likelihood function at iteration 5 = -8.097070E+003

Iterations stopped due to small change in likelihood function

Final Results - Iteration 32

Δ

IND1	0.67340	0.31616	0.38755	0.52412	0.53030	0.38971
IND2	0.31616	0.77832	0.47127	0.56726	0.54171	0.50187
IND3	0.38755	0.47127	0.91072	0.76829	0.66199	0.64640
IND4	0.52412	0.56726	0.76829	1.24542	0.88364	0.81782
IND5	0.53030	0.54171	0.66199	0.88364	1.05646	0.84356
IND6	0.38971	0.50187	0.64640	0.81782	0.84356	0.98722

Standard errors of Δ

IND1	0.08003	0.05328	0.07256	0.02999	0.02811	0.04341
IND2	0.05328	0.05757	0.06998	0.02542	0.03289	0.03656
IND3	0.07256	0.06998	0.07252	0.02966	0.03284	0.03565
IND4	0.02999	0.02542	0.02966	0.02844	0.03044	0.03913
IND5	0.02811	0.03289	0.03284	0.03044	0.03030	0.03518
IND6	0.04341	0.03656	0.03565	0.03913	0.03518	0.03859

Δ (as correlations)

IND1 ,π_0	1.000	0.437	0.495	0.572	0.629	0.478
IND2 ,π_1	0.437	1.000	0.560	0.576	0.597	0.573
IND3 ,π_2	0.495	0.560	1.000	0.721	0.675	0.682
IND4 ,π_3	0.572	0.576	0.721	1.000	0.770	0.738
IND5 ,π_4	0.629	0.597	0.675	0.770	1.000	0.826
IND6 ,π_5	0.478	0.573	0.682	0.738	0.826	1.000

τ_β

INTRCPT1	YEAR
INTRCPT2 ,β_{00}	INTRCPT2 ,β_{10}
0.14824	0.01268
0.01268	0.00935

Standard Errors of τ_β

INTRCPT1	YEAR
INTRCPT2 ,β_{00}	INTRCPT2 ,β_{10}
0.03286	0.00626
0.00626	0.00218

τ_β (as correlations)

INTRCPT1/INTRCPT2,β_{00}	1.000	0.341
YEAR/INTRCPT2,β_{10}	0.341	1.000

The value of the log-likelihood function at iteration 32 = -7.980254E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3 ,γ_{000}	-0.824938	0.054960	-15.010	59	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3 ,γ_{100}	0.755026	0.014229	53.062	59	<0.001

Statistics for the current model

Deviance = 15960.507331
Number of estimated parameters = 26

Output for Random Effects Model with Homogeneous Level-1 Variance

Summary of the model specified

Level-1 Model

$MATH_{mij} = (IND1_{mij})*MATH_{1ij}^* + (IND2_{mij})*MATH_{2ij}^* + (IND3_{mij})*MATH_{3ij}^* + (IND4_{mij})*MATH_{4ij}^* + (IND5_{mij})*MATH_{5ij}^* + (IND6_{mij})*MATH_{6ij}^*$

$MATH_{tij}^* = \pi_{0ij} + \pi_{1ij}*(YEAR_{tij}) + \varepsilon_{tij}$

Level-2 Model

$\pi_{0ij} = \beta_{00j} + r_{0ij}$
$\pi_{1ij} = \beta_{10j} + r_{1ij}$

Level-3 Model

$\beta_{00j} = \gamma_{000} + u_{00j}$
$\beta_{10j} = \gamma_{100} + u_{10j}$

$Var(\varepsilon_{ij}) = Var(\mathbf{A}\mathbf{r}_{ij} + e_{ij}) = \boldsymbol{\Delta} = \mathbf{A}\boldsymbol{\tau}_{\pi}\mathbf{A'} + \sigma^2\mathbf{I}$

A

IND1	1.00000	-2.50000
IND2	1.00000	-1.50000
IND3	1.00000	-0.50000
IND4	1.00000	0.50000
IND5	1.00000	1.50000
IND6	1.00000	2.50000

The value of the log-likelihood function at iteration 1 = -7.980254E+003
The value of the log-likelihood function at iteration 2 = -8.271230E+003
The value of the log-likelihood function at iteration 3 = -8.163134E+003
The value of the log-likelihood function at iteration 4 = -8.163116E+003

Iterations stopped due to small change in likelihood function

Final Results - Iteration 5

	Parameter	Standard Error
σ^2	0.30144	0.006598

τ_π

INTRCPT1,r_0	0.64046	0.04679
YEAR,r_1	0.04679	0.01126

Standard errors of τ_π

INTRCPT1,r_0	0.02515	0.00499
YEAR,r_1	0.00499	0.00197

τ_π (as correlations)

INTRCPT1,r_0	1.000	0.551
YEAR,r_1	0.551	1.000

Δ

IND1	0.77832	0.49553	0.51417	0.53282	0.55146	0.57011
IND2	0.49553	0.82687	0.55533	0.58523	0.61513	0.64503
IND3	0.51417	0.55533	0.89793	0.63765	0.67880	0.71996
IND4	0.53282	0.58523	0.63765	0.99150	0.74247	0.79489
IND5	0.55146	0.61513	0.67880	0.74247	1.10758	0.86981
IND6	0.57011	0.64503	0.71996	0.79489	0.86981	1.24618

Δ (as correlations)

IND1 ,π_0	1.000	0.618	0.615	0.607	0.594	0.579
IND2 ,π_1	0.618	1.000	0.644	0.646	0.643	0.635
IND3 ,π_2	0.615	0.644	1.000	0.676	0.681	0.681
IND4 ,π_3	0.607	0.646	0.676	1.000	0.709	0.715
IND5 ,π_4	0.594	0.643	0.681	0.709	1.000	0.740
IND6 ,π_5	0.579	0.635	0.681	0.715	0.740	1.000

τ_β

INTRCPT1	YEAR
INTRCPT2 ,β_{00}	INTRCPT2 ,β_{10}
0.16532	0.01705
0.01705	0.01102

Standard Errors of τ_β

INTRCPT1	YEAR
INTRCPT2 ,β_{00}	INTRCPT2 ,β_{10}
0.03641	0.00720
0.00720	0.00252

τ_β (as correlations)

INTRCPT1/INTRCPT2,β_{00}	1.000	0.399
YEAR/INTRCPT2,β_{10}	0.399	1.000

The value of the log-likelihood function at iteration 5 = -8.163116E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3 , γ_{000}	-0.779305	0.057829	-13.476	59	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3 , γ_{100}	0.763028	0.015262	49.996	59	<0.001

Statistics for the current model

Deviance = 16326.231108
Number of estimated parameters = 9

Output for Random Effects Model with Heterogeneous Level-1 Variance

Summary of the model specified

Level-1 Model

$\text{MATH}_{mij} = (\text{IND1}_{mij})^*\text{MATH}_{1ij}^* + (\text{IND2}_{mij})^*\text{MATH}_{2ij}^* + (\text{IND3}_{mij})^*\text{MATH}_{3ij}^* + (\text{IND4}_{mij})^*\text{MATH}_{4ij}^* + (\text{IND5}_{mij})^*\text{MATH}_{5ij}^* + (\text{IND6}_{mij})^*\text{MATH}_{6ij}^*$

$\text{MATH}_{tij}^* = \pi_{0ij} + \pi_{1ij}^*(\text{YEAR}_{tij}) + \varepsilon_{tij}$

Level-2 Model

$\pi_{0ij} = \beta_{00j} + r_{0ij}$
$\pi_{1ij} = \beta_{10j} + r_{1ij}$

Level-3 Model

$\beta_{00j} = \gamma_{000} + u_{00j}$
$\beta_{10j} = \gamma_{100} + u_{10j}$

$Var(\varepsilon_{ij}) = Var(\boldsymbol{A}\boldsymbol{r}_{ij} + e_{ij}) = \boldsymbol{\Delta} = \boldsymbol{A}^*\boldsymbol{\tau}_{\boldsymbol{\pi}}^*\boldsymbol{A}' + diag(\sigma^2{}_1,...,\sigma^2{}_6)$

A

IND1	1.00000	-2.50000
IND2	1.00000	-1.50000
IND3	1.00000	-0.50000
IND4	1.00000	0.50000
IND5	1.00000	1.50000
IND6	1.00000	2.50000

The value of the log-likelihood function at iteration 1 = -8.163116E+003
The value of the log-likelihood function at iteration 2 = -8.072345E+003
The value of the log-likelihood function at iteration 3 = -8.070198E+003
The value of the log-likelihood function at iteration 4 = -8.070086E+003
The value of the log-likelihood function at iteration 5 = -8.070080E+003

Iterations stopped due to small change in likelihood function

Final Results - Iteration 7

	σ^2	Standard Error
IND1	0.34891	0.059597
IND2	0.38314	0.020556
IND3	0.31846	0.014915
IND4	0.37849	0.015840
IND5	0.20344	0.011466
IND6	0.15546	0.014216

τ_π

INTRCPT1,r_0	0.62722	0.04769
YEAR,r_1	0.04769	0.01386

Standard errors of τ_π

INTRCPT1,r_0	0.02499	0.00495
YEAR,r_1	0.00495	0.00205

τ_π (as correlations)

INTRCPT1,r_0	1.000	0.511
YEAR,r_1	0.511	1.000

$\mathbf{\Delta}$

IND1	0.82432	0.48844	0.50148	0.51451	0.52755	0.54058
IND2	0.48844	0.89848	0.54224	0.56913	0.59603	0.62293
IND3	0.50148	0.54224	0.90146	0.62376	0.66452	0.70528
IND4	0.51451	0.56913	0.62376	1.05687	0.73300	0.78762
IND5	0.52755	0.59603	0.66452	0.73300	1.00493	0.86997
IND6	0.54058	0.62293	0.70528	0.78762	0.86997	1.10778

$\mathbf{\Delta}$ (as correlations)

IND1 ,π_0	1.000	0.568	0.582	0.551	0.580	0.566
IND2 ,π_1	0.568	1.000	0.603	0.584	0.627	0.624
IND3 ,π_2	0.582	0.603	1.000	0.639	0.698	0.706
IND4 ,π_3	0.551	0.584	0.639	1.000	0.711	0.728
IND5 ,π_4	0.580	0.627	0.698	0.711	1.000	0.825
IND6 ,π_5	0.566	0.624	0.706	0.728	0.825	1.000

τ_β

INTRCPT1 INTRCPT2 ,β_{00}	YEAR INTRCPT2 ,β_{10}
0.16531	0.01552
0.01552	0.00971

Standard Errors of τ_β

INTRCPT1 INTRCPT2 ,β_{00}	YEAR INTRCPT2 ,β_{10}
0.03637	0.00677
0.00677	0.00225

τ_β (as correlations)

INTRCPT1/INTRCPT2, β_{00}	1.000	0.387
YEAR/INTRCPT2, β_{10}	0.387	1.000

The value of the log-likelihood function at iteration 7 = -8.070079E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For INTRCPT2, β_{00}					
INTRCPT3, γ_{000}	-0.781960	0.057792	-13.531	59	<0.001
For YEAR slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	0.751231	0.014452	51.983	59	<0.001

Statistics for the current model

Deviance = 16140.158919
Number of estimated parameters = 14

Summary of Model Fit

Model	Number of Parameters	Deviance
1. Unrestricted	26	15960.50733
2. Homogeneous σ^2	9	16326.23111
3. Heterogeneous σ^2	14	16140.15892

Model Comparison	χ^2	*d.f.*	*p*-value
Model 1 vs Model 2	365.72378	17	<0.001
Model 1 vs Model 3	179.65159	12	<0.001
Model 2 vs Model 3	186.07219	5	<0.001

11 Special Features

11.1 Latent variable analysis

Researchers may be interested in studying the randomly varying coefficients not only as outcomes, but as predictors as well. For instance, in a two-level repeated measures study of adolescents' tolerance of deviant behaviors, a user may choose to use the level-1 coefficient capturing the level of tolerance at the beginning of the study to predict the coefficient tapping the linear growth rate.

Treating these coefficients as latent variables, the HLM2, HLM3, HMLM, HMLM2 modules allow researchers to study direct as well as indirect effects among them and to assess their impacts on coefficients associated with observed covariates in the model. Furthermore, using HMLM with unrestricted covariance structures, one may use latent variable analysis to run regressions with missing data.

Below are two examples of latent variable analysis via Windows mode. See Appendix F for batch and interactive modes.

11.1.1 A latent variable analysis using HMLM: Example 1

The first example employs the National Youth Survey data sets described in Section 10.1. The MDM file is NYS.MDM, the level-1 data file is NYS1.SAV, and the level-2 file is NYS2.SAV. Figure 11.1 displays a linear growth model with gender as a covariate. The command file that contains the model specification information is NYS2.MLM.

We use π_0, the level of tolerance at age 11, to predict π_1, the linear growth rate, controlling for gender. Note that FEMALE must be in the model for both π_0 and π_1 to control for gender fully. Note also that π_0 and π_1 are latent variables, that is, they are free of measurement error, which is contained in *e*. Furthermore, we assess whether the effect of gender on the linear growth rate may change after controlling for the initial status at age 11. We select the homogeneous level-1 variance option for this model. Thus, using HLM2 will yield identical results in this case.

Below are the steps for setting up a latent variable analysis.

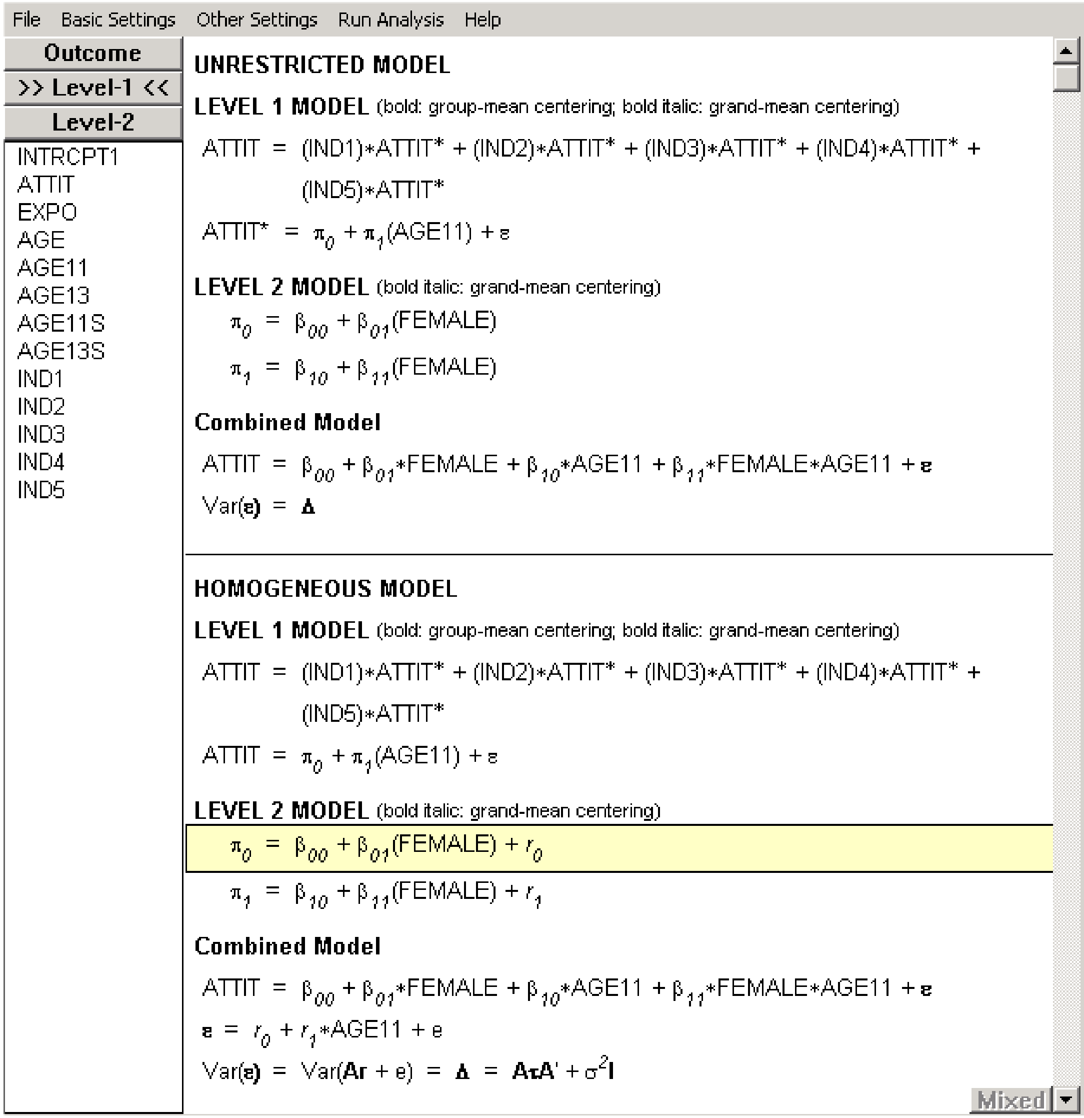


Figure 11.1 Model screen for the NYS example

To set up a latent variable analysis

1. After specifying the model, select the **Estimation Settings** option from the **Other Settings** menu.

2. Choose **Latent Variable Regression** to open the **Latent Variable Regression** dialog box (Figure 9.2 shows an example for the NYS example).

3. Select the predictor(s) and outcome(s) by clicking the selection buttons in front of them (for our example, select INTRCPT1, π_0, as the predictor and AGE11, π_1, as the outcome).

Select HMLM output to illustrate latent variable regression follows.

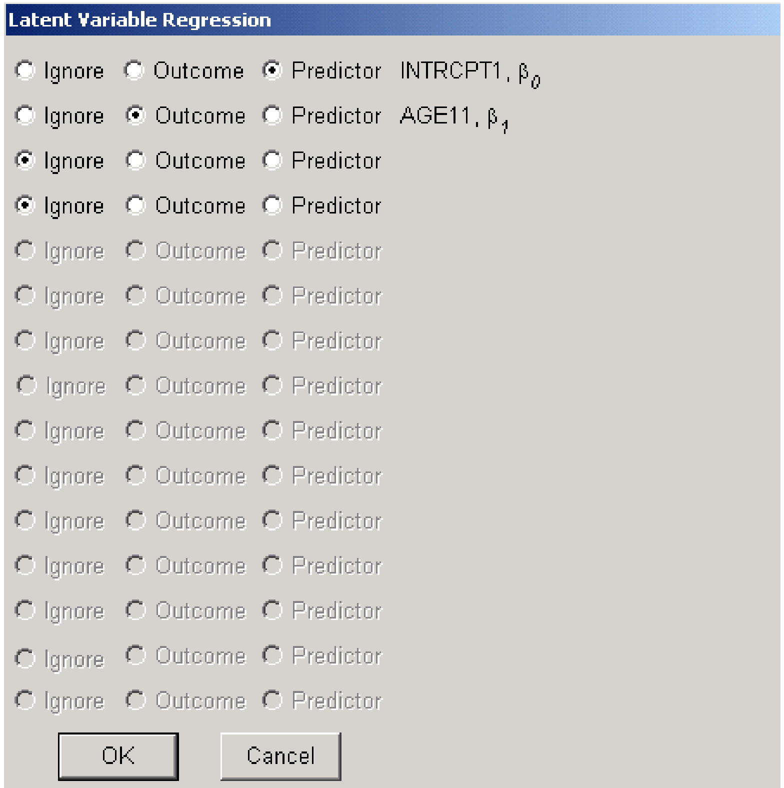


Figure 11.2 Latent Variable Regression dialog box for the NYS example

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, β_{00}	0.221755	0.015961	13.894	237	<0.001
FEMALE, β_{01}	-0.048274	0.022926	-2.106	237	0.035
For AGE11 slope, π_1					
INTRCPT2, β_{10}	0.070432	0.006781	10.386	237	<0.001
FEMALE, β_{11}	-0.012003	0.009826	-1.222	237	0.222

The results indicate that there is a significant linear growth rate in the attitude toward deviant behaviors (coefficient = 0.070432, s.e. = 0.006781) for males. Also, there is no gender effect on the linear growth rate.

Latent Variable Regression Results

The model specified (in equation format)

$$\pi_1 = \beta_{10}^* + \beta_{11}^* \text{ (FEMALE)} + \beta_{12}^* \text{ (} \pi_0 \text{)} + r_1^*$$

Outcome	Predictor	Estimated Coefficient	Standard Error	t-ratio	p-value
AGE11,r_1,π_1	INTRCPT2 ,β_{10}^*	0.024765	0.024807	0.998	0.319
	FEMALE ,β_{11}^*	-0.002062	0.013058	-0.158	0.875
	π_0,β_{12}^*	0.205934	0.105410	1.954	0.050

The results indicate that, controlling for gender, the initial status at age 11 has a marginally significant effect on the linear growth rate (coefficient = 0.205934, s.e. = 0.105410). There is no statistically significant partial gender effect, however. Indeed, the gender effect on π_1 appears somewhat reduced after controlling π_0.

Latent Variable Regression: Comparison of Original and Adjusted Coefficients

Outcome	Predictor	Original Coefficient	Adjusted Coefficient	Difference	Standard Error of Difference
AGE11,r_1,π_1	INTRCPT2	0.07043	0.02477	0.045667	0.024311
	FEMALE	-0.01200	-0.00206	-0.009941	0.006941

This table lists the original coefficients, the adjusted coefficients, and the difference between the two for the intercept and the gender effect. For the variable FEMALE, the "original coefficient" describes the total association, the "adjusted coefficient" describes the direct association, and the "difference" is the indirect association between gender and the linear growth rate, respectively.

Var(r^*)
AGE11,r_1 0.00196

An estimate of the variance of r_1^*, the residual variance in π_1, controlling both FEMALE and π_0, is also given.

As mentioned earlier, a latent variable analysis using HLM2 will reproduce identical results. The same procedures generalize to three-level applications (HMLM2, HLM3, & HGLM) to model randomly varying level-2 coefficients as outcome variables. See Raudenbush and Sampson (1999) for an example that implemented a latent variable analysis with a three-level model. In the study, they investigated the extent to which neighborhood social control mediated the association between neighborhood social composition and violence in Chicago.

11.1.2 A latent variable analysis using HMLM: Example 2

In this example, we illustrate how to use latent variable analysis to run regression with missing data with an artificial data set. We are interested in estimating regression coefficients that relate two predictors to the outcome. There are three intended measures, an outcome (OUTCOME) and two predictors (PRED1 and PRED2) for 15 participants in the data. Some participants are missing one or two measures. To use HMLM to run regression with missing data, we first re-organize the data and re-conceive the three measures for each participant j as "occasions of measurement. " If the data are complete, each case has R = 3 occasions. If participant j is missing one value, there will only be 2 occasions for that participant, and if participant j is missing 2 values, there will be only 1 occasion for that case. The measure is then re-conceived as MEASURE_{ij}, that is, the value of the datum collected at occasion *i* for participant *j*, with $i = 1, 2, ..., n_j$, and with $n_j \le R = 3$. If the data are complete for participant *j*, then:

$$\text{MEASURE}_{1j} = \text{OUTCOME}_j,$$
$$\text{MEASURE}_{2j} = \text{PRED}_{1j},$$
$$\text{MEASURE}_{3j} = \text{PRED}_{2j}.$$

Three indicators IND_{1j}, IND_{2j}, and IND_{3j} indicating whether MEASURE_{ij} is OUTCOME_j, PRED_{1j}, or PRED_{2j} are added to the data set.

Data for the first three participants are shown in Fig. 11.3.

	id	measures	ind1	ind2	ind3
1	1	48.92	1.00	.00	.00
2	1	41.86	.00	1.00	.00
3	1	60.41	.00	.00	1.00
4	2	56.06	.00	1.00	.00
5	2	52.99	.00	.00	1.00
6	3	59.49	1.00	.00	.00

Data View | Variable View

Figure 11.3 First three participants for Example 2

Note that Participant 1 has complete data, Participant 2 has data on PRED1 and PRED2 but not the outcome, and the Participant 3 has data only on OUTCOME.

Data on the measures and the three indicators constitute the level-1 data file, MISSING1.SAV, for the example. The level-2 file, MISSING2.SAV, contains a dummy variable, DUMMY, which is not to be used in the analysis. A MDM file, MISSING.MDM, is created. Figure 11.4 displays the model specified with unrestricted covariance structure for the missing data example. The file that contains the file specification information is MISSING1.MLM.

To regress OUTCOME (IND1) on PRED1 (IND2) and PRED2 (IND3), select IND1 as the outcome and IND2 and IND3 as predictors in the **Latent Variable Regression** dialog box.

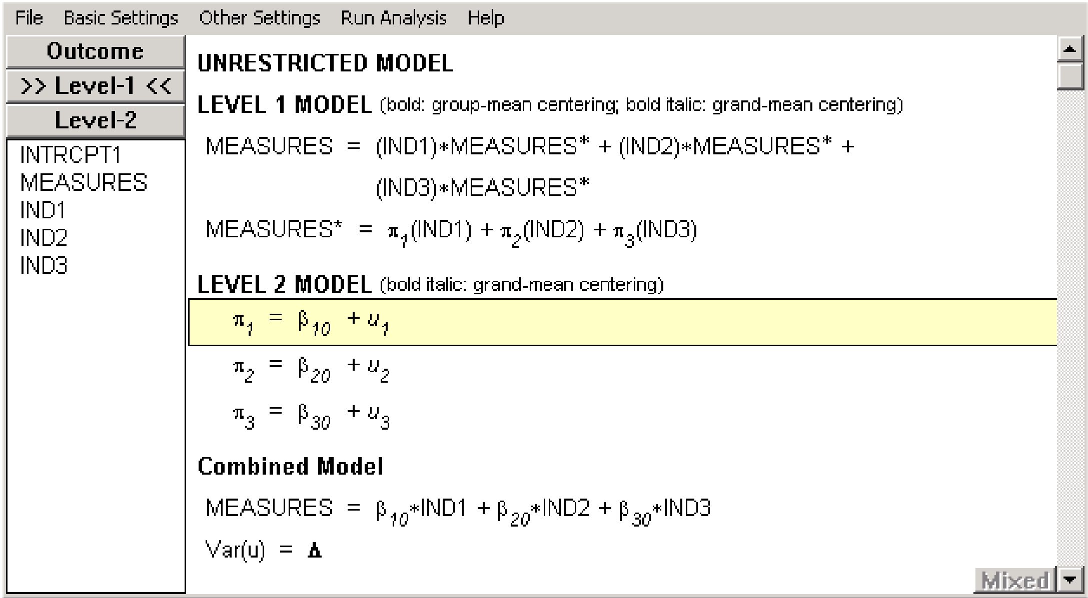


Figure 11.4 Model window for the missing data example

The following selected output (example MISSING1.MLM) gives the latent variable regression results.

Latent Variable Regression Results

The model specified (in equation format)

$$\pi_1 = \beta_{10}^* + \beta_{11}^* (\pi_2) + \beta_{12}^* (\pi_3) + r_1^*$$

Outcome	Predictor	Estimated Coefficient	Standard Error	*t*-ratio	*p*-value
IND1 , π_1	INTRCPT2 , β_{10}^*	-23.966159	14.173726	-1.691	0.117
	π_2, β_{11}^*	0.879462	0.232665	3.780	0.003
	π_3, β_{12}^*	0.544410	0.220194	2.472	0.029

Latent Variable Regression: Comparison of Original and Adjusted Coefficients

Outcome	Predictor	Original Coefficient	Adjusted Coefficient	Difference	Standard Error of Difference
IND1 , π_1	INTRCPT2	52.25565	-23.96616	76.221809	14.285875

Var(r^*)
IND1 33.51133

The results indicate that π_2 (associated with IND2) and π_3 (associated with IND3) have statistically significant effects on IND1 (OUTCOME)[4].

[4]Raudenbush and Bryk (*Hierarchical Linear Models*, 2002) have shown that using this approach with complete data replicated the results of SPSS regression analysis for the regression coefficients. As HMLM adopts the full maximum likelihood estimation approach and the SPSS uses the restricted maximum likelihood approach, the two sets of standard errors estimated differ by a factor

11.2 Applying HLM to multiply-imputed data

A satisfactory solution to the missing data problem involves multiple, model-based imputation (Rubin, 1987, Little & Rubin, 1987, Schafer, 1997). A multiple imputation procedure produces *M* "complete" data sets. Users can apply HLM2 and HLM3 to these multiply-imputed data to produce appropriate estimates that incorporate the uncertainty resulting from imputation.

There can be multiply-imputed values for the outcome or one covariate, or for the outcome and/or covariates.

HLM has two methods to analyze multiply-imputed data. They both use the same equations to compute the averages, so the method chosen depends on the data you are analyzing.

"Plausible Values" as described in Sections 11.2.1 and 11.2.3. This method is usually preferable for data sets that have only one variable (outcome or predictor) for which you have several plausible values. In this case, you need to make one MDM file containing *all* of the plausible values, plus any other variables of interest.

"Multiple Imputation" as described in Section 11.2.4. This method is necessary if you have more than one variable for which you have multiply-imputed data. This method also requires a different way of setting up MDM files. Here, you have to create as many MDMs as you have plausible vales. When making these MDMs, you should use the same level-2 file (and level-3 file if using HLM3), but several level-1 files are needed.

Those variables that are not multiply imputed should be the same in all these level-1 files. The variables that *are* multiply imputed should be separated into the separate level-1 files, but they *must* have the same variable names across these level-1 files, since the same model is run on each of these MDMs.

11.2.1 Data with multiply-imputed values for the outcome or one covariate

HLM2 and HLM3 enable users to produce correct HLM estimates when using data sets that contain two or more values or plausible values for the outcome variable or one covariate. One such data set is the National Assessment of Educational Progress (NAEP), an U.S. Department of Education achievement test given to a national sample of fourth, eighth, and twelfth graders.

Due to the use of balanced incomplete block (BIB) spiraling in the administration of the NAEP assessment battery, special procedures and calculations are necessary when estimating any population parameters and their standard errors with data sets such as NAEP. Every student was not tested on the same items, so item response theory (IRT) was used to estimate proficiency scores for each individual student. This procedure estimated a range or distribution of plausible values for each student's proficiency rather than an individual observed score. NAEP drew five plausible

of

$\sqrt{\frac{J}{J-Q-1}}$, where in this case $J = 15$ and $Q = 2$.

values at random from the conditional distribution of proficiency scores for each student. The measurement error is due to the fact that these scores are estimated, rather than observed.

In general, these plausible values are used to produce parameter estimates in the following way.

- Each parameter is estimated for each of the five plausible values, and the five estimates are averaged.
- Then, the standard error for this average estimate is calculated using the approach recommended by Little & Schenker (1995).
- This formula essentially combines the average of the sampling error from the five estimates with the variance between the five estimates multiplied with a factor related to the number of plausible values. The result is the measurement error.

In an HLM analysis, with either two- or three-levels, the parameter estimates are based on the average parameter estimates from separate HLM analyses of the five plausible values. That is, a separate HLM analysis is conducted on each of the five plausible values.

Without HLM, these procedures could be performed by producing HLM estimates for each plausible value, and then averaging the estimates and calculating the standard errors using another computer program. These procedures are tedious and time-consuming, especially when performed on many models, grades, and dependent variables.

HLM takes the plausible values into account in generating the HLM estimates. For each HLM model, the program runs each of the five (or the number specified) plausible values internally, and produces their average value and the correct standard errors. There will seem to be one estimate, but the five HLM estimates from the five plausible values are produced and their average and measurement error calculated correctly, thus ensuring an accurate treatment of plausible value data. The output is similar to the standard HLM program output, except that all the components are averaged over estimates derived from the five plausible values. In addition, the output from the five plausible value runs is available in a separate output file.

11.2.2 Calculations performed

The program conducts a separate HLM analysis for each plausible value. The output of the separate HLM analyses is written to files with consecutive numbers, for example, OUT.1, OUT.2, OUT.3, etc. Then, HLM calculates the average of the parameter estimates from the separate analyses and computes the standard errors. The output of the average HLM parameter estimates and their standard errors is found in the output file with the extension AVG.

11.2.2.1 Average parameter estimates

The following parameter estimates are averaged by HLM:

- The fixed effects (gammas)
- The reliabilities
- The parameter variances (tau) and its correlations
- The chi-square values to test whether the parameter variance is zero
- The standard errors for the variance-covariance components (full maximum likelihood estimates)

- Multivariate hypothesis testing for fixed effects

11.2.2.2 Standard error of the gammas

The standard error of the averaged fixed effects (gammas) is estimated as described below. The Student's *t*-value is calculated by dividing the average gamma by its standard error, and the probability of the *t*-value is estimated from a standard *t*-distribution table.

The standard error of the gammas consists of two components – sampling error and measurement error. The following routine provided in the NAEP *Data Files User Guide* (Rogers, *et al.*, 1992) is used to approximate the component of error variance due to the error in imputations and to add it to the sampling error.

Let $\hat{\theta}_m \; (m = 1,...,M)$ represent the *m*-th plausible value. Let $\hat{t}_m$ represent the parameter estimate based on the *m*-th plausible value. Let U_m represent the estimated variance of $\hat{t}_m$.

- Five HLM runs were conducted based on each plausible value $\hat{\theta}_m$. The parameter estimates from these runs were averaged:

$$t^* = \frac{\sum_{m=1}^{M} \hat{t}_m}{M} \tag{0.0}$$

- The variances of the parameters from these runs were averaged:

$$U^* = \frac{\sum_{m=1}^{M} U_m}{M} \tag{0.0}$$

- The variance of the *m* estimates, $\hat{t}_m$, was estimated:

$$B_m = \frac{\sum_{m=1}^{M} \left(\hat{t}_m - t^*\right)^2}{\left(M-1\right)} \tag{0.0}$$

- The final estimate of the variance of the parameter estimate is the sum of the two components:

$$V = U^* + \left(1 + M^{-1}\right) B_m \tag{0.0}$$

where the degrees of freedom is computed:

$$d.f. = (M-1)(1+r)^2,$$

where

$$r = \frac{U^*}{B\left(1+\frac{1}{M}\right)}.$$

The square root of this variance is the standard error of the gamma, and it is used in a standard Student's t formula to evaluate the statistical significance of each gamma.

11.2.3 Working with plausible values in HLM

Below is the procedure for running a plausible value analysis via Windows mode:

To run a plausible value analysis

1. After specifying the model, select the **Estimation Settings** option from the **Other Settings** menu.
2. Choose **Plausible Values** to open the **Select Plausible Value Outcome Variables** dialog box (See Figure 11.5 for an example).
3. Select the first plausible value (either the outcome or a covariate) from the **Choose first variable from level 1 equation** drop-down menu.
4. Double-click the other plausible values from the **Possible choices** box.
5. Click **OK**.

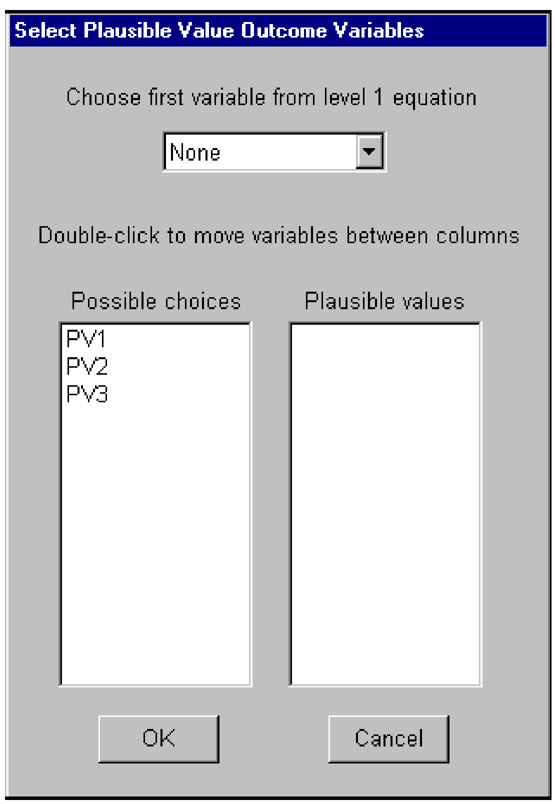


Figure 11.5 Select Plausible Value Outcome Variables dialog box

11.2.4 Data with multiply-imputed values for the outcome and covariates

There may be multiply-imputed values for both the outcome and the covariates. To apply HLM to such data, it is necessary to prepare as many MDM files as the number of imputed data sets. Thus, if there are five imputed data sets, five MDM files with identical variable labels need to be prepared. To run these models in batch mode, refer to Section F.3 in Appendix F.

Below are the commands for running an analysis with multiply-imputed data sets via Windows mode.

To run an analysis with multiply-imputed data sets

1. After specifying the model, select the **Estimation Settings** option from the **Other Settings** menu.
2. Choose **Multiple Imputation** to open the **Multiple Imputation MDM files** dialog box (See Figure 11.6 for an example).
3. Enter the names of the MDM files that contain the multiply-imputed data either by typing into the **File #** edit boxes or clicking **Browse** to open them.
4. Click **OK**. Model specification follows the usual format.

The calculations involved are very similar to the ones mentioned in Section 11.2.2.

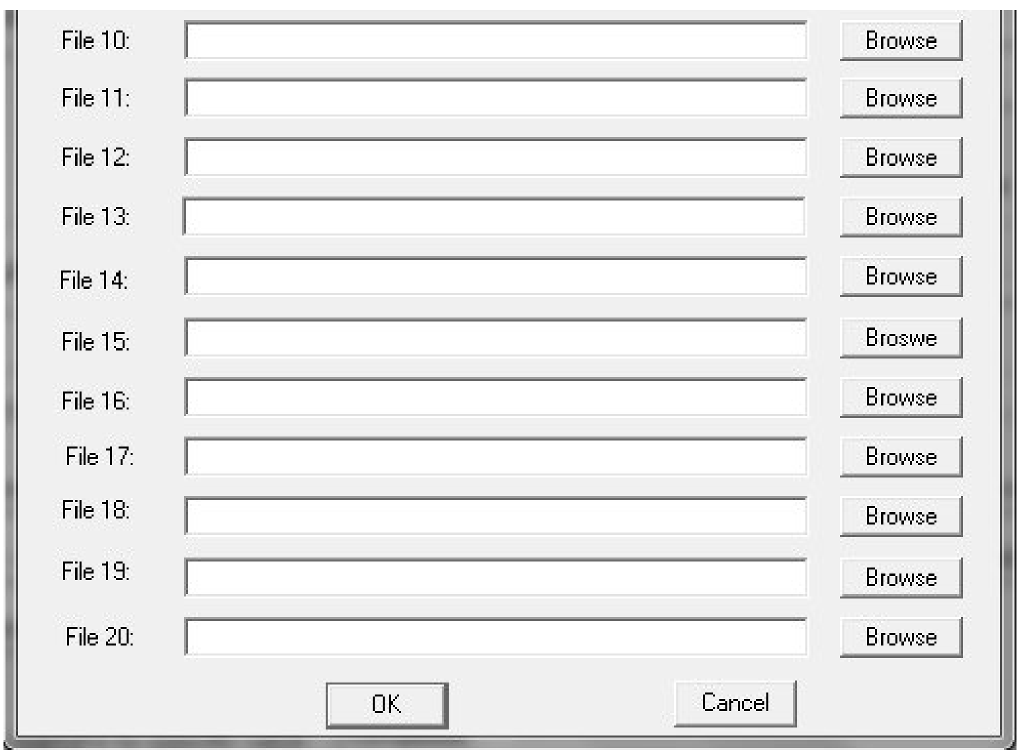


Figure 11.6 Multiple Imputation MDM files dialog box

11.3 "V-Known" models for HLM2

The V-known option in HLM2 is a general routine that can be used for applications where the level-1 variances (and covariances) are known. Included here are problems of meta-analysis (or research synthesis) and a wide range of other possible uses as discussed in Chapter 7 of *Hierarchical Linear Models*. The program input consists of Q random level-1 statistics for each group and their associated error variances and covariances.

We illustrate the use of the program with the following data from the meta-analysis of teacher expectancy effects described on pp. 210-216 of *Hierarchical Linear Models*. Here we show the process of V-known analysis in its most generic form, which requires using the interactive mode. See Section 11.3.4 for an easier alternative method for Q = 1 using the Windows interface.

```
 1     0.030      0.016   2.000
 2     0.120      0.022   3.000
 3    -0.140      0.028   3.000
 4     1.180      0.139   0.000
 5     0.260      0.136   0.000
 6    -0.060      0.011   3.000
 7    -0.020      0.011   3.000
 8    -0.320      0.048   3.000
 9     0.270      0.027   0.000
10     0.800      0.063   1.000
11     0.540      0.091   0.000
12     0.180      0.050   0.000
13    -0.020      0.084   1.000
14     0.230      0.084   2.000
15    -0.180      0.025   3.000
16    -0.060      0.028   3.000
17     0.300      0.019   1.000
18     0.070      0.009   2.000
19    -0.070      0.030   3.000
```

11.3.1 Data input format

Unlike the standard HLM2 program, the V-known routine uses only a single data input file. It consists of the following information:

1. The first field is the unit ID in character format.
2. This is followed by the Q statistics from each unit. In the teacher expectancy effects meta-analysis, Q equals one, the experiment effect size. (The effect size estimate appears in the third column of Table 7.1 in *Hierarchical Linear Models.*)
3. Next are the $Q(Q + 1)/2$ error variances and covariances associated with the set of Q statistics. These variance-covariance elements must be specified in row-column sequence from the lower triangle of the matrix, *i.e.*, $V_{11}, V_{12}, V_{22}, ..., V_{Q,Q-1}, V_{QQ}$. For the meta-analysis application only a single error variance is needed. (Note the values in the third column above are the squares of the standard errors that appear in the fourth column of Table 7.1.)
4. Last are the potential level-2 predictor variables. In the teacher expectancy effects meta-analysis, there was only one predictor, the number of weeks of prior contact. (See column 2 of Table 7.1).

The Q statistics, their error variances and covariances, and the level-2 predictors must be ordered as described above and have a numeric format.

11.3.2 Creating the MDM file

The V-known program must be implemented in batch or interactive mode; it is not available in Windows mode.

We present below an example of an HLM2 session that creates a multivariate data matrix file using the V-known routine on the teacher expectancy effects data.

```
C:\HLM>HLM2

Will you be starting with raw data?  y
Is the input file a v-known file? y
How many level-1 statistics are there? 1
How many level-2 predictors are there? 1
Enter 8 character name for level-1 variable number 1: EFFSIZE

Enter 8 character name for level-2 variable number 1:  WEEKS
Input format of raw data file (the first field must be the character ID)
format: (a2,3f12.3)
What file contains the data? expect.dat

Enter name of MDM file:  expect.MDM
19 groups have been processed
```

The file, EXPECT.DAT, contains the input data displayed above and the resulting multivariate data matrix are saved in the EXPECT.MDM file. Note that the input format has been specified for the character ID, the level-1 statistic (EFFSIZE), the associated variance, and the level-2 predictor (WEEKS).

11.3.3 Estimating a V-known model

Once the MDM file has been created, it can be used to specify and estimate a variety of models as in any other HLM2 application. The example below illustrates interactive use of the V-known program (example EXPECT.HLM).

```
C:\HLM>hlm2 expect.MDM

                    SPECIFYING AN HLM MODEL
Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
 For  EFFSIZE enter  1

     level-1 predictor? (Enter 0 to end)  1
```

```
Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
 For  EFFSIZE enter  1

     level-1 predictor? (Enter 0 to end)  1

Level-2 predictor variable specification
Which level-2 variables do you wish to use?

 The choices are:
 For    WEEKS enter  1

Which level-2 predictors to model  EFFSIZE?
      Level-2 predictor? (Enter 0 to end)  1

                       ADDITIONAL PROGRAM FEATURES

Select the level-2 variables that you might consider for
inclusion as predictors in subsequent models.
 The choices are:
 For    WEEKS enter  1

Which level-2 variables to model  EFFSIZE?
     Level-2 variable? (Enter 0 to end)  0

Do you want to run this analysis with a heterogeneous sigma^2? n
Do you wish to use any of the optional hypothesis testing procedures? n

                          OUTPUT SPECIFICATION
Do you want a residual file? n
How many iterations do you want to do? 10000
Do you want to see OLS estimates for all of the level-2 units? n
 Enter a problem title: Teacher expectancy meta-analysis
     Enter name of output file: expect.lis

Computing . . ., please wait

  Problem Title: Teacher expectancy meta-analysis

  The data source for this run  = expect.MDM
  The command file for this run =
  Output file name             = expect.lis
  The maximum number of level-2 units = 19
  The maximum number of iterations = 10000
  Method of estimation: restricted maximum likelihood
  Note: this is a v-known analysis

  The outcome variable is INTRCPT1

  The model specified for the fixed effects was:

 ----------------------------------------------------
   Level-1                 Level-2
   Effects                Predictors
-----------------------------------------------------
         EFFSIZE, B1      INTRCPT2, G10
                           WEEKS, G11

 The model specified for the covariance components was:
```

Variance(s and covariances) at level-1 externally specified

Tau dimensions
EFFSIZE slope

Summary of the model specified (in equation format)

Level-1 Model

Y1 = B1 + E1

Level-2 Model

B1 = G10 + G11*(WEEKS) + U1

STARTING VALUES

Tau(0)
EFFSIZE,B(null)   0.02004

Estimation of fixed effects
(Based on starting values of covariance components)

Fixed Effect	Coefficient	Standard Error	Approx. T-ratio	d.f.	P-value
For EFFSIZE, B1					
INTRCPT2, G10	0.433737	0.109700	3.954	17	0.001
WEEKS, G11	-0.168572	0.046563	-3.620	17	0.002

The value of the likelihood function at iteration 1 = -3.414348E+001
The value of the likelihood function at iteration 2 = -3.350241E+001
The value of the likelihood function at iteration 3 = -3.301695E+001
The value of the likelihood function at iteration 4 = -3.263749E+001
The value of the likelihood function at iteration 5 = -3.121675E+001
.
.
The value of the likelihood function at iteration 7849 = -2.979898E+001
The value of the likelihood function at iteration 7850 = -2.979898E+001
The value of the likelihood function at iteration 7851 = -2.979897E+001
The value of the likelihood function at iteration 7852 = -2.979897E+001

Iterations stopped due to small change in likelihood function

******* ITERATION 7853 *******

Tau
EFFSIZE,B   0.00001

Tau (as correlations)
EFFSIZE,B   1.000

Random level-1 coefficient	Reliability estimate
EFFSIZE, B	0.000

The value of the likelihood function at iteration 7853 = -2.979897E+001

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For EFFSIZE, B1					
INTRCPT2, G10	0.408572	0.087146	4.688	17	0.000
WEEKS, G11	-0.157963	0.035943	-4.395	17	0.000

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
EFFSIZE,	U	0.00283	0.00001	17	16.53614	>.500

Statistics for current covariance components model

Deviance = 59.59795
Number of estimated parameters = 2

In general, the HLM2 results for this example closely approximate the more traditional results that would be obtained from a graphical examination of the likelihood function. (For this particular model, the likelihood mode is at zero.) Note that the value of the likelihood was still changing after 7850 iterations. Often, HLM2 converges after a relatively small number of iterations. When the number of iterations required is large, as in this case, this indicates that the estimation is moving toward a boundary condition (in this example it is a variance estimate of zero for Tau). This can be seen by comparing the starting value estimate for Tau, 0.02004, with the final estimate of 0.00001. (For a further discussion see p. 202 of *Hierarchical Linear Models.*)

11.3.4 V-known analyses where Q = 1

There is an alternative and appealing method for analysis for V-known analyses when Q =1. This may be accomplished as follows:

1. Select the **Estimation Settings** option from the **Other Settings** menu.
2. Use the pull down menus to select the variable that represents the known level-1 variance.

This may be accomplished in either the two-level or the three-level HLM programs.

11.4 Spatial dependence models for HLM2

The spatial dependence option in HLM2 allows researchers to handle nested data collected in spatial settings. In addition to the clustering effects, the spatial HLM2 models accommodate dependence induced by contiguity or proximity in geographic locations. This type of models has applications for clustered data collected from contiguous geographic locations such as school districts, counties, neighborhoods, and countries. Verbitsky-Savitz and Raudenbush (2009), for example, applied these models to exploit the spatial dependence of neighborhood social processes to considerably improve the precision and validity of assessment of neighborhoods.

Below is an example of a spatial HLM2 model.

11.4.1 A spatial analysis using HLM2

This example uses data collected by the Project of Human Development in Chicago Neighborhoods (Sampson, Raudenbush, & Earl, 1997) on 7,729 residents living in 342 neighborhoods. It is an unconditional model with a ten-item collective efficacy scale, defined as the fusion of social cohesion and informal social control, as the outcome.

For spatial HLM2 models, the level-1 and level-2 models have the same structure as those described in Section 2.5. These two data files for the example, linked by level-2 neighborhood cluster IDs, are RESIDENT.SAV and NEIGHBOR.SAV. In the level-1 data file, there is one variable, collective efficacy (COLLEFF). In the level-2 data file, a dummy variable is included. The spatial dependence analysis requires another data file with information on spatial contiguity. The information allows the program to create a spatial weight matrix, W, which is a binary contiguity matrix indicating that sites are contiguous to each other. ROOK.SAV, contains such information for our illustrative example. The variables followed by the neighborhood cluster IDs are:

- N1 – N10 (the first to the tenth adjoining neighborhoods, if any)
- COUNT (the total number of contiguous neighborhoods)

The data for the first ten neighborhoods are displayed in Fig 11.7. Note that neighborhood 1 (that is, the neighborhood with ID = 1) shares a common boundary with two neighborhoods, specifically, neighborhoods 2 and 3. In contrast, neighborhood 2 shares a boundary with 4 neighborhoods, neighborhoods 8, 6, 3, and 1.

	id	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	count
1	1	2	3	.	.	.	.	.	.	.	.	2
2	2	8	6	3	1	.	.	.	.	.	.	4
3	3	16	5	4	2	1	.	.	.	.	.	5
4	4	17	5	3	.	.	.	.	.	.	.	3
5	5	7	6	3	4	17	.	.	.	.	.	5
6	6	8	7	5	2	.	.	.	.	.	.	4
7	7	15	11	5	17	6	.	.	.	.	.	5
8	8	11	10	2	6	.	.	.	.	.	.	4
9	9	334	252	250	241	238	.	.	.	.	.	5
10	10	13	12	8	.	.	.	.	.	.	.	3

Figure 11.7 First ten cases in ROOK.SAV

The file SPATIAL.MDMT stores the commands for creating the two-level multivariate data matrix file, SPATIAL.MDM. The procedure is very similar to those described in Section 2.5.1. An extra step needed is to instruct the program to include spatial dependence information with the following procedure:

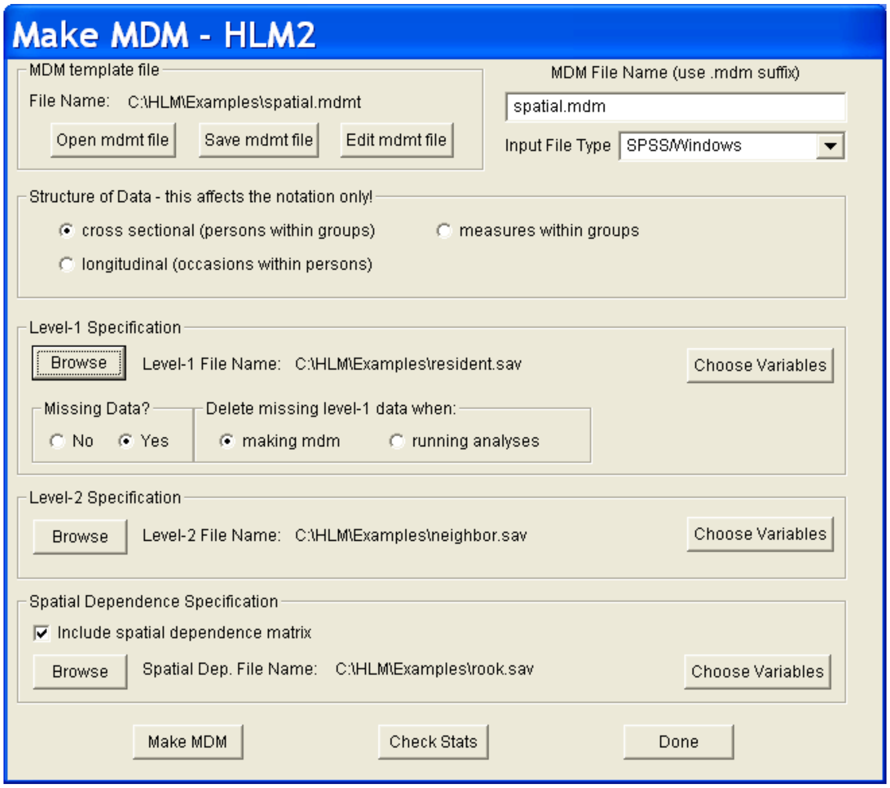


Figure 11.8 Make HLM – Dialog Box

At the **Make MDM – HLM2** dialog box,

1. Check the box for **Include spatial dependence matrix** (see Figure 11.8).
2. Click **Browse** to select ROOK.SAV.
3. Click **Choose Variables** to include the ID and the variables N1-N10 and COUNT.

The file SPATIAL.HLM contains the commands for setting up the unconditional model. The procedure follows the steps outlined and illustrated in Section 2.5.2.5. An additional step is to instruct HLM2 to run the model as a spatial dependence model by the following procedure:

1. Open the **Other Settings** menu and select the **Estimation Settings**.
2. Check the box for **Run as spatial dependence model** (see Figure 11.9).

5 An exception is that only the intercept (β_0) can be specified as random.

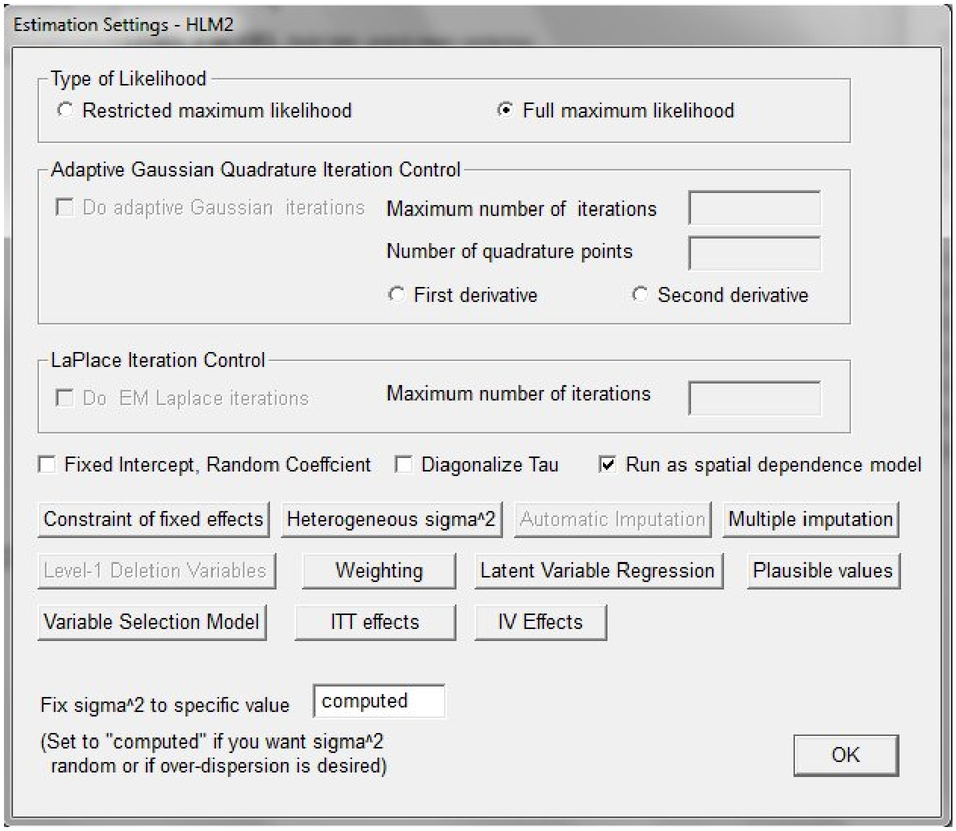


Figure 11.9 Estimation Settings – HLM2

The model window for our illustrative example in Figure 11.10 gives the model specifications.

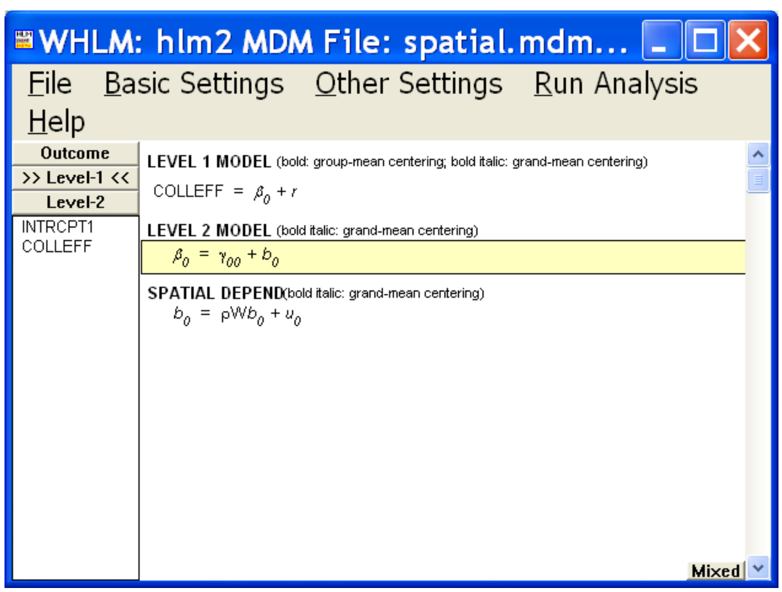


Figure 11.10 Specification of the spatial dependence uncondtional model

Note that a model for the spatial dependence model is given:

$$b_0 = \rho W b_0 + u$$

where as described in Verbitsky-Savitz and Raudenbush (2009),

- b_0 is a vector of level-2 random spatially autoregressive effects;
- ρ is a spatial correlation parameter with zero indicating no spatial dependence and positive or negative values indicating whether a site is typically surrounded by other sites with similar or different values on the outcome;
- W is the spatial weight matrix used in the analysis. As discussed earlier, it is constructed from ROOK.SAV; and
- u is the level-2 error.

A spatial dependence analysis using HLM2 provides two sets of results, one for regular HLM and the other HLM with spatial dependence. A comparison test of the fit of these models is performed and the result is given. Below is a partial output of the results of the unconditional model.

Here are the partial results for the regular HLM:

Iterations stopped due to small change in likelihood function

σ^2 = 0.42136

Standard error of σ^2 = 0.00693

τ

INTRCPT1,β_0 0.08904

Standard error of τ

INTRCPT1,β_0 0.00850

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.799

The value of the log-likelihood function at iteration 6 = -7.911855E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	3.433243	0.018056	190.142	341	<0.001

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	3.433243	0.018056	190.144	341	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, u_0	0.29839	0.08904	341	1870.37148	<0.001
level-1, r	0.64913	0.42136			

Statistics for the current model

Deviance = 15823.710765
Number of estimated parameters = 3

HLM with Spatial Dependence Model Results - Iteration 135
The value of the log-likelihood function at iteration 135 = -7.835990E+003
Iterations stopped due to small change in likelihood function

σ^2 = 0.42149

τ

INTRCPT1,β 0.03477

ρ

INTRCPT1,β 0.81701

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	3.404181	0.056443	60.312	341	<0.001

Statistics for the current model

Deviance = 15671.980461
Number of estimated parameters = 4

Regular HLM vs. HLM with spatial dependence model comparison test

χ2 statistic = 151.73030
Degrees of freedom = 1
p-value = <0.001

Average Level-2 Variance = 0.088502

Average Level-2 Covariance = 0.005961

The average level-2 variance is the average of the neighborhood-specific variance. These depend on τ, but also on the magnitude of the spatial dependence correlation, ρ, and the configuration of neighborhoods near that neighborhood. The average level-2 covariance is the average covariance between pairs of contiguous neighborhoods.

Two features of the results are noteworthy:

- The result of the comparison test provides evidence that the HLM with spatial dependence provides a better fit, as indicated by the χ^2 statistic of 151.73, $df = 1$, $p < .001$.
- A comparison of the standard errors for $\hat{\gamma}_{00}$ the regular HLM and HLM with spatial dependence (.018 vs .056) suggests that, given $\hat{\rho}$ is equal to .8, that there is an underestimation of the standard errors when spatial dependence is ignored.

Users can also obtain spatial empirical Bayes estimates of the neighborhood collective efficacy measures by following the procedure as specified in Section 2.5.4.2. Figure 11.11 gives the ten records of the residual file for the uncondtional model.

	l2id	nj	u_intrcp	b_intrcp	fvintrcp
1	1	39	-.213	-.420	3.404
2	2	43	-.168	-.292	3.404
3	3	15	.021	-.214	3.404
4	4	41	-.128	-.203	3.404
5	5	15	.101	-.021	3.404
6	6	13	-.099	-.155	3.404
7	7	39	-.113	-.138	3.404
8	8	14	.192	.178	3.404
9	9	45	.287	.328	3.404
10	10	17	.281	.376	3.404

Data View | Variable View

Figure 11.11 Level-2 Residual File

U_INTRCP and B_INTRCP are the two Empirical Bayes for the regular HLM and the HLM with spatial dependence. For a discussion of the properties of the empirical Bayes estimator that exploits spatial dependence, see Verbitsky-Savitz and Raudenbush (2009).

11.4.2 Other outcome variables

Spatial dependence models handles continuously distributed as well as discrete outcomes, including binary outcomes, counted data, ordered categories, and multinomial outcomes.

12 Conceptual and Statistical Background for Cross-classified Random Effect Models (HCM2)

All of the applications discussed thus far have involved a strictly hierarchical data structure. Such nesting structures would occur, for example, in a study of neighborhood and school effects on child development in which all children living in the same neighborhood attended the same school, with multiple neighborhoods per school. In this case we would have children at level 1 nested within neighborhoods at level 2 and neighborhoods nested within schools at level 3. Alternatively, we might have a nested structure in which every child attending a given school lived in the same neighborhood, with multiple schools per neighborhood. In this case, we would have children nested within schools nested within neighborhoods. HLM3 can be used to accommodate such three-level nested data structures. However, we typically find, in fact, that children who reside in a specific neighborhood can enroll in one of several schools, and each school might draw students from several neighborhoods. In this case, the data gathered will no longer have a purely nested structure. Instead, a cross-classification of students by two higher-level factors, neighborhoods and schools, arises. To handle this more complex data structure while modeling the developmental influences of neighborhoods and schools requires the use of cross-classified random effects models (HCM2).

Chapter 12 of *Hierarchical Linear Models* discusses two applications of cross-classified random effects models, one with cross-sectional, and the other with longitudinal data. The first application is from a study of neighborhood and school effects on educational attainment in Scotland (Garner & Raudenbush, 1991). Some of the children in this study enrolled in schools located in neighborhoods that were different from the ones they resided in. These students were thus cross-classified by neighborhoods and schools. The second case is an assessment of the effects of classrooms on children's cognitive growth during the primary school years (Raudenbush, 1993) using longitudinal data collected from the Immersion Study (Ramirez, Yuen, Ramey, & Pasta, 1991). As there were changes in classroom memberships among the students during the course of the investigation, the repeated assessments on cognitive growth were cross-classified by teachers. A similar data structure was displayed in Sampson, Sharkey and Raudenbush's (2008) longitudinal study on the impact of concentrated disadvantage on the verbal ability of African American children. During the seven years of data collection, some of the participants moved to live in different neighborhoods. Consequently, the repeated measures of their verbal ability were cross-classified by children and time-varying neighborhoods.

12.1 The general cross-classified random effects models

A general random cross-classified model consists of two sub-models: the level-1 or within-cell model and level-2 or between-cell model. The cells refer to the cross-classifications by the two higher-level factors. For example, if the research problem consists of data on students cross-classified by neighborhoods and schools, a cell consists of a set of students who live in the same neighborhood and attend the same school. The level-1 or within-cell model will represent the relationships among the student-level variables for those students while the level-2 or between-cell model will capture the influences of neighborhood- and school-level factors. Formally, there are $i = 1, 2, ..., n_{jk}$ level-1 units (*e.g.*, students) nested within each cell cross-classified by $j = 1,..., J$

units of the first higher-level factor (*e.g.*, neighborhoods), designated as rows, and $k = 1,..., K$ units of the second higher-level factor (*e.g.*, schools), designated as columns. For a graphical representation of this data layout in Garner and Raudenbush (1991), see Table 12.1 in Chapter 12 of *Hierarchical Linear Models.*

In HLM7, HCM2 handles continuously distributed as well as discrete outcomes, including binary outcomes, counted data, ordered categories, and multinomial outcomes. We use the continuous outcome models in the following discussion. The logic of HGLM, as described and illustrated in Chapter 5, applies and extends to analyses with any of the four types of discrete outcomes with HCM2.

12.1.1 Level-1 or "within-cell" model

We represent in the level-1 or within-cell model the outcome for case i nested within row j and column k of the two-way classification:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}a_{1ijk} + \pi_{2jk}a_{2ijk} + \ldots + \pi_{pjk}a_{pijk} + e_{ijk} \tag{12.1}$$

where

π_{0jk} is the intercept, the expected value of Y_{ijk} within cell jk when all explanatory variables are set to zero;

π_{pjk} are the level-1 coefficients of predictors a_{pijk}, for $p = 1,\ldots,P$;

e_{ijk} is the level-1 or within-cell random effect; and

σ^2 is the variance of e_{ijk}, that is the level-1 or within-cell variance. Here we assume that the random term $e_{ijk} \sim N(0,\sigma^2)$.

12.1.2 Level-2 or "between-cell" model

Each of the π_{pjk} coefficients in the level-1 or within-cell model becomes an outcome variable in the level-2 or between-cell model:

$$\begin{aligned}\pi_{pjk} = {} & \theta_{p0} + \left(\beta_{p1} + b_{p1j}\right)X_{1k} + \left(\beta_{p2} + b_{p2j}\right)X_{2k} + \ldots + \left(\beta_{pQ_p} + b_{pQ_pj}\right)X_{Q_pk} + \\ & \left(\gamma_{p1} + c_{p1k}\right)W_{1j} + \left(\gamma_{p2} + c_{p2k}\right)W_{2j} + \ldots + \left(\gamma_{pR_p} + c_{pR_pk}\right)W_{R_pj} + \\ & b_{p0j} + c_{p0k}\end{aligned} \tag{12.2}$$

where

θ_{p0} is the model intercept, the expected value of π_{pjk} when all explanatory variables are set to zero;

β_{pq} are the fixed effects of column-specific predictors X_{qk}, $q = 1,...,Q_p$;

b_{pqj} are the random effects associated with column-specific predictors X_{qk}. They vary randomly over rows $j = 1,..., J$;

γ_{pr} are the fixed effects of row-specific predictors $W_{rj}, r = 1,..., R_p$;

c_{prk} are the random effects associated with row-specific predictors W_{rj}. They vary randomly over columns $k = 1,\ldots, K$, and;

b_{p0j} and c_{p0k} are residual row and column random effects, respectively, on π_{pjk}, after taking into account X_{qk} and W_{rj}. We assume that $b_{p0j} \sim N(0, \tau_{pb00})$, $c_{p0k} \sim N(0, \tau_{pc00})$, and that the effects are independent of each other.

The vector of random row effects b_{pqj} ($p = 0,\ldots,P$; $q = 0,\ldots,Q_p$;) is assumed multivariate normal with a mean zero and a full covariance matrix τ. Similarly the vector of random column effects c_{prk} ($p = 0,\ldots,P$; $r = 0,\ldots,R_p$;) is assumed multivariate normal with mean vector zero and full covariance matrix Δ.

12.2 Parameter estimation

For continuous outcomes, three kinds of parameter estimates are available in HCM2: empirical Bayes estimates of random coefficients; maximum-likelihood estimates of the fixed regression coefficients; and maximum likelihood estimate of the variance-covariance components. The estimation procedure uses a full maximum likelihood approach (Raudenbush, 1993).

For discrete outcomes, the parameter estimates of the fixed regression coefficients are based on the method of penalized quasi-likelihood. Unlike HGLM, however, unit-specific but not population-averaged results are available.

12.3 Hypothesis testing

As in the case of HLM2, HCM2 routinely prints standard errors and *t*-tests for each of the fixed level-2 coefficients as well as a chi-square test of homogeneity for each random effect. In addition, optional "multivariate hypothesis tests" are available in HCM2. Multivariate tests in the case of continuous outcomes parallel those described in Section 2.8.8. For discrete outcomes, hypothesis testing parallels those described in Section 5.10.

13 Working with HCM2

13.1 An example using HCM2 in Windows mode

HCM2 analyses can be executed in Windows, interactive, and batch modes. We describe a Windows execution below. We consider interactive and batch execution in Appendix G. A number of special options are presented at the end of the chapter.

Chapter 12 in *Hierarchical Linear Models* presents a series of analyses of data from a study of neighborhood and school contribution to educational attainment in Scotland (Garner & Raudenbush, 1991). We use the data from the study, provided along with the HLM software, to illustrate the operation of the HCM2 program.

13.1.1 Constructing the MDM file from raw data

In constructing the MDM file, there are the same range of options for data input as for HLM2. Similar to HLM3, HCM2 requires two IDs, one for each higher-level unit, and the IDs have to be sorted. The two higher-level units in our example are neighborhoods and schools. Whereas the user can choose either higher-level unit as the row or column factor, we adopt the convention that the data are arranged such that the level with more units becomes the row factor and the level with fewer units becomes the column factor. Thus, we will designate the neighborhood ($N = 542$) as the row factor and school ($N = 17$) as the column factor.

13.1.1.1 SPSS input

Data input requires a level-1 file (student-level file), a level-2 row-factor (neighborhood-level) file, and a level-2 column-factor (school-level) file.

Level-1 file. The level-1 or within-cell file, ATTAINW.SAV has 2,310 students and 8 variables. The two IDs are NEIGHID for neighborhoods and SCHID for schools. The variables are:

- ATTAIN (a measure of educational attainment)
- P7VRQ (Primary 7 verbal reasoning quotient)
- P7READ (Primary 7 reading test scores)
- DADOCC (father's occupation scaled on the Hope-Goldthorpe scale in conjunction with the Registrar General's social-class index (Willms, 1986))
- DADUNEMP, an indicator for father's unemployment status (1 if unemployed, 0 otherwise)
- DADED, an indicator for father's educational level (1 if schooling past the age of 15, 0 otherwise)
- MOMED, an indicator for mother's educational level (1 if schooling past the age of 15, 0 otherwise)
- MALE, an indicator for student gender (1 if male, 0 if female)

Data for the first 15 observations are shown in Figure 13.1. Note that five students from Neighborhood 26 and one from Neighborhood 27 attended School 20. These first six observations provided information about two neighborhood-by-school combinations or cells. One of the next nine students living in Neighborhood 29 attended School 18 and the other eight went to School 20. They provided data for two cross-classified neighborhood-by-school cells (see Table 12.1 in *Hierarchical Linear Models*, p. 374, for a display of the organization of the data by counts in each neighborhood-by-school cell).

	neighid	schid	attain	p7vrq	p7read	dadocc	dadunemp	daded	momed	male
1	26.00	20.00	-1.33	-1.03	-.87	-3.45	.00	.00	.00	1.00
2	26.00	20.00	-1.33	-10.03	-27.87	-3.45	.00	.00	.00	.00
3	26.00	20.00	1.52	1.97	11.13	-9.09	.00	.00	.00	.00
4	26.00	20.00	.56	2.97	6.13	2.32	.00	.00	.00	.00
5	26.00	20.00	1.52	17.97	17.13	16.20	.00	.00	.00	1.00
6	27.00	20.00	-.13	3.97	-.87	-3.45	.00	.00	1.00	.00
7	29.00	18.00	.03	8.97	6.13	16.20	.00	1.00	1.00	1.00
8	29.00	20.00	-1.33	-17.03	-23.87	-3.45	.00	.00	.00	1.00
9	29.00	20.00	-1.33	-8.03	-4.87	-3.45	.00	.00	.00	.00
10	29.00	20.00	.16	1.97	-4.87	-11.49	.00	.00	.00	.00
11	29.00	20.00	.56	-.03	-5.87	-3.45	.00	.00	.00	.00
12	29.00	20.00	.16	4.97	11.13	-11.49	.00	.00	.00	1.00
13	29.00	20.00	.74	-4.03	.13	-3.45	.00	.00	.00	1.00
14	29.00	20.00	-1.33	-15.03	-25.87	-3.45	.00	.00	.00	.00
15	29.00	20.00	-.36	-8.03	-13.87	-3.45	.00	.00	.00	.00

Figure 13.1 First 16 cases in the ATTAINW.SAV dataset

13.1.1.2 Level-2 row-factor file

For our neighborhood example, the level-2 row-factor (neighborhood) level file, ATTAINR.SAV, consists data on 1 variable for 542 neighborhoods. The variable is DEPRIVE (a scale measuring social deprivation, which incorporates information on the poverty concentration, health, and housing stock of a local community).

Figure 13.2 shows data from the first 4 neighborhoods.

	neighid	deprive
1	26	-.55
2	27	.15
3	28	-.51
4	29	-.08

Data View

Figure 13.2 First 4 cases in the ATTAINR.SAV data set

13.1.1.3 Level-2 column-factor file

The level-2 column-factor (neighborhood) file, ATTAINCO.SAV, has 17 schools and 1 variable. The variable is DUMMY, a dummy variable. Figure 13.3 shows data for the first 4 schools.

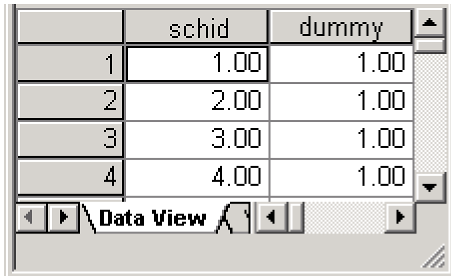


Figure 13.3 First 4 cases in the ATTAINCO.SAV data set

The steps for the construction of the MDM for HCM2 are similar to the ones described earlier. Select **HCM2** in the **Select MDM type** dialog box (see Figure 2.5). Note that the program can handle missing data at level 1 or within-cell only. The MDM template file, ATTAIN.MDMT, contains a log of the input responses used to create the MDM file, ATTAIN.MDM, using ATTAINW.SAV, ATTAINR.SAV, and ATTAINCO.SAV. Figure 13.4 displays the dialog box used to create the MDM file. Figures 13.5 to 13.7 show the dialog boxes for the within-cell file, ATTAINW.SAV, the row-factor file, ATTAINR.SAV, and the column-factor file, ATTAINCO.SAV.

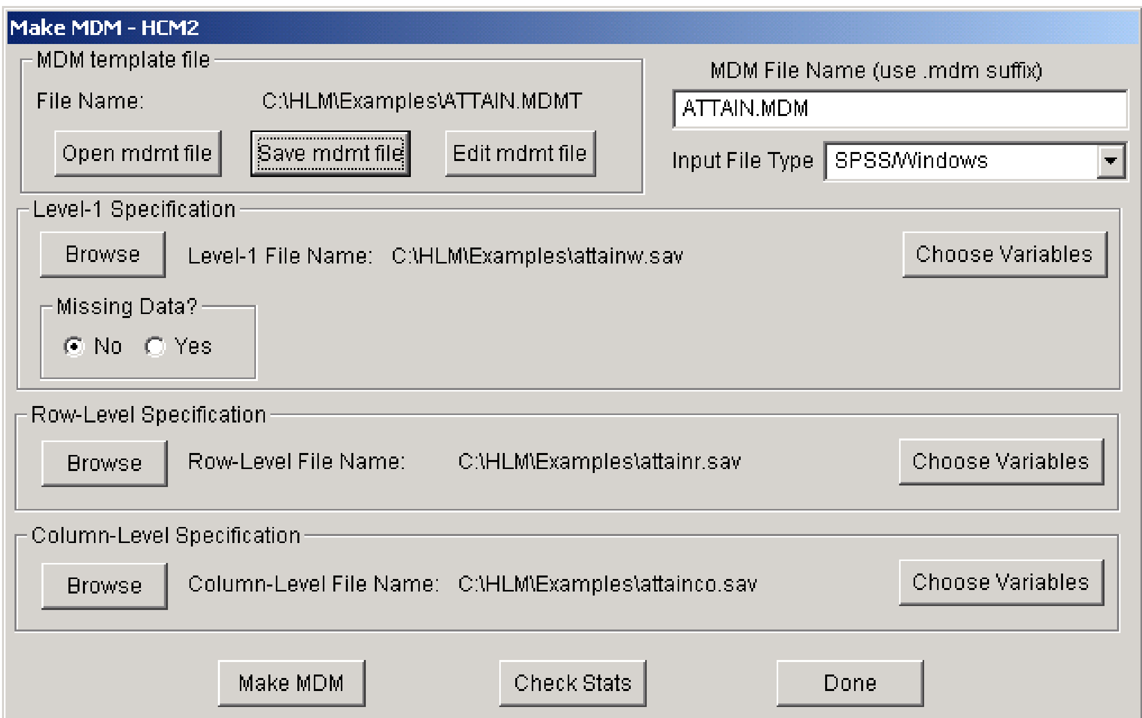


Figure 13.4 Make MDM – HCM2 dialog box for ATTAIN.MDMT

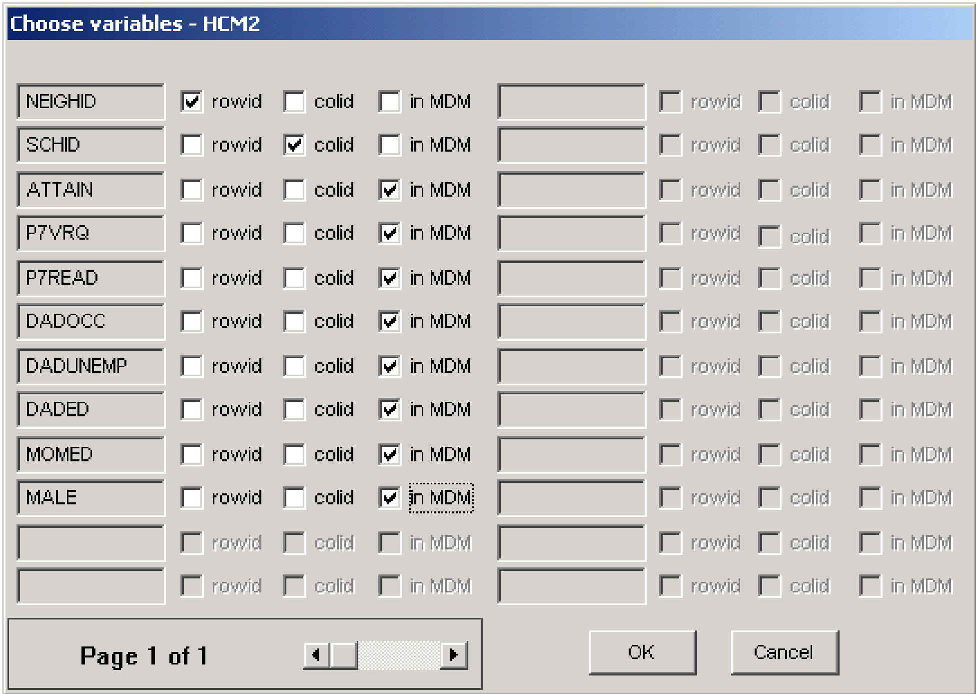


Figure 13.5 Choose variables – HCM2 dialog box for level-1 or within-cell file, ATTAINW.SAV

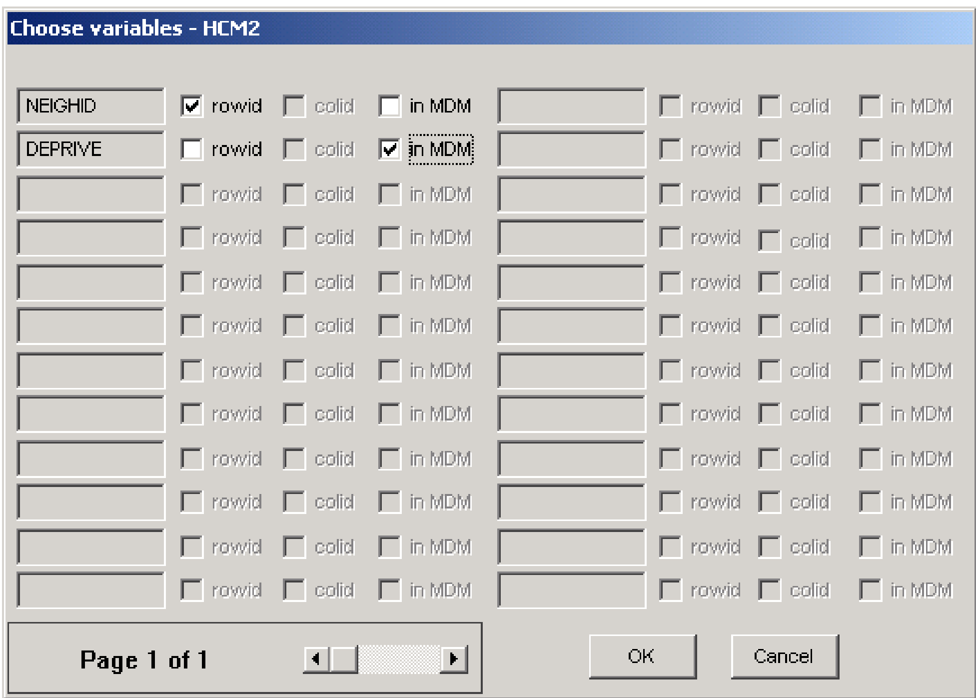


Figure 13.6 Choose variables – HCM2 dialog box for level-1 or row-factor file, ATTAINR.SAV

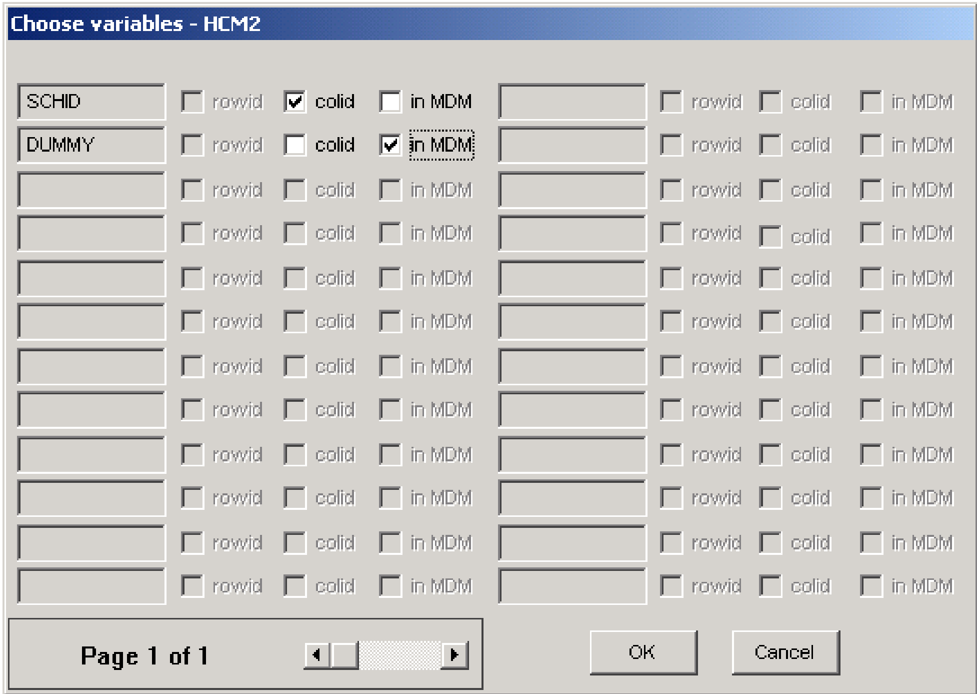


Figure 13.7 Choose variables – HCM2 dialog box for level-1 or column-factor file, ATTAINCO.SAV

13.2 Executing analyses based on the MDM file

Once the MDM file is constructed, it can be used as input for the analysis. Model specification has three steps:

1. Specification of the level-1 or within-cell model. In our example, we shall model educational attainment (ATTAIN) as the outcome. We first formulate an unconditional model that includes no predictor variables at any level. In the second or conditional model, we use prior measures of cognitive skill, verbal reasoning quotient and reading achievement, father's employment status and occupation and father's and mother's education to predict attainment.
2. Specification of the row- or column-factor prediction model. In the second or conditional model, we shall predict each student's intercept with social deprivation.
3. Specification of the residual row, column, and cell-specific effects as random or non-random, the effects associated with row-specific predictors as varying randomly or fixed over columns, and the effects associated with column-specific predictors as varying randomly or fixed over rows. We shall test whether the association between social deprivation (a row-specific predictor) and attainment varies over schools in the third model.

Following the three steps above, we first specify a model with no student-, neighborhood-, or school-level predictors. The purpose is to estimate the components of variation that lie between neighborhoods, between schools, and within cells.

1. From the **WHLM** window, open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open the existing MDM file (ATTAIN.MDM in our example).
3. Click on the name of the outcome variable (ATTAIN in our example). Click **Outcome variable**. The specified model will appear in equation format (see Figure 13.8).

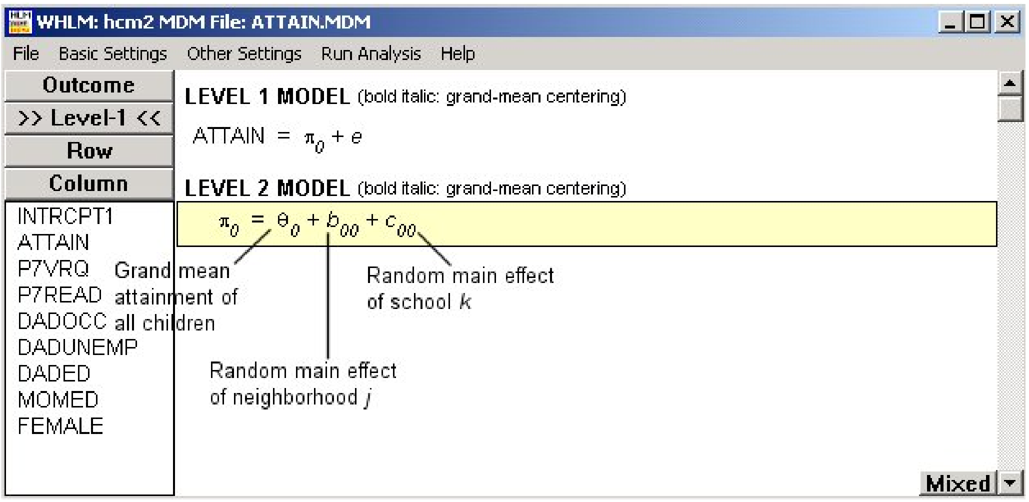


Figure 13.8 Unconditional model for the attainment example

The results of the analysis are given below.

Problem Title: Unconditional model

The data source for this run = ATTAIN.MDM
The command file for this run = attain1.hlm
Output file name = hcm2.html
The maximum number of level-1 units = 2310
The maximum number of row-level units = 524
The maximum number of column-level units = 17
The maximum number of iterations = 100

Method of estimation: full maximum likelihood
The maximum number of iterations = 100
Z-structure: independent

The outcome variable is ATTAIN

Summary of the model specified

Level-1 Model

$$\text{ATTAIN}_{ijk} = \pi_{0jk} + e_{ijk}$$

Level-2 Model

$$\pi_{0jk} = \theta_0 + b_{00j} + c_{00k}$$

For starting values, data from 2310 level-1, 524 row-level and 17 column-level records were used

Final Results - iteration 21

Iterations stopped due to small change in likelihood function

σ^2 = 0.79909

τ_{rows}

INTRCPT1
ICPTROW,b_{00j}
0.14105

τ_{columns}

INTRCPT1
ICPTCOL,c_{00k}
0.07546

The intra-neighborhood correlation, the correlation between outcomes of two students who live in the same neighborhood but attend different schools, is estimated to be:

$$
\begin{aligned}
C\hat{o}rr\left(Y_{ijk}, Y_{ijk'\varepsilon}\right) &= \frac{\hat{\tau}_{b00}}{\hat{\tau}_{b00} + \hat{\tau}_{c00} + \hat{\sigma}^2} \\
&= \frac{0.141}{0.141 + 0.075 + 0.799} \\
&= 0.139.
\end{aligned}
$$

Thus, about 13.9% of the total variance lies between neighborhoods.

The intra-school correlation is the correlation between outcomes of two students who attend the same school but live in different neighborhoods:

$$
\begin{aligned}
C\hat{o}rr\left(Y_{ijk}, Y_{i'j'k}\right) &= \frac{\hat{\tau}_{c00}}{\hat{\tau}_{b00} + \hat{\tau}_{c00} + \hat{\sigma}^2} \\
&= \frac{0.075}{0.141 + 0.075 + 0.799} \\
&= 0.074.
\end{aligned}
$$

That is, about 7.4% of the variation lies within schools.

The intra-cell correlation is the correlation between outcomes of two students who live in the same neighborhood and attend the same school:

$$
\begin{aligned}
C\hat{o}rr\left(Y_{ijk}, Y_{ijk'\varepsilon}\right) &= \frac{\hat{\tau}_{b00} + \hat{\tau}_{c00}}{\hat{\tau}_{b00} + \hat{\tau}_{c00} + \sigma^2} \\
&= \frac{0.141 + 0.075}{0.141 + 0.075 + 0.799} \\
&= 0.212.
\end{aligned}
$$

Thus, according to the fitted model, about 21% of the variance lies between cells.

The value of the log-likelihood function at iteration 21 = -3.178356E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTERCEPT, θ_0	0.075357	0.072226	1.043	1769	0.297

Final estimation of row and level-1 variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTROW, b_{00j}	0.37556	0.14105	523	904.83225	<0.001
level-1, e	0.89392	0.79909			

Final estimation of column level variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTCOL, c_{00k}	0.27470	0.07546	16	120.45262	<0.001

Statistics for the current model

Deviance = 6356.711470
Number of estimated parameters = 4

13.3 Specification of a conditional model with the effect associated with a row-specific predictor fixed

The above example involves a model that is unconditional at all levels. In this model we set up a level-1 and a row-factor prediction model.

To set up the level-1 model:

At the model specification dialog box, select P7VCR, P7READ, DADOCC, DADUNEMP, DADED, MOMED, and MALE and grand-mean center all the predictors. Figure 13.9 shows the model with the level-1 predictors. In the interest of parsimony, given the small cell sizes and within-neighborhood sizes, all level-1 coefficients are fixed. (To specify any of them as randomly varying, select the equation containing a specific regression coefficient, π_p, and click on b_{p0}).

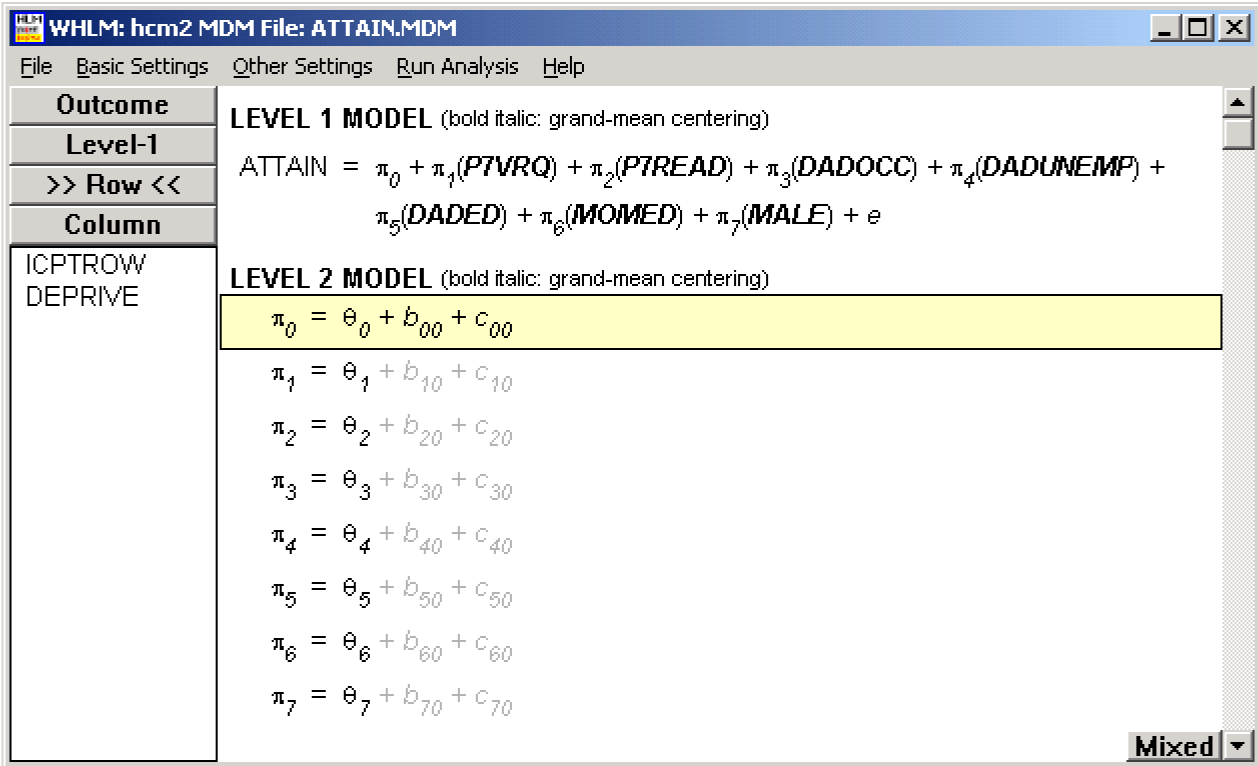


Figure 13.9 Level-1 Prediction Model for the Attainment Study

To set up the level-2 row-factor prediction model:

Select the equation containing π_0. A list box for row-factor variables (>>**Row**<<) will appear. Click DEPRIVE and apply the grand-mean centering scheme. In the level-2 model, we treated the association between social deprivation and educational attainment as fixed across all schools. We relax this assumption in our next model. Figure 13.10 displays the conditional model. Note that c_{01} is disabled.

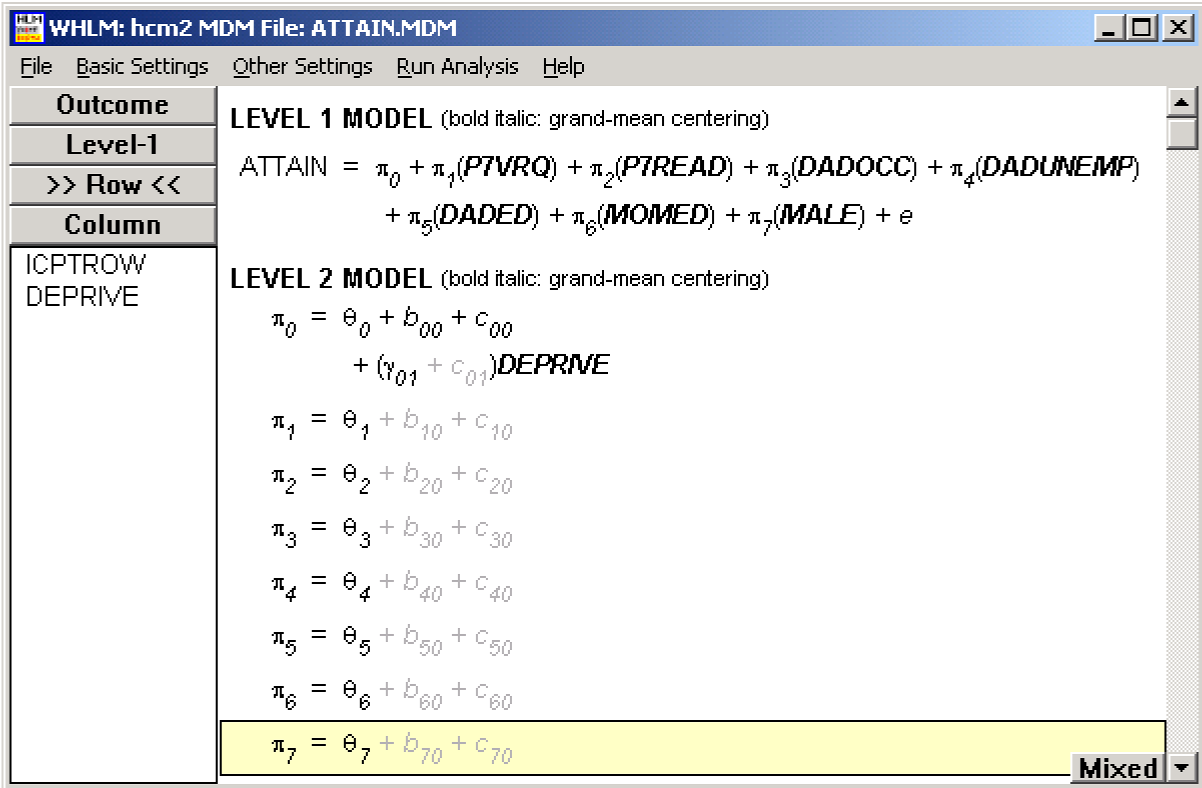


Figure 13.10 Conditional Model for the Attainment Study, with Social Deprivation Effect Fixed

The results of the analysis are given below.

Problem Title: Conditional Model, with social deprivation effect fixed
The data source for this run = ATTAIN.MDM
The command file for this run = ATTAIN2.hlm
Output file name = hcm2.html
The maximum number of level-1 units = 2310
The maximum number of row-level units = 524
The maximum number of column-level units = 17
The maximum number of iterations = 100
Method of estimation: full maximum likelihood
The maximum number of iterations = 100
Z-structure: independent

The outcome variable is ATTAIN

Summary of the model specified

Level-1 Model

$$\mathrm{ATTAIN}_{ijk} = \pi_{0jk} + \pi_{1jk}*(\mathrm{P7VRQ}_{ijk}) + \pi_{2jk}*(\mathrm{P7READ}_{ijk}) + \pi_{3jk}*(\mathrm{DADOCC}_{ijk}) + \pi_{4jk}*(\mathrm{DADUNEMP}_{ijk}) + \pi_{5jk}*(\mathrm{DADED}_{ijk}) + \pi_{6jk}*(\mathrm{MOMED}_{ijk}) + \pi_{7jk}*(\mathrm{MALE}_{ijk}) + e_{ijk}$$

Level-2 Model

$\pi_{0jk} = \theta_0 + b_{00j} + c_{00k} + (\gamma_{01})$*DEPRIVE$_j$
$\pi_{1jk} = \theta_1$
$\pi_{2jk} = \theta_2$
$\pi_{3jk} = \theta_3$
$\pi_{4jk} = \theta_4$
$\pi_{5jk} = \theta_5$
$\pi_{6jk} = \theta_6$
$\pi_{7jk} = \theta_7$

P7VRQ P7READ DADOCC DADUNEMP DADED MOMED MALE have been centered around the grand mean.
DEPRIVE has been centered around the grand mean.

For starting values, data from 2310 level-1, 524 row-level and 17 column-level records were used

Final Results - iteration 34

Iterations stopped due to small change in likelihood function

σ^2 = 0.45891

τ_{rows}
INTRCPT1
ICPTROW,b_{00j}
0.00014

$\tau_{columns}$
INTRCPT1
ICPTCOL,c_{00k}
0.00389

The value of the log-likelihood function at iteration 34 = -2.384802E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTERCEPT,θ_0	0.094740	0.021133	4.483	1769	<0.001
DEPRIVE, γ_{01}	-0.156676	0.025178	-6.223	522	<0.001
For P7VRQ, π_1					
INTERCEPT,θ_1	0.027556	0.002263	12.176	1769	<0.001
For P7READ, π_2					
INTERCEPT,θ_2	0.026291	0.001749	15.028	1769	<0.001
For DADOCC, π_3					
INTERCEPT,θ_3	0.008165	0.001359	6.008	1769	<0.001

For DADUNEMP, π_4					
INTERCEPT,θ_4	-0.120771	0.046779	-2.582	1769	0.010
For DADED, π_5					
INTERCEPT,θ_5	0.144426	0.040782	3.541	1769	<0.001
For MOMED, π_6					
INTERCEPT,θ_6	0.059440	0.037381	1.590	1769	0.112
For MALE, π_7					
INTERCEPT,θ_7	-0.056058	0.028401	-1.974	1769	0.049

Final estimation of row and level-1 variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTROW,b_{00j}	0.01184	0.00014	522	548.81015	0.202
level-1, *e*	0.67743	0.45891			

Final estimation of column level variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTCOL,c_{00k}	0.06239	0.00389	15	36.38151	0.002

Statistics for the current model

Deviance = 4769.604659
Number of estimated parameters = 12

Several features of the results are remarkable:

- Several level-1 covariates are significantly related to educational attainment, with especially large effects for P7READ and P7VRQ.
- The residual level-1 variance is estimated to be 0.459, implying that 43% of the unconditional level-1 variance (estimated at 0.799) is accounted for by the covariates.
- Controlling these level-1 effects, a highly significant negative effect of social deprivation appears ($\hat{\gamma}_{01}$ = -0.157, t = –6.22).
- The residual variation between neighborhoods, τ_{b00}, (estimated at 0.0001), and between schools, τ_{c00} (estimated at 0.004) are close to zero; compare to the unconditional variance estimates (0.141 and 0.075). The level-2 neighborhood variance component was substantially reduced.

13.4 Specification of a conditional model with the effect associated with the row-specific predictor random

In the previous model, the relationship between social deprivation and attainment was assumed invariant across schools. Now we test the tenability of this assumption.

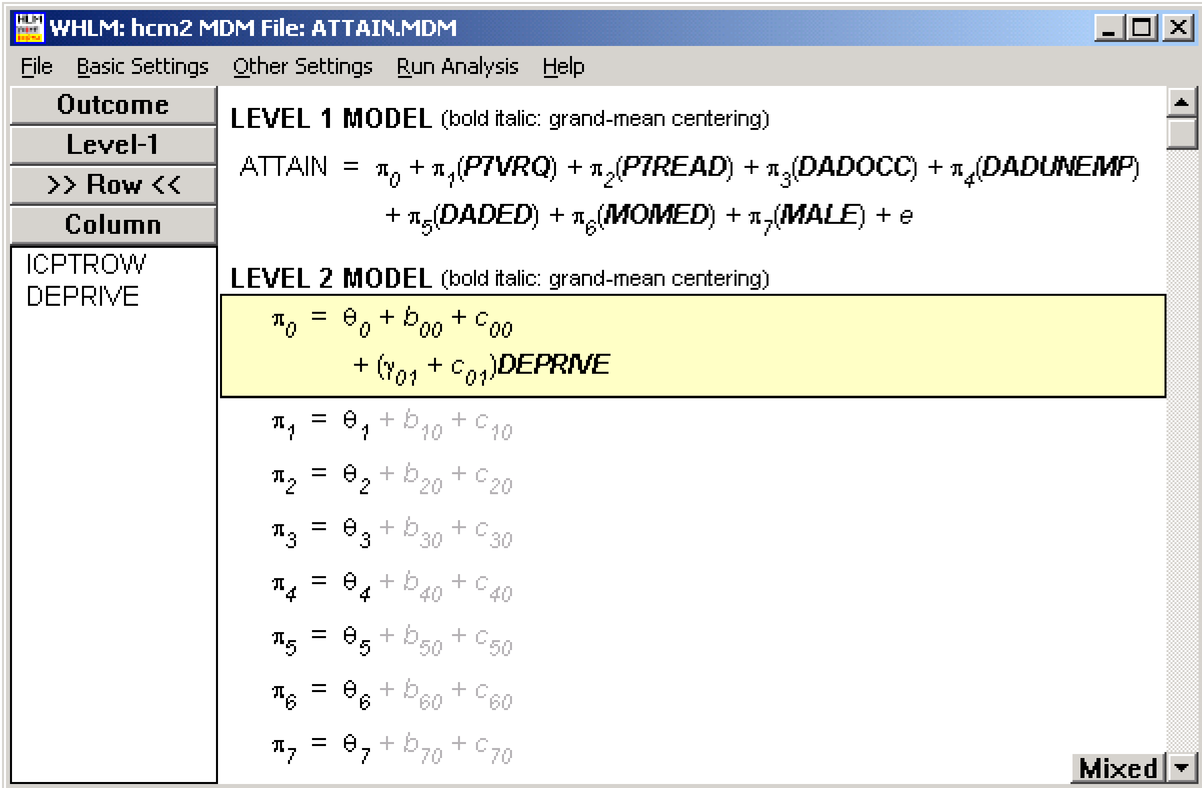


Figure 13.11 Conditional Model for the Attainment Study, with Social Deprivation Effect Random

To specify the effect of the row-specific predictor random, select the equation containing π_0. Click on c_{01}. Figure 13.11 displays the conditional model with the social deprivation effect specified as random. We compare the model deviance of this model against the one estimated in the last analysis. The procedure is the same as described in Section 2.9.6.

The results of the analysis are given below.

σ^2 = 0.45519

τ_{rows}

INTRCPT1 ICPTROW,b_{00j}
0.00371

τ_{columns}

INTRCPT1 ICPTCOL,c_{00k}	INTRCPT1 DEPRIVE,c_{01k}
0.00391	0.00159
0.00159	0.00067

The point estimate of the variance of the unique contribution of school k to the association between social deprivation and attainment is .001 and that of the covariance between the effect with the school random effect is .002.

$\tau_{columns}$ (as correlations)

INTRCPT1/ ICPTCOL, c_{00k}	1.000	0.984
INTRCPT1/ DEPRIVE, c_{01k}	0.984	1.000

The value of the log-likelihood function at iteration 865 = -2.384254E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTERCEPT, θ_0	0.092434	0.021354	4.329	1752	<0.001
DEPRIVE, γ_{01}	-0.159051	0.026763	-5.943	522	<0.001
For P7VRQ, π_1					
INTERCEPT, θ_1	0.027636	0.002263	12.211	1752	<0.001
For P7READ, π_2					
INTERCEPT, θ_2	0.026242	0.001750	14.992	1752	<0.001
For DADOCC, π_3					
INTERCEPT, θ_3	0.008112	0.001360	5.964	1752	<0.001
For DADUNEMP, π_4					
INTERCEPT, θ_4	-0.120306	0.046759	-2.573	1752	0.010
For DADED, π_5					
INTERCEPT, θ_5	0.142622	0.040753	3.500	1752	<0.001
For MOMED, π_6					
INTERCEPT, θ_6	0.060870	0.037358	1.629	1752	0.103
For MALE, π_7					
INTERCEPT, θ_7	-0.056139	0.028383	-1.978	1752	0.048

Final estimation of row and level-1 variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTROW, b_{00j}	0.06087	0.00371	522	545.30137	0.232
level-1, *e*	0.67468	0.45519			

Final estimation of column level variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ ICPTCOL, c_{00k}	0.06255	0.00391	15	32.32912	0.006
INTRCPT1/ DEPRIVE, c_{01k}	0.02582	0.00067	15	9.67718	>0.500

Statistics for the current model

Deviance = 4768.508277
Number of estimated parameters = 14

Model comparison test

χ^2 statistic = 1.09638
Degrees of freedom = 2
p-value = >.500

The result of the deviance test is not significant. There is no evidence that the association between neighborhood social deprivation and attainment varies over schools. Not surprisingly, the standard error for $\hat{\gamma}_{01}$, the social deprivation effect, remains nearly unchanged, as do all inferences about the fixed effects.

13.5 Other program features

HCM models provide options for multivariate hypothesis tests for the fixed effects and the variance-covariance components. A "no-intercept" model is available for the level-1, level-2 and between-cell models. Figure 13.12 displays the **Basic Model Specifications - HCM2** dialog box.

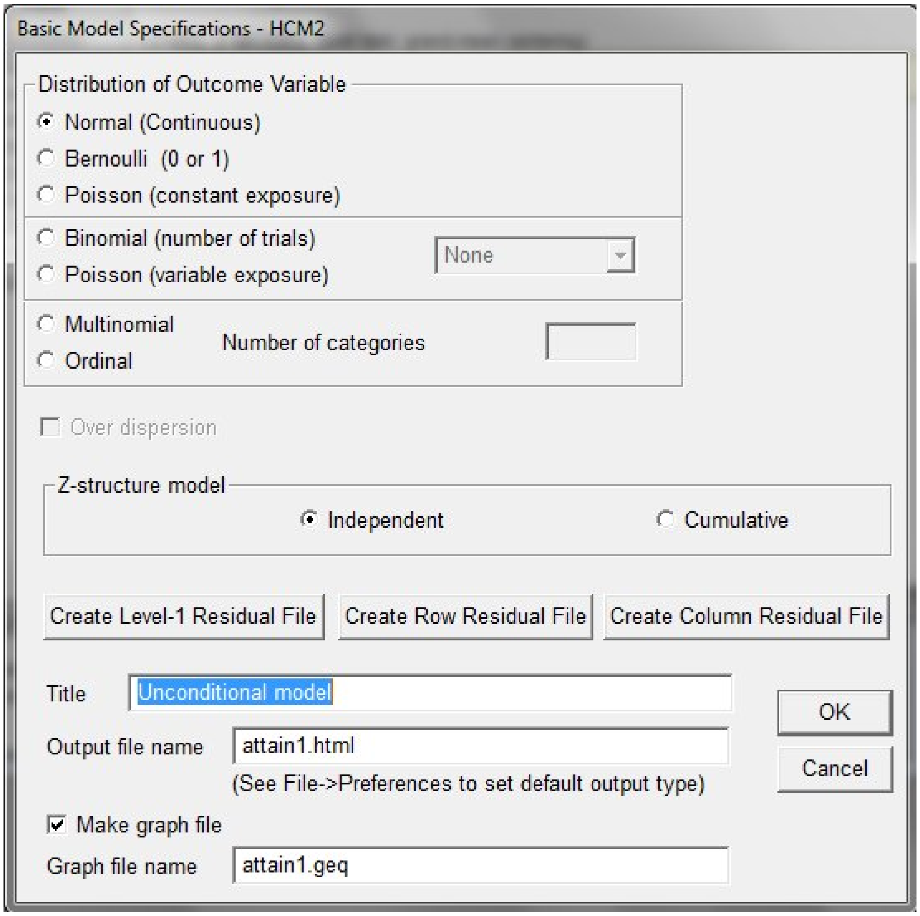


Fig 13.12 The Basic Model Specifications – HCM2 dialog box

The options are similar to the corresponding dialog box for HLM2 (see Section 2.5.2). Unlike HLM2, the user has the option to create a level-1, row and column residual file. There is an option unique to HCM2. When modeling longitudinal, repeated measures, it is possible to select a cumulative effect model to allow carry-over treatment effects by specifying a cumulative Z-structure model. See *Hierarchical Linear Models*, p. 390, for an example. HCM2 also allows users to diagonalize the τ s for rows and columns and weigh the cases within cells and rows (see Fig 13.13).

The **Fixed Intercept, Random Coefficient** option on **the Estimation Settings** dialog box is used to invoke the fiting of fixed intercepts random coefficients in models as discussed in Chapter 19. The **Diagonalize Tau** options constrain the variance-covariance matrices to diagonal matrices; in other words no covariation between random coefficients are assumed or estimated if this option is checked.

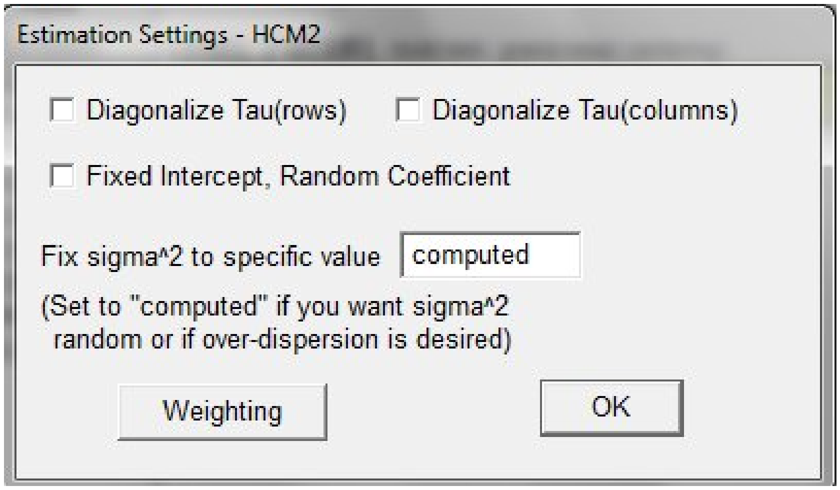


Figure 13.13 The Estimation Settings – HCM2 dialog box

14 Conceptual and Statistical Background for Three-Level Hierarchical and Cross-classified Random Effects Models (HCM3)

The HCM2 models discussed in the previous chapters allow researchers to analyze data that display structures in which the lower-units are cross-classified by two higher-level factors. Suppose, however, that one of the higher-level factors is itself nested within a yet-higher level factor. The three-level hierarchical and cross-classified random effects models (HCM3) represent this case, where level-1 units are cross-classified by two higher-level factors, with units from *one* of the higher-level factors nested within a next higher-level unit.

Hong and Raudenbush (2008) used three-level hierarchical and cross-classified random effects models to investigate how schools and their teachers may contribute to student growth, taking into account also the student-level variables. In their study, students were moving over time across teachers and the teachers were nested within schools. We can say that the repeated measures (level-1) were cross-classified by students (rows) and teachers (columns) with teachers nested within schools (clusters). The model is sufficiently flexible to allow the students also to change schools over the course of the study. In general, we may say that level-1 observations are crossed by *rows* and *columns* and the columns are nested within *clusters*.

14.1 The general 3-level hierarchical and cross-classified random effects models

A general three-level hierarchical and cross-classified model consists of three sub-models: level-1 or within-cell, level-2 or between-cell, and a level-3 or between-cluster model. As in HCM2, the cells refer to the cross-classifications by rows and columns. The columns, however, are nested within clusters.

For example, if the research problem consists of repeated developmental data on students cross-classified by student and teachers, with teachers clustered within schools, the level-1 or within-cell model will represent the relationship between time and development for each child. The level-2 or between-cell model will capture the influences of student- and teacher-level predictors, and the level-3 or between-cluster model will examine the effects of school-level variables. Formally, there are i=1,2,…, n_{jkl} level-1 units (*e.g.*, repeated measurement of student achievement) nested within cells cross-classified by j = 1,…, J rows (*e.g.*, students) and k = 1,…., K columns, with columns with cluster l = 1, …., L.

Here is an example of a data layout for three waves of developmental data (n_{jkl} = 3) for J = 4 students crossed by K = 9 teachers, with the teachers nested within L = 3 schools:

Table 14.1 Organization of data of the HCM3 example

	School_1			School_2			School_3		
	Teacher_{11}	Teacher_{21}	Teacher_{31}	Teacher_{12}	Teacher_{22}	Teacher_{32}	Teacher_{13}	Teacher_{23}	Teacher_{33}
Stud 1	Y_{1111}	Y_{2121}	Y_{3131}						
Stud 2				Y_{1212}	Y_{2222}	Y_{3232}			
Stud 3							Y_{1313}		Y_{3333}
Stud 4	Y_{1411}				Y_{2422}				Y_{3433}

Table 14.1 indicates that the repeated assessments are cross-classified by students and teachers, with teachers clustered within schools. Student 1 stayed in school 1 over three years of observation, changing teachers each year. Similarly Student 2 stayed in school 2 while changing teachers each year. Student 3 stayed in the same school, but was not observed during year 2. Student 4 had all three observations, but changed schools after year 1 and year 2.

HCM3 can handle continuously distributed as well as binary outcomes. We use the continuous outcome models in the following discussion. The logic of HGLM, as described and illustrated in Chapter 7, applies and extends to analyses with binary outcomes with HCM3.

14.1.1 Level-1 or "within-cell" model

We represent in the level-1 or within-cell model the outcome for case i in individual cells cross-classified by level-2 units j and k, with unit k nested within cluster l.

$$\begin{aligned} Y_{ijkl} &= \pi_{0jkl} + \pi_{1jkl}a_{1ijkl} + \pi_{2jkl}a_{2ijkl} + \cdots + \pi_{pjkl}a_{pijkl} + e_{ijkl} \\ &= \pi_{0jkl} + \sum_{p=1}^{P} \pi_{pjkl}a_{pijkl} + e_{ijkl} \end{aligned} \tag{14.1}$$

where

π_{0jkl} is the intercept, the expected value of Y_{ijkl} when all explanatory variables are set to zero;

π_{pjkl} are level-1 coefficients of predictors a_{pjkl} (p=1, 2, …, P) for case i in cell jkl;

e_{ijkl} is the level-1 or within-cell random effect, and;

σ^2 is the variance of e_{ijkl}, that is the level-1 or within-cell variance. Here we assume that the random term $e_{ijkl} \sim N(0,\sigma^2)$.

14.1.2 Level-2 or "between-cell" model

Each of the π_{pjkl} coefficients in the level-1 or within-cell model becomes an outcome variable in the level-2 or between-cell model:

$$
\begin{aligned}
\pi_{pjkl} &= \theta_{pl} + (\beta_{p1l} + b_{p1j})X_{1kl} + (\beta_{p2l} + b_{p2j})X_{2kl} + \cdots + (\beta_{pQ_pl} + b_{pQ_pj})X_{Q_pkl} + \\
&\quad (\gamma_{p1l} + c_{p1kl})W_{1jl} + (\gamma_{p2l} + c_{p2kl})W_{2jl} + \cdots + (\gamma_{pR_pl} + c_{pR_pkl})W_{R_pjl} + \\
&\quad b_{p0j} + c_{p0kl} \\
&= \theta_{pl} + \sum_{q=1}^{Q_p} (\beta_{pql} + b_{pqj})X_{qkl} + \sum_{r=1}^{R_p} (\gamma_{prl} + c_{prkl})W_{rjl} + b_{p0j} + c_{p0kl}
\end{aligned} \tag{14.2}
$$

where

θ_{pl} is the level-2 model intercept, the expected value of π_{pjkl} when all explanatory variables are set to zero;

β_{pql} are the level-2 coefficients of column-specific predictors $X_{qkl}, q = 1,...,Q_p$,

b_{pqj} are the random effects associated with column-specific predictors X_{qk}. They vary randomly over rows $j = 1,..., J$;

γ_{prl} are the level-2 coefficients of row-specific predictors $W_{rjl}, r = 1,...,R_p$;

c_{prkl} are the random effects associated with row-specific predictors W_{rjl}. They vary randomly over columns $k = 1,\ldots, K_l$ and clusters $l = 1,\ldots, L$; and

b_{p0j} and c_{p0kl} are residual row- and column-specific random effects, respectively, on π_{pjkl}, after taking into account X_{qkl} and W_{rjl}.

The vector of row random effects, containing $b_{p0j}, \ldots, b_{PQj}$ is assumed multivariate normal with a mean zero and a full covariance matrix τ. Similarly the vector with elements $c_{p0kl}, \ldots, c_{PRkl}$ is assumed multivariate normal with mean vector zero and full covariance matrix Δ.

14.1.3 Level-3 model

Each of the level-2 coefficients become an outcome variable at level 3:

$$
\begin{aligned}
\theta_{pl} &= \delta_{p00} + (\delta_{p01} + b_{p01j})Z_{1l} + (\delta_{p02} + b_{p02j})Z_{2l} + \cdots + (\delta_{p0S_{p0}} + b_{p0S_{p0}j})Z_{S_{p0}l} + d_{p0l} \\
&= \delta_{p00} + \sum_{s=1}^{S_{p0}} (\delta_{p0s} + b_{p0sj})Z_{sl} + d_{p0l} \\
\beta_{pql} &= \delta_{pq0} + (\delta_{pq1} + b_{pq1j})Z_{1l} + (\delta_{pq2} + b_{pq2j})Z_{2l} + \cdots + (\delta_{pqS_{pq}} + b_{pqS_{pq}j})Z_{S_{pq}l} + d_{pql} \\
&= \delta_{pq0} + \sum_{s=1}^{S_{pq}} (\delta_{pqs} + b_{pqsj})Z_{sl} + d_{pql} \\
\gamma_{prl} &= \delta_{pr0} + (\delta_{pr1} + b_{pr1j})Z_{1l} + (\delta_{pr2} + b_{pr2j})Z_{2l} + \cdots + (\delta_{prS_{pr}} + b_{prS_{pr}j})Z_{S_{pr}l} + d_{prl} \\
&= \delta_{pr0} + \sum_{s=1}^{S_{pr}} (\delta_{prs} + b_{prsj})Z_{sl} + d_{prl}
\end{aligned} \tag{14.3}
$$

where

δ_{p00} is the intercept, the expected value of θ_{pl} when all explanatory variables are set to zero;

δ_{p0s} are the coefficients of cluster-specific predictors Z_{sl} for θ_{pl};

δ_{pq0} is the intercept, the expected value of β_{pql} when all explanatory variables are set to zero;

δ_{pqs} are the coefficients of cluster-specific predictors Z_{sl}, $s = 1,...,S_{pq}$ for β_{pql} ;

b_{pqsj} are the random effects associated with cluster-specific predictors Z_{sl}. They vary randomly over rows $j = 1,..., J$;

δ_{pr0} is the intercept, the expected value of γ_{prl} when all explanatory variables are set to zero;

δ_{prs} are the coefficients of cluster-specific predictors Z_{sl} for γ_{prl};

b_{prsj} are the random effects associated with cluster-specific predictors Z_{sl}. They vary randomly over rows $j = 1,..., J$; and

d_{p0l}, d_{pql}, and d_{prl} are residual random effects. We assume these to be multivariate normal in distribution with zero means and variances $\tau_{p0}, \tau_{pq}, \tau_{pr}$, respectively.

14.2 Parameter estimation

Three kinds of parameter estimates are available in HCM3. For continuous outcomes, empirical Bayes estimates of random effects, maximum-likelihood estimates of the level-3 coefficients, and maximum likelihood estimates of variance-covariance parameters are available. In nonlinear models, the level-3 coefficients are estimated via penalized quasi-likelihood. Unlike HGLM, however, only unit-specific and not population-averaged results are available.

14.3 Hypothesis testing

As in the case of HLM2, HCM3 routinely prints standard errors and *t*-tests for each of the fixed level-2 coefficients as well as a chi-square test of homogeneity for each random effect. In addition, optional “multivariate hypothesis tests“ are available in HCM3. Multivariate tests in the case of continuous outcomes parallel those described in Section 2.8.8. For binary outcomes, hypothesis testing parallels those described in Section 5.10.

15 Working with HCM3

15.1 An example using HCM3 in Windows mode

HCM3 analyses can be executed in Windows, interactive, and batch modes. We describe a Windows execution below. We consider interactive and batch execution in Appendix H. A number of special options are presented at the end of the chapter.

To illustrate the operation of the program, we use the data from Hong and Raudenbush's (2008) study on the effects of time-varying instructional treatments (intensive vs. conventional math instruction) on student achievement.

15.1.1 Constructing the MDM file from raw data

In constructing the MDM file, there is the same range of options for data input as for HCM2. HCM3 requires three IDs, one for each of two level-2 factors, and one for the level-3 clusters. The two level-2 factors in our examples are student and teacher. As teachers (N = 498) were clustered within schools (N = 67) and the model allows students (N = 4216) to change schools, we will designate teacher as the column factor and student as the row factor.

Note: The level-1 file is to be sorted on ascending row (student) IDs, and, in this file, sorting by column IDs within clusters. The level-2 row file is to be sorted on ascending row (student) IDs. The level-2 column file is to sorted by column IDs within clusters. The cluster file is to be sorted by cluster IDs.

15.1.2 Statistical package input

Data input requires a level-1 file (a time-series student achievement data file in our example), a level-2 row-factor (student-level) file, a level-2 column-factor (teacher-level) file, and a level-3 cluster-level (neighborhood-level) file. Our illustration uses SPSS file input, but the procedure for all other statistical packages is analogous.

Level-1 file. The level-1 or within-cell file, GROWTH.SAV has 7342 repeated measures collected on 4216 students. Figure 15.1 shows the time series data for the first four students. Following the school, student, and teacher ID fields are students' values on six variables:

- MATH
  A Stanford Achievement Test math test score.
- YEAR (year of the study minus 2)
  This variable can take on values of -1, 0, and 1 for the three years of data collection from grade 3 to grade 5.
- G4D1 is an indicator that that takes on a value of 1 if a child receives intensive math instruction in grade 4 and if the outcome is observed at grade 4. This will be used to assess the effect of grade-4 intensive math instruction on grade-4 outcome.

- G4D21 is an indicator that a child receives intensive math instruction in grade 4 and if the outcome is observed at grade 5. This will be used to test the effect of grade-4 intensive math instruction on grade-5 outcome for those who do not receive intensive math instruction in grade 5.
- G5D22

  An indicator that a child receives intensive math instruction in grade 5 and if the outcome is observed at grade 5. This will be used to test the effect of intensive math instruction in grade 5 on grade 5 outcome for those who did not have intensive math instruction in grade 4.
- TWOWAY

  A product term of a two-way interaction between G4D21 with G5D22. It will thus be an indicator that the child received intensive math instruction in both grades 4 and 5 and if the outcome is observed at grade 5. This will test whether having intensive math instruction in both years has an effect that exceeds the sum of the separate effects.

	schid	studid	tchrid	math	year	g4d1	g4d21	g5d22	twoway
1	15	1	83	609.00	-1.00	0.00	0.00	0.00	0.00
2	15	1	104	592.00	0.00	0.00	0.00	0.00	0.00
3	15	1	135	638.00	1.00	0.00	0.00	0.00	0.00
4	15	2	83	586.00	-1.00	0.00	0.00	0.00	0.00
5	15	2	104	592.00	0.00	0.00	0.00	0.00	0.00
6	15	2	185	629.00	1.00	0.00	0.00	1.00	0.00
7	15	3	114	561.00	0.00	0.00	0.00	0.00	0.00
8	15	3	135	584.00	1.00	0.00	0.00	0.00	0.00
9	15	4	83	593.00	-1.00	0.00	0.00	0.00	0.00
10	15	4	114	621.00	0.00	0.00	0.00	0.00	0.00

Figure 15.1 First 10 records in the GROWTH.SAV dataset

We see that student 1 attended school 15 and was taught by teachers 83, 104, and 135. None of the teachers adopted intensive math instruction. In addition, student 3 had data for the second and third year only.

Level-2 row-factor file. The level-2 row-factor units in the illustration are 4216 students. The data are stored in the file STUDENT.SAV. The level-2 data for the first ten children are listed in Figure 15.2. The file has one dummy variable.

	studid	dummy
1	1	0.00
2	2	0.00
3	3	0.00
4	4	0.00
5	5	0.00
6	6	0.00
7	7	0.00
8	8	0.00
9	9	0.00
10	10	0.00

Figure 15.2 First 10 cases in the STUDENT.SAV dataset

Level-2 column-factor file. The level-2 column-factor (teacher) file, TEACHER.SAV, has two IDs and a dummy variable. The first ID is the level-3 (*i.e.*, school) ID and the second ID is the level-2 column factor (*i.e.*, teacher) ID. Figure 15.3 lists the data for the first ten records.

	schid	tchrid	dummy
1	1	3963	0.00
2	1	3974	0.00
3	1	4025	0.00
4	2	9253	0.00
5	2	9263	0.00
6	2	9273	0.00
7	2	9304	0.00
8	2	9315	0.00
9	2	9325	0.00
10	2	9335	0.00

Figure 15.3 First 10 cases in the TEACHER.SAV data set

Level-3 file. The level-3 (school) level file, SCHOOL.SAV, has the level-3 (school) ID and a dummy variable. Figure 15.4 lists the data for the first ten records.

	schid	dummy
1	1	0.00
2	2	0.00
3	3	0.00
4	4	0.00
5	5	0.00
6	6	0.00
7	7	0.00
8	8	0.00
9	9	0.00
10	10	0.00

Figure 15.4 First 10 cases in the SCHOOL.SAV data set

In sum, there are six variables at level 1 and one dummy variable for each of the level-2 row- and column-factor files and the level-3 file. The steps for the construction of the MDM for HCM3 are similar to the ones described in Section 2.5.1.1. The user will select **HCM3** in the **Select MDM type** dialog box (see Figure 2.5). Note that the program can handle missing data at level 1 or within cell only. The MDM template file, GROWTH.MDMT, contains a log of the input responses used to create the MDM file, GROWTH.MDM, using GROWTH.SAV, STUDENT.SAV, TEACHER.SAV, and SCHOOL.SAV. Figure 15.5 displays the dialog box used to create the MDM file. Figures 15.6 show the dialog boxes for the level-1 file.

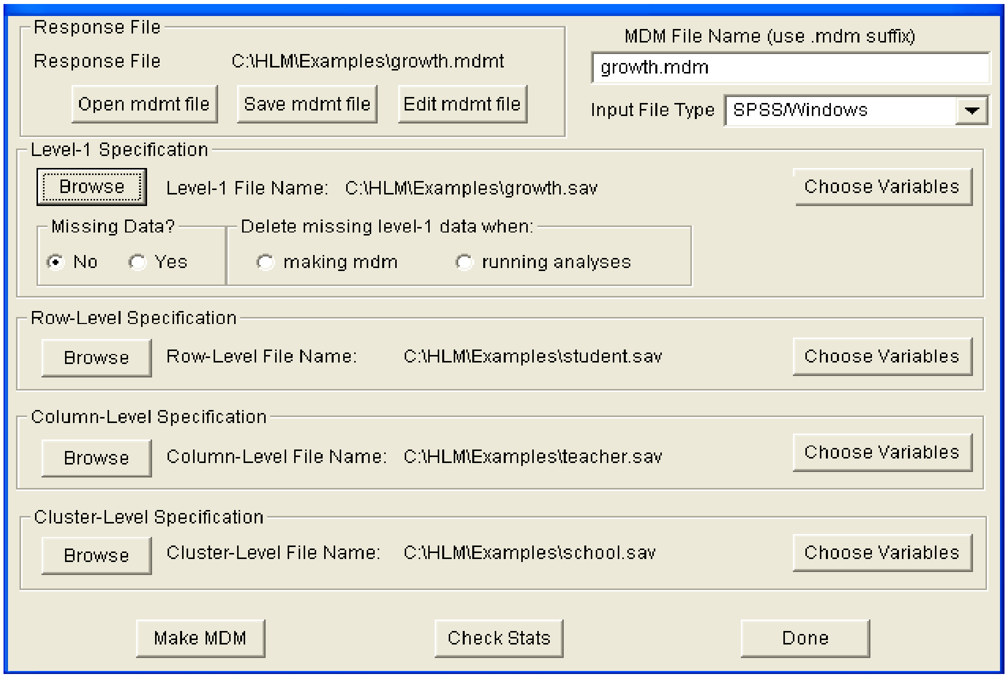


Figure 15.5 Make MDM - HCM3 dialog box for GROWTH.MDMT

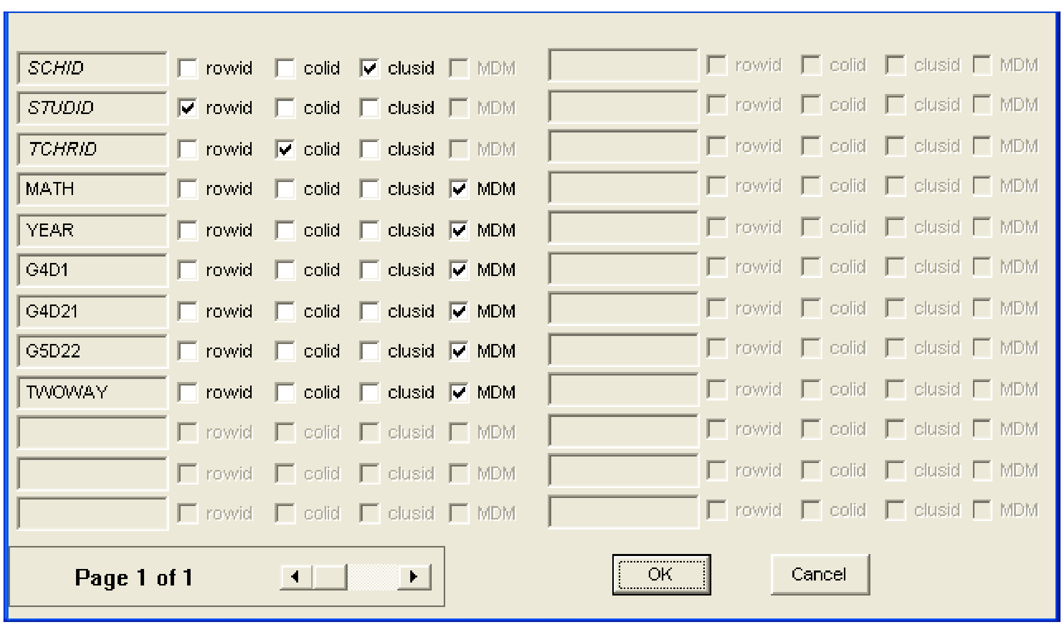


Figure 15.6 Choose variables - HCM3 dialog box for level-1 file, GROWTH.SAV

15.2 Executing analyses based on the MDM file

Once the MDM file is constructed, it can be used as input for the analysis. Model specification has five steps:

1. Specification of the level-1 model. In our case we shall model mathematics achievement (MATH) as the outcome, to be predicted by YEAR, G4D1, G4D21, G5D22, and TWOWAY. Hence, the level-1 model will have six coefficients for each student: the intercept and the partial slopes for the five variables. For longitudinal analysis, it is possible to select a cumulative effect model to allow carry-over treatment effects by specifying a cumulative Z-structure model (see *Hierarchical Linear Models*, p.390); we use this option in the analysis.
2. Specification of the level-2 row- or column-factor prediction model. Here each level-1 coefficient – the intercept and the five slopes in our example – becomes an outcome variable. One may use variables on student and teacher characteristics (not supplied with the example data files) to predict each of these level-1 coefficients.
3. Specification of row- or column effects as random or non-random. We shall model the intercept and the YEAR slope as varying randomly over rows and columns.
4. Specification of the level-3 prediction model. Here each level-2 coefficient becomes an outcome, and one may select school variables (not included in the example data files) to predict school-to-school in these level-2 coefficients.
5. Specification of the level-2 coefficients as random or non-random. We let two of the six level-2 intercepts vary over schools.

Following the five steps above, we specify a model to study the effects of time-varying instructional treatments on student achievement. The Windows execution is very similar to the one for HCM2 as described in Section 13.4. The command file, GROWTH1.MLM, contains the model specification input responses. Figure 15.7 displays the model specified.

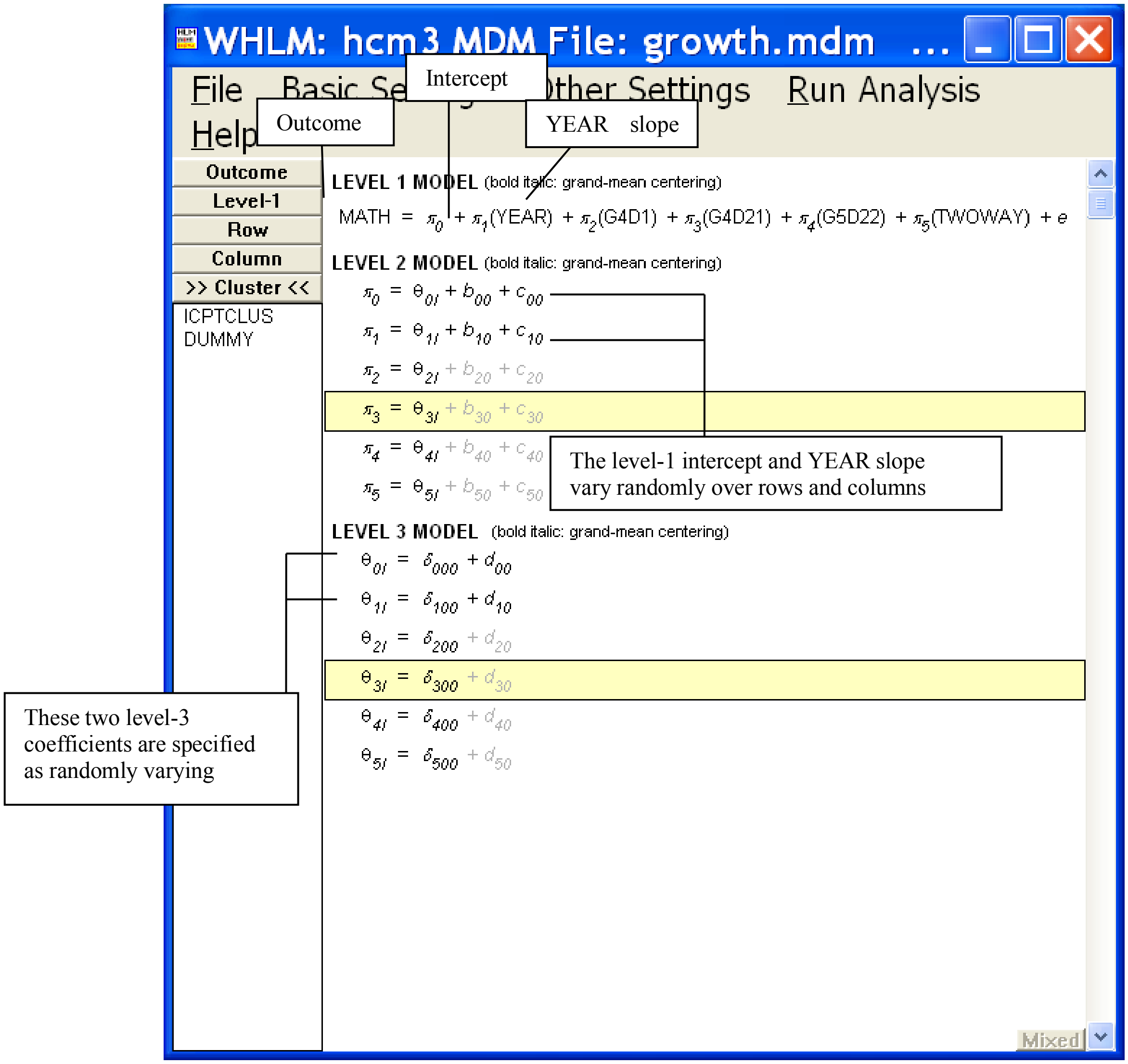


Figure 15.7 Unweighted model for the growth example

The results of the analysis are given below.

Problem Title: Unweighted model
The data source for this run = growth.mdm
The command file for this run = growth1.hlm
Output file name = growth1.html
The maximum number of level-1 units = 7342
The maximum number of row units = 4216
The maximum number of column units = 498
The maximum number of cluster units = 67
The maximum number of iterations = 100
Method of estimation: full maximum likelihood
Z-structure: cumulative across columns
Data design: (row by column) within clusters

The outcome variable is MATH

Summary of the model specified

Level-1 Model

$$MATH_{ijkl} = \pi_{0jkl} + \pi_{1jkl}*(YEAR_{ijkl}) + \pi_{2jkl}*(G4D1_{ijkl}) + \pi_{3jkl}*(G4D21_{ijkl}) + \pi_{4jkl}*(G5D22_{ijkl}) + \pi_{5jkl}*(TWOWAY_{ijkl}) + e_{ijkl}$$

Level-2 Model

$\pi_{0jkl} = \theta_{0l} + b_{00jl} + c_{00kl}$
$\pi_{1jkl} = \theta_{1l} + b_{10jl} + c_{10kl}$
$\pi_{2jkl} = \theta_{2l}$
$\pi_{3jkl} = \theta_{3l}$
$\pi_{4jkl} = \theta_{4l}$
$\pi_{5jkl} = \theta_{5l}$

Level-3 Model

$\theta_{0l} = \delta_{000} + d_{00l}$
$\theta_{1l} = \delta_{100} + d_{10l}$
$\theta_{2l} = \delta_{200}$
$\theta_{3l} = \delta_{300}$
$\theta_{4l} = \delta_{400}$
$\theta_{5l} = \delta_{500}$

For starting values, data from 5299 level-1 records, 2173 rows, 498 column, and 65 cluster records were used

Final Results - iteration 485

Iterations stopped due to small change in likelihood function

σ^2 = 304.82130

τ_π

	YEAR
θ_0,b_{00}	θ_1,b_{10}
769.17514	-18.09880
-18.09880	21.22623

τ_π (as correlations)

1.000	-0.142
-0.142	1.000

τ_β

	YEAR
θ_0,c_{00}	θ_1,c_{10}
133.52764	-24.04565
-24.04565	48.79836

τ_β (as correlations)

1.000	-0.298
-0.298	1.000

τ_γ

	YEAR
θ_0,d_{00}	θ_1,d_{10}
169.31794	28.10279
28.10279	29.76755

τ_γ (as correlations)

1.000	0.396
0.396	1.000

The value of the log-likelihood function at iteration 485 = -3.536565E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1					
For INTERCEPT					
θ_0,δ_{000}	609.850986	1.962504	310.751	66	<0.001
For YEAR					
For INTERCEPT					
θ_1,δ_{100}	21.064011	1.140716	18.466	66	<0.001
For G4D1					
For INTERCEPT					
θ_2,δ_{200}	2.753381	2.371599	1.161	7338	0.246#
For G4D21					
For INTERCEPT					
θ_3,δ_{300}	0.231710	3.584218	0.065	7338	0.949#
For G5D22					
For INTERCEPT					
θ_4,δ_{400}	7.507799	2.332107	3.219	7338	0.002#
For TWOWAY					
For INTERCEPT					
θ_5,δ_{500}	1.160337	4.322456	0.268	7338	0.788#

The p-vals above marked with a “#” should regarded as a rough approximation.

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1					
For INTERCEPT					
θ_0,δ_{000}	609.850986	1.954775	311.980	66	<0.001
For YEAR					
For INTERCEPT					
θ_1,δ_{100}	21.064011	1.112653	18.931	66	<0.001
For G4D1					
For INTERCEPT					
θ_2,δ_{200}	2.753381	2.927131	0.941	7338	0.347#
For G4D21					
For INTERCEPT					
θ_3,δ_{300}	0.231710	4.389057	0.053	7338	0.958#
For G5D22					
For INTERCEPT					
θ_4,δ_{400}	7.507799	3.019164	2.487	7338	0.013#
For TWOWAY					
For INTERCEPT					
θ_5,δ_{500}	1.160337	6.470068	0.179	7338	0.858#

The p-vals above marked with a “#” should regarded as a rough approximation.

Final estimation of row and level-1 variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
θ_0,b_{00}	27.73401	769.17514	2172	11413.58016	<0.001
YEAR/θ_1,b_{10}	4.60719	21.22623	2172	2177.42726	0.463
level-1, *e*	17.45913	304.82130			

Note: The chi-square statistics reported above are based on only 2173 of 4216 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of column level variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
θ_0,c_{00}	11.55542	133.52764	429	539.50878	<0.001
YEAR/θ_1,c_{10}	6.98558	48.79836	429	0.01770	>0.500

Note: The chi-square statistics reported above are based on only 495 of 498 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of cluster level variance components:

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
θ_0,d_{00}	13.01222	169.31794	64	256.96222	<0.001
YEAR/θ_1,d_{10}	5.45596	29.76755	64	136.92770	<0.001

Note: The chi-square statistics reported above are based on only 65 of 67 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

As reported by Hong and Raudenbush (2008), no significant causal effect of Grade 4 treatment on Grade 4 outcomes. A positive and significant effect of Grade 5 treatment on Grade 5 outcome, $\hat{\delta}_{400} = 7.51$ ($SE = 3.019$, $t = 2.487$) [6].

Statistics for the current model

Deviance = 70731.304874
Number of estimated parameters = 16

15.3 Other program features

HCM3 models provide options similar to those of HCM2. It also allows users to diagonalize the τ_π, τ_β, and τ_γ when estimating the variance components if interests focus only on the diagonal elements of any of the three matrices. In addition, design weights are allowed for level-1, level-2 row factor and level-3 units.

---

[6] We used an improved algorithm here and thus the results are a bit different from those published in Hong and Raudenbush (2008).

16 Conceptual and Statistical Background for Hierarchical Linear Model with Cross-Classified Random Effects (HLMHCM)

In HCM2, level-1 units are nested within cells and cross-classified by two higher-level factors. HLMHCM adds a level within the cells. For example, we may have a growth model for each of a set of students, all of whom live in the same neighborhood and attend the same school. We would say that level-1 units (repeated measures) are nested within level-2 units (children); level-2 units are crossed by rows (neighborhoods) and columns (schools). Another example might involve repeated item responses at a given time for a student encountering a given teacher. The level-1 units are the item responses, nested within occasions (level-2) crossed by rows (students) and columns (teachers).

16.1 The general hierarchical linear model with cross-classified random effects

A general hierarchical HLMHCM has three sub-models: a level-1 model and a level-2 model within each cell; and a level-3 model or between-cell model that incorporates row and column effects.

Formally, there are $m = 1,2,\ldots, n_{ijk}$ level-1 units (*e.g.*, repeated measurement of student achievement) nested within level-2 (*e.g.*, students) $I = 1,\ldots, n_{jk}$ nested within cells cross-classified by $j = 1,\ldots, J$ rows (*e.g.*, neighborhoods) and $k = 1,\ldots, K$ columns (*e.g.*, schools).

Here is an example of a data layout for three waves of developmental data ($n_{ijk} = 3$) nested within $J = 10$ students nested within cells cross-classified by $J = 3$ neighborhoods (rows) and $K = 3$ schools (columns):

Table 16.1 Organization of data of the HLMHCM example

	School_1	School_2	School_3
Neighborhood_1	Y_{1111}, Y_{2111}, Y_{3111} of Stud 1 Y_{1211}, Y_{2211}, Y_{3211} of Stud 2	Y_{1311}, Y_{2311}, Y_{3311} of Stud 3	
Neighborhood_3	Y_{1411}, Y_{2411}, Y_{3411} of Stud 4	Y_{1511}, Y_{2511}, Y_{3511} of Stud 5 Y_{1611}, Y_{2611}, Y_{3611} of Stud 6	Y_{1711}, Y_{2711}, Y_{3711} of Stud 7
Neighborhood_3	Y_{1811}, Y_{2811}, Y_{3811} of Stud 8 Y_{1911}, Y_{2911}, Y_{3911} of Stud 9		Y_{11011}, Y_{21011}, Y_{31011} of Stud 10

Table 16.1 indicates that the repeated developmental data are nested within individual students nested within cells cross-classified by neighborhoods and schools. Note that unlike in HCM3, the students *never* leave the neighborhood or school of origin.

16.1.1 Level-1 or "within-unit" model

We represent in the level-1 model the outcome Y for response m of the level-2 unit i cross-classified by row j and column k.

$$
\begin{aligned}
Y_{mijk} &= \psi_{0ijk} + \psi_{1ijk}a_{1ijk} + \psi_{2ijk}a_{2ijk} + \cdots + \psi_{pijk}a_{pijk} + \xi_{mijk} \\
&= \psi_{0ijk} + \sum_{p=1}^{P}\psi_{pijk}a_{pijk} + \xi_{mijk}
\end{aligned}
\tag{16.1}
$$

where

ψ_{0ijk} is the intercept, the expected value of Y_{mijk} when all explanatory variables are set to zero;

ψ_{pijk} are level-1 coefficients of predictors a_{pijk} (p=1,2,…,P) ;

ξ_{mijk} is the level-1 random effect; and

σ^2 is the variance of ξ_{mijk}, that is the level-1 variance. Here we assume that the random term $\xi_{mijk} \sim N(0,\sigma^2)$.

16.1.2 Level-2 or "between-unit" or "within-cell" model

Each of the ψ_{pijk} (p=0,1,…,P) coefficients in the level-1 model becomes an outcome variable in the level-2 or within-cell model:

$$
\begin{aligned}
\psi_{pijk} &= \pi_{p0jk} + \pi_{p1jk}\alpha_{p01k} + \pi_{p2jk}\alpha_{p02k} + \cdots + \pi_{pQ_pjk}\alpha_{Q_pjk} + e_{pijk} \\
&= \pi_{p0jk} + \sum_{q=1}^{Q_p}\pi_{pqjk}\alpha_{p0jk} + e_{pijk}
\end{aligned}
\tag{16.2}
$$

π_{p0jk} is the intercept, the expected value of ψ_{pijk} when all explanatory variables are set to zero;

π_{pqjk} are level-1 coefficients of predictors α_{p0jk} (p=1,2,…,P);

e_{pijk} is the level-2 or within-cell random effect, and

τ is the variance-covariance matrix of e_{pijk}, that is the level-2 variance. Here we assume that the random term $e_{pijk} \sim N(0,\tau)$. The vector containing elements e_{pijk} is assumed multivariate normal with a mean zero and a full covariance matrix, τ.

16.1.3 Level-3 model or "between-cell" model

Each of the π_{pqjk} (q = 0, 1, …, Q_p) coefficients in the level-2 or within-cell model becomes an outcome variable in the level-3 or between-cell model:

$$
\begin{aligned}
\pi_{pqjk} = {} & \theta_{pq0} + (\beta_{pq1} + b_{pq1j})X_{1k} + (\beta_{pq2} + b_{pq2j})X_{2k} + \cdots + (\beta_{pqR_p} + b_{pqR_pj})X_{R_qj} + \\
& (\gamma_{pq1} + c_{pq1k})W_{1j} + (\gamma_{pq2} + c_{ps2k})W_{2j} + \cdots + (\gamma_{pqS_p} + c_{pqS_pk})W_{S_pj} + \\
& b_{pq0j} + c_{pq0k}
\end{aligned}
$$

(16.3)

where

θ_{pq}, β_{pqr} are the fixed effects of column-specific predictors $X_{qk}, r = 1, \ldots, R_p$,

b_{pqrj} are the random effects associated with column-specific predictors X_{rk}. They vary randomly over rows $j = 1,..., J;$

γ_{pqs} are the fixed coefficients of row-specific predictors W_{sj}, $s = 1,..., S_p$;

c_{pqsk} are the random effects associated with row-specific predictors W_{sj}. They vary randomly over columns $k = 1,\ldots, K;$ and

b_{pqrj}, and c_{pqsk} are residual row- and column-specific random effects, respectively, on π_{pqjk}, after taking into account X_{rk} and W_{sj}.

The vector containing elements b_{pqrj} is assumed multivariate normal with a mean zero and a full covariance matrix Ω. Similarly, the vector with elements c_{pqsk} is assumed multivariate normal with mean vector zero and full covariance matrix Δ.

16.2 Parameter estimation

Three kinds of parameter estimates are available in HLMHCM. For continuous outcomes, empirical Bayes estimates of random effects, maximum-likelihood estimates of the level-3 coefficients, and maximum likelihood estimates of variance-covariance parameters are available. In nonlinear models, the level-3 coefficients are estimated via penalized quasi-likelihood. Unlike HGLM, however, only unit-specific and not population-averaged results are available.

16.3 Hypothesis testing

As in the case of HLM2, HLMHCM routinely prints standard errors and *t*-tests for each of the fixed level-3 coefficients as well as a chi-square test of homogeneity for each random effect. In addition, optional “multivariate hypothesis tests“ are available in HLMHCM. Multivariate tests in the case of continuous outcomes parallel those described in Section 2.8.8. For discrete outcomes, hypothesis testing parallels those described in Section 7.10.

17 Working with HLMHCM

17.1 An example using HLMHCM in Windows mode

HLMHCM analyses can be executed in Windows, interactive, and batch modes. We describe a Windows execution below. We consider interactive and batch execution in Appendix I. A number of special options are presented at the end of the chapter.

Chapter 8 in *Hierarchical Linear Models* and Chapter 4 of this manual provide examples of HLM3 analyses of repeated measures data nested within students within schools collected by the US *Sustaining Effects Study* and by an urban school effects study, respectively. To illustrate the operation of the HLMHCM program, we perform another achievement growth analysis. Unlike the previous examples, however, this analysis considers not only the school but the neighborhood contexts within which the students resided in as well. The data were obtained from 567 students from 224 schools in 74 urban neighborhoods in which repeated achievement measures are nested within students cross-classified by schools and neighborhoods. We chose a similar set of covariates to allow users to compare and contrast these set of models with those HLM3 models executed in Chapter 4.

17.1.1 Constructing the MDM file from raw data

In constructing the MDM file, there is the same range of options for data input as for HLM2. HLMHCM requires three IDs, one for the level-2 (students in our illustration) units, and one for the units of each of the higher-level factors (school and neighborhood), and **the level-2 IDs have to be sorted**. As there are more schools than neighborhoods in our example, we follow the convention adopted for HCM2 and designate school as the row factor and neighborhood as the column factor.

17.1.2 Statistical package input

Data input requires a level-1 within-unit file (a time-series student achievement data file in our example), a level-2 or between unit (student-level) file, a level-3 row-factor (school-level) file, and a level-3 column-factor (neighborhood-level) file.

Level-1 file. The level-1 or within-cell file, GROWTH.SAV has 2008 observations collected on 567 students beginning at grade one and followed up annually thereafter for six years. Figure 17.1 shows the time series data for the first three students. All of them have complete data; typically there are three or four observations per child. Following the student ID field are that student's values on two variables:

- AGE8
  The age of the child minus 8 at each testing occasion. Therefore, it is 0 at age 8, 1 at age 9, etc.
- MATH
  A math test score in an IRT metric.

	studid	age8	math
1	1	-0.420	2.100
2	1	0.580	3.000
3	1	1.580	4.300
4	1	2.580	6.200
5	1	3.580	7.300
6	1	4.580	8.100
7	2	-0.053	2.600
8	2	0.947	2.933
9	2	1.947	3.500
10	2	2.947	4.000
11	2	3.947	5.400
12	2	4.947	7.200
13	3	-0.299	2.700
14	3	0.701	3.400
15	3	1.701	4.400
16	3	2.701	5.200
17	3	3.701	7.100
18	3	4.701	8.500

Figure 17.1 First 18 records in the GROWTH.SAV dataset

We see that the first student was about seven and a half years old (AGE8 = –0.420) during the first data collection wave with a math score of 2.1.

Level-2 file. The level-2 units in the illustration are 567 students. The data are stored in the file STUDENT.SAV. The level-2 data for the first eight children are listed in Figure 17.2. The first ID is the level-3 row-factor (*i.e.*, school) ID, the second ID is the level-3 column factor (*i.e.*, neighbor) ID, and the third ID is the level-2 (*i.e.*, student) ID. **Note that the level-2 files must be sorted in the same order of level-2 ID.**

There are three variables:

- FEMALE (1 = female, 0 = male)
- BLACK (1 = African-American, 0 = other)
- HISPANIC (1 = Hispanic, 0 = other)

We see, for example, that student 1 who attended school 175 and resided in neighborhood 68 is a African-American male (FEMALE = 0, BLACK = 1, HISPANIC = 0).

	schid	neighid	studid	female	black	hispanic
1	175	68	1	0.000	1.000	0.00
2	99	68	2	0.000	1.000	0.00
3	99	68	3	0.000	1.000	0.00
4	109	23	4	1.000	0.000	0.00
5	165	72	5	0.000	1.000	0.00
6	207	72	6	1.000	1.000	0.00
7	6	72	7	1.000	1.000	0.00
8	148	72	8	0.000	0.000	0.00

Figure 17.2 First 10 cases in the STUDENT.SAV dataset

Level-3 row-factor file. The level-3 row-factor (school) level file, SCHOOL.SAV, consists of data on 1 variable for 224 schools. The variable is SCHPOV, which is an indicator of school poverty, as measured by the percentage of the total number of students enrolled in free or subsidized lunch programs.

We see that the first school, school 1, has 91% of its students enrolled in free or subsidized lunch programs.

	schid	schpov
1	1	91.400
2	2	78.900
3	3	63.200
4	4	93.000
5	5	71.800
6	6	92.600
7	7	96.200
8	8	83.400

Figure 17.3 First 8 cases in the SCHOOL.SAV data set

Level-3 column-factor file. The level-3 row-factor (neighborhood) level file, NEIGH.SAV, consists of data on 1 variable for 74 neighborhoods. The variable is DISADV (a scale measuring social deprivation, which incorporates information on the poverty concentration, health, and housing stock of a local community). A measure of neighborhood disadvantage, constructed through an oblique factor analysis from the 1990 decennial census data, tapped the level of poverty and unemployment, and the percentage of families that were headed by females and percentage on welfare (Sampson & Raudenbush, 1999; Sampson, Raudenbush, & Earls, 1997).

	neighid	disadv
1	1	0.449
2	2	-0.553
3	3	-0.382
4	4	-0.563
5	5	-0.313
6	6	-0.566
7	7	0.160
8	8	1.222

Figure 17.4 First 8 cases in the NEIGH.SAV data set

In sum, there are two variables at level 1, three at level 2, and one for each of the level-3 factors.

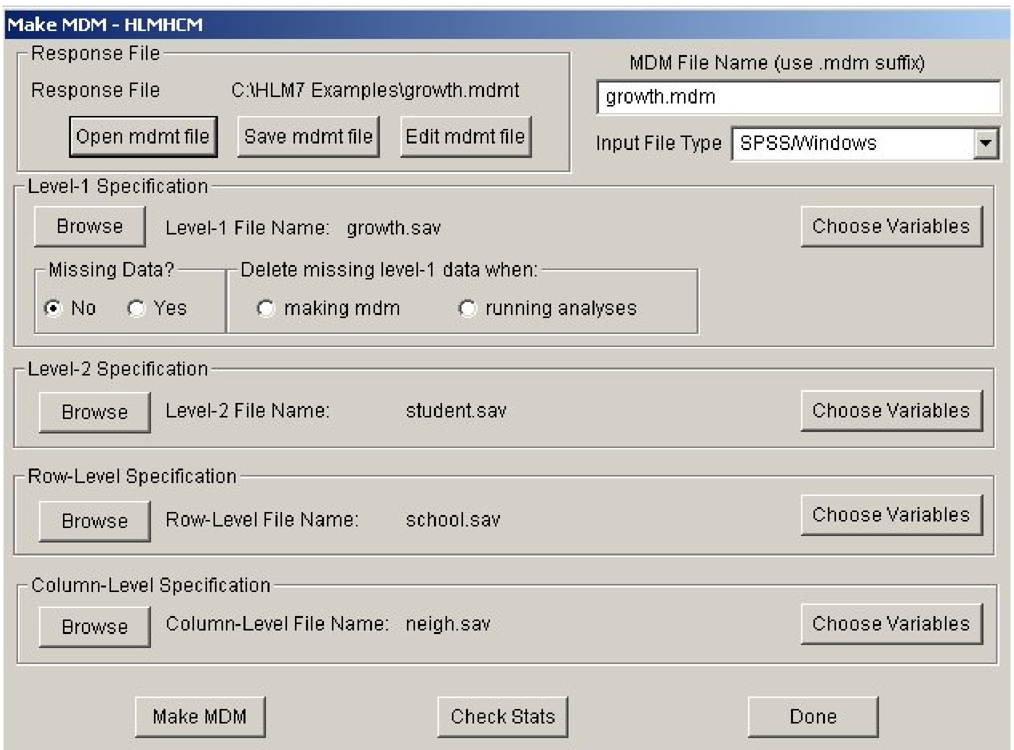


Figure 17.5 Make MDM- HLMHCM dialog box for GROWTH.MDMT

The steps for the construction of the MDM for HLMHCM2 are similar to the ones described in Section 2.5.1.1. The user will select **HLMHCM** in the **Select MDM type** dialog box (see Figure 2.5). Note that the program can handle missing data at level 1 or within cell only. The MDM template file, GROWTH.MDMT, contains a log of the input responses used to create the MDM file, GROWTH.MDM, using GROWTH.SAV, STUDNET.SAV, SCHOOL.SAV, and NEIGH.SAV. Figure 18.5 displays the dialog box used to create the MDM file. Figures 17.6 to 17.9 show the dialog boxes for the level-1 file, GROWTH.SAV, the level-2 file, STUDENT.SAV, the level-3 row file, SCHOOL.SAV, and the level-3 column file, SCHOOL.SAV.

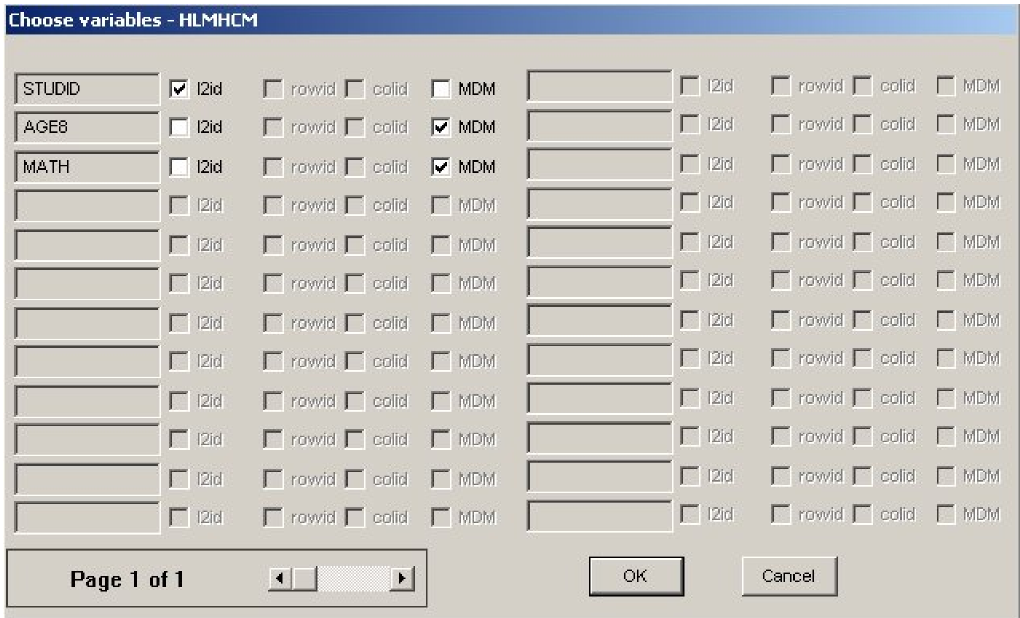


Figure 17.6 Choose variables HLMHCM dialog box for level-1 file, GROWTH.SAV

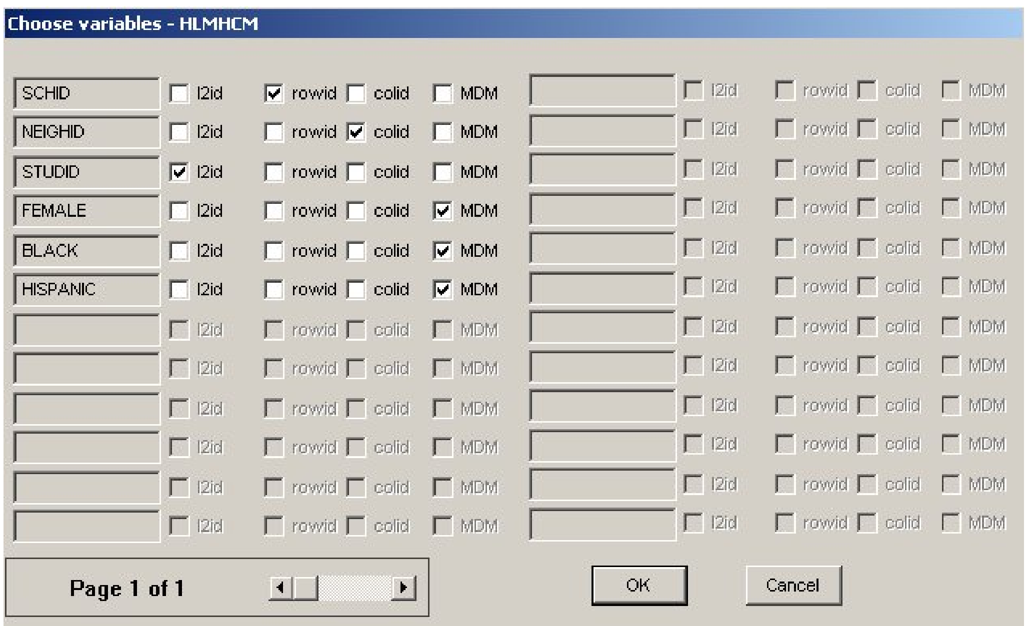


Figure 17.7 Choose variables HLMHCM dialog box for level-2 file, STUDENT.SAV

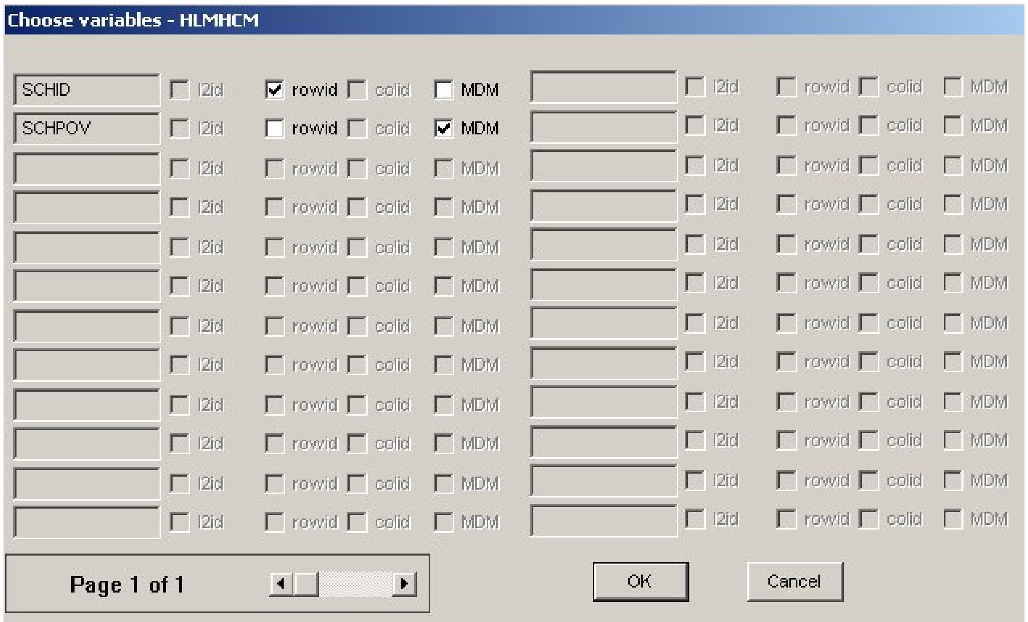


Figure 17.8 Choose variables HLMHCM dialog box for level-3 row-factor file, SCHOOL.SAV

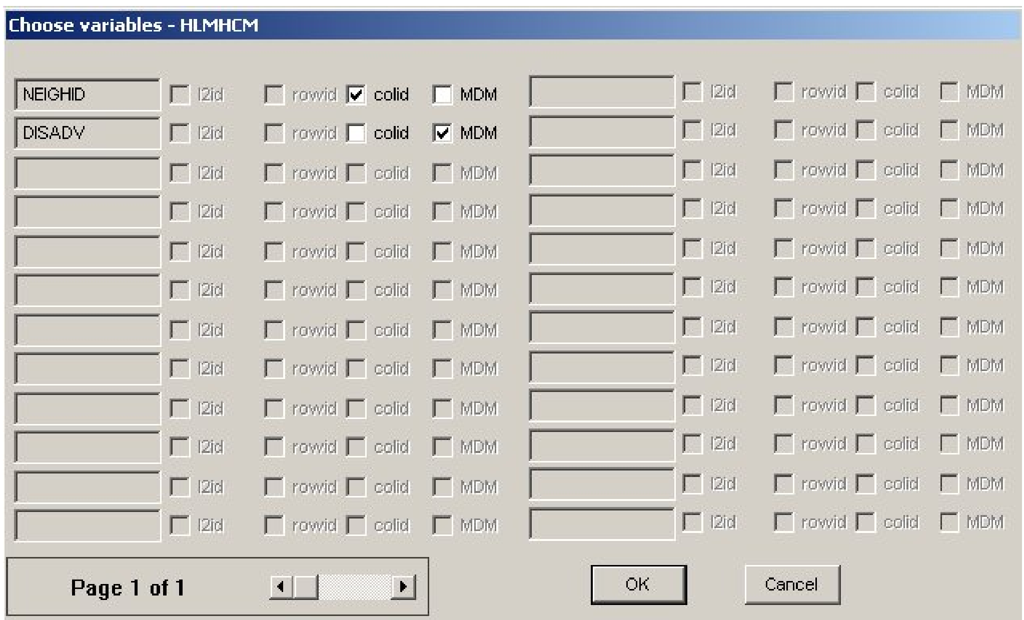


Figure 17.9 Choose variables HLMHCM dialog box for level-3 column-factor file, NEIGH.SAV

17.2 Executing analyses based on the MDM file

Once the MDM file is constructed, it can be used as input for the analysis. Model specification has three steps:

1. Specification of the level-1 model. In our case we shall model mathematics achievement (MATH) as the outcome, to be predicted by AGE8. Hence, the level-1 model will have two coefficients for each student: the intercept and the AGE slope.
2. Specification of the level-2 prediction model. Here each level-1 coefficient – the intercept and the AGE8 slope in our example – becomes an outcome variable. We may select certain student characteristics to predict each of these level-1 coefficients. In principle, the level-2 parameters

then describe the distribution of growth curves cross-classified by schools and neighborhoods.

3. Specification of level-1 coefficients as random or non-random across level-two units. We shall model the intercept and the AGE8 slope as varying randomly across the students cross-classified by schools and neighborhoods.
4. Specification of the level-3 row- and/or column-factor prediction model. Here each level-2 coefficient becomes an outcome, and we can select row- and/or column-factor variables to predict school-to-school and neighbor-to-neighbor variation in these level-2 coefficients. In principle, this model specifies how schools and neighborhoods differ with respect to the distribution of growth curves within them.
5. Specification of the residual row and column as random or non-random, the effects associated with row-specific predictors as varying randomly or fixed over columns, and the effects associated with column-specific predictors as varying randomly or fixed over rows. We shall test whether the associations between neighborhood disadvantage (a column-specific predictor) and growth parameters vary over schools.

Following the five steps above, we first specify a model with no student-, neighborhood-, or school-level predictors. The Windows execution is very similar to the one for HCM2 as described in Section 11.2. The command file, GROWTH1.HLM, contains the model specification input responses. Figure 17.10 displays the model specified.

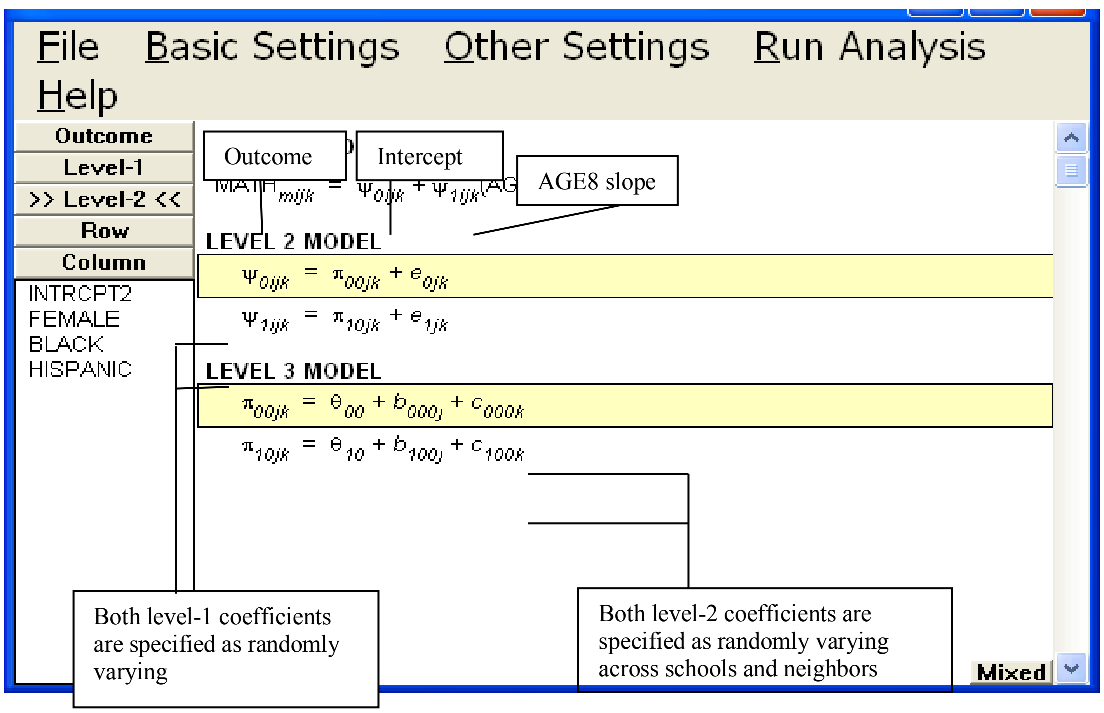


Figure 17.10 Unconditional model for the growth example

The results of the analysis are given below.

Specifications for this HLM-HCM run

Problem Title: UNCONDITIONAL LINEAR GROWTH MODEL

The data source for this run = growth.mdm
The command file for this run = growth1.hlm
Output file name = growth1.html
The maximum number of level-1 units = 2008

The maximum number of level-2 units = 567
The maximum number of row units = 224
The maximum number of column units = 74
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Level-1 Model

$$MATH_{mijk} = \psi_{0ijk} + \psi_{1ijk}*(AGE8_{mijk}) + \varepsilon_{mijk}$$

Level-2 Model

$$\psi_{0ijk} = \pi_{00jk} + e_{0jk}$$
$$\psi_{1ijk} = \pi_{10jk} + e_{1jk}$$

Row/Column Model

$$\pi_{00jk} = \theta_{00} + b_{000j} + c_{000k}$$
$$\pi_{10jk} = \theta_{10} + b_{100j} + c_{100k}$$

For starting values, data from 1967 level-1, 526 level-2, 219 rows, and 74 column records were used

Final Results - iteration 814

Iterations stopped due to small change in likelihood function

σ^2 = 0.16452

τ

INTRCPT1	AGE8
INTRCPT2,e_0	INTRCPT2,e_{1jk}
0.27574	0.07972
0.07972	0.03283

τ (as correlations)

1.000	0.838
0.838	1.000

Note that the estimated correlation between true status at AGE = 8 and true rate of change is estimated to be 0.838 for students in the same cell cross-classified by schools and neighborhoods.

Ω

INTRCPT1	AGE8
INTRCPT2	INTRCPT2
ICPTROW,b_{000}	ICPTROW,b_{100}
0.10927	-0.00606
-0.00606	0.00580

Ω (as correlations)

1.000	-0.241
-0.241	1.000

Note that the estimated correlation between true school mean status at AGE = 8 and true school-mean rate of change is estimated to be -0.241.

Δ

INTRCPT1	AGE8
INTRCPT2	INTRCPT2
ICPTCOL,c_{000}	ICPTCOL,c_{100}
0.02840	0.01363
0.01363	0.00720

Δ (as correlations)

1.000	0.954
0.954	1.000

Note that the estimated correlation between true neighborhood mean status at AGE = 8 and true neighborhood-mean rate of change is estimated to be 0.954.

The value of the log-likelihood function at iteration 814 = -1.917348E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2, INTERCEPT, θ_{00}	2.257403	0.042925	52.589	274	<0.001
For AGE8, π_1					
INTRCPT2, INTERCEPT, θ_{10}	0.880177	0.016734	52.598	274	<0.001

The above table indicates that the average growth rate is significantly positive at 0.880 logits per year, $t = 52.598$.

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, e_0	0.52510	0.27574	268	4818.18751	<0.001
AGE8, e_{1jk}	0.18119	0.03283	268	1465.94774	<0.001
σ^2,ε	0.40561	0.16452			

Note: The chi-square statistics reported above are based on only 526 of 567 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

The results above indicate significant variability among children cross-classified by schools and neighborhoods in terms of mean status at AGE = 8 ($\chi^2 = 4818.18751$, $df = 268$) and in terms of yearly rate of change ($\chi^2 = 1465.94774$, $df = 268$).

Final estimation of row level variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ INTRCPT2/ ICPTROW,b_{000}	0.33055	0.10927	224	87.39230	>0.500
AGE8/ INTRCPT2/ ICPTROW,b_{100}	0.07616	0.00580	224	201.21512	>0.500

The results above indicate there is no significant variability among schools in terms of mean status at AGE = 8 (χ^2 = 87.39230, *df* = 224) and in terms of yearly rates of change (χ^2 = 201.21512, *df* = 224).

Final estimation of column level variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/INTRCPT2/ ICPTCOL,c_{000}	0.16851	0.02840	73	1316.77855	<0.001
AGE8/INTRCPT2/ ICPTCOL,c_{100}	0.08484	0.00720	73	831.88840	<0.001

The results above indicate significant variability among neighbors in terms of mean status at AGE = 8 (χ^2 = 1316.77855, *df* = 73) and in terms of yearly rates of change (χ^2 = 831.88840, *df* = 73).

Statistics for the current model

Deviance = 3834.695088
Number of estimated parameters = 12

17.3 Specification of a level-2 and level-3 conditional model, with the effect associated with a column-specific predictor fixed

The above example involves a model that is unconditional at all levels. In this model we set up a level-2 and a row-factor prediction model.

To set up the level-2 model:

Select the equation containing ψ_{pijk} to be modeled, a list box for level-2 variables (>>**Level-2**<<) will appear. Figure 17.12 shows the models with BLACK and HISPANIC as the level-2 predictors. In the interest of parsimony, all level-2 coefficients are fixed. (To specify either of them as randomly varying, select the equation containing a specific regression coefficient, π_{pqjk}, and click on b_{pqrj} and/or c_{pqsk}).

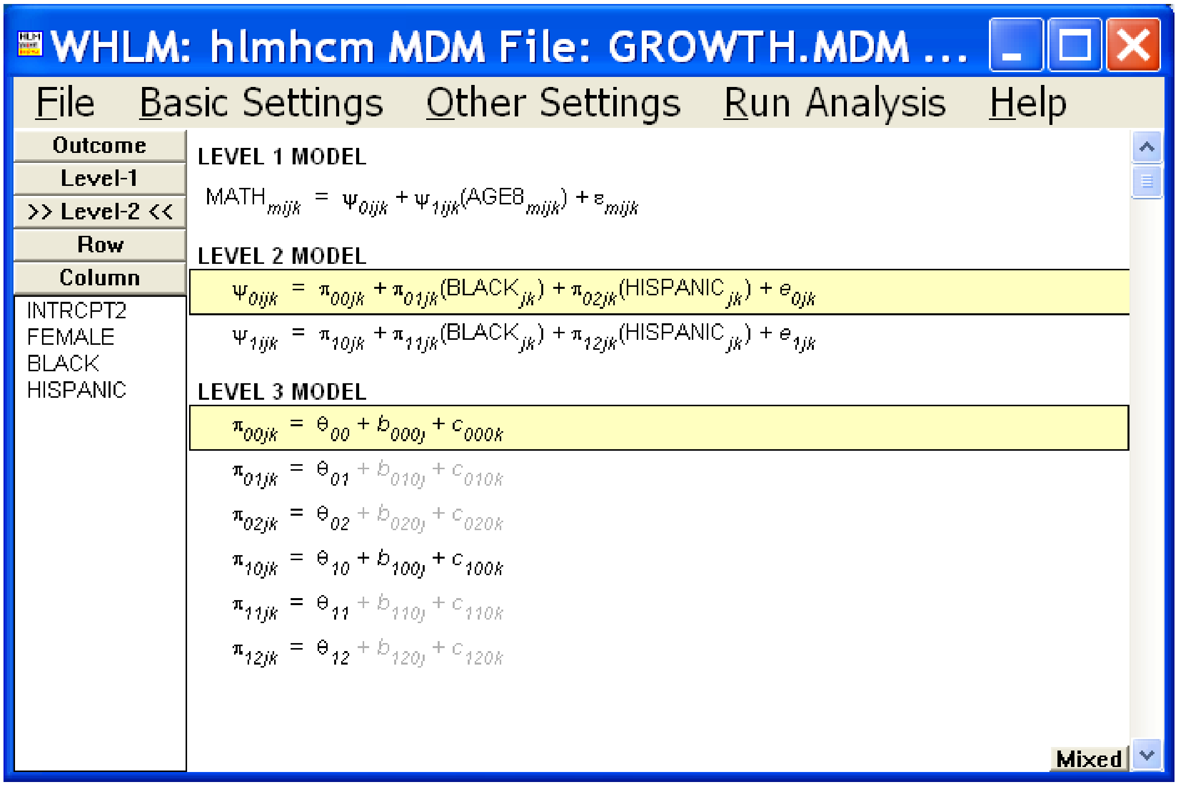


Figure 17.12 Level-2 prediction model for the growth study

To set up the level-3 row or/and column-factor prediction model:

Select the equation containing π_{pqjk} to be modeled, a list box for level-3 row-factor variables (>>**Row**<<) will appear. To display level-3 column-factor variables, click on **Column** and the corresponding list box of variables. Figure 17.13 shows the level-3 column-factor prediction model with DISADV as the covariate. In the level-3 model, we treated the association between neighborhood disadvantage and the growth parameters as fixed across all schools. Note that b_{001j} and b_{101j} are disabled. We relax this assumption in our next model.

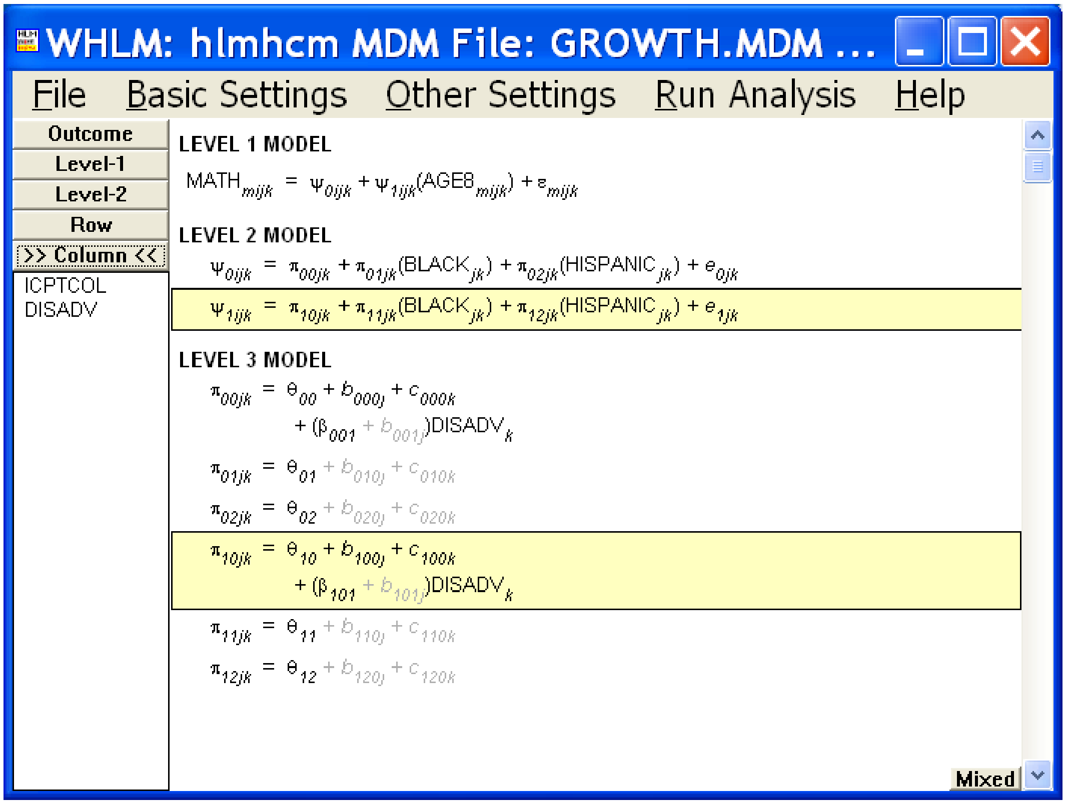


Figure 17.13 Conditional model for the growth study, with neighborhood disadvantage effect fixed

The results of the analysis are given below.

Specifications for this HLM-HCM run

Problem Title: CONDITIONAL LINEAR GROWTH MODEL, WITH NEIGHBORHOOD DISADVANTAGE

The data source for this run = growth.mdm
The command file for this run = growth2.hlm
Output file name = growth1.html
The maximum number of level-1 units = 2008
The maximum number of level-2 units = 567
The maximum number of row units = 224
The maximum number of column units = 74
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Level-1 Model

$$\mathrm{MATH}_{mijk} = \psi_{0ijk} + \psi_{1ijk}*(\mathrm{AGE8}_{mijk}) + \varepsilon_{mijk}$$

Level-2 Model

$$\psi_{0ijk} = \pi_{00jk} + \pi_{01jk}*(\mathrm{BLACK}_{jk}) + \pi_{02jk}*(\mathrm{HISPANIC}_{jk}) + e_{0jk}$$
$$\psi_{1ijk} = \pi_{10jk} + \pi_{11jk}*(\mathrm{BLACK}_{jk}) + \pi_{12jk}*(\mathrm{HISPANIC}_{jk}) + e_{1jk}$$

Row/Column Model

$\pi_{00jk} = \theta_{00} + b_{000j} + c_{000k}$
$+ \text{DISADV}_k{*}(\beta_{001})$
$\pi_{01jk} = \theta_{01}$
$\pi_{02jk} = \theta_{02}$
$\pi_{10jk} = \theta_{10} + b_{100j} + c_{100k}$
$+ \text{DISADV}_k{*}(\beta_{101})$
$\pi_{11jk} = \theta_{11}$
$\pi_{12jk} = \theta_{12}$

For starting values, data from 1967 level-1, 526 level-2, 219 rows, and 74 column records were used

Final Results - iteration 1300

Iterations stopped due to small change in likelihood function

σ^2 = 0.16386

τ

INTRCPT1	AGE8
INTRCPT2,e_0	INTRCPT2,e_{1jk}
0.27546	0.08088
0.08088	0.03538

τ (as correlations)

1.000	0.819
0.819	1.000

Ω

INTRCPT1	AGE8
INTRCPT2	INTRCPT2
ICPTROW,b_{000}	ICPTROW,b_{100}
0.09506	-0.00711
-0.00711	0.00320

Ω (as correlations)

1.000	-0.408
-0.408	1.000

Δ

INTRCPT1	AGE8
INTRCPT2	INTRCPT2
ICPTCOL,c_{000}	ICPTCOL,c_{100}
0.01332	0.00656
0.00656	0.00338

Δ (as correlations)

1.000	0.979
0.979	1.000

The value of the log-likelihood function at iteration 1300 = -1.900326E+003

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
INTRCPT2,					
INTERCEPT, θ_{00}	2.639580	0.090173	29.272	270	<0.001
DISADV, γ_{001}	-0.001726	0.050288	-0.034	222	0.973
BLACK,					
INTERCEPT, θ_{01}	-0.443355	0.103660	-4.277	270	<0.001
HISPANIC,					
INTERCEPT, θ_{02}	-0.468207	0.098680	-4.745	270	<0.001
For AGE8, π_1					
INTRCPT2,					
INTERCEPT, θ_{10}	0.933753	0.035488	26.312	270	<0.001
DISADV, γ_{101}	-0.050330	0.020853	-2.414	222	0.016
BLACK,					
INTERCEPT, θ_{11}	-0.105109	0.040518	-2.594	270	0.010
HISPANIC,					
INTERCEPT, θ_{12}	-0.036124	0.038978	-0.927	270	0.354

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1, e_0	0.52484	0.27546	268	6019.63723	<0.001
AGE8, e_{1jk}	0.18811	0.03538	268	1363.77540	<0.001
σ^2,ε	0.40480	0.16386			

Note: The chi-square statistics reported above are based on only 526 of 567 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Final estimation of row level variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/ INTRCPT2/ ICPTROW, b_{000}	0.30832	0.09506	224	79.66634	>0.500
AGE8/ INTRCPT2/ ICPTROW, b_{100}	0.05653	0.00320	224	182.46985	>0.500

Final estimation of column level variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/INTRCPT2/ ICPTCOL, c_{000}	0.11543	0.01332	73	2085.34935	<0.001
AGE8/INTRCPT2/ ICPTCOL, c_{100}	0.05810	0.00338	73	1337.03181	<0.001

Statistics for the current model

Deviance = 3800.651318
Number of estimated parameters = 18

The results suggest that:

- Compared to their reference group (non-Black and non-Hispanic); African and Hispanic American students on average had a lower mathematics score at age 8 than did white students. Also, African American students had a significantly lower growth rate in mathematics achievement ($\hat{\theta}_{11}$ = -0.105, t = -2.594) than did white students.
- Neighborhood disadvantage had a negative association with the growth rate of the reference group ($\hat{\gamma}_{101}$ = -0.050, t = -2.414).
- The column level variance at level 3 of each growth parameter was substantially reduced (> 50%). The residual variation between neighborhoods in c_{000} (estimated at 0.01332) and in c_{100} (estimated at 0.00338) are less than half of those in the unconditional models (0.02840 and 0.00720).

17.4 Other program features

HLMHCM models provide options for multivariate hypothesis tests for the fixed effects and the variance-covariance components. A “no-intercept” option is available for the level-1, level-2, and the level-3 row and column models. In addition to continuous outcomes, they handle binary and count outcomes. HLMHCM also allows users to diagonalize the τ, Ω , and Δ when interests focus only on the diagonal elements.

18 Graphing Data and Models

HLM2 and HLM3 provide the ability to make data-based and model-based graphs. Data-based graphs allow examination of univariate and bivariate distributions. Model-based graphs, which can be produced by the HLM2, HLM3, HMLM, HMLM2 and HCM2 modules of WHLM, facilitate visualization and presentation of analytic results for the whole or a subset of the population of interest. They also enable users to check the tenability of underlying model assumptions.

18.1 Data – based graphs – two level analyses

18.1.1 Box-and-whisker plots

We first illustrate how to use box-and-whisker plots to display univariate distributions of level-1 variables for each level-2 unit, with and without a level-2 classification variable. Using the HS&B data (see Section 2.5.1.1), we display graphical summaries of the mathematics achievement variable, MATHACH, and simultaneously show differences in the student scores within a school and among schools.

To prepare box-and-whisker plots

1. From the **HLM** window open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open HSB.MDM.
3. Open the **File** menu, choose **Graph Data … box-whisker plots** to open an **Choose Y for box plot** dialog box (see Figure 18.1).
4. Select MATHACH in the **Y-axis** drop-down list box.
5. Choose the number of groups to be used for graphing. There are three options: (a) **First ten groups**; (b) **Random sample of spec'd prob** (specified probability) and (c) **All groups** (n = total number of groups) for users to choose from in the **Number of groups** drop-down list box. The selection of option (b) requires the user to specify the proportion or percent of the level-2 units to be included. to do so, enter a probability into the text box for **Probability (0 to 1)**. In our example, we randomly select 10 percent of the schools to illustrate. we select **Random sample of spec'd prob** from the **Number of groups** drop-down list box. Enter 0.1 into the text box for **Probability (0 to 1)** to indicate that 10 percent or a proportion of .1 of the schools will be used.
6. Specify the arrangement of the plots by either (a) the original order of the groups as they appear in the data set or (b) the median in an ascending order. Click on the selection button for **median** in the **Sort by** section to arrange the box-and-whisker plots of MATHACH by median in an ascending order (see Figure 18.2).

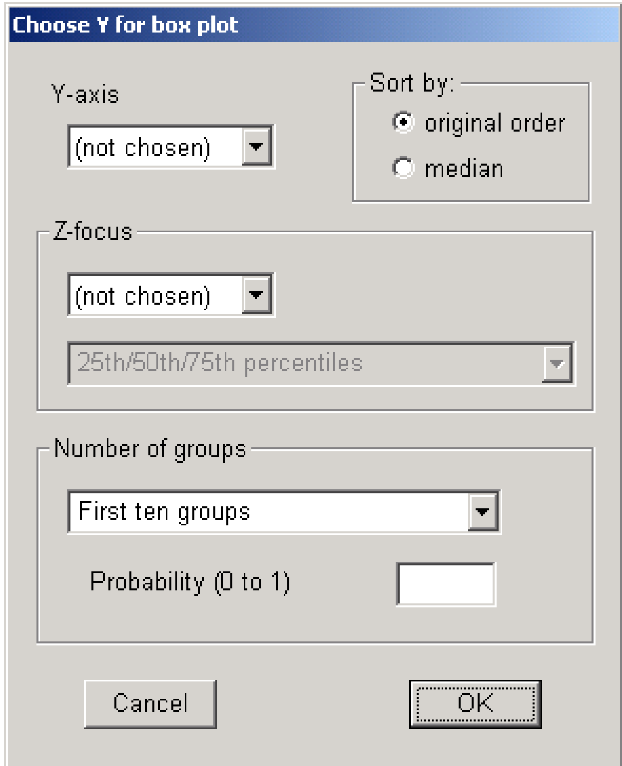


Figure 18.1 Choose Y for box plot dialog box

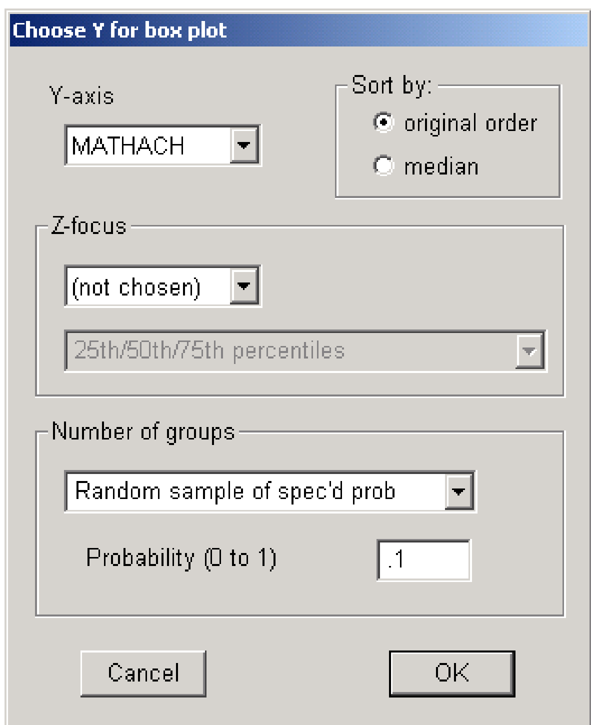


Figure 18.2 Choose Y for box plot dialog box for the MATHACH example

7. Click **OK** to display the plots (see Figure 18.3).

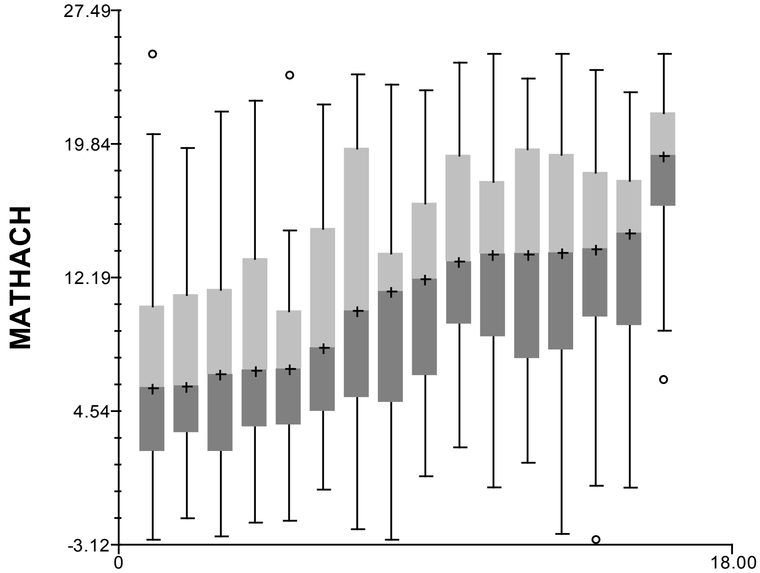


Figure 18.3 Box and whisker plot for MATHACH

The figure gives side-by-side graphical summaries of the distributions of MATHACH for the sixteen schools sorted by median. The x-axis denotes number of schools in the display and the y-axis mathematics achievement. The plot tells us that the first school from the left has a median score of about 6.05, which is the lowest school median in this group. The distribution of the scores of the students in this school is positively skewed and there is an outlier at the upper end.

The third and the fourth schools from the left have similar distributions of mathematics scores. Compared to the distribution of the scores of the adjacent school on the right, however, the scores of these two schools display greater variability, as defined by the lengths of the boxes or interquartile ranges. In addition, there is an outlier at the upper end of the distribution for the fifth school. The highest median mathematics score among the 16 schools was 19.08.

8. (Optional) WHLM allows users to list the raw data of a specific group that is graphically summarized in one of the box-and-whisker plots as well. To see the data of a specific level-2 unit, click on one of the box-and-whisker plots (near the median is usually a good place) in Figure 18.3, which brings up the following dialog box:

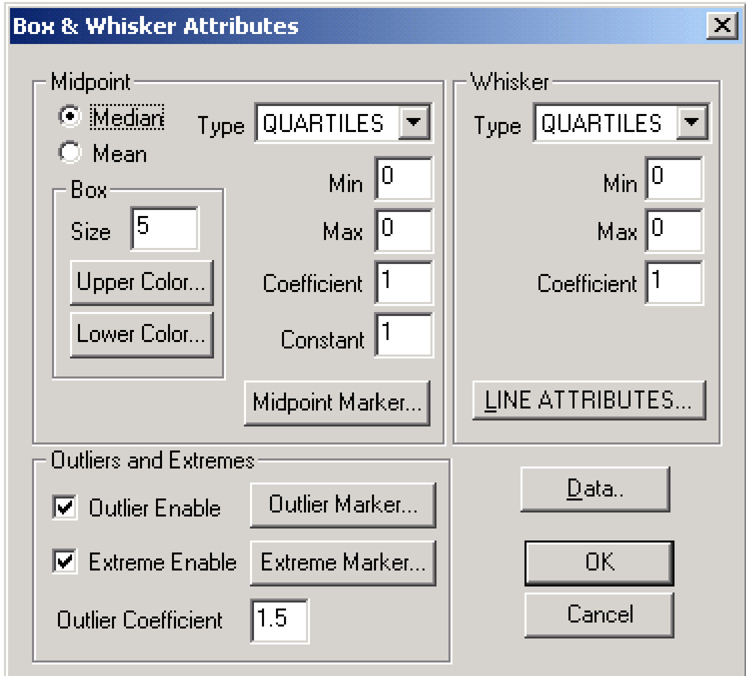


Figure 18.4 Box & Whisker Attributes dialog box

For a description of the options, see Table 18.1.

Click **Data** and then a dialog box containing the data of a specific group will appear. In our example, we examine the raw scores of the school with the highest median (see Figure 18.5). The title bar of Figure 18.5 tells us the level-2 ID of the box-and-whisker plot we selected is 3427. # is a zero-based counter for group plots.

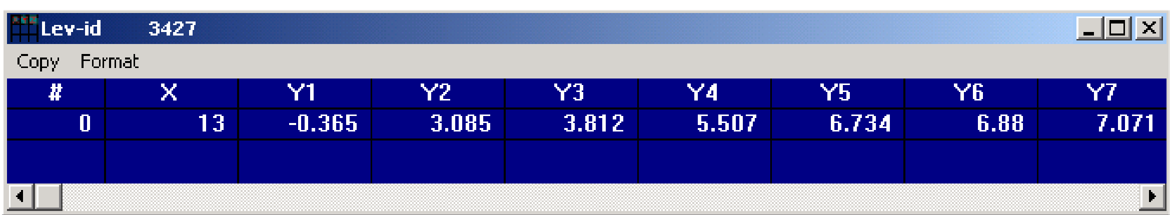


Figure 18.5 Data for School 3427 dialog box

As the box-and-whisker plots are plotted individually in the example, it is 0. X tells us that the data are from the thirteenth school displayed on the plot. Y1 to Y11 list the mathematics scores for the first eleven students in School 3427. Move the bottom scroll box to the left to display more scores for the other students.

9. (Optional) To edit the graph, open the **Edit** menu and choose **Graph Parameters...**. The user can change attributes such as size and color of the graph, border, and plotting area. By choosing **Copy graph** or **Copy current page** (when there are more than one pages of graphs), users can directly copy and paste the graph or current page into a word processing or graphics document.
10. (Optional) To print the graph, open the **File** menu, select **Print current page** or **Print selected graph** when there are more than one graph. Users can choose **Printing Options...** to change printing parameters such as choice of background, border type, aspect ratio (the ratio of the x-axis length to the y-axis length, the default is 5/3), and printing style.

Table 18.1 Definitions and options in the Box & Whisker Attributes dialog box

	Key terms	Function	Option	Definition
1	Midpoint	Specify the type of average used	2 choices	1. Median 2. Mode
2	Box Size	Specify the width in units of the axis that the box width is parallel to.		
3	Min, Max, and Coefficient for box or whisker and Constant for box	Min and Max specify the box percentage minimum and maximum when the box or whisker Type is PERCENT. The coefficient is the box or whisker coefficient by which the selected range value will be multiplied. The Constant is the box constant, valid when the box Type is CONSTANT.		
4	Midpoint marker	Display a **Marker Attributes** dialog box that allows the user to specify the shape, color, size, and style of the midpoint marker.		
5	Line attributes	Display a **Line Parameters** dialog box that allows the user to specify the thickness, color, and style of the whisker.		

11. To save the graph for future use by opening the **File** menu and choose **Save as metafile**. A **Save as** dialog box will open. Enter a filename for the file and click **OK**. The file can be saved as an Enhanced Metafile (.EMF) (default and preferred as it holds more information than the other option) or Windows Metafile (.WMF). Users can use word processing programs to insert the graph file into the text. For example, to insert the saved .EMF file into Word, choose **Insert-...Picture-...From File** from Word's main menu.
12. (Optional) To make modifications to the specifications, select **Graph Settings**. The **Equation Graphing** dialog box will appear. We are going to illustrate this by adding a level-2 classification variable next.

To include a level-2 classification variable

13. After choosing the **Y-Axis** variable, select the level-2 classification variable in the **Z-focus** drop-down list box. There are two types of level-2 classification variables, categorical and continuous. For categorical variables, WHLM will classify the plots with the levels of the variables. For continuous variables, users can choose either to dichotomize them using median splits, or trichotomize them into three groups: (a) 0 to 24th percentile; (b) 25th to 75th percentile; and (c) 76th percentile and above. These two options, available when a continuous classification variable is chosen, can be found in the lower **Z-focus** drop-down list box. In our example, we will choose school sector, Catholic vs. public school, as the classification variable. To continue working on the plot we have just made, click **Graph Settings** to open the **Equation Graphing** dialog box. Select SECTOR in the **Z-focus** dialog box. The following graph will be displayed (see Figure 18.6).

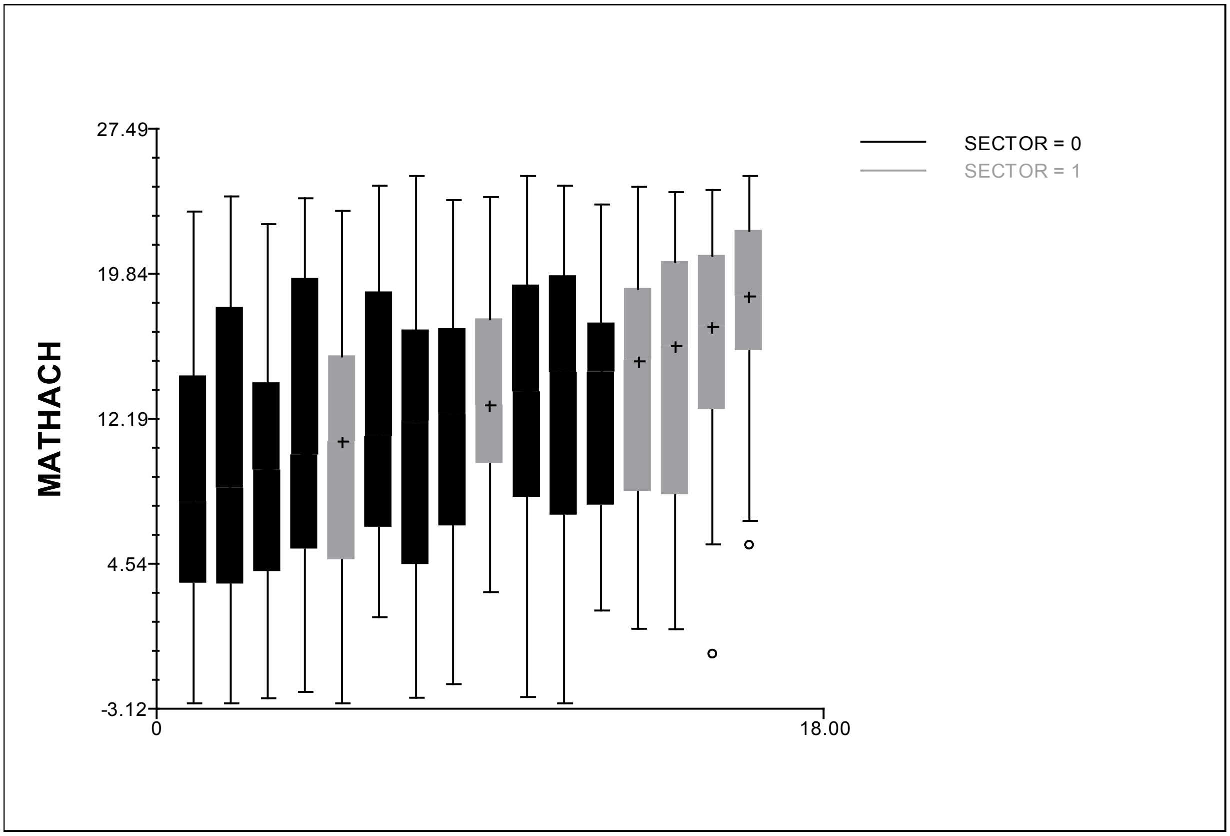


Figure 18.6 Box-and-whisker plots for MATHACH for a random sample of schools as classified by school sector

In the graph, the box-and-whisker plots for Catholic and public schools are coded differently (red for Catholic and blue for public schools). The colored graphs (not showed here) suggest that the three schools that have the highest median mathematics scores are Catholic schools. The school with the lowest average belongs to the public sector.

Users can edit the legends by clicking on them in the graph above to open the **Legend Parameters** dialog box (see Figure 18.7), which allows them to make changes in the titles of the legends, their sizes and font types, and the display of the legend box. For example, one may like to change SECTOR = 0 in the text box of Figure 18.7 to PUBLIC = 0 and SECTOR = 1 to CATHOLIC = 1.

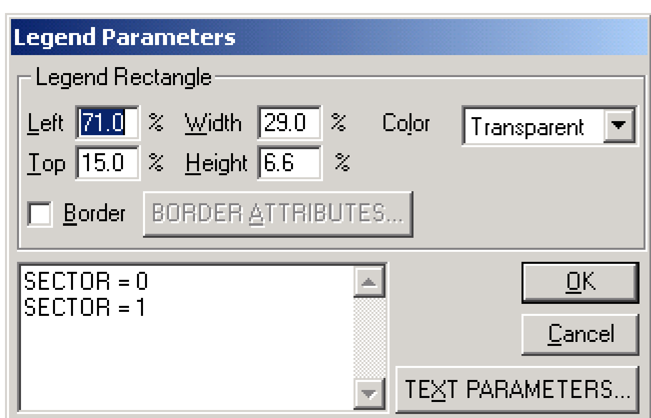


Figure 18.7 Legend Parameters dialog box

18.1.2 Scatter plots

In the previous section, we illustrated how to graphically summarize and compare univariate distributions of level-1 variables, with and without a level-2 classification variable. Now we demonstrate how to use data-based scatter plots to explore bivariate relationships between level-1 variables for individual or a group of level-2 units, with and without controlling level-2 variables. We will continue to use the HS&B data set and we are going examine the relationships between MATHACH and SES for a group or individual schools, with and without controlling for the sector of the school.

To prepare a scatter plot

1. From the HLM window, open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open HSB.MDM.
3. Open the **File** menu, choose **Graph Data …. line plots, scatter plots** to open a **Choose X and Y variables** dialog box (see Figure 18.8).
4. Select SES from the **X-axis** drop-down list box.
5. Select MATHACH from the **Y-axis** drop-down list box.
6. Select number of groups. In this example, select **Random sample of spec'd prob** and enter .2 into the textbox to select 20 percent of the schools.

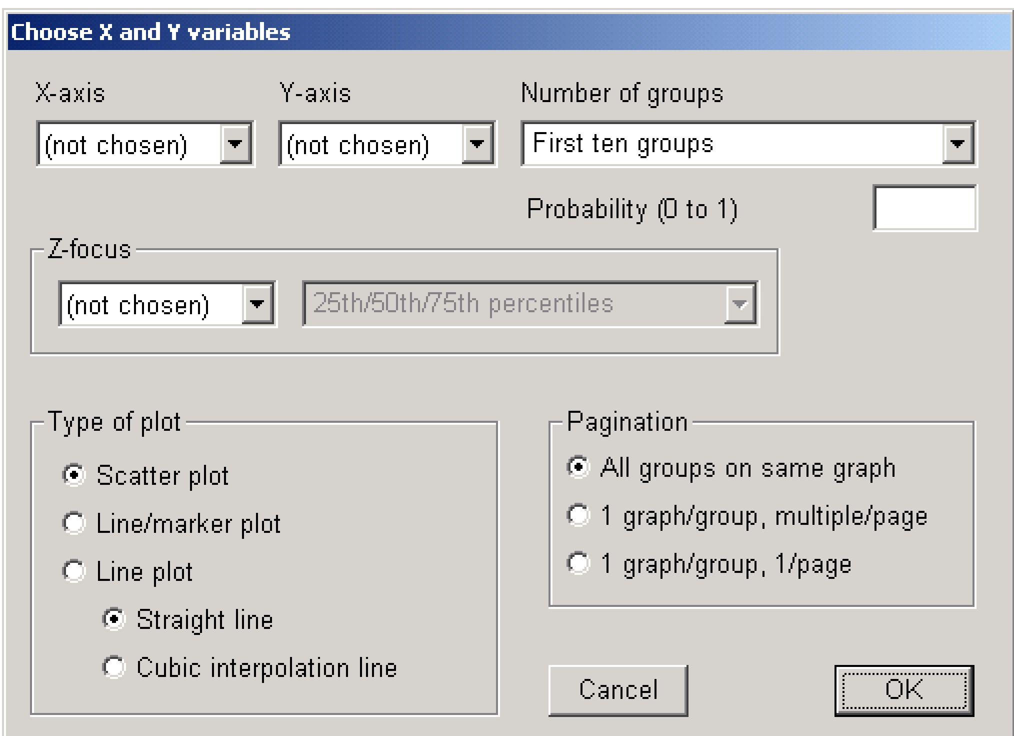


Figure 18.8 Choose X and Y variables dialog box

7. Select type of plot. Users can select one of the two major types of plots: (a) scatter plot; and (b) line plot with and without markers or asterisks showing where the data points are. Click the selection button for **Scatter plot** (default) for this example.
8. Select type of pagination. There are three options: (a) all groups on the same graph (default); (b) one graph per groups and to display a maximum of eight graphs on one page, and (c) 1

graph per group and to be displayed on multiple pages. In this example, we will display the bivariate relationship between SES and MATHACH for all the selected schools on a single graph. We choose the option **All groups on same graph** accordingly.

9. Click **OK** to make the scatter plot. This gives us the following graph (see Figure 18.9), indicating a moderate positive association between SES and MATHACH, and suggesting that both variables have "ceilings" (upper limits).
10. For more information on the editing, printing, saving and modification options, see Steps 11 to 13 in Section 18.1.1.

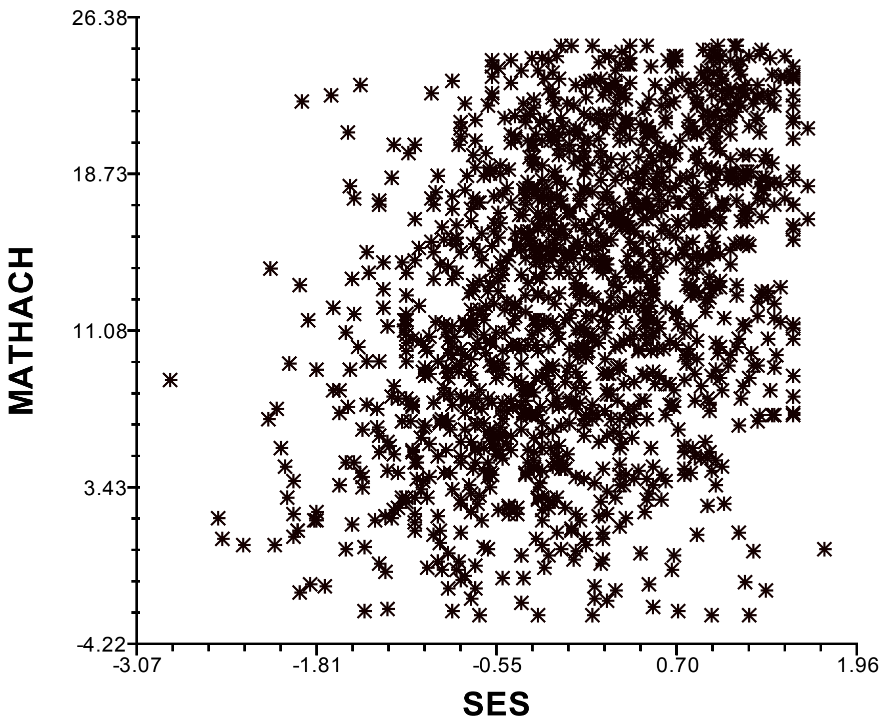


Figure 18.9 Scatter plot for the 20% random sample of cases

To include a level-2 classification variable

11. After specifying the variables for the x- and y-axis, select the controlling variable from the **Z-focus** drop-down list box. As in the case for the box-and-whisker plots, users can choose either a categorical and continuous controlling variable (see Step 14 in Section 18.1.1). In our example, we will choose school sector, Catholic vs. public school, as the controlling variable. To continue working on the scatter plot we have just made, click **Graph Settings** to open the **Equation Graphing** dialog box. Select SECTOR in the **Z-focus** dialog box. The following graph will be displayed (see Figure 18.10).

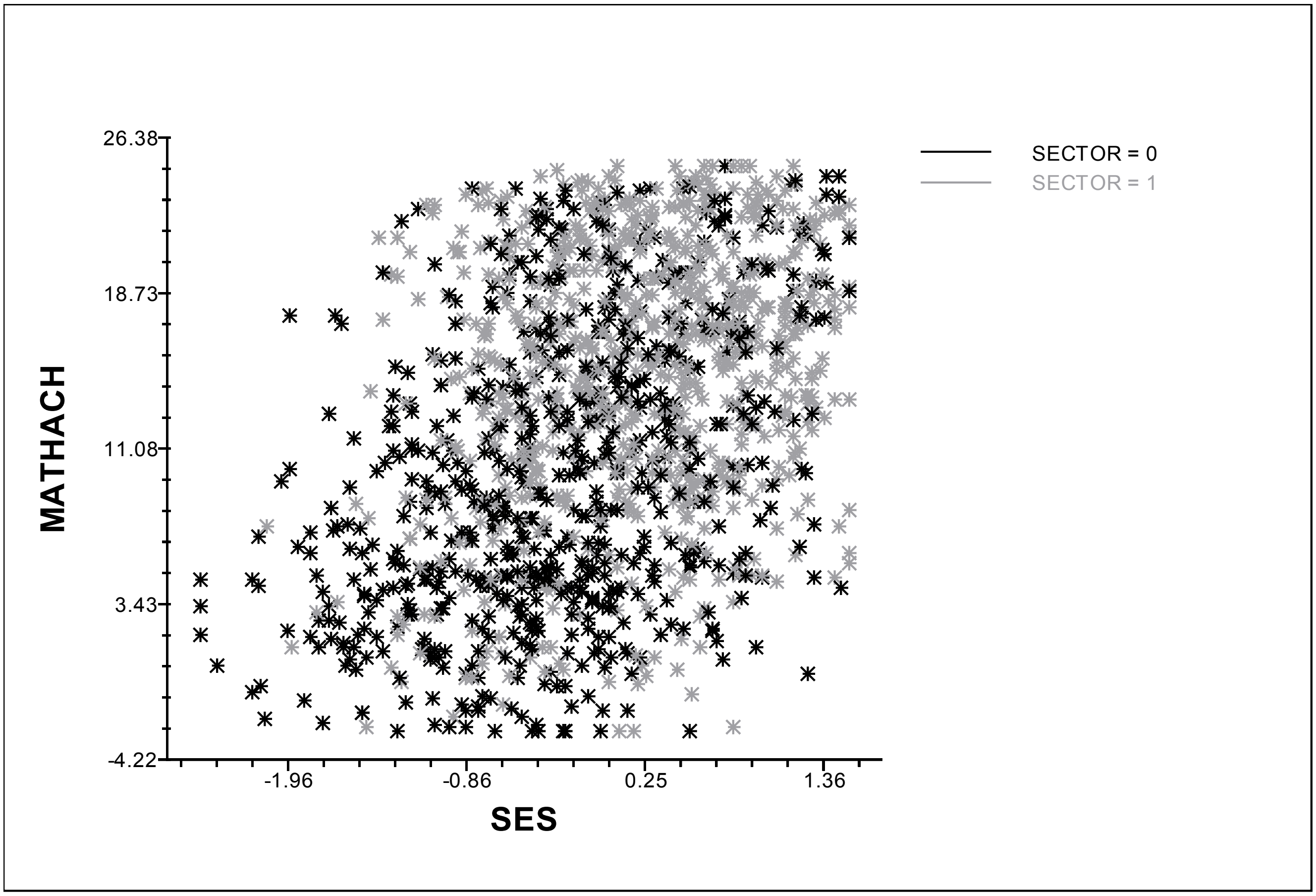


Figure 18.10 Scatter plot for random sample by SECTOR

The color-coded scatter plot shows that there is not in general a radical difference in the SES-MATHACH relationship for the two types of schools.

It may be helpful to use a different pagination option to help us to discern the relationships for these two groups of school. Instead of having all the groups on the same graph, we select the **1 graph/group, multiple/page** pagination option. This gives us Figure 18.11, where we see how the two groups of schools vary in their SES and MATHACH distributions. Note, for example, that school 8946 has high levels of SES and that in school 4325, the association between SES and MATHACH appears a bit stronger than in several of the other schools. WHLM puts a maximum of 8 groups in a window. We can page back and forth using the **->** and **<-** buttons in the lower right corner of the window to display the scatter plots for other schools.

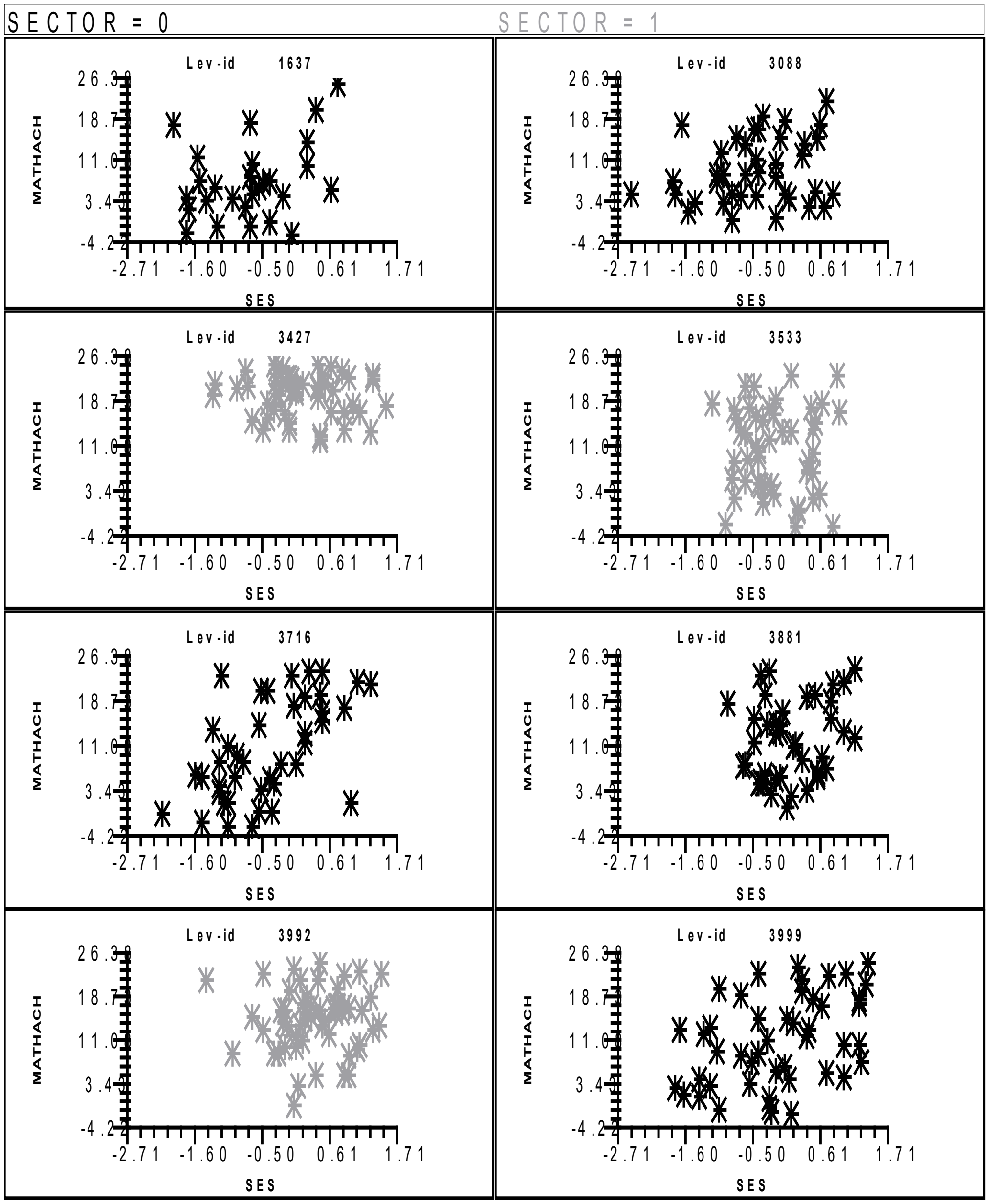


Figure 18.11 Scatter plots for individual schools on one page

As an elaboration of this, we can also choose on the **Graph Settings** dialog box to have each group's plot in a separate graph by choosing **1 graph/group, 1/page**, as shown below:

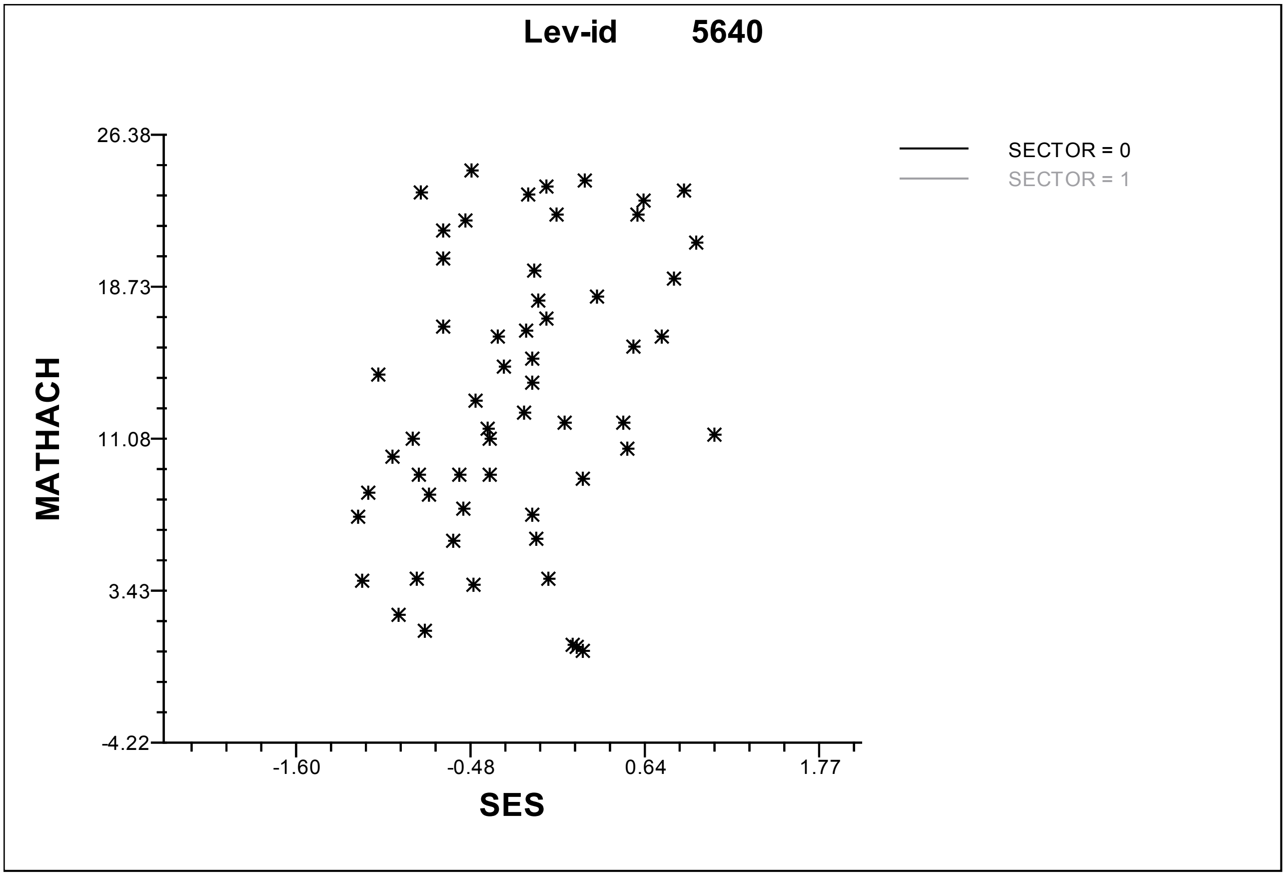


Figure 18.12 Individual scatter plot for school 7697

18.1.3 Line plots – two-level analyses

In scatter plots, observations on a pair of level-1 variables are plotted to examine their association, with and without a level-2 controlling variable. In line plots, level-1 repeated measures observations are joined by lines to describe changes or developments over time during the course of the research study. We illustrate this type of plot with data from two studies of children's vocabulary development (Huttenlocher, Haight, Bryk, and Seltzer, 1991, see also *Hierarchical Linear Models*, pp. 170-179). Twenty-two children were observed in the home on three to seven occasions at 2 to 4-month intervals during their second year of birth. A measure of the child's vocabulary size at each measurement occasion was derived from these observations. In this example, the level-1 file, VOCABI1.SAV has

- AGE Age in months
- VOCAB Vocabulary size
- AGE12 Age in months minus 12
- AGE12Q AGE12*AGE12

The level-2 data file, VOCABL2.SAV, consists of 22 children and an indicator variable for gender

- MALE An indicator for gender (1 = male, 0 = female)

To prepare a scatter plot

1. From the HLM window, open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open VOCAB.MDM.
3. Open the **File** menu, choose **Graph Data...line plots, scatter plots** to open an **Choose X and Y variables** dialog box (see Figure 18.8).
4. Select AGE from the **X-axis** drop-down list box.
5. Select VOCAB from the **Y-axis** drop-down list box.
6. Select number of groups. In this example, we include all the children in the display by selecting **All groups (n = 22)** in **Number of groups** drop-down list box.
7. Select type of line plot and method of interpolation. Users can select line plots with and without markers or asterisks showing where the data points are. The two types of interpolation are linear and cubic. In linear interpolations, the data points are simply joined by straight line segments. Cubic interpolations may be chosen to provide a smoother function and more continuity between the segments. For our example, suppose we want a line plot with no markers that is graphed with the linear interpolation method. Click the selection button for **Straight line**.
8. Select type of pagination. In this example, we want to have the trajectories for all children on the same graph and select **All groups on same graph pagination** option accordingly. When all the choices are made, the **Choose X and Y variables** dialog box should look like the one shown in Figure 18.13.

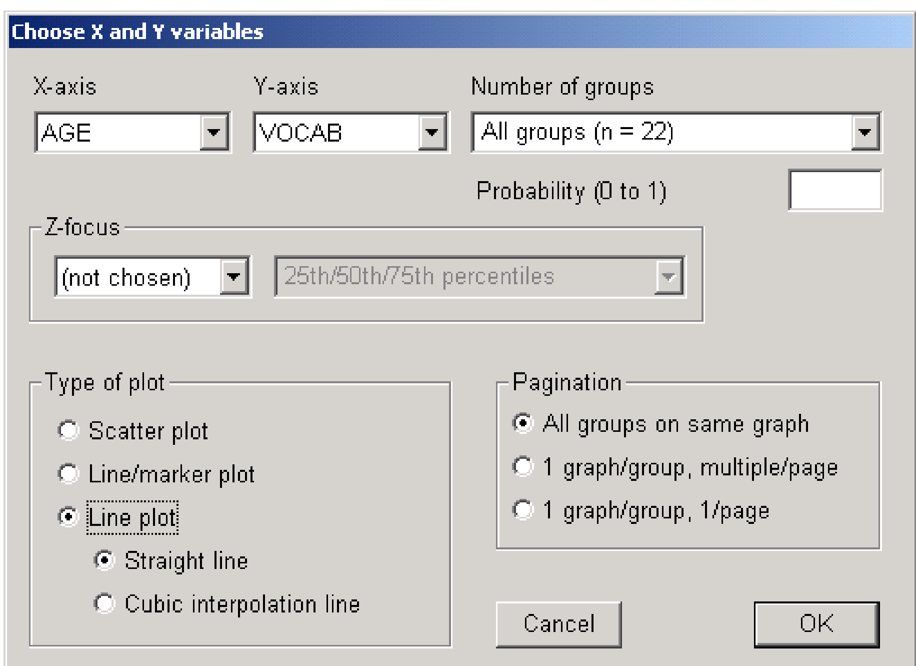


Figure 18.13 Choose X and Y variables dialog box for line plot of VOCAB and AGE

9. Click **OK** to make the line plot. The following graph will appear.

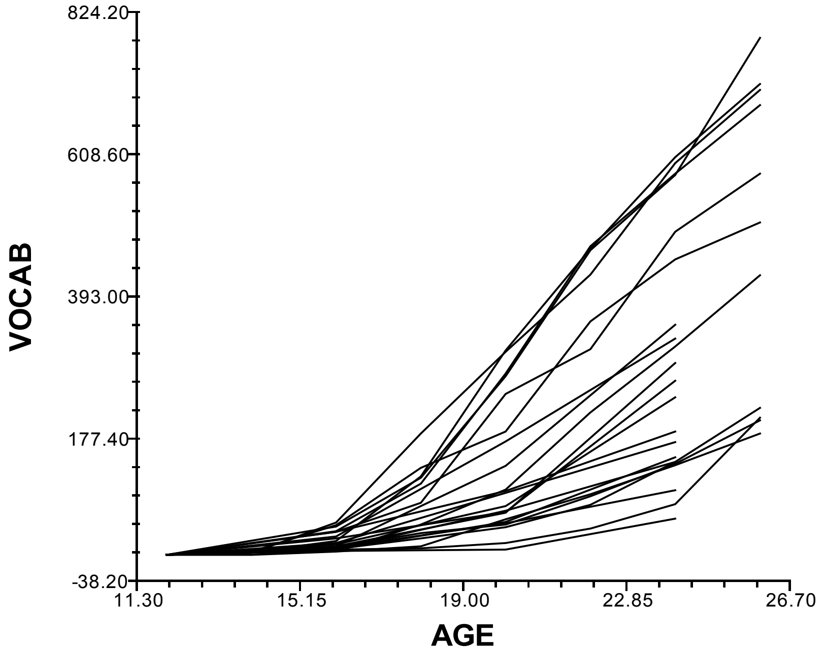


Figure 18.14 Line plot of the vocabulary score vs. the age of the child

We see that, for all children, vocabulary size is near zero at around a year of age (12 – 15 months) and that for each child, vocabulary size increases, typically quite rapidly during the second year of life.

To include a classifying level-2 variable

Now we want to look at the difference between boys and girls. On the menu of the graph dialog box, click **Graph Settings**. Here we choose the level-2 variable FEMALE as a **Z-focus** variable. For illustrative purposes, we will use the cubic interpolation method this time by clicking the selection button for **Cubic interpolation line**. The colored version of the following graph shows that girls' vocabulary tends to grow more rapidly than that of boys, on average.

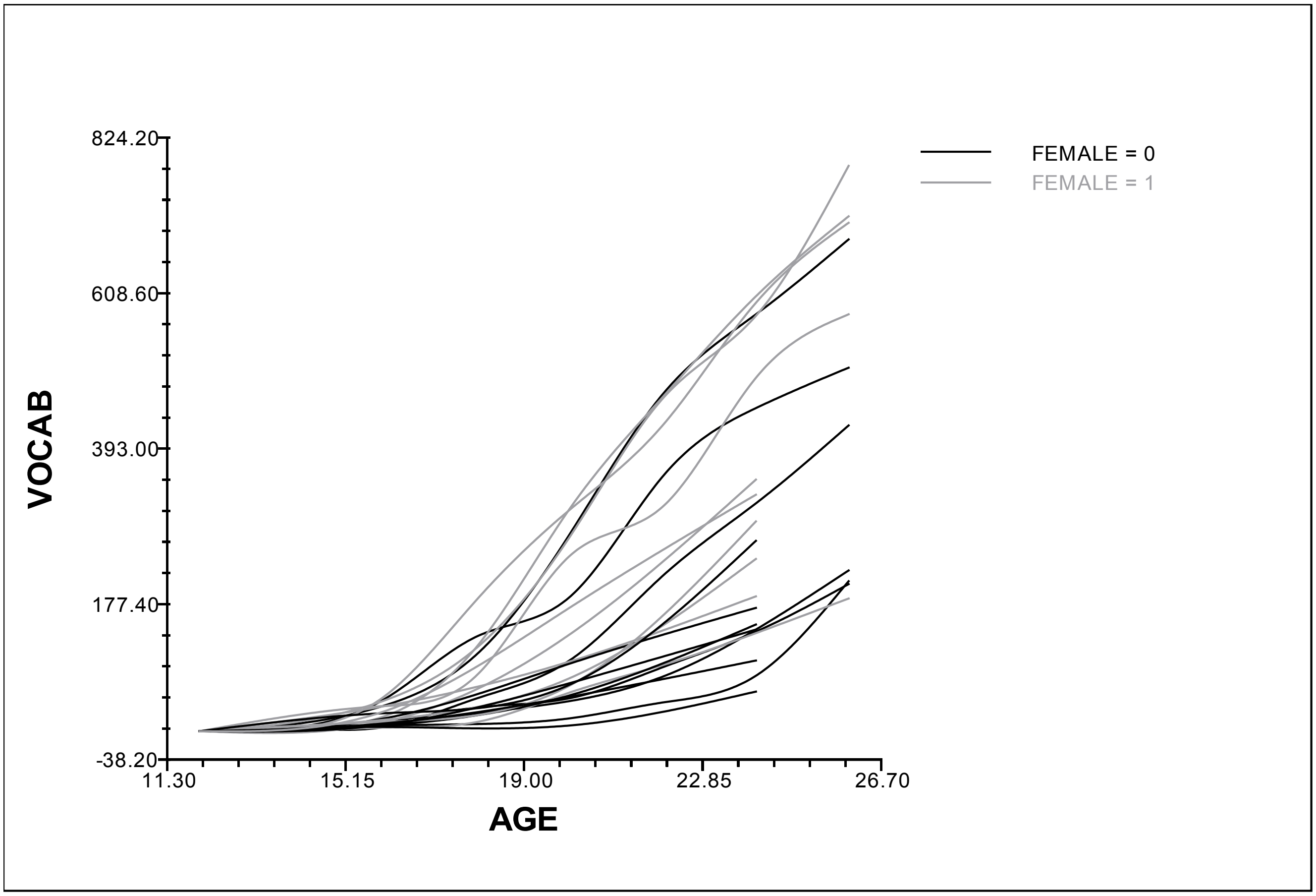


Figure 18.15 Cubic interpolation line plot of the difference between boys and girls

18.2 Model-based graphs – two level

18.2.1 Model graphs

WHLM provides graphing options to display the relationships between the outcome and the predictor(s) based on the final analytic results. The options allow us to visually represent the results of the models for the whole or a subset of population, and to graphically examine underlying model assumptions as well. Below we provide a 2-level example of a growth curve analysis of pro-deviant attitude for fourteen-year-old youth over a period of five years with data from the National Youth Survey (Elliot, Huizinga, & Menard, 1989; Raudenbush & Chan, 1993). In our example, the level-1 file, NYSW2.SAV, has 1,066 observations collected from interviewing annually fourteen-years-old youths beginning at 1976:

- ATTIT A nine-item scale assessing attitudes favorable to deviant behavior

  Subjects were asked how wrong (very wrong, wrong, a little bit wrong, not wrong at all) they believe it is for someone their age, for example, to damage and destroy property, use marijuana, use alcohol, sell hard drugs, or steal.

  The measure was positively skewed; so a logarithmic transformation was performed to reduce the skewness.
- AGE16 Age of participant at a specific time minus 16
- AGE16S = AGE16 * AGE16

The level-2 data file, NYSB2.SAV, consists of 241 youths and three variables per participant.

- FEMALE An indicator for gender (1 = female, 0 = male)
- MINORITY An indicator for ethnicity (1 = minority, 0 = other)
- INCOME Income

At level-1, we formulate a polynomial model of order 2 using AGE16 and AGE16S (see Figure 18.16) with FEMALE and MINORITY as covariates at level-2 modeling π_0, the expected pro-deviant attitude score at age 16 for subject *j*; π_1 and π_2, which are the expected average linear and quadratic growth rate for pro-deviant attitude score respectively. The procedure for setting up the model is given in 2.5.2. We will ask WHLM to graph the predicted values of pro-deviant attitude scores at different ages for different gender-by-ethnicity groups.

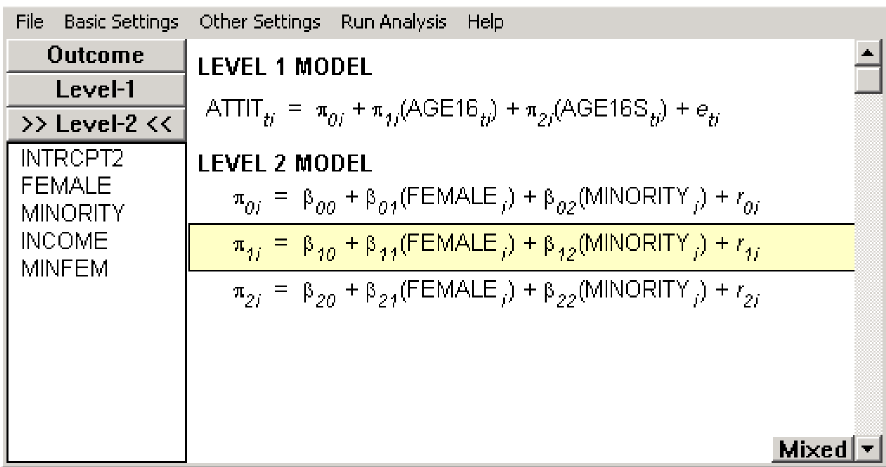


Figure 18.16 A polynomial model of order 2 with FEMALE and MINORITY as level-2 covariates

To prepare the graph

1. After running the model, select **Basic Settings** to open the **Basic Model Specifications – HLM2** dialog box.
2. Enter a name for the graphics file. The default name is grapheq.geq.
3. Enter a title and name the output filename, save the command file, and run the analysis as described in section 2.5.2.
4. Open the **File** menu and choose **Graph Equations**. An **Equation Graphing dialog** box will open (see Figure 18.17). Table 18.2 lists the definitions and options in the **Equation Graphing** dialog box.

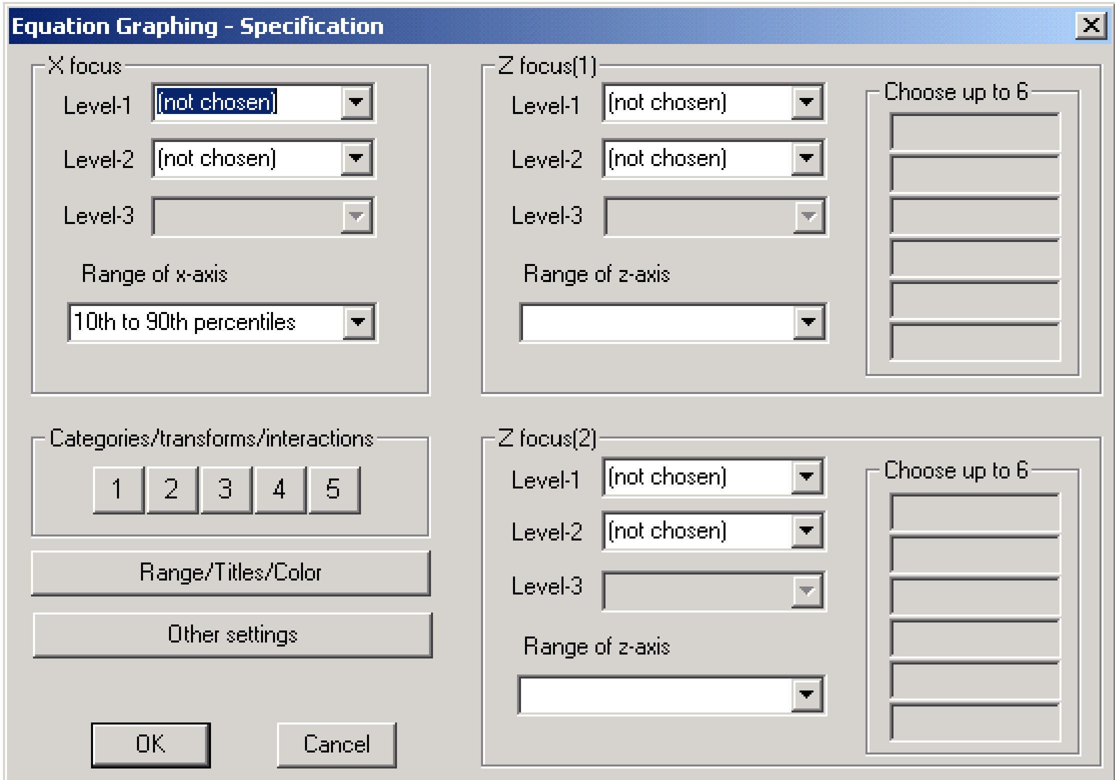


Figure 18.17 Equation Graphing – Specification dialog box

We now proceed to select the predictor variables and specify their ranges or values, and choose the graphing functions and the various attributes of the plot for the polynomial model represented in Figure 18.16, as described in Steps 5 to 14 below.

5. Select AGE16 in the **X focus Level 1** drop-down list box to graph pro-deviant attitude score as a function of age.
6. Select **Entire range** in the **Range of x-axis** drop-down list box to include the entire range of age on the x axis in the graph.

Table 18.2 Definitions and options in the Equation Graphing dialog box

	Key terms	Function	Option	Definition
1	X focus	Specify the variable to be displayed on x-axis	2 choices	1. Level-1 predictor 2. Level-2 predictor
2	Range of x-axis	Specify the maximum and minimum values of X to be displayed	5 choices	1. 10^{th} to 90^{th} percentiles 2. 5^{th} to 95^{th} percentiles 3. 25^{th} to 75^{th} percentiles 4. +/- 2 s.e.'s 5. Entire range
3	Categories/ transforms/ interactions	Define the reference category for categorical variables with more than two levels, and specify the relationship between the transformed\interaction and the original variables	5 choices	1. define categorical variable (for variable with more than two levels) 2. interaction 3. power of x/z 4. square root 5. natural log
4	Range/ Titles/ Color	Specify the maximum and minimum values of X and Y to be displayed (defaults are values computed). Enter legend titles for X and Y. Enter graph title. Select screen color	2 choices	1. Black and white 2. Color
5	Other Settings	Specify graphing function Predictors not in graph Use fixed effects from These are only available for HGLM models, and Laplace is only available if Laplace was asked for in HGLM2/HGLM3 Bernoulli runs	2 choices 2 choices 3 choices	1. rough – original points 2. smooth – smoothed data 1. constant at grand mean (default) 2. constant at zero. 1. unit-specific PQL estimates 2. population-average estimates 3. unit-specific Laplace estimates
6	Z focus(1 or 2)	Specify the first or second classification variable for X	3 choices	1. Level-1 predictor 2. Level-2 predictor
7	Range of z-axis	Specify the specific values of Z focus to be included.	4 choices for continuous variables 2 choices for categorical variables	1. 25^{th} and 75^{th} percentiles 2. 25th/50th/75th percentiles 3. Averaged lower/upper quartiles 4. Choose up to 6 values (enter the six values into the textboxes) 1. Use the two actual values 2. Choose one or two values

7. Click 1 in the **Categories/transforms/interactions s**ection and select **power of x/z** for Polynomial relationships. An **Equation Graphing - power** dialog box will open (see Figure 18.18).

Figure 18.18 Equation Graphing – power dialog box

8. The textbox to the left of the equal sign is for the entry of the transformed variable. Select AGE16S in the drop-down list box (see Figure 18.19). The textbox to the right is for the entry of the original variable. AGE16 will appear in the drop-down list box as it is the only level-1 variable left. Enter 2 in the textbox for the **power** to be raised. Click **OK**.

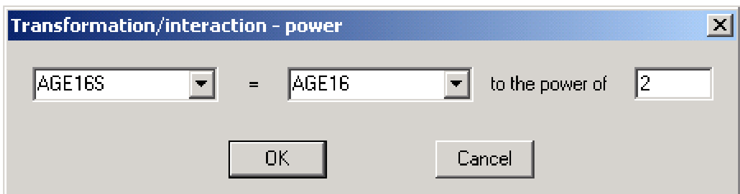


Figure 18.19 Equation for the transformed variable AGE16S

9. Click **Range/Legend/Color** to specify the ranges for x- and y-axis (the default values are those computed from the data), to enter legend and graph titles, and to select screen color (see Figure 18.20). Enter **Pro-deviant attitude score as a function of age, gender and ethnicity** in the textbox for **Graph title**. Click **OK**.

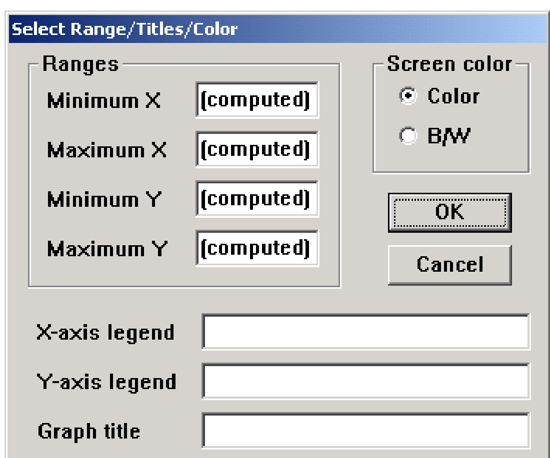


Figure 18.20 Select Range/Legend/Color dialog box

10. Click the **Other settings** button and click the selection button for **Smooth** in **For continuous x** section to display a set of smooth curves.
11. Select FEMALE in the **Z focus(1)** drop-down list box to graph pro-deviant attitude score as a function of age for male and female youths. **Use the two actual values** will appear in the

textbox for the **Range of z-axis** as FEMALE is an indicator variable. We will use this default option.

12. Select MINORITY in the **Z focus(2)** drop-down list box to graph pro-deviant attitude score as a function of age for minority and non-minority male and female youths. **Use the two actual values** will appear in the textbox for the **Range of z-axis** as MINORITY again is an indicator variable. We will use this default option. See Figure 18.21 for the specifications for this growth curve analysis example.
13. Click **OK**. A colored version of the plot (not displayed here) showing the relationship between pro-deviant attitude score and age for different gender-by-ethnicity groups will appear (see Figure 18.22). The curves indicate that there is a nonmonotonic and nonlinear relationship between pro-deviant attitude scores and age for minority and non-minority male youths over the five year period. Such a relationship, however, does not exist for minority and non-minority female youths.
14. For information on the editing, printing, saving, and modification options, see Steps 11 to 13 in section 18.1.1.

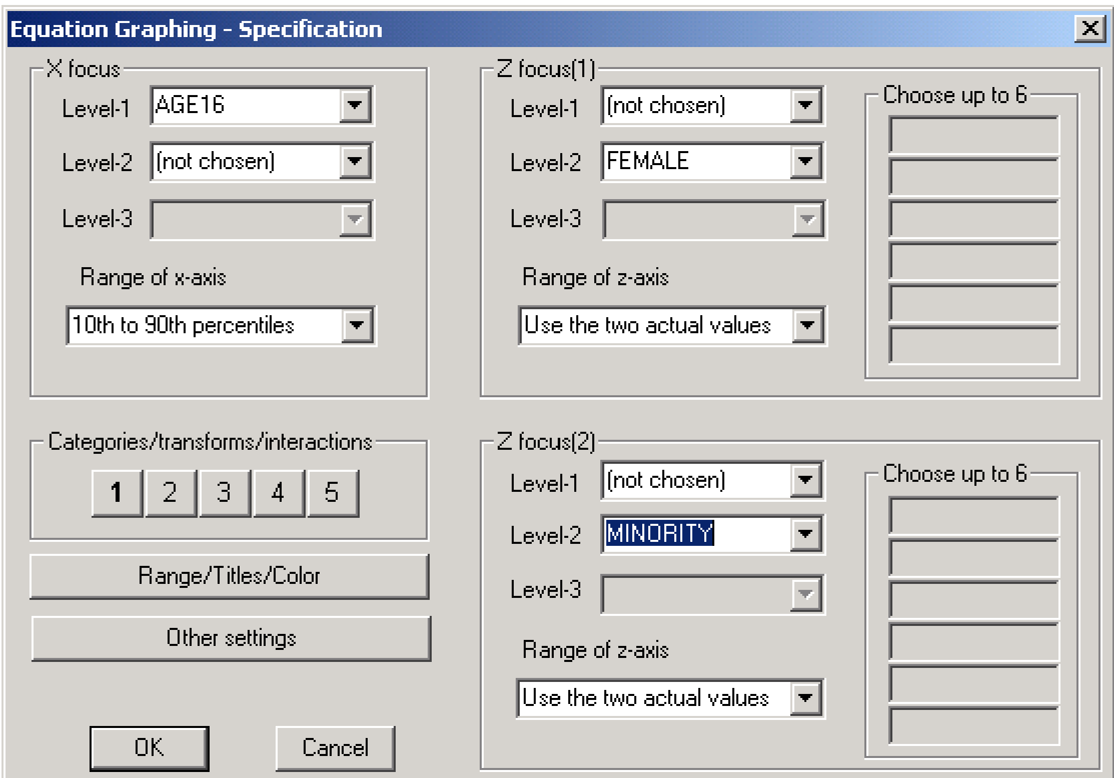


Figure 18.21 Specifications for the Growth Curve Analysis Example

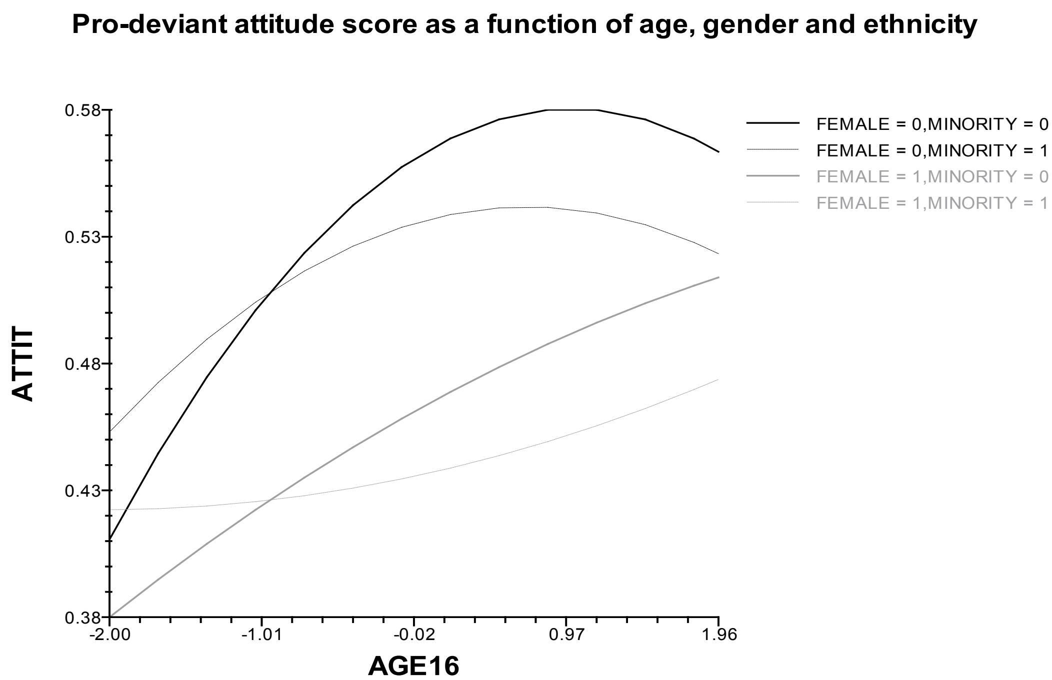


Figure 18.22 Plot showing the relationship between pro-deviant attitude score and age for different gender-by-ethnicity groups

18.2.2 Level-1 equation modeling

WHLM will also let us examine plots for individual level-2 units by just using the level-1 equation instead of the entire model. For this example, we will be using the vocabulary data, VOCAB.MDM described in section 18.1.2, and have run the following model:

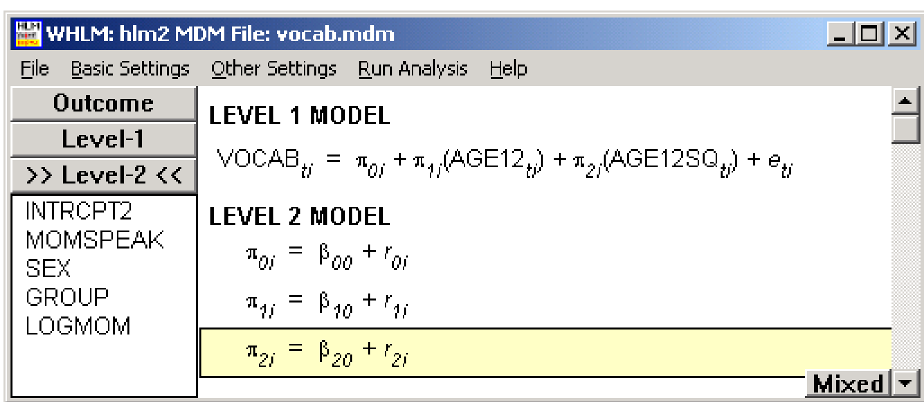


Figure 18.23 Model specification window for the vocabulary data

To perform the level-1 equation graphing

1. After the model is run, select **Graph Equations...Level-1 equation graphing** from the **File** menu, which will give us the following dialog box.

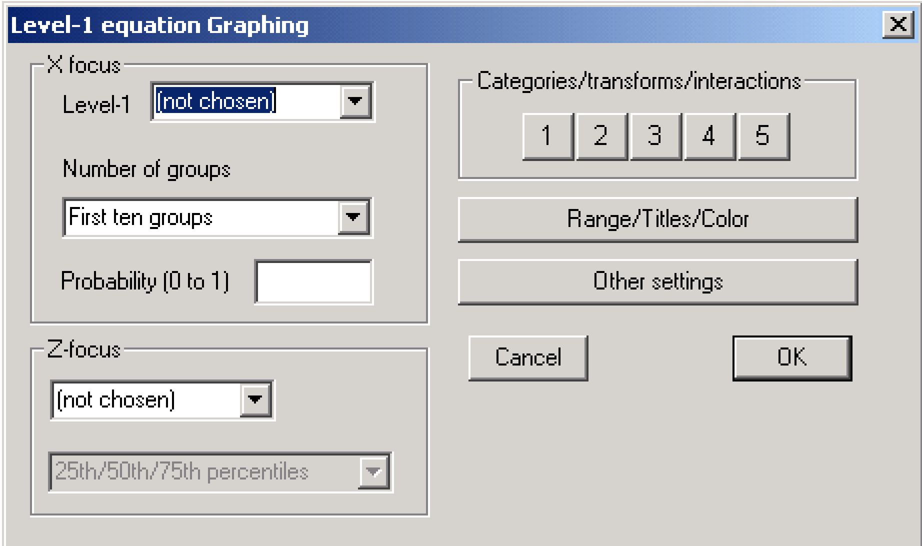


Figure 18.24 Level-1 equation Graphing dialog box

For the definition of **Number of groups**, see step 5 in section 18.1. Table 18.2 describes and explains the other options in the dialog box.

2. Select an **X focus** variable. In our example, we want the age of the child in months minus 12 to be the **X focus**. Choose AGE12 from the **X focus** drop-down list box.
3. Select number of groups. We will include all the children. Choose **All groups (n=22)** in the **Number of groups** drop-down list box.
4. Specify the relationship between the transformed and the original variable. The transformed variable is AGE12S and the original variable is AGE12. Click 1 in the **Categories/ transforms/interactions** section and select **power of x/z** for Polynomial relationships. A **Equation Graphing - power** dialog box will open. Select AGE12S from the drop-down list box to the left of the equal sign. AGE12 will appear in the drop-down list box as it is the only level-1 variable left. Enter 2 in the textbox for the **power** to be raised. Click **OK**.
5. (Optional) click **Range/Legend/Color** to specify the ranges for x- and y-axis (the default values are those computed from the data), to enter legend and graph titles, and to select screen color.
6. Click the **Other settings** button and click the selection button for **Smooth** in **For continuous x** section to display a set of smooth curves. Click **OK**.
7. Click **OK** and we get the following figure that shows vocabulary size accelerates during the second year of life. Note that the individual trajectories, as expected, are "smoother" than in the comparable data-based graphs in Figure 18.14 in Section 18.1.3.

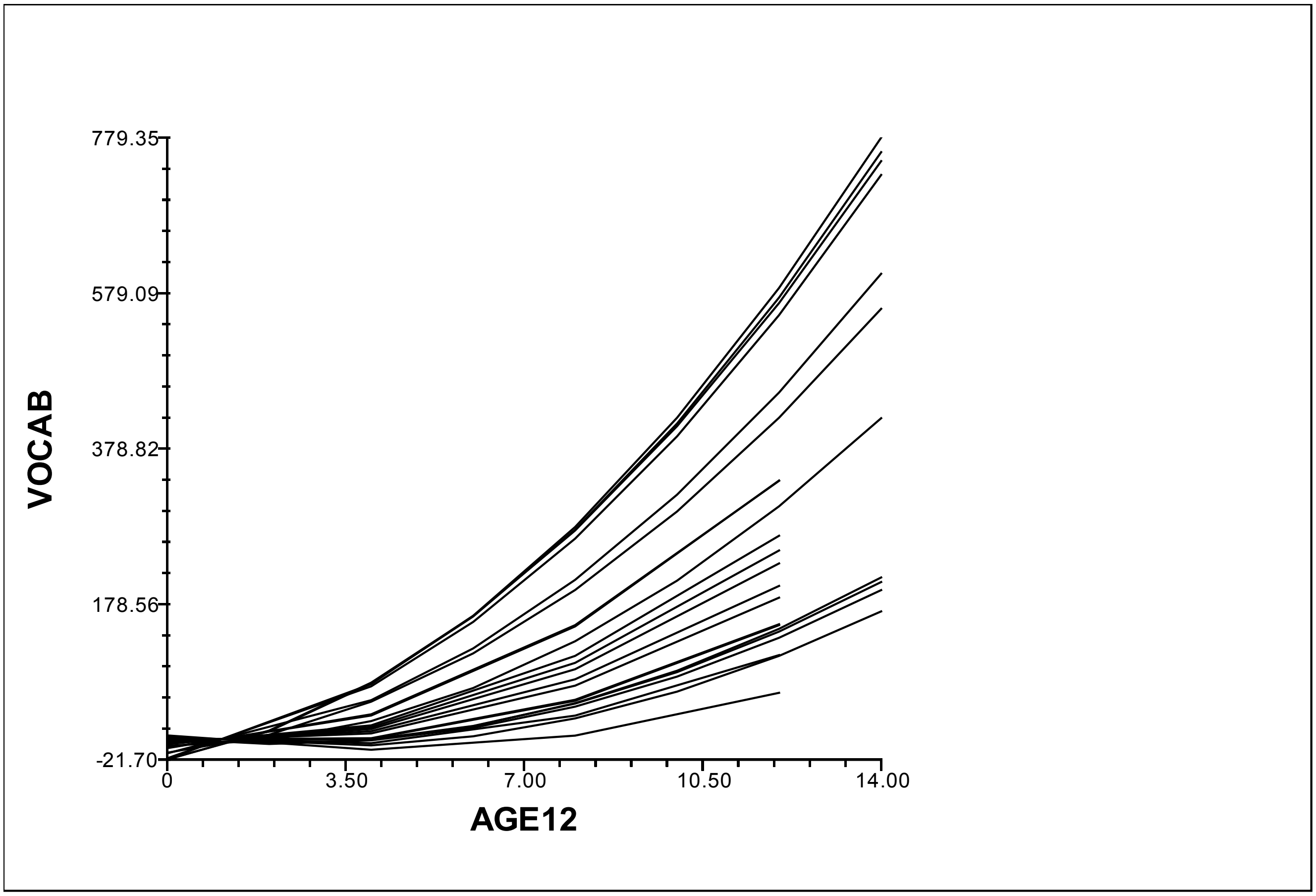


Figure 18.25 Predicted trajectories of vocabulary growth for individual children

To include a level-2 classification variable

8. Click **Graph Settings** on the menu bar to open the **Level-1 equation Graphing** dialog box.
9. Choose MALE from the **Z-focus** drop-down list box as the level-2 classification variable.
10. Click **OK**. The following figure will appear. A colored version of the graph (not shown here) indicates that girls on average have a greater acceleration rate in vocabulary growth over the course of the study.

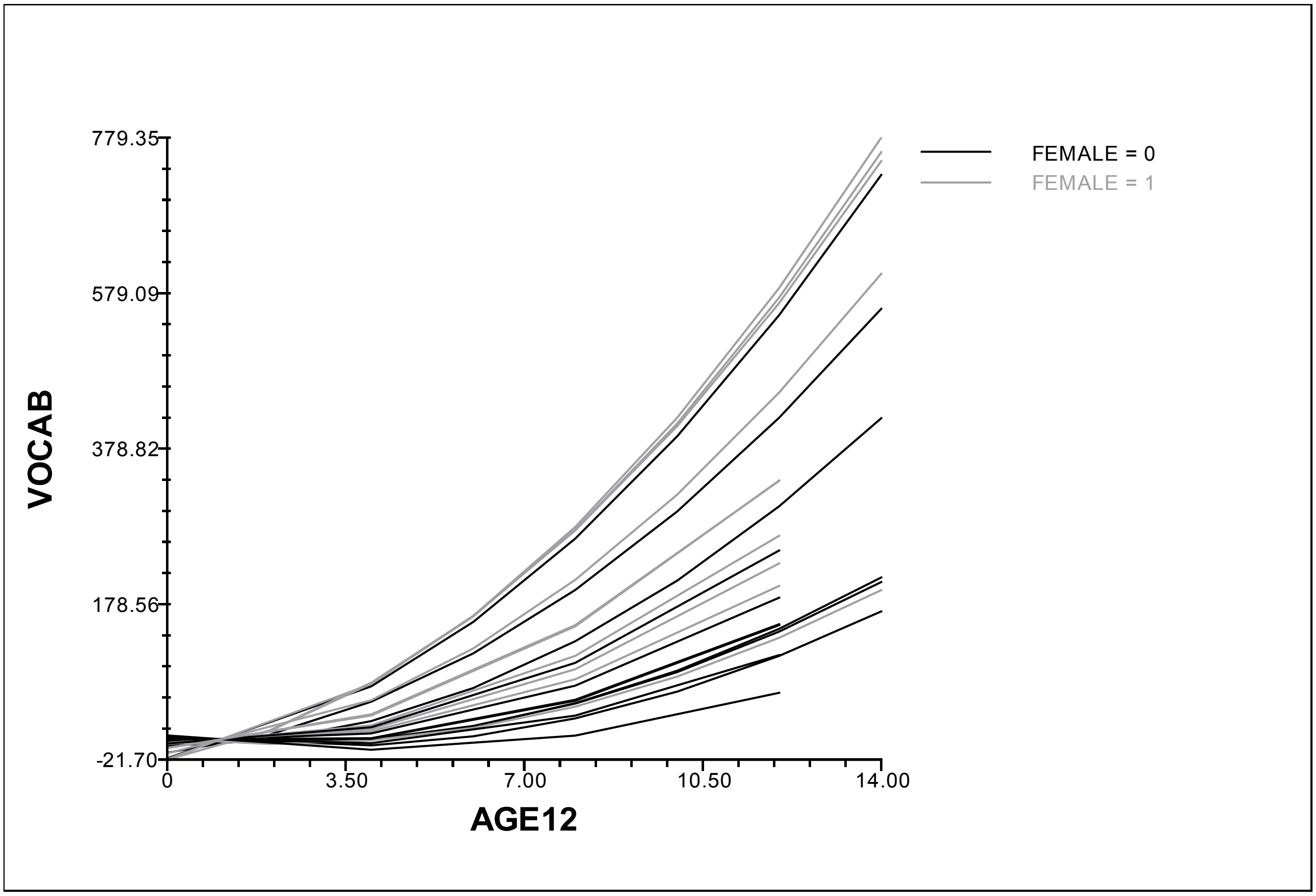


Figure 18.26 Predicted trajectories of vocabulary growth of individual children grouped by gender

18.2.3 Level-1 residual box-and-whisker plots

In addition to plotting predicted values for individual level-2 units using level-1 equations, users can also examine the distributions of the level-1 errors or residuals (see Equation 3.63 on p. 50 in *Hierarchical Linear Models*). The plots allow users to graphically examine the assumptions about the level-1 residuals and to identify cases for which the model provides a particularly poor fit. We continue to use VOCAB.MDM to illustrate this graphing procedure.

To prepare level-1 residual box-and-whisker plots

1. After the model is run, select **Graph Equations...Level-1 box whisker** from the **File** menu, which will give us the following dialog box.

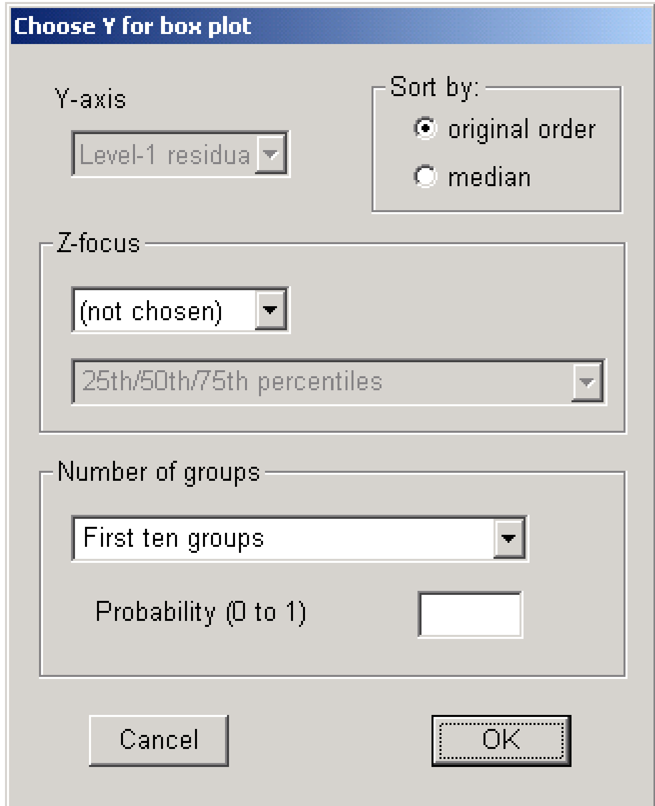


Figure 18.27 Choose Y for box plot dialog box

For definitions of the options in the dialog box, see Section 18.1.1. Note that the variable for Y-axis, level-1 residual has been pre-selected.

2. Select **All groups (n=22)** in the **Number of groups** to include all the 22 children in the display.
3. Click the selection button for **median** in the **Sort by** section to arrange the plots by median order.
4. Click **OK**. The following graph will appear.

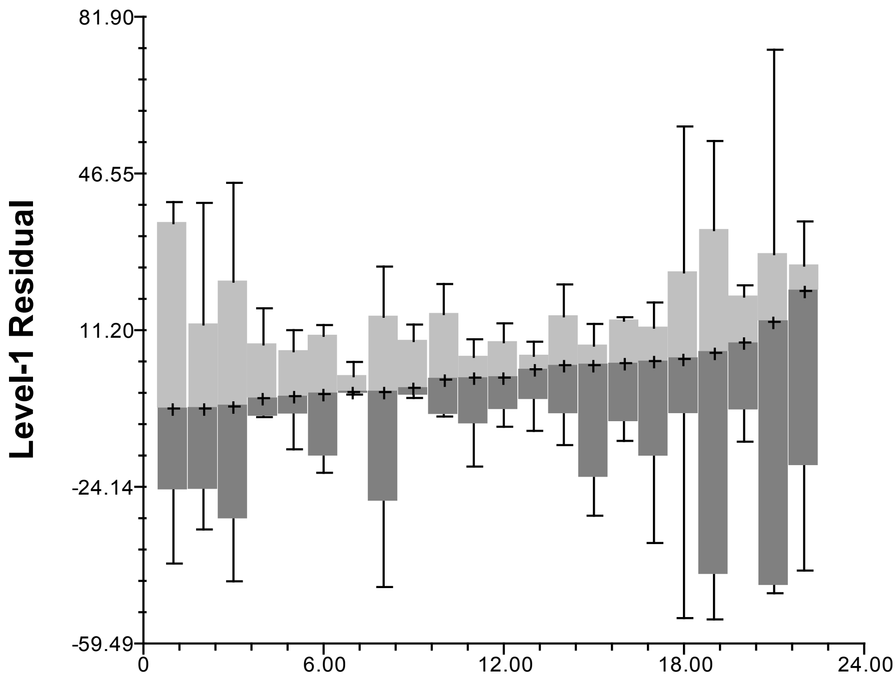


Figure 18.28 Box-and-whisker plots of the level-1 residuals

The box-and-whisker plots provide side-by-side graphical summaries of the level-1 residuals for each level-2 units. The plots suggests that the underlying model assumptions may not be tenable. First, quite a number of the distributions are highly asymmetric, such as the last one from the left. Thus, the normality assumption may not hold. There seems to be heterogeneity of variance as well, judging from the wide disparities in the box lengths. The nonconstant residual spread may suggest an omission of important effects from the model. However, there are no extreme values or outliers in any of the 22 plots. Note that this graphical analysis of level-1 residuals differs from the one performed in Section 2.5.4.1.2 in that it does not pool the residuals across level-2 units. In addition, WHLM has a statistical test for evaluating the adequacy of the homogeneity of level-1 variance assumption (see Section 2.8.8.2). See *Hierarchical Linear Models* pp. 263-267 for a discussion of the examination of assumptions about level-1 random effects.

5. (Optional) Users can look at the EB estimates for any child by clicking on the corresponding box-and-whisker plot. See Step 9 in Section 18.1.1.

6. (Optional) Users can choose to include a level-2 classification variable when examining the level-1 residuals. See Step 14 in Section 18.1.1.

18.2.4 Level-1 residual vs predicted value

Users can graphically assess the assumptions of constant error variance and linearity and probe for outlying cases by examining a scatter plot of level-1 residuals and predicted values. Using the same data and model of the previous two sections, we now plot the level-1 residual against its predicted value.

To prepare a level-1 residuals by predicted values scatter plot

- After the model is run, select **Graph Equations...Level-1 residual vs predicted value** from the **File** menu, which will give us the following dialog box.

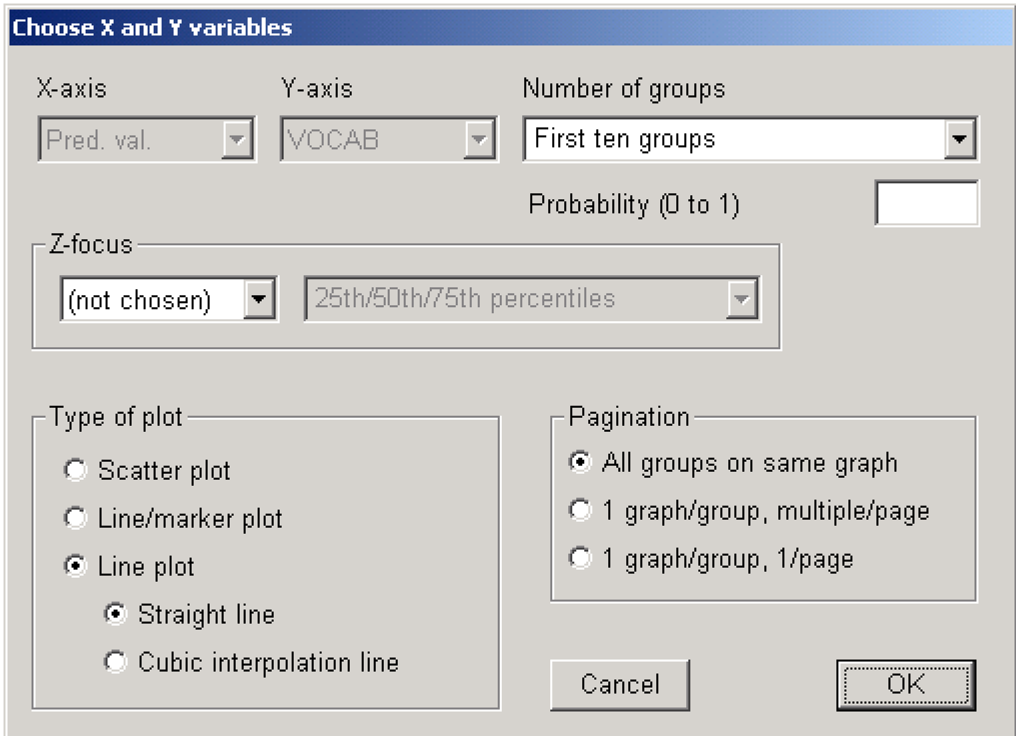


Figure 18.29 Choose X and Y variables

For definitions of the various options in the dialog box, see Section 18.1.2. Note that the **X-axis** variable, **Pred. val.** and **Y-axis** variable, **Level-1 residuals** have been pre-selected.

- Select **All groups (n=22)** in the **Number of groups** to include all the 22 children in the display.
- Click the selection button for **Scatter plot** in the **Type of plot** section to request a scatter plot of the predicted values by level-1 residuals.
- Select **All groups on same graph** in the **Pagination** section to display all the residuals pooled across the level-2 units. To examine the residuals for individual children, choose either of the other pagination options.
- Click **OK**.

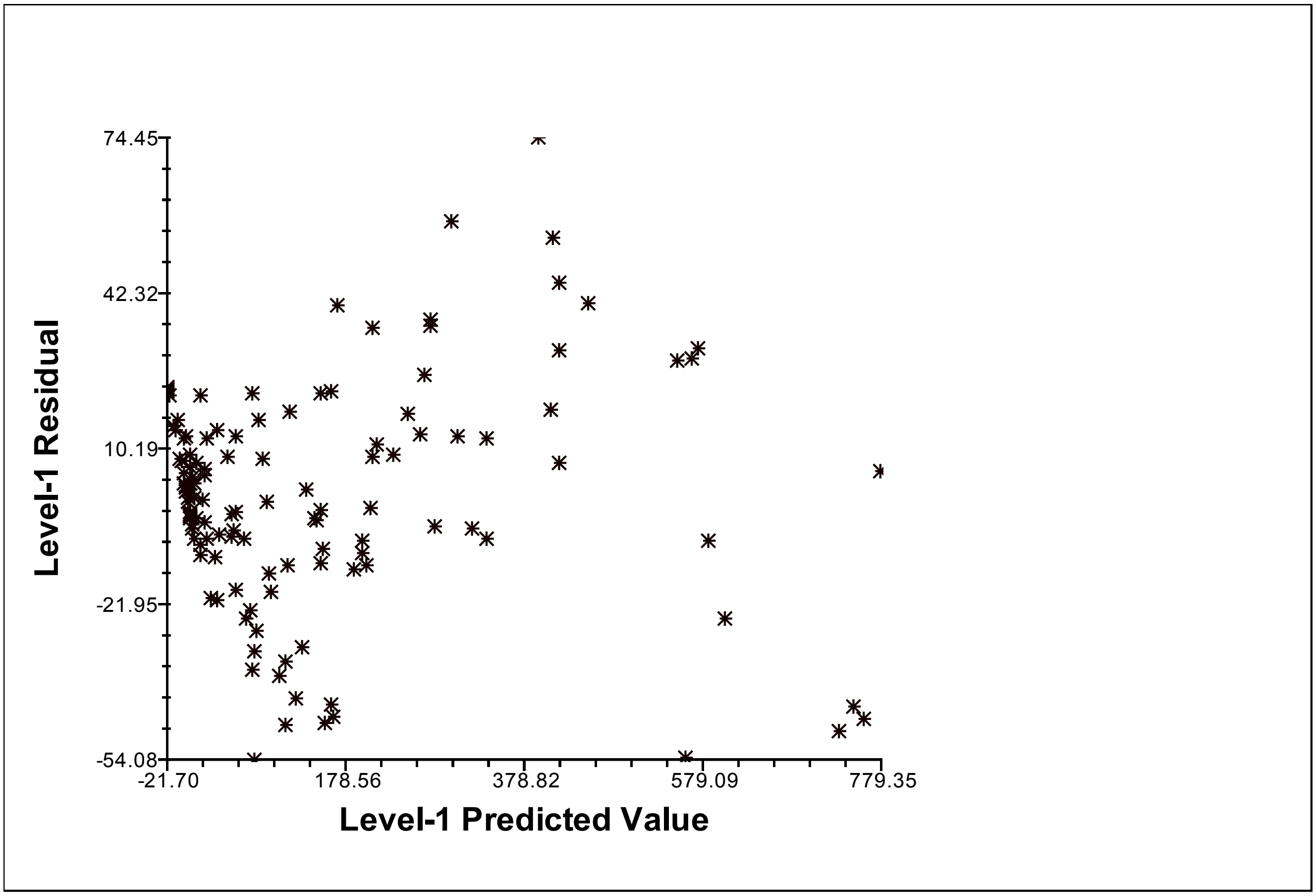


Figure 18.30 Plot of level-1 residuals by predicted values

The plot suggests that there is a tendency for the residual scatter to get narrower at the smallest predicted values and to get wider around the interval between 150 and 170. The residuals seem to follow a slightly curvilinear trend as well. They may suggest that there is a specification error in the model.

- (Optional) Users can choose to include a level-2 classification variable when examining the level-1 residuals. See Step 14 in Section 18.1.1.

18.2.5 Level-1 EB/OLS coefficient confidence intervals

We can also look at graphs of the estimated empirical Bayes (EB) or OLS estimates of randomly varying level-1 coefficient (see Section 1.3 and *Hierarchical Linear Models*, p. 47 and p. 49 for their computational formulae). This enables us to compare level-2 units with respect to these two types of estimates.

To prepare level-2 EB estimates of randomly varying level-1 coefficient confidence intervals

1. After the model is run, select **Graph Equations...Level-2 EB/OLS coefficient confidence intervals** from the **File** menu, which will give us the following dialog box:

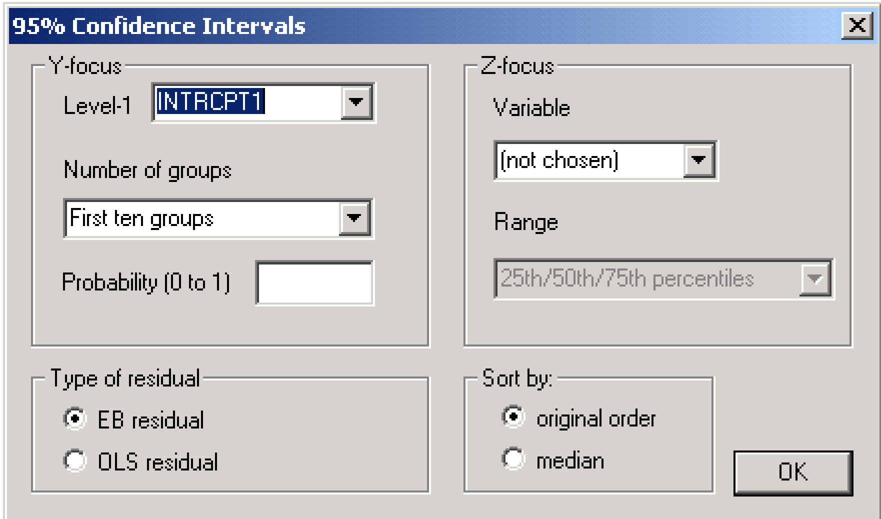


Figure 18.31 95% Confidence Intervals dialog box

For definitions about the various options regarding Y- and Z-focus and sorting, see Section 18.1.1.

2. Choose the randomly varying level-1 coefficient of interest. We will look at the coefficient for the quadratic term or acceleration rate of vocabulary growth in this example. Choose AGE12S from the **Y-focus** drop-down list box.
3. Select **All groups (n=22)** in the **Number of groups** to include all the 22 children in the display.
4. Click the **EB residual** button in the **Type of residual** section to select the empirical Bayes estimates.
5. Click **OK**. The following graph will appear.

The graph suggests that there is significant variation in the rate of acceleration in vocabulary growth in children during the second year of life. For instance, the confidence intervals of the EB estimates of the AGE12S coefficients for the last four children from the left did not overlap with those of the first eleven children.

6. Users can look at the actual empirical Bayes estimates and their 95% confidence intervals of individual level-2 units by clicking on the confidence interval plots.
7. (Optional) Users can choose to include a level-2 classification variable when examining the confidence interval plots. See Step 14 in Section 18.1.1.

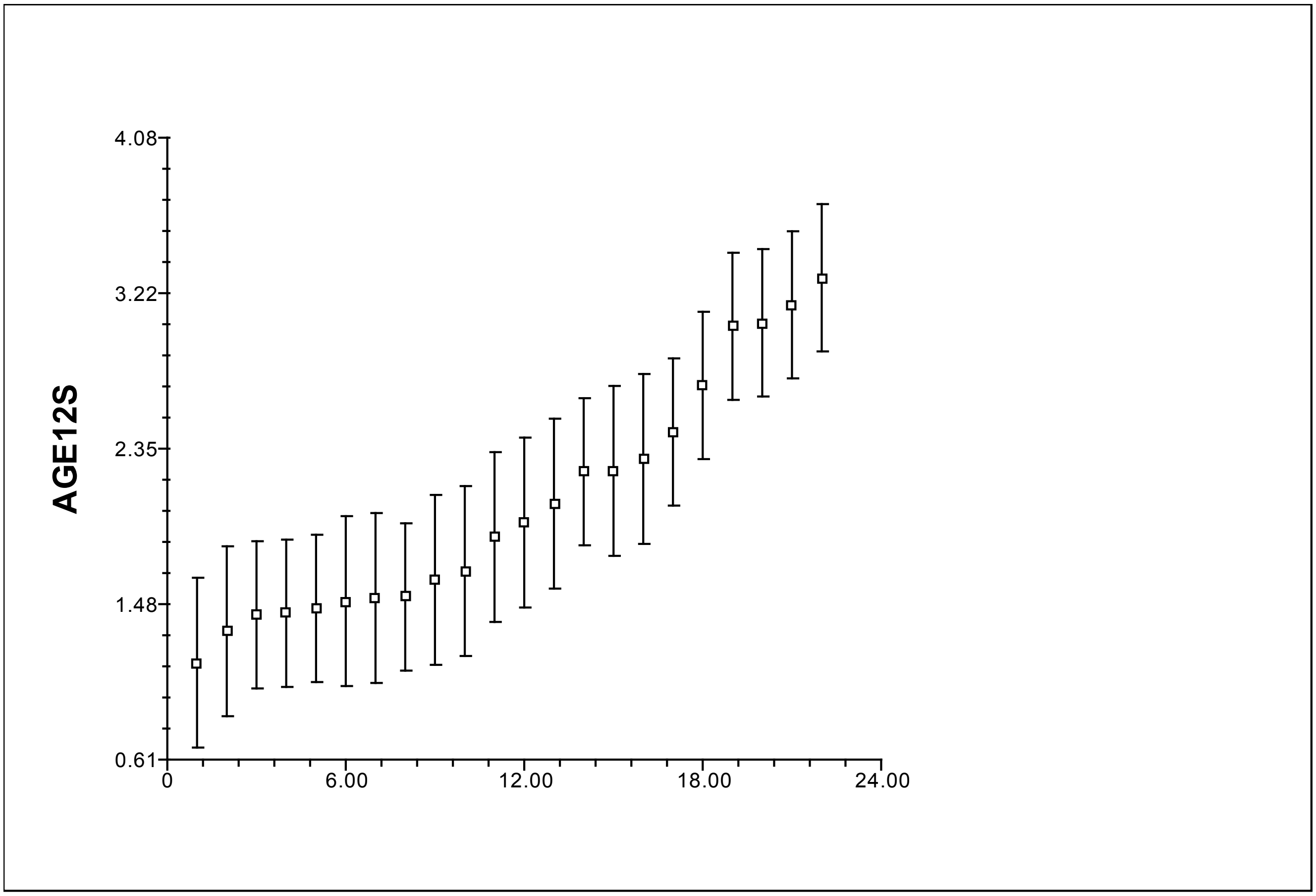


Figure 18.32 Confidence intervals of empirical Bayes estimates of the AGE12S coefficients

18.2.6 Graphing categorical predictors

Model graphs can be displayed in which predictor variables are categorical. Suppose, for example, that the variable ETHNICITY has three possible values: BLACK, HISPANIC, and WHITE and that this variable is represented by indicator variables for BLACK and HISPANIC, with WHITE serving as the reference category. To represent ethnicity as a predictor, click the first box under **Categories/transformations/interactions**. Next, click on **define categorical variable**. Then four boxes will appear:

1. Under the box **Choose first category from foci** click on the variable that is the first of the indicator variables in the model. In our example, this will be BLACK.
2. Under the box **Possible choices** click on any other indicators in the model that represent the categorical variable of interest; in our case, there is only one : HISPANIC.
3. Under **Name of reference** category, type in the name of the reference group; in our case, this will be WHITE.
4. Under **Category Name**, type the name of the categorical variable; in our case, this will be ETHNICITY.

Now click **OK** to continue.

18.3 Three-level applications

Graphing with 3-level data is very similar to the 2-level graphing. The only two differences are that users can (a) group the plots at either level 2 or 3, and (b) choose exclusively a level-2 or level-3 classifying or conditioning variable. To illustrate these two differences, we will use the EG.MDM as describe in Section 4.1. We will prepare line plots of the mathematics test score, MATH, to detect trends over the course of the six-year study, grouped by the level-3 units, schools, and classified by a level-3 variable, the socioeconomic composition of schools. The same logic applies to the sets of three-level model-based graphing procedure.

To prepare line plots with level-3 grouping

1. From the HLM window, open the **File** menu.
2. Choose **Create a new model using an existing MDM file** to open an **Open MDM File** dialog box. Open EG.MDM.
3. Open the **File** menu, choose **Graph Data...line plots, scatter plots** to open an **Choose X and Y variables** dialog box (see Figure 18.33).
4. Select YEAR from the **X-axis** drop-down list box.
5. Select MATH from the **Y-axis** drop-down list box.
6. Select number of groups. In this example, we want to include a random sample of 20 percent of the schools in the display. Select **Random sample of spec'd prob** from the **Number of groups** drop-down list box. Enter 0.2 into the textbox for **Probability (0 to 1)** to indicate that 10 percent or a proportion of .1 of the schools will be used.
7. Select type of plot and method of interpolation (see Step 7 in Section 18.1.3 for explanations). For our example, we want a line plot with no markers that is graphed with the linear interpolation method. Click the selection button for **Straight line**.
8. Select type of grouping at level 2 or level 3. In this example, we want to have the trajectories for individual schools (**Group at level 3**). Click **Group at level-3** selection button (default) in the **Grouping** section.
9. Select type of pagination. We want separate plots for individual schools and choose **1 graph/group, multiple page** option accordingly.

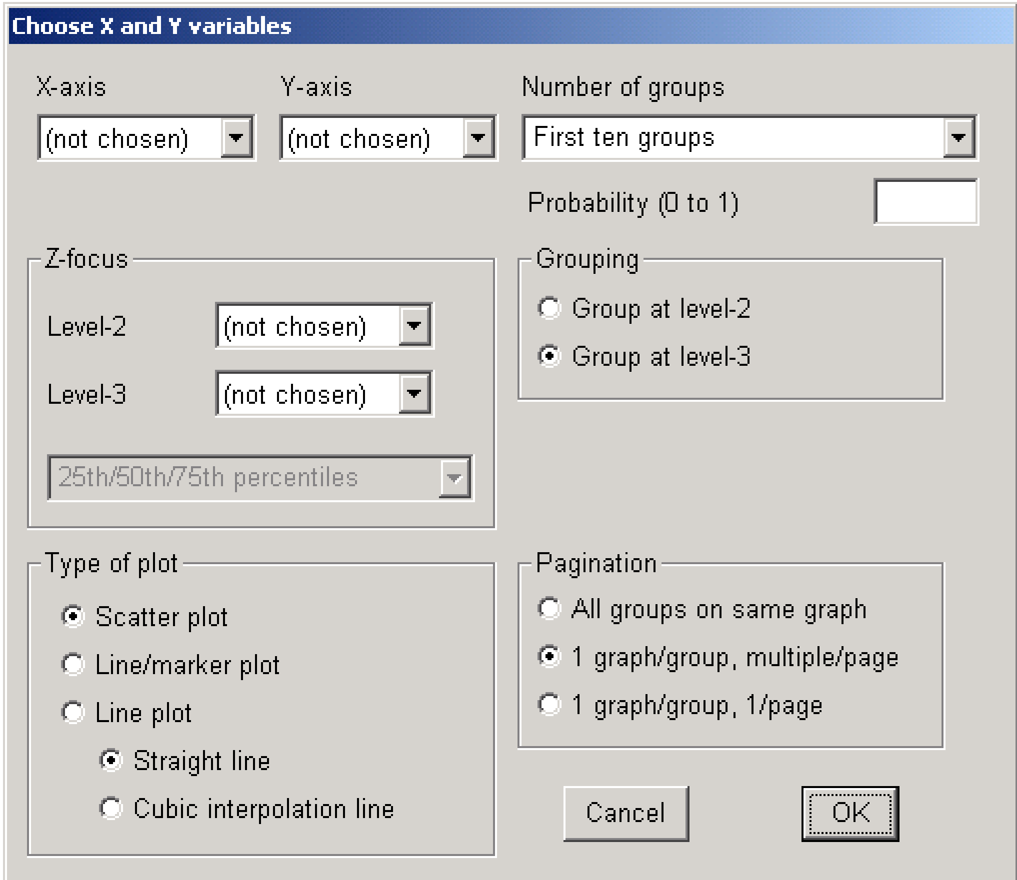


Figure 18.33 Choose X and Y variables dialog box

10. Click **OK**. The following graph will appear.

The eight line plots indicate the collection of students' growth trajectories of mathematics achievement within individual schools. The schools varied in their number of students. There was a generally positive average rate of growth across all schools.

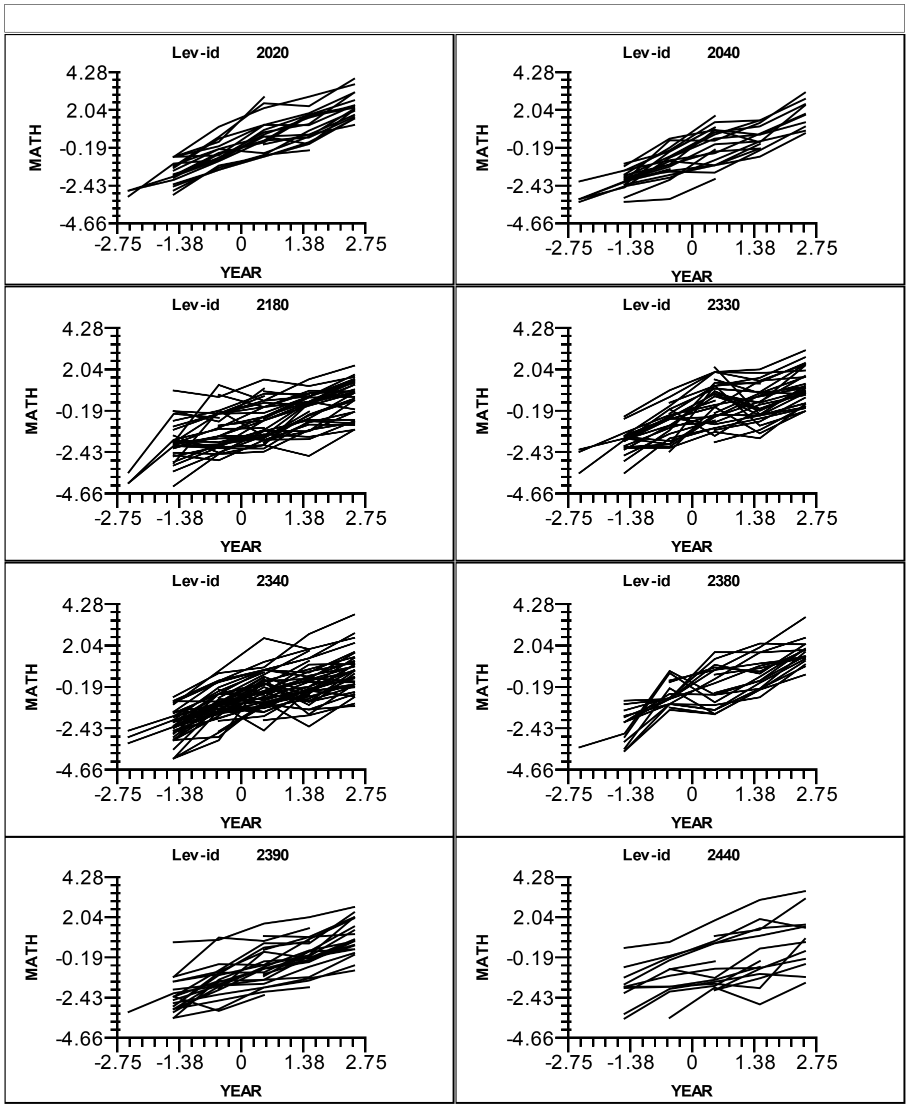


Figure 18.34 Line plots of MATH against YEAR for eight schools

To include a level-3 classification variable

11. Now we want to look at the trajectories as classified by the socioeconomic composition of the study body of a school. On the menu of the graph dialog box, click **Graph Settings**. Choose the level-3 variable LOWINC, the percent of students from low income families, as a **Z-focus** variable. As LOWINC is a non-dichotomous variable we have an additional choice that was not needed for our earlier dichotomous z-foci. In this case, we choose **Above/Below 50th percentile** from the combo box immediately below where we chose the LOWINC as the grouping variable.
12. Click **OK**. The following graph will appear.

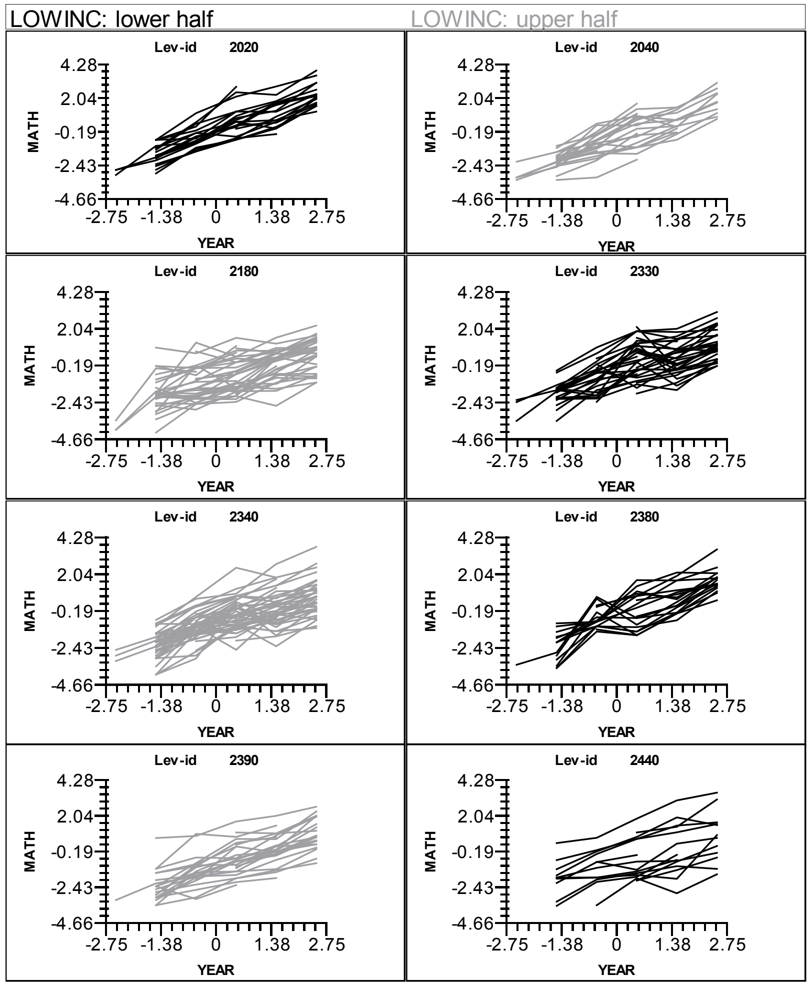


Figure 18.35 Line plots of MATHACH against YEAR for eight schools by LOWINC

This shows us that schools with a greater percent of students from low income families (upper high) tend to have lower mathematics achievement than do schools with less percent of poor students. Compared to their peers in School 2020, for instance, students in School 2330 generally have lower achievement across the six years.

19 The Fixed Intercepts Random Coefficient (FIRC) Model

19.1 Conceptual background for FIRC

Fixed effects models can serve as a useful tool for causal inference. In multilevel settings, for instance, they can help to remove unobservable confounding attributable to clusters in the analysis such as persons, schools, neighborhoods, states, or countries when the treatment assignment occurs *within* clusters. Many articles have considered the choice between a fixed effects model versus a random effects model using criteria such as the assumptions required for the estimators to be consistent (Raudenbush, 2009). HLM offers the option to estimate either or both classes of models. In addition, it allows researchers to combine features of both types of models with a fixed intercepts and a random treatment coefficient (FIRC) to improve causal inferences in multisite intervention (e.g., Bloom, Raudenbush, Weiss, & Porter, 2017), meta-analysis (e.g., Weiss et al., 2017), as well as panel studies (e.g., Raudenbush, 2009) by investigating heterogeneity of treatment effects across sites.

19.1.1 The fixed intercepts and a random treatment coefficient (FIRC) model

To illustrate the FIRC models, we first consider a) a random intercept and a fixed treatment coefficient HLM2, and b) a fixed intercepts and a fixed treatment coefficient HLM2 model and some of their key assumptions using a multisite trial example in which the level-1 individuals within each level-2 study site are randomly assigned to a treatment or control group.

19.1.1.1 A random intercept and a fixed treatment coefficient HLM2 model

The random intercept and fixed treatment coefficient HLM2 model, as described in Section 1.1, consists of two sub-models at level 1 and level 2. The level-1 model is represented as

$$Y_{ij} = \beta_{0j} + \beta_{1j} Treatment_{ij} + r_{ij} \tag{19.1}$$

where $Treatment_{ij}$ is an indicator variable for the treatment group membership of individual i in study site j with 1 = treatment, 0 = control; and r_{ij} is a random term and we assume $r_{ij} \sim N(0, \sigma^2)$.

The Level-2 model is represented as

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned} \tag{19.2}$$

where

γ_{00} is the overall mean of the control group;
u_{0j} is the level-2 random intercept effect and we assume $u_{0j} \sim N(0, \tau_{00})$. ; and
γ_{10} is the overall treatment effect.

The mixed or combined model is

$$Y_{ij} = \gamma_{00} + \gamma_{10} Treatment_{ij} + u_{0j} + r_{ij} \tag{19.3}$$

For the estimate of the treatment effect to be consistent, the fraction of persons assigned to the treatment, $\overline{Treatment}_j$ is assumed to be uncorrelated with site-specific random effects u_{0j}. Varying $\overline{Treatment}_j$ correlated with unobserved site characteristics can produce inconsistent parameter estimates (Bloom et al., 2017).

19.1.1.2 A fixed intercepts and a fixed treatment coefficient HLM2 model

The level-1 model remains the same, and the level-2 model is represented as

$$\begin{aligned} \beta_{0j} &= u_{0j} \\ \beta_{1j} &= \gamma_{10} \end{aligned} \tag{19.4}$$

where

> u_{0j} is a fixed constant. The key innovation in computation is recognizing that we can equivalently regard u_{0j} as a random effect for which we have no prior information, that is we assume $u_{0j} \sim N(0, \tau_{00} \to \infty)$, i.e., $\tau_{00}^{-1} = 0$; and
>
> γ_{10} is the treatment effects, assumed in this case to be constant across all level-2 units.

This model is also known as the fixed-effects model in the econometric literature. It estimates fixed site-specific intercepts (u_{0j}). This parameterization removes the between-site variability in any level-1 predictors. Thus the previously discussed assumptions for the fixed intercept and fixed coefficient model required for consistent estimation of the treatment effect can be relaxed.(7) The mixed or combined model is

$$Y_{ij} = \gamma_{1j} Treatment_{ij} + u_{0j} + r_{ij} \tag{19.5}$$

By re-formulating the fixed effects model as a random effects model with infinite variance, $u_{0j} \sim N(0, \tau_{00} \to \infty)$, i.e., $\tau_{00}^{-1} = 0$, HLM2 enables users to estimate the fixed effects model very simply without having to include dummy variables and without centering of variables (see Section 19.2.14 for details).

19.1.1.3 A fixed intercepts and a random treatment coefficient (FIRC) HLM2 model

Again the level-1 model remains the same. The level-2 model becomes

$$\begin{aligned} \beta_{0j} &= u_{0j} \\ \beta_{1j} &= \gamma_{01} + u_{1j} \end{aligned} \tag{19.6}$$

---

7 An alternative parameterization of the same model is to group-mean center $Treatment_{ij}$ (see Section 2.5.2).

There is an additional site-specific random effect, u_{1j}, associated with the treatment in FIRC in this model. We assume that $u_{0j} \sim N(0, \tau_{00} \to \infty)$, i.e., $\tau_{00}^{-1} = 0$ and $u_{1j} \sim N(0, \tau_{11})$. The combined model is

$$Y_{ij} = \gamma_{1j} Treatment_{ij} + u_{0j} + u_{1j} Treatment_{ij} + r_{ij} \qquad (19.7)$$

where u_{0j} is a site-specific fixed effect and u_{1j} is a random effect for treatment. The random site-specific program assignment effects allow researchers to investigate cross-site variation and to produce site-specific empirical Bayes estimates of impact. In addition, when the aim is to generalize to a population of clusters, the FIRC model also reduces the bias associated with the conventional site fixed effects model with fixed treatment effect (see Bloom, Raudenbush, Weiss, and Porter, 2017, winner of the Best Article Award in the *Journal of Research on Educational Effectiveness*).

19.1.2 Parameter estimation

Three kinds of parameters are available in HLM2 and HLM3 FIRC. Empirical Bayes estimates of random effects, maximum-likelihood estimates of the level-2 or level-3 coefficients in HLM2 and HLM3 respectively, and maximum likelihood estimates of variance-covariance parameters are available. When estimating the variance of the treatment effects, HLM2 and HLM3 allows the treatment/control heteroscedasticity. For example, the outcome variance for its treatment group members differ from that for its control group within sites (Raudenbush & Bloom, 2015) (see Section 2.8.5 for details).

19.1.3 Hypothesis testing

HLM2 and HLM3 output a chi-square test of homogeneity for each random effect. Also, users can use the likelihood ratio test to compare the fit of the various models.

19.2 Working with FIRC

HLM2 and HLM3 FIRC analyses can be executed in Windows, interactive, and batch modes. To illustrate the operation of the program, we use the data from the Tennessee's Student/Teacher Achievement Ratio study project (STAR) (Shin & Raudenbush, 2011), which was a statewide effort to study the effect of reduced class size on student academic performance in Tennessee. Windows model execution is illustrated.

We will first look at the effects of reduced of class size using a two-level model with students nested within school.

19.2.1 HLM2 Statistical package input

We will use SPSS file input in our example. There are two data files for the HLM2 FIRC analysis, one at the student level, and one at the school site level.

Level-1 file. The level-1 file, **STAR1.SAV** has math and reading proficiency data as well as the type of class of 5,786 students participated in STAR. The variables are:

- MATH a math test in IRT scale score metric
- CLASSTYP an indicator of class type (1 = small with 13-17 students, 0 = other)

Level-2 file. The level-2 file, **STAR3.sav** has data collected from 79 schools that the students attended. The variable is:

- SIZE school size

Using HLM2, the MDM file **STARHM2.MDM** is created.

19.2.1.1 Executing analyses based on the MDM File

We first illustrate a) a random intercept and a fixed treatment coefficient model, then b) a fixed intercepts and a fixed treatment coefficient model, finally followed by c) a fixed intercepts and a random treatment coefficient model. We summarize the results at the end of this section.

19.2.1.2 A random intercept and a fixed treatment coefficient model

The command file, **STARHM2A.HLM**, contains the model specification input responses for the fixed intercepts and a fixed treatment coefficient model. Figure 19.1 displays the model specified.

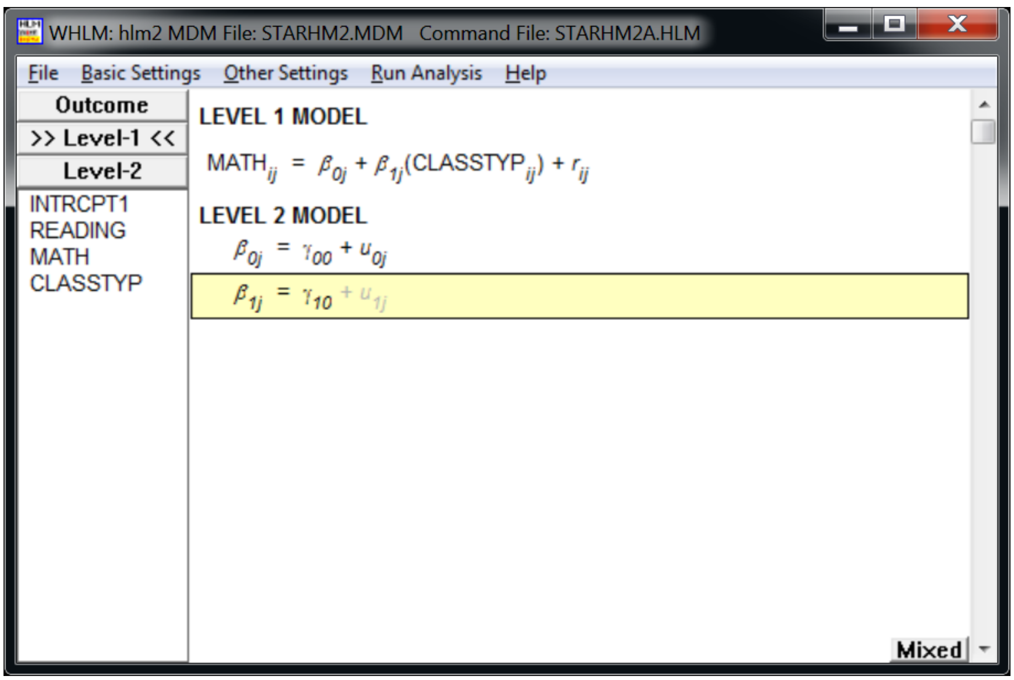


Figure 19.1 The random intercept and a fixed treatment coefficient model specification for the STARHM2 example

The results of the analysis are given below.

Problem Title: Random Intercept and A Fixed Treatment Coefficient Model

The data source for this run = STARHM2.MDM
The command file for this run =STARHM2A.HLM
Output file name = STARHM2A.HTML
The maximum number of level-1 units = 5786
The maximum number of level-2 units = 79
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified
Step 2 model

Level-1 Model

$MATH_{ij} = \beta_{0j} + \beta_{1j}*(CLASSTYP_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = \gamma_{00} + u_{0j}$
$\beta_{1j} = \gamma_{10}$

Mixed Model

$MATH_{ij} = \gamma_{00}$
$+ \gamma_{10}*CLASSTYP_{ij} + u_{0j}+ r_{ij}$

Final Results - Iteration 3
Iterations stopped due to small change in likelihood function

σ^2 = 1804.30900

Standard error of σ^2 = 33.77701

τ

INTRCPT1,β_0 458.63366

Standard error of τ
INTRCPT1,β_0 77.25120

Approximate confidence intervals of tau variances
INTRCPT1 : (327.960,641.373)

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.945

The value of the log-likelihood function at iteration 3 = -3.001734E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	483.000600	2.506586	192.693	78	<0.001
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	9.087321	1.232934	7.370	5706	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	483.000600	2.615873	184.642	78	<0.001
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	9.087321	2.340424	3.883	5706	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	21.41573	458.63366	78	1540.50368	<0.001
level-1, *r*	42.47716	1804.30900			

Statistics for the current model
Deviance = 60034.676525
Number of estimated parameters = 4

19.2.1.3 A fixed intercepts and a fixed treatment coefficient model

The command file, **STARHM2B.HLM**, contains the model specification input responses for the fixed intercepts and a fixed treatment coefficient model. A conventional way to specify such model is to include $J-1$ school site dummy variables into the model. HLM2 offers a simple step to set up the model.

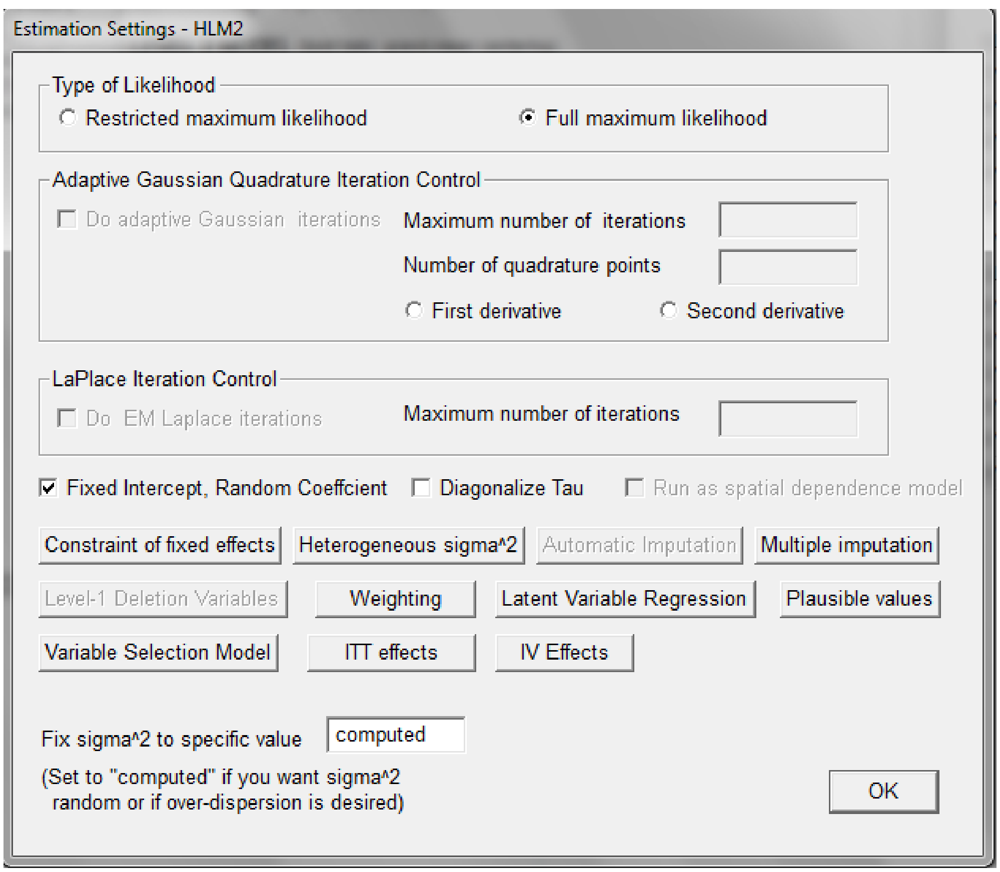


Figure 19.2 Estimation settings – HLM2 dialog box

After clicking **OK**, the fixed intercepts and fixed treatment coefficient will be displayed, as shown in Figure 19.3. Note that the level 2 model for β_{0j} is a no-intercept model.

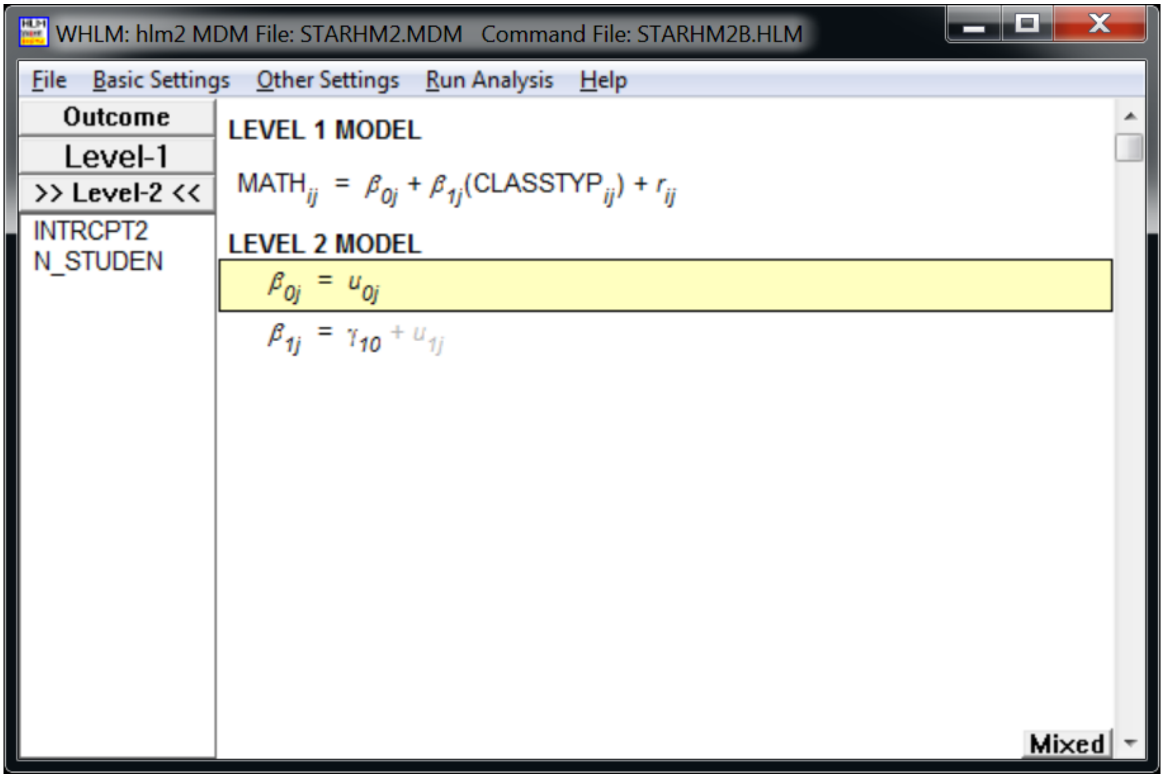


Figure 19.3 The fixed intercepts and a fixed coefficient model specification for the STARHM2B example

Here is the output:

Specifications for this HLM2 run

Problem Title: Random Intercept and Fixed Coefficient Model

The data source for this run = STARHM2.MDM
The command file for this run = STARHM2B.HLM
Output file name = hlm2.html
The maximum number of level-1 units = 5786
The maximum number of level-2 units = 79
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Step 2 model

Level-1 Model

$MATH_{ij} = \beta_{0j} + \beta_{1j}*(CLASS\_TY_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = u_{0j}$
$\beta_{1j} = \gamma_{10}$

Mixed Model

$MATH_{ij} =$
$+ \gamma_{10}{*}CLASSTYP_{ij} + u_{0j} + r_{ij}$

Final Results - Iteration 6

Iterations stopped due to small change in likelihood function

σ^2 = 1804.31836

The value of the log-likelihood function at iteration 6 = -2.973353E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	9.127153	1.233739	7.398	5706	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	9.127153	2.343758	3.894	5706	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
level-1, *r*	42.47727	1804.31836			

Statistics for the current model
Deviance = 59467.054530
Number of estimated parameters = 2

19.2.1.4 A fixed intercepts and a random treatment coefficient model

The command file, **STARHM2C.HLM**, contains the model specification input responses for the fixed intercepts and a random treatment coefficient model. Figure 19.4 displays the model specified.

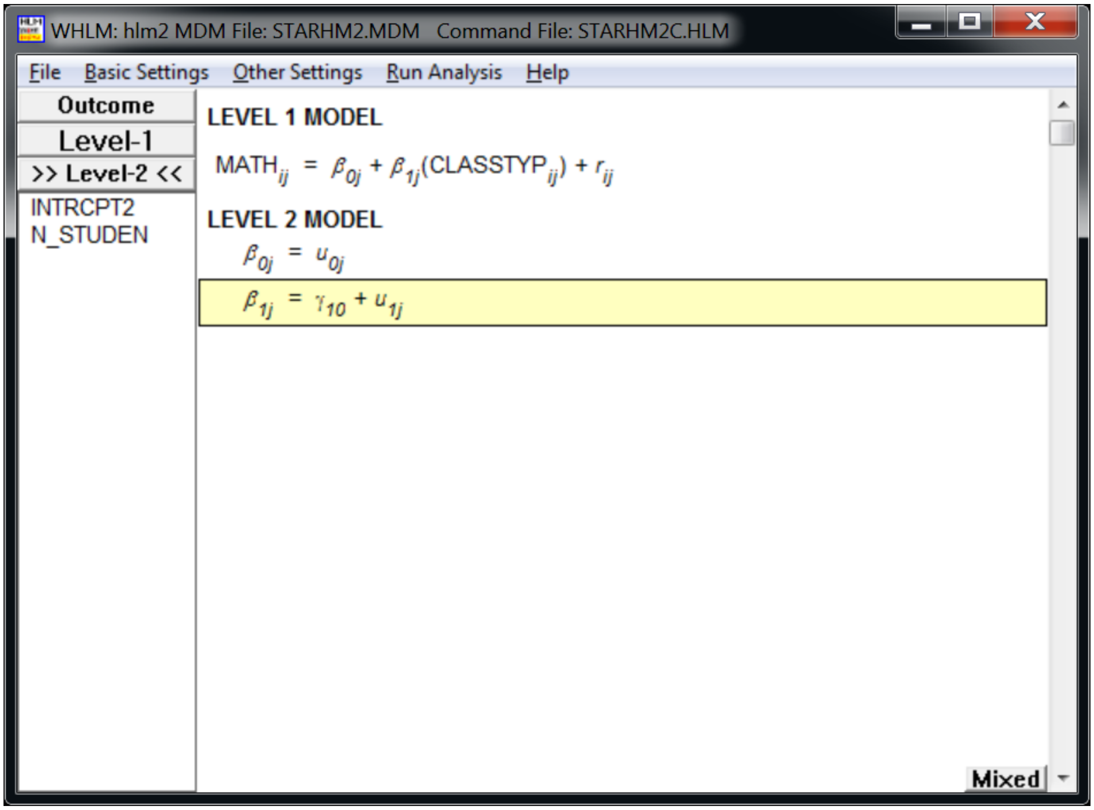


Figure 19.4 The fixed intercepts and a random treatment coefficient model specification for the STARHM2C example

Here is the output:

Problem Title: Fixed Intercepts and A Random Treatment Coefficient Model

The data source for this run = STARHM2.MDM
The command file for this run = STARHM2C.HLM
Output file name =STARHM2C.HTML
The maximum number of level-1 units = 5786
The maximum number of level-2 units = 79
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Step 2 model

Level-1 Model

$MATH_{ij} = \beta_{0j} + \beta_{1j}*(CLASSTYP_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

Mixed Model

$MATH_{ij} =$

$+ \gamma_{10}*CLASSTYP_{ij} + u_{0j} + u_{1j}*CLASSTYP_{ij} + r_{ij}$

Final Results - Iteration 11

Iterations stopped due to small change in likelihood function

σ^2 = 1742.81131

τ

CLASSTYP,β_1 301.75903

τ (as correlations)

CLASSTYP,β_1 1.000

Random level-1 coefficient	Reliability estimate
CLASSTYP,β_1	0.705

The value of the log-likelihood function at iteration 11 = -2.968392E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	8.538461	2.328306	3.667	78	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For CLASSTYP slope, β_1					
INTRCPT2, γ_{10}	8.538461	2.327321	3.669	78	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
CLASSTYP slope, u_1	17.37121	301.75903	78	280.89921	<0.001
level-1, *r*	41.74699	1742.81131			

Statistics for the current model
Deviance = 59367.836357
Number of estimated parameters = 3

19.2.1.5 Summary of the results

Table 19.1 summarizes the results for the three models.

Model / Estimate	A Random Intercept and A Fixed Treatment Coefficient	Fixed Intercepts and a Fixed Treatment Coefficient	Fixed Intercepts and a Random Treatment Coefficient
Average Treatment Effect	9.087321	9.127153	8.538461
Model-Based Standard Error of Average Treatment Effect	2.340424	1.233739	2.328306
Robust Standard Error of Average Treatment Effect[8]	2.340424	2.343758	2.327321
Variance of Treatment Effect	NA	NA	301.75903 (χ^2 = 280.89921, *df* = 78, $p < 0.001$)

Table 19.1 Summary of the treatment estimates from the three models

The results of the FIRC models with a minimum of assumptions suggest that there is evidence of cross-site variation in the program impact.

19.2.2 An example of HLM3 FIRC

The above illustrative example ignores a level of nesting--the classroom level, thus a three-level model with students nested within classrooms within schools will better accommodate the data structure.

19.2.2.1 HLM3 Statistical package input

There are three data files for the HLM3 FIRC analysis: the student-, classroom-, and school-level files.

Level-1 file. The level-1 file, **STAR1.SAV** has math and reading proficiency data of 5,786 students participated in STAR. The variables are:

- MATH a math test in IRT scale score metric
- READING a reading test in an IRT scale score metric

Level-2 file. The level-2 file, **STAR2.SAV** has class treatment type data collected from 325 classrooms that the students attended. The variable is:

- CLASSTYP an indicator of class type (1 = small with 13-17 students, 0 = other)

Level-3 file. The level-3 file, **STAR3.SAV** has data collected from 79 schools that the students attended. The variable is:

- SIZE school size

8 See Section 1.9 for a discussion of the robust standard errors.

Note that CLASSTYP is now a classroom-level variable. Using HLM3, the MDM file **STARHM3.MDM** is created.

19.2.2.2 An annotated example of HLM3 FIRC

The command file, **STARHM3A.HLM**, contains the model specification input responses for the fixed intercepts and a fixed treatment coefficient model. Figure 19.5 displays the model specified.

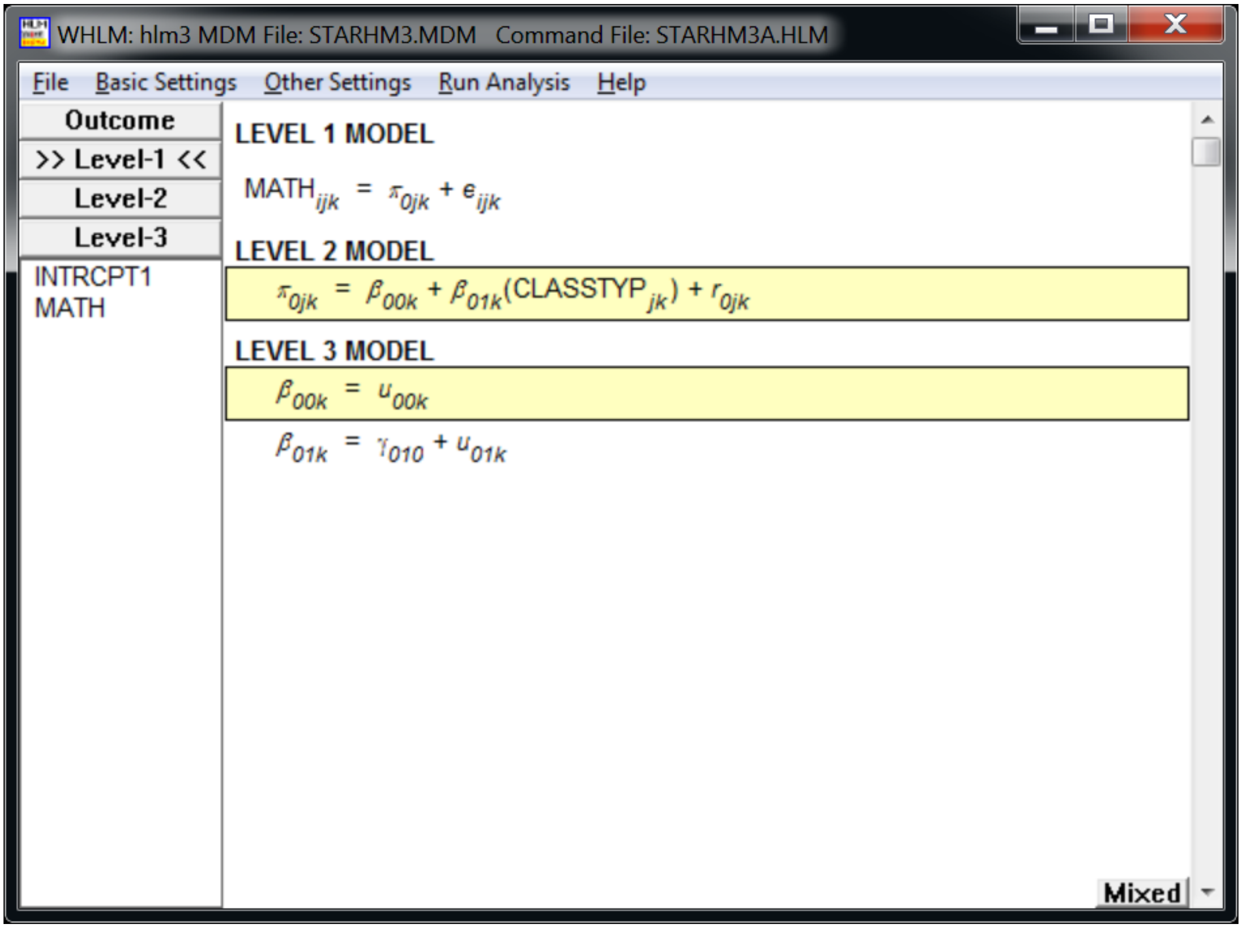


Figure 19.5 The fixed intercepts and a random treatment coefficient model specification for the STARHM3 example

Here is the output:

Specifications for this HLM3 run

Problem Title: Fixed Intercepts and a Random Treatment Coefficient Model

The data source for this run = STARHM3.MDM
The command file for this run = STARHM3A.HLM
Output file name = STARHM3A.HTML

The maximum number of level-1 units = 5786
The maximum number of level-2 units = 325
The maximum number of level-3 units = 79
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

The outcome variable is MATH

Summary of the model specified

Level-1 Model
$MATH_{ijk} = \pi_{0jk} + e_{ijk}$

Level-2 Model
$\pi_{0jk} = \beta_{00k} + \beta_{01k}*(CLASSTYP_{jk}) + r_{0jk}$

Level-3 Model
$\beta_{00k} = u_{00k}$
$\beta_{01k} = \gamma_{010} + u_{01k}$

Mixed Model
$MATH_{ijk} = \gamma*\text{CLASSTYP}_{jk} + r_{0jk} + u_{01k} + u_{01k}*\text{CLASSTYP}_{jk} + e_{ijk}$

For starting values, data from 5786 level-1 and 325 level-2 records were used

Final Results - Iteration 46

Iterations stopped due to small change in likelihood function

Standard errors for σ^2, τ_π, and τ_β are not computable.

σ^2 = 1597.25481

τ_π

INTRCPT1,π_0 262.45362

τ_β

INTRCPT1
CLASSTYP,β_{01}
68.18371
τ_β (as correlations)
INTRCPT1/CLASSTYP,β_{01} 1.000

Random level-2 coefficient	Reliability estimate
INTRCPT1/CLASSTYP,β_{01}	0.149

The value of the log-likelihood function at iteration 46 = -2.955771E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For CLASSTYP, β_{01}					
INTRCPT3, γ_{010}	8.744220	2.406371	3.634	78	<0.001

Final estimation of fixed effects (with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, π_0					
For CLASSTYP, β_{01}					
INTRCPT3, γ_{010}	8.744220	2.386153	3.665	78	<0.001

Final estimation of level-1 and level-2 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1,r_0	16.20042	262.45362	167	684.59376	<0.001
level-1, *e*	39.96567	1597.25481			

Final estimation of level-3 variance components

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	*p*-value
INTRCPT1/CLASSTYP,u_{01}	8.25734	68.18371	78	87.77158	0.210

Note that the between-school variance of the treatment effect is now 68.18, as compared to 301.76 when the classroom level was ignored.

Statistics for the current model
Deviance = 59115.428958
Number of estimated parameters = 5

20 Multivariate HLM2 from Incomplete Data based on Automated Multiple Imputation

20.1 Conceptual Background Regarding Automated Multiple Imputation

Missing data are a ubiquitous problem in most social sciences research. In multilevel studies, explanatory as well as outcome variables may be subject to missingness at any of the levels. It is extremely important to use an multilevel imputation model when the analysis model is a multilevel model. If one uses a single-level imputation procedure and then subjects the multiply imputed data to multilevel analysis, one can anticipate significant bias.

HLM offers a completely automated procedure to handle ignorable missing data in two-level models (Y. Shin & S. W. Raudenbush, 2013). The user first specifies a two-level HLM model. This model is flexible in that it can involve multiple outcome variables, including a mixture of level-1 and level-2 outcomes. The program then i) searches the variables that have missing values; ii) estimates a multivariate imputation model; iii) generates multiple imputed data sets, iv) analyzes each of these according the user's specified model; and averages the results using "Rubin's rules"(Rubin, 1987). Users can also specify a list of "auxiliary variables" at each of the two levels, ones that are not needed for the substantive analysis but that contain information about the missing data. These variables are used in the estimating the imputation model to improve the precision of the analysis and to improve robustness. The imputation model is estimated using full-information maximum likelihood.

20.1.1 Logic and assumptions of the approach

HLM implements the multiple imputation of missing data in multilevel studies developed by Shin and Raudenbush (2013). The key idea is to re-express a desired hierarchical model as the joint distribution of the outcomes and all variables that are subject to missingness, conditional on all of the covariates that are completely observed, and to estimate the joint model. We present an example of a general two-level random intercept model, as described in Shin (2013), to illustrate the logic and assumptions of the approach. Shin and Raudenbush (2013) provide the details of the estimation and of a more general framework that could handle "ignorable" missing data. The key assumption is that the data are missing at random ("MAR"- Little & Rubin, 2002).. MAR means that the missing pattern is conditionally independent of missing data given the observed data, and provides robust inferences when the observed data contain substantial information about the missing values.

20.1.1.1 A Random Intercept HLM2 model

A random intercept model with a level-1 and a level-2 predictor, as discussed in Section 1.1, can be expressed as consisting of two models.

Level-1 model. The level-1 model is

$$Y_{ti} = \pi_{0i} + \pi_{1i}a_{ti} + e_{ti} \tag{20.1}$$

where

π_{0i} $_{0j}$ is the intercept,

π_{1i} is the level-1 coefficient of predictor X_{ij}, and

e_{ti} is the level-1 random effect, and it is assumed that $e_{ti} \sim N(0,\sigma^2)$.

Level-2 model. The level-2 model is

$$\begin{aligned} \pi_{0i} &= \beta_{00} + \beta_{01}X_i + r_{0i} \\ \pi_{1i} &= \beta_{10} \end{aligned} \tag{20.2}$$

where

γ_{00}, γ_{01}, and γ_{10} are level-2 coefficients;
W_j is a level-2 predictor; and
u_{0j} is a level-2 random effect and it is assumed that $u_{0j} \sim N(0,\tau)$.

Mixed model. The mixed or combined model is

$$Y_{ti} = \beta_{00} + \beta_{01}X_i + \beta_{10}a_{ti} + r_{0i} + e_{ti} \tag{20.3}$$

When the outcome, Y_{ti}, and the level-1 and level-2 predictors, a_{ti} and X_i, are subject to missingness, there will be a total of seven possible missing data patterns for individual i in unit j, i.e., one, two, or all three values of (Y_{ti}, a_{ti}, X_i) could be missing. . The imputation model is

$$\begin{bmatrix} Y_{ti} \\ a_{ti} \\ X_i \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} b_{1i} \\ b_{2i} \\ b_{3i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1ti} \\ \varepsilon_{2ti} \\ 0 \end{bmatrix} \tag{20.4}$$

where

α_1, α_2, and α_3 are the means of (Y_{ti}, a_{ti}, X_i) ;

b_{1j}, b_{2j}, and b_{3j} are level-2 specific effects, and they are assumed to be multivariate normally distributed with a mean vector of zero and variance-covariance matrix T; and

ε_{1ti} and ε_{2ti} are level-1specific effects, and they are assumed to be bivariate normally distribution with a mean vector of zero and variance and covariance matrix $\sum$.

HLM estimates 20.4 in which all variables having missing values are regressed on all variables having complete data; it then uses the parameter estimates to generate M imputed data sets; it then analyzes each of these in turn and combines the results using the "Rubin's rules," as described in Section 11.2.1.

To improve the precision and robustness of the analysis, the user can specify a list of "auxiliary variables." These are variables that are not needed for the substantive analysis, but that contain information about the missing data. They are used in the estimation of the imputation model and therefore influence the imputed data sets. However, they are not included in the user's desired models. The auxiliary variables can include a mixture of level-1 and level-2 variables and may or may not themselves be subject to missingness.

The user can also output all of the multiply-imputed data sets for further analysis for analysis using HLM or another program. With the analysis of incomplete data routine, the program allows users to specify and analyze a general class of univariate and multivariate models in which there is an arbitrary number of outcome variables defined at either level 1 or level 2.

20.2 Working with Automated Multiple Imputation in HLM2

HLM2 analyses of incomplete data can be executed in Windows, interactive, and batch modes. To illustrate the operation of the program, we use the data from the Early Childhood Longitudinal Student Kindergarten Cohort (ECLS_K) of 1998 (Tourangeau, Nord, Lê, Sorongon, & Najarian, 2009). The study followed the children in fall kindergarten (K) of 1998, spring-K of 1999, fall-first grade (G1) of 1999, spring-G1 of 2000, spring-third grade (G3) of 2002, spring-fifth grade (G5) of 2004 and spring-eighth grade (G8) of 2007. Windows mode execution is illustrated.

20.2.1 An example using Analysis of Incomplete Multilevel Data in Windows mode

We first run a complete case analysis studying how income is related to the trajectories of mathematics and reading proficiency; then we illustrate a three-step procedure to perform an analysis of incomplete multilevel data (Shin, 2013; Shin & Raudenbush, 2007).

20.2.1.1 HLM2 Statistical package input

We will use SPSS file input in our example. There are two data files for the HLM2 analysis of incomplete multilevel data.

Level-1 file. The level-1 file, **ECLK981.SAV**, has 148,470 observations collected on 21,210 children between fall kindergarten and spring-eight grade.

There are three variables:

- MATH a math test in IRT scale score metric
- READING a reading test in an IRT scale score metric
- GRADE the grade level minus 3 of the child at each testing occasion. Therefore, it is 0 at Grade 3.

	CHILDID	MATH	READING	GRADE
1	1.00	39.54	36.58	-3.50
2	1.00	50.10	49.37	-3.00
3	1.00	.	.	-2.50
4	1.00	96.04	87.32	-2.00
5	1.00	.	.	.00
6	1.00	.	.	2.00
7	1.00	.	.	5.00
8	2.00	44.44	50.82	-3.50
9	2.00	58.93	83.50	-3.00
10	2.00	.	.	-2.50
11	2.00	76.33	155.63	-2.00
12	2.00	124.61	185.22	.00
13	2.00	152.24	192.70	2.00
14	2.00	168.77	206.63	5.00

Figure 20.1 First fourteen records from two children in the ECLK981.SAV dataset

Note that Child 1 in School 1 has missing data at the third and five to seventh occasions.

Level-2 file. The level-2 units are 21,210 children. The data are stored in the file **ECLK982.SAV**.

There are three variables of interest:

- INCOME income of the family
- PARSCR a parental occupational prestige score
- BLACK an indicator for ethnicity (1 = African American, 0 = other)

In creating the MDM file, we inform WHLM that there are missing data at level 1 and instruct the program to delete cases while performing analyses. The response file, **ECLK98.MDMT**, contains a log of the input responses for creating the MDM file, **ECLK98.MDM**. Below are the descriptive statistics:

```
LEVEL-1 DESCRIPTIVE STATISTICS

 VARIABLE NAME         N        MEAN         SD        MINIMUM       MAXIMUM
      MATH          94543      69.07       44.41        10.51        172.20
    READING         92208      87.39       53.05        21.01        208.90
      GRADE        148470      -0.57        2.88        -3.50          5.00


                      LEVEL-2 DESCRIPTIVE STATISTICS

 VARIABLE NAME         N        MEAN         SD        MINIMUM       MAXIMUM
     INCOME         14439       5.31        6.44         0.00         14.18
     PARSCR         20122      39.47       21.34         0.00         77.50
      BLACK         21210       0.15        0.36         0.00          1.00
```

At Level 1, there are 36% missing data on math proficiency, at level 2, there are 32% of income missing.

20.2.1.2 An annotated example of HLM2 analyses with complete cases

We first perform a complete case analysis studying income inequality in the average level and growth of mathematics and reading proficiency. To set up a model with both the mathematics and reading outcomes, after selecting MATH as the outcome, we add READING as an additional outcome, as shown in Fig. 20.2.

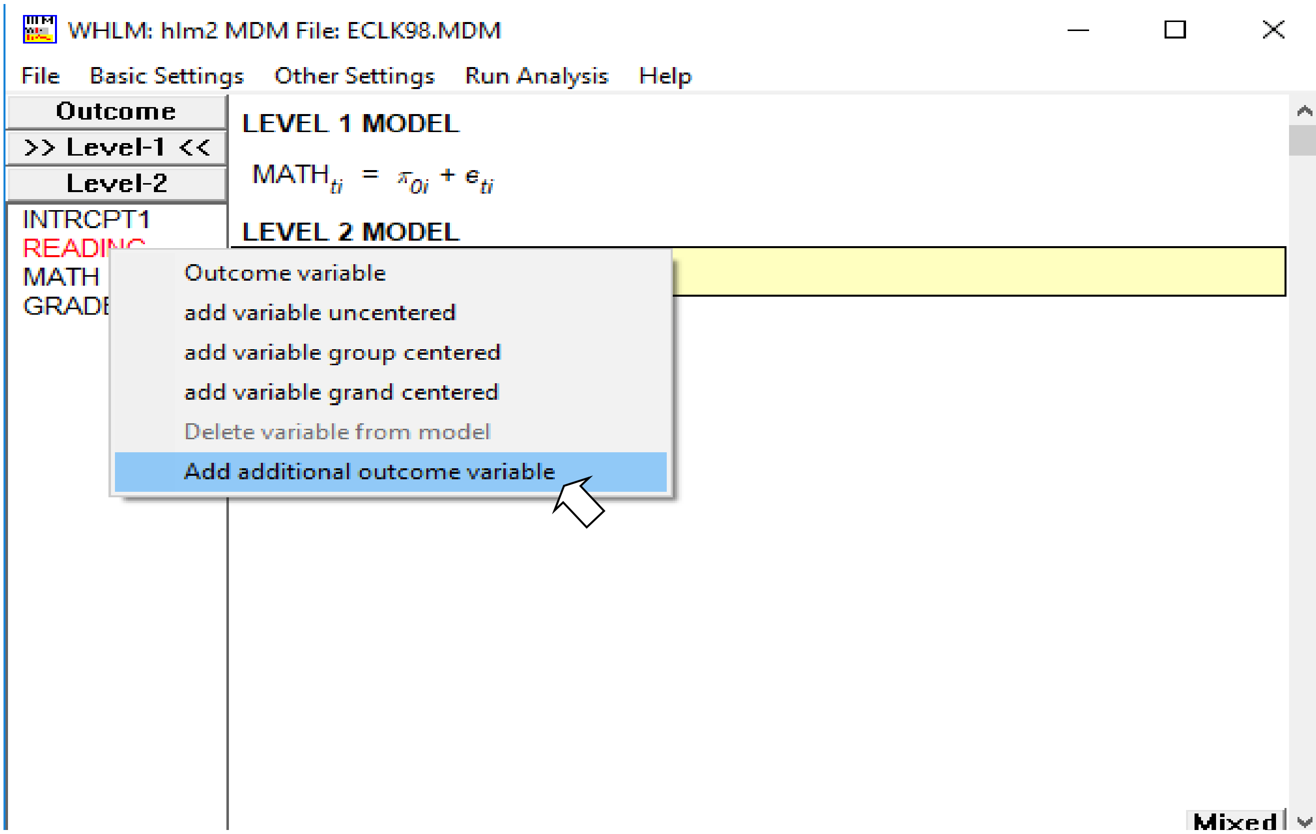


Figure 20.2 Model window for the bivariate outcome model

Then we add the level-2 variable INCOME to predict the trajectories of the mathematics and reading proficiency. The model is displayed in Figure 20.3.

LEVEL 1 MODEL

$$\text{READING}_{ti} = \pi_{0i} + \pi_{1i}(\text{GRADE}_{ti}) + e_{ti}$$

$$\text{MATH}_{ti} = \pi_{2i} + \pi_{3i}(\text{GRADE}_{ti}) + e_{ti}$$

LEVEL 2 MODEL

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{SQRTINC}_i - \overline{\text{SQRTINC}}_{.}) + r_{0i}$$

$$\pi_{1i} = \beta_{10}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(\text{SQRTINC}_i - \overline{\text{SQRTINC}}_{.}) + r_{2i}$$

$$\pi_{3i} = \beta_{30}$$

Figure 20.3 Model window for the income inequality model

Specifications for this HLM2 run

Problem Title: complete case analysis

The data source for this run = ecls_growth2.mdm
The command file for this run comp_case.hlm
Output file name = complete_case.html
The maximum number of level-1 units = 148470
The maximum number of level-2 units = 21210
The maximum number of iterations = 100
Method of estimation: full maximum likelihood

Note that the data include 148,470 level-1 records and 21,210 level-2 records

Summary of the model specified
Step 2 model
Level-1 Model

$READING_{ti} = \pi_{0i} + \pi_{1i}*(GRADE_{ti}) + e_{ti}$

$MATH_{ti} = \pi_{2i} + \pi_{3}*(GRADE_{ti}) + e_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00} + \beta_{01}*(SQRTINC_i) + r_{0i}$

$\pi_{1i} = \beta_{10}$

$\pi_{2i} = \beta_{20} + \beta_{21}*(SQRTINC_i) + r_{2i}$

$\pi_{3i} = \beta_{30}$

SQRTINC has been centered around the grand mean.

Run-time deletion has reduced the number of level-1 records to 66244

Run-time deletion has reduced the number of level-2 records to 14227

Note that in using only the cases with complete data, the sample size has been reduced from 148,470 to 66,244 level-1 records and from 21,210 to 14,227 level-2 records.

Multivariate Results - Iteration 11

NOTE: level-1 and level-2 slopes have been duplicated across all level-2 equations.

Iterations stopped due to small change in likelihood function

Σ

READING /INTRCPT2	400.82401	223.10221
MATH /INTRCPT2	223.10221	226.33769

Standard errors of Σ

READING /INTRCPT2	2.47540	1.63818
MATH /INTRCPT2	1.63818	1.39974

Σ (as correlations)

READING /INTRCPT2	1.000	0.741
MATH /INTRCPT2	0.741	1.000

Tau

READING	179.35413	128.60219
MATH	128.60219	148.38418

Standard errors of Tau

READING	3.27465	2.51379
MATH	2.51379	2.42009

Tau (as correlations)

READING	1.000	0.788
MATH	0.788	1.000

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For READING /INTRCPT2					
INTRCPT2	88.020042	0.399905	220.103	21208	<0.001
SQRTINC	2.274971	0.047271	48.127	21208	<0.001
GRADE	17.440996	0.030265	576.270	21209	<0.001
For MATH /INTRCPT2					
INTRCPT2	70.532434	0.343039	205.610	21208	<0.001
SQRTINC	1.914606	0.040636	47.116	21208	<0.001
GRADE	14.632042	0.022925	638.269	21209	<0.001

The value of the log-likelihood function at iteration 11 = -5.563652E+005

20.2.1.3 An annotated example of HLM2 with analysis of incomplete data using multiple, model-based imputation

The results show that income is positively associated with both the mean of math and reading proficiency at Grade 3 as well as the linear growth of both proficiencies. This complete case analysis assumed that the data were missing completely at random (“MCAR” - Rubin, 1976). MCAR is a very strong assumption indicating that the missing data are a simple random sample from the complete data. When this strong assumption was untenable, such analysis, in general, is inefficient and will generally result in biased inferences.

20.3 An annotated example of HLM2 analysis of incomplete data

HLM analysis of incomplete data has three major steps (Shin, 2013; Shin & Raudenbush, 2007):

1. Specify the desired models given incompletely observed multilevel data as described in Section 20.2.1.2.

The model specified by the user is the same as that just estimated. However, both the mathematics and reading proficiency variables at level 1 and the INCOME variable at level 2 have missing data. The GRADE variable has no missing data. With the analysis of incomplete multilevel data, HLM automatically reparameterizes the models as the joint distribution of the math and reading outcomes and the variables subject to missingness conditional on the completely observed variables. HLM then efficiently estimates the joint distribution using maximum likelihood under the assumption of multivariate normality.

2. Generate multiply-imputed complete data based on the ML estimates of the joint model. The procedure consists of

a. Opening the **Other Settings** menu and selecting the **Estimation Settings** to open the **Estimation Settings – HLM2** dialog box (See Figure 20.4).

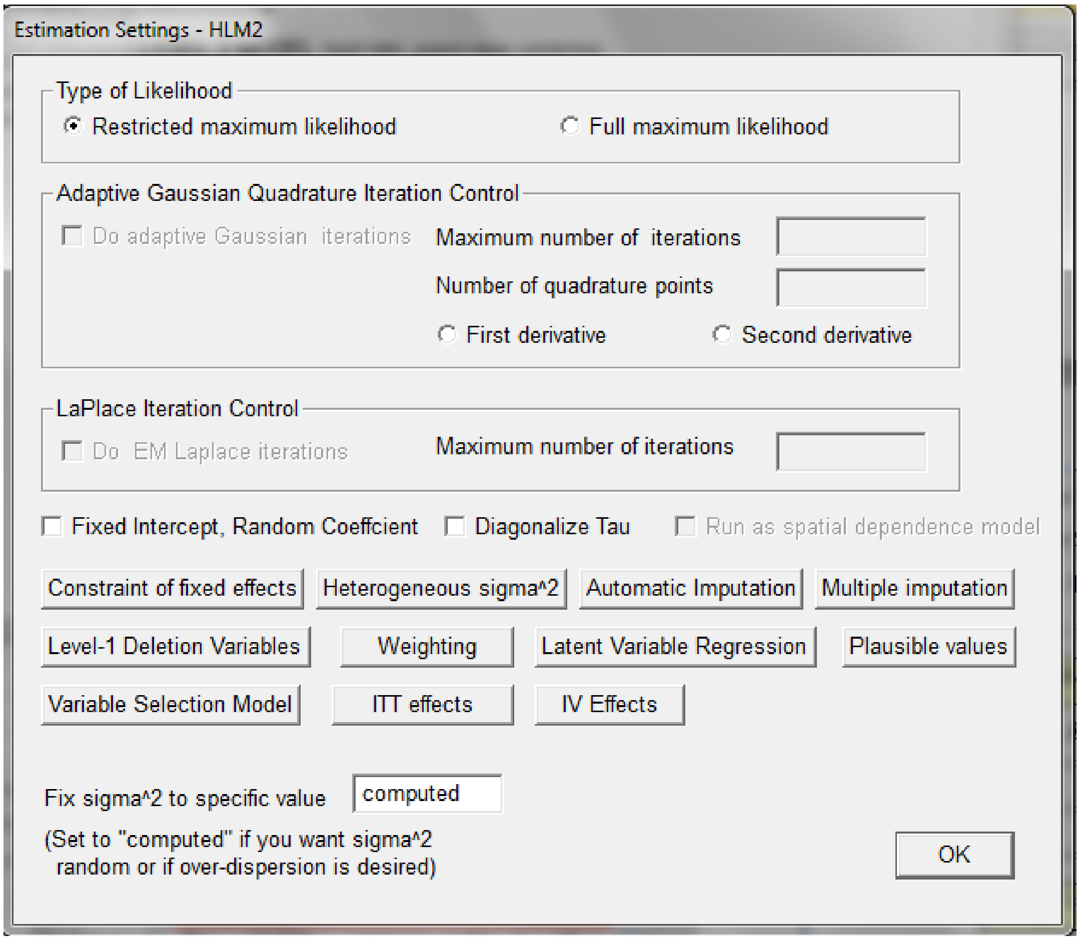


Figure 20.4 Estimation Settings – HLM2 dialog box

Click **Automatic Imputation** to open the **Automatic Multiple Imputation** dialog box (see Figure 20.5). We have selected 10 data sets.

We also choose 2 “augmentation” variables, sometimes called auxiliary variables to improve the imputations. “PARSCAR” is a measure of occupational status and should be a good predictor of income. “BLACK” is an indicator for African-American background.

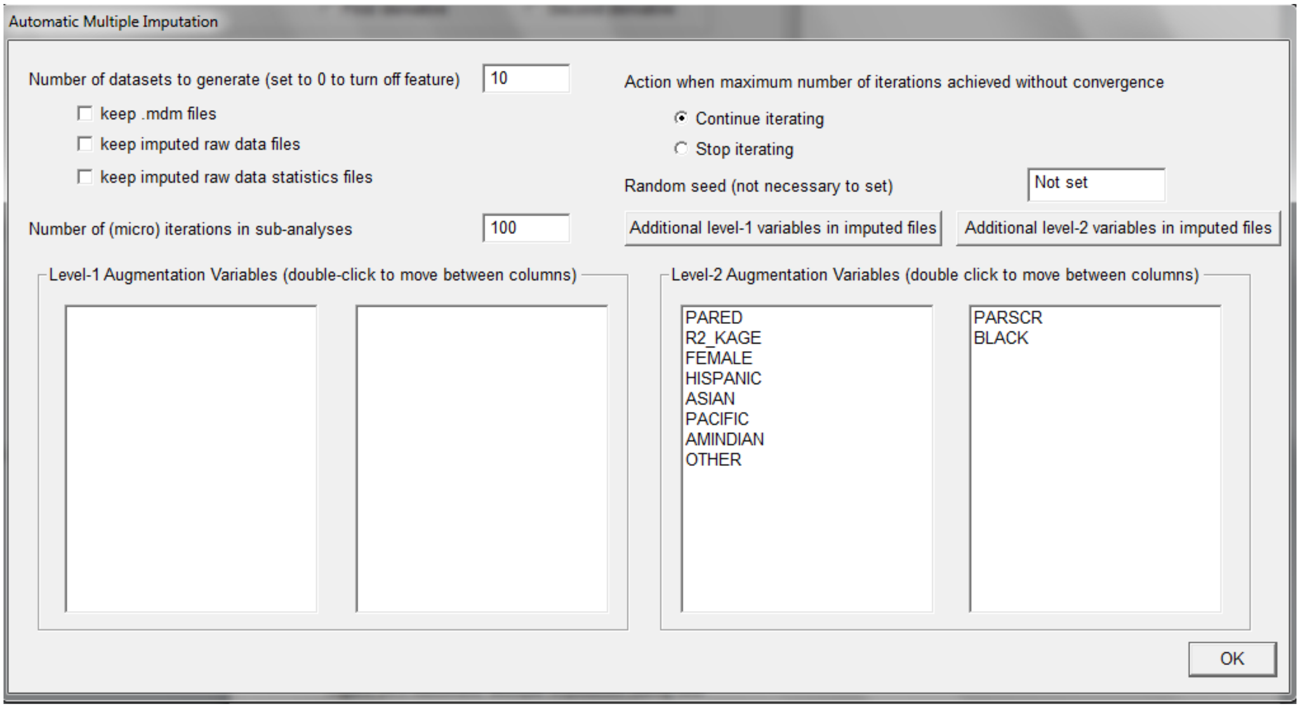


Figure 20.5 Automatic Multiple Imputation dialog box

c. Enter the number of datasets to generate. We enter **10** as a pilot for this example. There is an option to set a random seed number. Users have the option to save the multivariate data matrix files for each data set and keep imputed raw data files that they may like to analyze further using HLM or another program. They can also ask for a record of imputed raw data statistics (see Figure 20.5).

3. Analyze the desired model by complete-data analysis given the multiple imputation. Click **OK** on the **Automatic Multiple Imputation** dialog box, then click **OK** on the **Estimation Settings – HLM2** dialog box. Save and run the model.

The results of the analysis are given below.

Specifications for this HLM2 run

Problem Title: multiple imputation analysis with augmentation

The data source for this run = ecls_growth2.mdm
The command file for this run = mult_imput_aug.hlm
Output file name = mult_imp_aug_avg.html
The maximum number of level-1 units = 148470
The maximum number of level-2 units = 21210
The maximum number of iterations = 100
Method of estimation: full maximum likelihood
Automatic imputation random number seed: -1563333359
Summary of the model specified

Level-1 Model

$READING_{ti} = \pi_{0i} + \pi_{1i}*(GRADE_{ti}) + e_{ti}$
$MATH_{ti} = \pi_{2i} + \pi_{3}*(GRADE_{ti}) + e_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00} + \beta_{01}*(SQRTINC_{i}) + r_{0i}$
$\pi_{1i} = \beta_{10}$
$\pi_{2i} = \beta_{20} + \beta_{21}*(SQRTINC_{i}) + r_{2i}$
$\pi_{3i} = \beta_{30}$

Run-time deletion has reduced the number of level-2 records to 21177

Imputation Model Results - Iteration 12

NOTE: level-1 and level-2 slopes have been duplicated across all level-2 equations.

Iterations stopped due to small change in likelihood function

Σ*

READING /INTRCPT2	384.59710	212.27386
MATH /INTRCPT2	212.27386	218.08554

Standard errors of Σ*

READING /INTRCPT2	2.01342	1.32284
MATH /INTRCPT2	1.32284	1.13236

Σ (as correlations)

READING /INTRCPT2	1.000	0.733
MATH /INTRCPT2	0.733	1.000

Tau*

READING	219.61749	162.47932	18.88858	89.97725
MATH	162.47932	172.81037	15.57391	72.71537
INTRCPT1/ SQRTINC	18.88858	15.57391	8.49343	24.58023
INTRCPT1/ PARSCR	89.97725	72.71537	24.58023	271.33135

Standard errors of Tau*

READING	3.11204	2.39569	0.43120	2.16650
MATH	2.39569	2.24130	0.36548	1.83209
INTRCPT1/ SQRTINC	0.43120	0.36548	0.09854	0.41885
INTRCPT1/ PARSCR	2.16650	1.83209	0.41885	2.70453

Tau (as correlations)

READING	1.000	0.834	0.437	0.369
MATH	0.834	1.000	0.407	0.336
INTRCPT1 / SQRTINC	0.437	0.407	1.000	0.512
INTRCPT1 / PARSCR	0.369	0.336	0.512	1.000

Final estimation of fixed effects (Imputation model)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For READING /INTRCPT2					
INTRCPT2	105.186989	0.137141	767.001	106050	<0.001
BLACK	-8.957698	0.348835	-25.679	106050	<0.001
GRADE	17.364407	0.025382	684.125	21209	<0.001
For MATH /INTRCPT2					
INTRCPT2	85.276019	0.116036	734.912	106050	<0.001
BLACK	-9.810738	0.295823	-33.164	106050	<0.001
GRADE	14.585079	0.019155	761.426	21209	<0.001
For SQRTINC					
INTRCPT2	7.892670	0.024780	318.504	21208	<0.001
BLACK	-2.151829	0.068689	-31.327	21208	<0.001
For PARSCR					
INTRCPT2	42.878649	0.125462	341.766	21208	<0.001
BLACK	-9.272788	0.327137	-28.345	21208	<0.001

The value of the log-likelihood function at iteration 12 = -9.006426E+005

Note that the two outcomes, reading and math (at level 1) are in the multivariate model as outcomes. Also note that the predictor SQRTINC, which is subject to missingness, is also an outcome. In addition, the augmentation variable PARSCR is an additional outcome because HLM has found that it is subject to missingness. All outcomes are regressed on GRADE and BLACK, which are completely observed.

The results for the user specified model are below.

Final Imputation Model Results - 10 Imputations

Σ

READING /INTRCPT2	385.72825	211.59869
MATH /INTRCPT2	211.59869	217.84582

Standard errors of Σ

READING /INTRCPT2	1.29524	1.28592
MATH /INTRCPT2	0.54992	1.23742

Σ (as correlations)

READING /INTRCPT2	1.000	0.730
MATH /INTRCPT2	0.730	1.000

Tau

READING	187.73964	137.45631
MATH	137.45631	154.13887

Standard errors of Tau

READING	2.37041	1.86367
MATH	1.86367	1.80465

Tau (as correlations)

READING	1.000	0.808
MATH	0.808	1.000

Final estimation of fixed effects

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For READING /INTRCPT2					
INTRCPT2	86.624227	0.293843	294.797	4814	<0.001#
SQRTINC	2.229091	0.036345	61.331	12	<0.001#
GRADE	17.371634	0.081062	214.301	9	<0.001#
For MATH /INTRCPT2					
INTRCPT2	69.271059	0.581186	119.189	14	<0.001#
SQRTINC	1.882094	0.081095	23.209	12	<0.001#
GRADE	14.593083	0.056957	256.214	10	<0.001#

The p-vals above marked with a "#" should regarded as a rough approximation.

Note the small degrees of freedom. This reflects the large amount of missing data, particularly on income. The degrees of freedom can be increased by increasing M, the number of data sets.

20.3.1 Cross-Level Interactions

When working with the missing data program, cross-level interactions are presented differently than has been standard in HLM.

Consider for example the univariate model

$$Y_{ti} = \pi_{0i} + \pi_{1i}a_{ti} + e_{ti}$$
$$\pi_{0i} = \beta_{00} + \beta_{01}X_i + r_{0i}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}X_i$$

Suppose that the predictor X is subject to missingness. It must therefore be put on the left side of the imputation model. However, modeling X as a predictor of π_{1i} induces an interaction between a and X as we can see by inspecting the mixed model

$$Y_{ti} = \beta_{00} + \beta_{01}X_i + \beta_{10}\pi_{0i} + \beta_{11}X_i a_{ti} + r_{0i} + e_{ti}\,.$$

Because X is missing, so is $X_i a_{ti}$, so it must also be put on the left. Technically, many such interaction terms will not follow a normal distribution. However, the robustness of the procedure to failure of normality can typically be improved by centering both predictors a and X.

Here is how HLM will represent the model results:

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value	Fraction of Missing Info.
For INTRCPT1, π_0						
INTRCPT2, β_{00}	84.780324	0.381040	222.497	71	<0.001	0.282
SQRTINC, β_{01}	2.475539	0.047462	52.158	60	<0.001	0.301
For GRADE slope, π_1						
INTRCPT2, β_{10}	15.345154	0.098909	155.145	18	<0.001	0.470
For CPROD1 slope, π_2						
INTRCPT2, β_{20}	0.249093	0.012099	20.588	16	<0.001	0.483

CPROD1 is the generated cross-product of GRADE*SQRTINC.

A Using HLM2 in interactive and batch mode

This appendix describes and illustrates how to use HLM2 in batch mode to construct MDM files, to execute analyses based on the MDM file, and to specify a residual file to evaluate model fit. It also lists and describes command keywords and options. References are made to appropriate sections in the manual where the procedures are described in greater details.

A.1 Using HLM in batch and/or interactive mode

HLM users can control which questions come to the screen by means of a command file. At one extreme, the command file is virtually empty and questions regarding every possible optional procedure or output will come to the screen. At the other extreme, the command file specifies the answer to every question that might arise, in which case the analysis is performed completely in batch mode. In between the two extremes are a large number of possibilities in which various questions are answered in the command file while other questions come to the screen. Hence, the execution can be partly batch and partly interactive.

An example of a command file for the Intercept and Slopes-as-Outcomes Model for the HS&B data is shown below. The italicized comments provide a brief description of each command function. A complete overview of each of the keywords and related options in this command file appears in the Section A.2.

```
#This command file was run with HSB.MDM                        Indicates which MDM was used.
NUMIT:100                                            Sets the maximum number of iterations.
STOPVAL:0.0000010000          Sets the criteria for automatically stopping the iterations.
NONLIN:N                                                Switch to do a non-linear analysis.
LEVEL1:MATHACH=INTRCPT1+SES,1+RANDOM                            Specifies the level-1 model.
LEVEL2:INTRCPT1=INTRCPT2+SECTOR+MEANSES+RANDOM/SIZE,PRACAD,DISCLIM,HIMNTY
LEVEL2:SES=INTRCPT2+SECTOR+MEANSES+RANDOM/SIZE,PRACAD,DISCLIM,HIMNTY
                                              Specifies the level-2 model and other level-2
                                         predictors for possible inclusion in subsequent models
                                                          for both intrcpt1 and the ses slope.
LEVEL1WEIGHT: NONE                                             Specifies level-1 weight variable.
LEVEL2WEIGHT: NONE                                             Specifies level-2 weight variable.
RESFIL:N                                              Controls whether a residual file is created.
HETEROL1VAR:N                                      Specifies an analysis with a heterogeneous sigma2.
ACCEL:5                                                Controls frequency of use of accelerator.
LVR:N                                                Specifies a latent variable regression model.
LEV1OLS:10                                         Controls the number of level-1 OLS regressions
                                                                                      printed out.
MLF: N                                                   Specifies restricted maximum likelihood.
HYPOTH:N                                       Disables some optional hypothesis testing procedures.
FIXTAU:3                                           Alternative options for generating starting values.
CONSTRAIN:N                                   Estimates a model with constrained level-2 coefficients.
OUTPUT:HSB1.OUT                                              File where HLM2 output will be saved.
FULLOUTPUT: Y                                              Controls amount of output in output file.
   TITLE: Intercept and Slopes-as-Outcome Model                           Title on page 1 of output.
```

An user can rename the file with or without modification with a plain text (ASCII) editor for subsequent batch-mode application. For instance, he or she may request the program to print out all the level-1 OLS regressions by changing the LEV1OLS:10 to LEV1OLS:160 and rename the file to HSB2.MLM. The user can execute the analysis by typing:

```
HLM2 HSB.MDM HSB2.MLM
```

at the system prompt. As the run is fully specified in the command file HSB2.MLM, no questions will come to the screen during its execution. This is full batch mode. The user may choose a fully interactive execution mode or an execution mode that is partly interactive and partly batch. With partly interactive, partly batch mode, some specification occurs in the command file; the program prompts the user with questions for the remaining program features. Some users may find this a useful way to suppress some the questions relating to less often used features of the programs. Fully interactive mode is invoked when one of the programs is invoked without a second argument, i.e.,

```
HLM2 HSB.MDM
```

In this case, all of the possible questions will be asked with the exception relating to type of estimation used. (mlf:y must be specified in the command file).

A.2 Using HLM2 in batch mode

A command file consists of a series of lines. Each line begins with a keyword followed by a colon, after the colon is the option chosen by the user, i.e.,

```
KEYWORD:OPTION
```

For example, HLM2 provides several optional hypothesis-testing procedures, described in detail in the Sections 2.9.2 to 2.9.4. Suppose the user does not wish to use these optional procedures in a given analysis. Then the following line would be included in the command file:

```
HYPOTH:N
```

The keyword HYPOTH concerns the optional hypothesis testing procedures; the option chosen, 'N', indicates that the user does not wish to employ these procedures. Alternatively, the user might include the line:

```
HYPOTH:Y
```

This prompts HLM2 to activate the optional hypothesis testing menu during model specification in the interactive mode. Lines beginning with a pound (#; also called hash mark) are ignored and may be used to put comments in the command file.

HLM2, by default, has set up the following options unless the user specifies an alternative command file.

STOPVAL:0.0000010000	*Sets convergence criterion to be 0.000001.*
ACCEL:5	*Use accelerator once after five iterations.*
FIXTAU:3	*Use the "standard" computer-generated values for the variances and covariances.*
MLF:N	*Use the restricted maximum likelihood approach.*

Table A.1 presents the list of keywords and options recognized by HLM2. Examples with detailed explanation follow.

Table A.1 Keywords and options for the HLM2 command file

Keyword	Function	Option	Definition
LEVEL1	Level-1 model specification	INTRCPT1	Level-1 intercept
		+VARNAME	Level-1 predictor (no centering)
		+VARNAME,1	Level-1 predictor centered around group (or level-2 unit) mean
		+VARNAME,2	Level-1 predictor centered around grand mean
LEVEL2	Level-2 model specification	INTRCPT2	Level-2 intercept
		+VARNAME	Level-2 predictor (no centering)
		+VARNAME,2	Level-2 predictor centered around grand mean
		+/VARLIST	List after the slash level-2 variables for exploratory analysis and "t-t-enter" statistics on subsequent runs
NUMIT	Maximum number of iterations	POSITIVE INTEGER	
ACCEL	Controls iteration acceleration	INTEGER $\geq$ 3	Selects how often the accelerator is used. Default is 5.
LEV1OLS	Number of units for which OL equations should be printed	POSITIVE INTEGER	Default is 10.
CONSTRAIN	Constraining of gammas	N	No constraining
		Y	Yes: two or more gammas will be constrained

The program will prompt the user interactively to set the constraints. Alternatively, constraints can be set in the command file. For example, suppose the following coefficients were estimated: $\gamma_{01}, \gamma_{11}, \gamma_{20}, \gamma_{21}$ and we wish to specify $\gamma_{20} = \gamma_{21}$, we add the following command line:

```
CONSTRAIN: 0,0,1,1.
```

For the following coefficients: $\gamma_{00}, \gamma_{01}, \gamma_{02}, \gamma_{10}, \gamma_{11}, \gamma_{12}$, the command line:

```
CONSTRAIN: 0,1,2,0,1,2
```

will have the following result: $\gamma_{01} = \gamma_{11}$ and $\gamma_{02} = \gamma_{12}$.

Note that all coefficients sharing the value "0" are free to be estimated independently.

HYPOTH	Select optional hypothesis testing menu	Y	Yes: send optional hypothesis testing menu to the screen during interactive mode use.
		N	No. (Note, during batch execution, HYPOTH:N should be selected to suppress screen prompt. Select desired options through keywords below.
GAMMA#	Specifies a particular multivariate contrast to be tested.		In any single run, HLM2 will test up to 5 multivariate hypotheses. Each hypothesis may consist of up to 5 contrasts.

Table A.1 Keywords and options for the HLM2 command file (continued)

Each contrast is specified by its own line in the command file. The contrast associated with the first hypothesis is specified with the keyword GAMMA1. For example, the contrast shown in Fig 2.37 can be specified by adding the following lines:

```
GAMMA1:0.0,1.0,0.0,0.0,0.0,0.0
GAMMA1:0.0,0.0,0.0,0.0,1.0,0.0
```

For the second hypothesis, the keyword is GAMMA2 and for the third it is GAMMA3 (See Section 2.9.2 for further discussion and illustration.)

HOMVAR	Test homogeneity of level-1 variance	N Y	No Yes
DEVIANCE	Deviance statistic from prior analysis	POSITIVE REAL NUMBER	-2 * log-likelihood at maximum-likelihood estimate
DF	Degrees of freedom associated with deviance statistics from prior analysis (use only if "DEVIANCE" has been specified)	POSITIVE INTEGER	
FIXTAU	Method of correcting unacceptable starting values	1 2 3 4 5	Set all off-diagonal elements to 0 Manually reset starting values Automatic fix-up (default) Terminate run Stop program even if starting values are acceptable; display starting values and then allow user to manually reset them.
HETERO1VAR		N VARLIST	No Variable list
FIXSIGMA2	Controls σ^2	N REAL NUMBER>0	Default: does not restrict σ^2. Fixes σ^2 to the specified value.
LEVEL1WEIGHT LEVEL2WEIGHT	Specifies design weights	Variable name	Allows specification of design weights at the respective levels. Example level1weight:weight1
LEVEL1DELETION	Level-1 deletion list	VARLIST	This keyword only comes into play when the user has opted for deleting data at analysis time while making the MDM file. By default in such cases, deletion is done on the variables in the model. See section 2.9.2.2 for more details.
STOPVAL	Convergence criterion for maximum likelihood estimation	POSITIVE REAL NUMBER	Example: 0.000001. Can be specified to be more (or less) restrictive.

Table A.1 Keywords and options for the HLM2 command file (continued)

MLF	Controls maximum likelihood estimation method	N Y	No Yes, full maximum likelihood. Produces standard errors of T and σ^2
FIRC	Controls fixed intercept, random coefficient feature	N Y	No Use feature as documented in [insert reference]
RESFIL1	Create level-1 residual file	N Y/[vl1]/[vl2]	No Yes – this may be followed by two '/'s denoting the two levels that can be in the residual file. By default, all the variables in the model will be present in the residual file, this can be added to put additional variables. Vl1 and vl2 are lists of comma-separated variables
RESFIL2	Create a residual file	Y N /VARLIST	Yes No List after the slash additional level-2 variables to be included in the residual file.
RESFIL1NAME, RESFIL2NAME	Name of residual file	FILENAME	The names, respectively of the level-1 and level-2 residual files.
RESFILTYPE	Type of residual file	SYSTAT SAS SPSS STATA FREEFORMAT	Selects program type to be used in subsequent analysis of residual file. SPSS and Stata residual files are written out as .sav and .dta files. Free format files are written out in ASCII format with the first line of the file being the variable names
PRINTVARIANCE-COVARIANCE	Output files containing the variance-covariance matrices of Tau and Gammas	N Y A	No Yes Append the files in consecutive runs.
TITLE			Program label up to 64 characters.
OUTPUT	Filename of file that contains output	FILENAME	Will be written to disk; output will overwrite a file of same name.
FULLOUTPUT	Amount of desired output	Y N	Full (traditional) output Reduced output only containing header page and final results

The following keywords are specific to nonlinear, latent variable, and multiply imputed data analysis:

NONLIN	Selects a nonlinear analysis	BERNOULLI POISSON BINOMIAL, COUNTVAR POISSON, COUNTVAR MULTINOMIAL, COUNTVAR ORDINAL, COUNTVAR	These options are explained in detail in Chapter 8.
MACROIT	Maximum number of macro iterations	POSITIVE INTEGER	Used in non-linear models
MICROIT	Maximum number of micro iterations	POSITIVE INTEGER	Used in non-linear models
STOPMACRO	Convergence criterion for change in parameters across macro iterations	POSITIVE INTEGER	

Table A.1 Keywords and options for the HLM2 command file (continued)

STOPMICRO	Convergence criterion for micro iterations	POSITIVE INTEGER	Note same function as STOPVAL in a linear analysis.
LAPLACE	Requests Laplace-6 iterations	N Y,# where '#' is the maximum number of iterations desired	No Yes, with # iterations; uses a sixth order approximation to the likelihood based on a Laplace transform for Bernoulli models. See Sections 7.6.3 and 8.8.2 for details.
EMLAPLACE	Requests EM-Laplace iterations	N Y,# where '#' is the maximum number of iterations desired	No Yes, with # iterations. Uses third order approximation.
LVR	Performs a latent variable regression	N I P O	No Ignore Predictor Outcome (See Section 11.1 for details)
PLAUSVALS	Selects a list of plausible values for multiple imputation application	VARLIST	See Section 11.2.1 for details.

The following keywords are specific to multiple imputation

Autoimpute	Requests auto-imputation of missing data	Positive integer	Number of imputed datasets to create
Autoimputeiter	Controls behavior of the automatic imputation. Format like NUMIT: autoimputeriter:#,[y/n]	# default number of joint model iterations	N, stop at this number of joint model Y: continue until joint model converges
Autoimputekeep 3 choices here. Example autoimputkeep:y,y,y	The first choice tells the program to keep the created mdm files	The second choice tells the program to keep the imputed data files	The third choice controls keeping the stats files of the generated mdm files
Autoimputseed	Not often used. Specifies the random number seed	Positive integer	
Level1-Augvars	List of level-1 variables used to augment the joint model. Will not be used in user specified model		
Level2-Augvars	List of level-2 variables used to augment the joint model. Will not be used in user specified model		

The following keywords are specific to multivariate models

Mhlm	Specifies the model will be multivariate. Must be specified before the model itself.	Y N	Use the feature Regular univariate model
Level2outcome	Only used in multivariate models and tells the program to use a level-2 variable as an outcome	Specifies model. The format is the same as LEVEL2 above	

A.3 Printing of variance and covariance matrices for fixed effects and level-2 variances

The variance-covariance matrices of estimates of fixed effects and variance-covariance parameters based on HLM2 or HLM3 can be saved by checking the "print variance-covariance matrices" in the Output Settings dialog box accessed via the **Other Settings** menu. The keyword PRINTVARIANCE-COVARIANCE facilitates the same purpose in batch mode.

The following gives a description of the files containing critical statistics and their variances that are provided by the program upon request.

Let r = number of random effects at level-1.
f = number of fixed effects
p = number of outcomes in a latent variable run
pm = number of alphas in a latent variable run

1. For HLM2:

- TAUVC.DAT contains tau in r columns of r rows and then the inverse of the information matrix (the standard errors of tau are the square roots of the diagonals). The dimensions of this matrix are $r*(r+1)/2 \times r*(r+1)/2$.
- GAMVC.DAT contains the gammas and the gamma variance-covariance matrix. After the gammas, there are f more rows of f entries containing the variance-covariance matrix.
- GAMVCR.DAT contains the gamma and the gamma variance-covariance matrix used to compute the robust standard errors. After the gammas, there are f rows of f entries containing the variance-covariance matrix.

2. For HGLM:

- TAUVC.DAT contains tau for the final unit-specific results in r columns of r rows and then the inverse of the information matrix (the standard errors of tau are the square roots of the diagonals). The dimensions of this matrix are $r*(r+1)/2 \times r*(r+1)/2$.
- GAMVCUS.DAT contains the final unit-specific gammas and the gamma variance-covariance matrix. The gammas are in the first line and this line has f entries. Then there are f more rows of f entries containing the variance and covariance matrix.
- GAMVCPA.DAT contains the final unit-specific gammas and the gamma variance-covariance matrix. The gammas are in the first line and this line has f entries. Then there are f more rows of f entries containing the variance and covariance matrix.
- GAMVCPAR.DAT contains the final unit-specific gammas and the gamma variance-covariance

matrix used to compute the population-averaged robust standard errors. The gammas are in the first line and this line has f entries. Then there are *f* more rows of *f* entries containing the variance and covariance matrix.

3. For Bernoulli models, if Laplace iterations are requested:

GAMVCL.DAT contains the gammas and the variance-covariance matrix used to compute the Laplace standard errors. There are *f* rows of *f* entries containing the variance and covariance matrix.

4. For latent variable regression:

LVRALPHA.DAT contains *pm* lines each containing an alpha and its standard error. The order is the same as in the output table. The final p lines of *p* columns contain the $Var(u^{*})$ matrix printed in the output.

5. For plausible values analysis:

GAMVC.DAT (and GAMVCR.DAT and TAUVC.DAT) are from the last run and TAUVCPC.DAT, GAMVCPV.DAT, and GAMVCPVR.DAT are the PV average files.

All of the above files are created with an n(F15.7 1X) format. That is, each entry is fifteen characters wide with even decimal places, followed by a space (blank character).

If the value of r or $r*(r+1)/2$ exceeds 60, the line is split into two or more pieces.

B Using HLM3 in Interactive and Batch Mode

This appendix describes and illustrates how to use HLM3 in batch mode to construct MDM files, and to execute analyses based on the MDM file. It also lists and defines command keywords and options unique to HLM3. References are made to appropriate sections in the manual where the procedures are described in greater details.

As in the case of HLM2, formulation, estimation, and testing of models using HLM3 in two ways: Windows mode (PC users only), or batch mode. However, batch mode can be considerably faster once the user becomes skilled in working with the program.. The degree to which the execution is automated (via batch mode) is controlled by the command file, as in the case of HLM2.

B.1 Using HLM3 in batch mode

The command file structure for HLM3 closely parallels that of HLM2. Each line begins with a keyword followed by a colon. After the colon is the option chosen by the user, i.e.,

```
KEYWORD:OPTION
```

As with HLM2, a pound sign ("#") as the first character of a line can be used to introduce a comment into the command file.

The following keywords have the same definitions and options in HLM3 as in HLM2 (Table A.1)

ACCEL	CONSTRAIN	DEVIANCE	DF	FIXTAU	FIXSIGMA2
GAMMA#	HYPOTH	LAPLACE	MACROIT	NONLIN	PRINTVARIANCE-COVARIANCE
NUMIT	OUTPUT	PLAUSVALS	RESFIL1	RESFIL1	RESFIL1NAME
TITLE	RESFILTYPE	FIXSIGMA2	STOPMACRO	STOPMICRO	LEVEL1DELETION
OUTPUT	FULLOUTPUT	FIRC	MICROIT	STOPMACRO	RESFIL2NAME

The following keywords are available only for HLM2:

LEV1OLS HOMVAR HETERO1VAR MLF LVR

Table of keywords and options

Table B.1 presents the list of keywords and options unique to HLM3.

Table B.1 Keywords and options unique to the HLM3 command file

Keyword	Function	Option	Definition
LEVEL1	Level-1 model specification	INTRCPT1	Level-1 intercept
		+VARNAME	Level-1 predictor (no centering)
		+VARNAME,1	Level-1 predictor centered around unit mean a.jk
		+VARNAME, 2 (Note: variable names may be specified in either upper or lower case.)	Level-1 predictor centered around grand mean a...
LEVEL2	Level-2 model specification	INTRCPT2	Level-2 intercept
		+VARNAME	Level-2 predictor (no centering)
		+VARNAME,1	Level-2 predictor centered around group mean, $X_{.k}$
		+VARNAME, 2	Level-2 predictor centered around grand mean, $X_{..}$
		/VARLIST	Comma separated list after the slash level-2 variables for exploratory analysis and "t-to-enter" statistics on subsequent runs. A slash without a subsequent variable suppresses the interactive prompt.
LEVEL3	Level-3 model specification	INTRCPT3	Level-3 intercept (must be included in the level-2 model)
		+ VARNAME	Level-3 predictor (no centering)
		+ VARNAME, 2	Level-3 predictor centered around grand mean, W.
		/VARLIST	List after the slash level-3 variables for exploratory analysis and "t-to-enter" statistics on subsequent runs. A slash without a subsequent variable suppresses the interactive prompt.
RESFIL3	Create a level-3 residual file	Y	Yes
		N	No
		/ VARLIST	List after the slash additional level-3 variables to be included in the residual file. Works just like RESFIL2
RESFIL3 NAME	Name of residual file	FILENAME	Changes the default
FIXTAU2	Method of correcting unacceptable starting values for T_{π}	1	Set all off-diagonal elements to 0
		2	Manually reset starting values
		3	Automatic fix-up (default)
		4	Terminate run
		5	Stop program even if the starting values are acceptable; display starting values and then allow user to manually reset them.

FIXTAU3	Method of correcting unacceptable starting values for T_β	1	Set all off-diagonal elements to 0
		2	Manually reset starting values
		3	Automatic fix-up (default)
		4	Terminate run
		5	Stop program even if the starting values are acceptable; display starting values and then allow user to manually reset them.
LVR-BETA	Performs a latent variable regression	N	No
		P,O	P for predictor(s); O for outcomes (s) See Section 11.1 for details.
DOFISHER	Turns on/off Fisher estimation	Y	Use Fisher
		N	Do not use Fisher
FISHERTYPE	Controls type of Fisher acceleration	0	Same as DOFISHER:N
		1	Use 1st derivate Fisher
		2	Use 2nd derivative Fisher(default) See section 4.5.

B.3 Printing of variance and covariance matrices

HLM3 can provide the following files upon request.

Note that adding the command line

```
PRINTVARIANCE-COVARIANCE:Y
```

to the command file will request HLM3 to print out statistics for both tau(pi) as well as tau(beta).

Let r = number of random effects at level-1
$r2$ = number of random effects at level-2

1. For HLM3:

- TAUVC.DAT contains tau (tau(pi)) in r columns of r rows, the next $r2$ lines are the tau(beta), and then the inverse of the information matrix (the standard errors of tau[s] are the square roots of the diagonals). The dimensions of this matrix are $(r*(r+1)/2+r2*(r2+1)/2)\times(r*(r+1)/2+r2*(r2+1)/2)$.

2. For three-level HGLM:

- TAUVC.DAT has the same format as the one for HLM3. The tau(s) are the final unit-specific results.

The files for the gammas have the identical structure as those for two-level models.

All files are created with an n(F15.7,1X) format. That is, each entry is fifteen characters wide with seven decimal places, followed by a space (blank character).

If the value of r or f or $r*(r+1)/2+r2*(r2+1)/2$ exceeds 60, the line is split into two or more pieces.

C Using HLM4 in Batch Mode

Unlike the older modules (HLM2, HLM3, etc.), HLM4 does not have interactive modes to create the MDM or specify a model. If the windows interface is not available, these file must be created with an ASCII editor and submit them to obtain results.

C.1 Example: Creating an MDM file from raw data

The first thing that needs creating is an MDM template file (usually suffixed with .mdmt), which tells HLM4 how to read the raw data. Here is the MDMT file from section 6.1.1:

```
#HLM4 MDM CREATION TEMPLATE
mdmtype:3
rawdattype:spss
l1fname:C:\HLM\Examples\measure.sav
l2fname:C:\HLM\Examples\occas.sav
l3fname:C:\HLM\Examples\tchr.sav
l4fname:C:\HLM\Examples\sch.sav
l1missing:n
timeofdeletion:now
mdmname:literacy.mdm
*begin l1vars
level4id:SCHID
level3id:TCHRID
level2id:OCCASID
EXPERTIS
STDERR
*end l1vars
*begin l2vars
level4id:SCHID
level3id:TCHRID
level2id:OCCASID
OCCASION
ARTIFACT
*end l2vars
*begin l3vars
level4id:SCHLID
level3id:TCHRID
COACH
NEWTCHR
PDPART
SCMT
Y2ENT
Y3ENT
*end l3vars
*begin l4vars
level4id:SCHID
CHGCOACH
*end l4vars
```

The file is broken into two sections. The first is to declare the filenames of the raw data and other characteristics of the MDM file to be made, the second chooses the variables to be included at the various levels. Below is the first part with explanation in parentheses:

```
#HLM4 MDM CREATION TEMPLATE                           (Required to be exactly like this.)
mdmtype:3                              (This declares the structure of the data, and only
                                 affects the notation used in the output. Possible values are
1                              for cross sectional, 2 for longitudinal, 3 for cross sectional
                                    with measurement model at level 1 and 4 for longitudinal
                                                            with measurement model at level 1)
```

```
rawdattype:spss               (This declares the type of input data. Possible values are
                                  spss, sas (version 5 transport file), stata, and ascii)
l1fname:measure.sav                  The next four lines declare the names and locations
                                                                  of the four input files.)
l2fname:occas.sav
l3fname:tchr.sav
l4fname:sch.sav
l1missing:n                          (This declares whether or not there are missing data
                                  at level-1. Possible values are n for not missing, or y
                                                              for missing data present.)
timeofdeletion:now              (This may be n[ow], where all level-1 cases with missing
                               data on selected variables will be deleted, or a[nalysis]
                                    where the missing data will be left in and deleted at
                                                  run-time based on the model specified.)
mdmname:literacy.mdm                                   (Specifies the name of the mdm file.)
```

The second part of the mdmt file specifies which variables are ID variables, and which ones go into the mdm file as possible analysis variables. The structure looks like this:

```
*begin l1vars
level4id:SCHID
level3id:TCHRID
level2id:OCCASID
[list of level-1 variables, one per line]
*end l1vars
*begin l2vars
level4id:SCHID
level3id:TCHRID
level2id:OCCASID
[list of level-2 variables, one per line]
*end l2vars
*begin l3vars
level4id:SCHLID
level3id:TCHRID
[list of level-3 variables, one per line]
*end l3vars
*begin l4vars
level4id:SCHID
[list of level-4 variables, one per line]
*end l4vars
```

The IDs must be specified in the order shown, and must all be of the same type, either numeric (preferable) or alphanumeric(not advised).

Once the mdmt file is created, the file must be submitted to HLM4:

```
C:\HLM> HLM4 –r literacy.mdmt
```

The results on the screen should then be examined to make sure the data were read correctly. These descriptive statistics will also be contained in a file named HLM4MDM.STS.

C.2 Example: Creating an HLM file and running the model

The next step is to create a file that specifies the desired model. (This is usually suffixed with a .hlm) For example, we will use the model shown in section 6.2.

```
nonlin:n
numit:100
stopval:0.0000010000
level1:EXPERTIS=STDERR+RANDOM
level2:STDERR=INTRCPT2+OCCASION+ARTIFACT+random
level3:INTRCPT2=INTRCPT3+random
level4:INTRCPT3=INTRCPT4+CHGCOACH+random
level3:OCCASION=INTRCPT3+random
level4:INTRCPT3=INTRCPT4+CHGCOACH+random
level3:ARTIFACT=INTRCPT3
level4:INTRCPT3=INTRCPT4+CHGCOACH+random
fixsigma2:1.000000
fixtaupi:3
fixtaubeta:3
fixtaugamma:3
accel:5
level1weight:none
level2weight:none
level3weight:none
level4weight:none
hypoth:n
resfil1:n
resfil2:n
resfil3:n
resfil4:n
title:Unconditional model for literacy program
output: literacy1.txt
fulloutput:y
```

The above is very similar to an HLM3 model file, with the exception of the model specification at the top where an extra level is shown. Here is the model part that better demonstrates the nested nature of the model specification (the shown indentation will not run):

```
level1:EXPERTIS=STDERR+RANDOM
       level2:STDERR=INTRCPT2+OCCASION+ARTIFACT+random
          level3:INTRCPT2=INTRCPT3+random
             level4:INTRCPT3=INTRCPT4+CHGCOACH+random
          level3:OCCASION=INTRCPT3+random
             level4:INTRCPT3=INTRCPT4+CHGCOACH+random
          level3:ARTIFACT=INTRCPT3
             level4:INTRCPT3=INTRCPT4+CHGCOACH+random
```

The basic rule here is that for each level-1 variable in the model, there needs to be a `level2` line, for each level-2 variable, a `level3` file, and for each level-3 variable, a `level4` line. The order is not arbitrary and must follow the pattern above.

Assuming that the above file is named literacy1.hlm, then the following command should be run:

```
C:\HLM> HLM4 LITERACY.MDM LITERACY1.HLM
```

Given the HLM file above the output would then be in literacy1.txt. Note that if html output is desired, a .html suffix should be specified on the output: line rather than .txt.

Table C.1 presents the list of keywords and options unique to HLM4 relative to HLM3.

Table C.1 Keywords and options unique to the HLM4 command file

Keyword	Function	Option	Definition
LEVEL1	Level-1 model specification	INTRCPT1 +VARNAME +VARNAME,2 +VARNAME,3 +VARNAME,4 +VARNAME,G *(Note: variable names may be specified in either upper or lower case.)*	Level-1 intercept Level-1 predictor (no centering) Level-1 predictor centered around level-2 mean Level-1 predictor centered around level-3 mean Level1 predictor centered around level-4 mean Level-1 predictor centered around grand mean a...
LEVEL2	Level-2 model specification	INTRCPT2 +VARNAME +VARNAME,3 +VARNAME,4 +VARNAME,G	Level-2 intercept Level-2 predictor (no centering) Level-2 predictor centered around level-3 mean Level-2 predictor centered around level-4 mean Level-2 predictor centered around grand mean
LEVEL3	Level-3 model specification	INTRCPT3 +VARNAME +VARNAME,4 +VARNAME,G	Level-3 intercept (must be included in the level-2 model) Level-3 predictor (no centering) Level-3 predictor centered around level-4 mean Level-3 predictor centered around grand mean
LEVEL4	Level-4 model specification	INTRCPT4 +VARNAME +VARNAME,G	Level-3 intercept (must be included in the level-2 model) Level-3 predictor (no centering) Level-3 predictor centered around grand mean
RESFIL4	Create a level-3 residual file	Y N / VARLIST	Yes No List after the slash additional level-4 variables to be included in the residual file. Works just like RESFIL2
RESFIL4NAME	Name of residual file	FILENAME	Changes the default
FIXTAU4	Method of correcting unacceptable starting values for $\mathbf{T}_\gamma$	1 2 3 4 5	Set all off-diagonal elements to 0 Manually reset starting values Automatic fix-up (default) Terminate run Stop program even if the starting values are acceptable; display starting values and then allow user to manually reset them.

The command file structure for HLM3 closely parallels that of HLM2. Each line begins with a keyword followed by a colon. After the colon is the option chosen by the user, *i.e.*,

```
KEYWORD:OPTION
```

As with HLM2, a pound sign ("#") as the first character of a line can be used to introduce a comment into the command file.

The following keywords have the same definitions and options in HLM3 as in HLM2 (Table A.1)

```
ACCEL      CONSTRAIN  DEVIANCE      DF         MACROIT    MICROIT    NONLIN     NUMIT
OUTPUT     RESFIL1    RESFIL1NAME   RESFIL2    RESFIL2NAME           RESFILTYPE FIXSIGMA2
STOPMACRO  STOPMICRO  STOPVAL       TITLE      LEVEL1DELETION        FULLOUTPUT
```

D Using HGLM in Interactive and Batch Mode

This appendix describes and illustrates how to use HGLM in interactive and batch mode to execute analyses based on the MDM files. References are made to appropriate sections in the manual where the procedures are described in greater details.

D.1 Example: Executing an analysis using THAIUGRP.MDM

Here is an example of an HLM2 session in the interactive mode. At the system command line prompt, we first type the program name – HLM2 – followed by the name of the multivariate data matrix file – THAIUGRP.MDM. The program now takes the user directly into the model specification process.

```
C:\HLM> HLM2 THAIUGRP.MDM

Do you want to do a non-linear analysis? Y

Enter type of non-linear analysis:
```

See Chapter 5 for details regarding type of non-linear analysis.

```
1) Bernoulli (0 or 1)
2) Binomial (count)
3) Poisson  (constant exposure)
4) Poisson  (variable exposure)
5) Multinomial
6) Ordinal

type of analysis: 1
```

As mentioned, with one binary outcome per level-1 unit, the model choice is "1" (Bernoulli).

If "2"(Binomial) is chosen, the user will be asked:

```
For the non-linear analysis, which variable indicates the number of trials?
```

If "4"(Poisson (variable exposure)) is chosen, the user will be asked:

```
For the non-linear analysis, which variable indicates the exposure?
```

If "5"(Multinomial) or "6"(Ordinal) is chosen, the user will be asked:

```
How many categories does the "OUTCOME" have?

Enter maximum number of macro iterations: 25
Enter maximum number of micro iterations: 20
```

Specifying 25 macro iterations sets an upper limit; if, after the 25th iteration the algorithm has not converged. The program will nonetheless terminate and print the results at that iteration. Similarly, setting 20 as the number of micro iterations insures that, after 20 micro iterations, the *current macro iteration will terminate even if the micro iteration convergence criterion has not been met.*

```
Do you wish to allow over-dispersion at level 1? N
```

An answer of "Y" here allows a user to estimate a level-1 dispersion parameter σ^2. If the assumption of no dispersion holds, $\sigma^2 = 1.0$. If the data are over-dispersed, $\sigma^2 > 1.0$; if the data are under-dispersed, $\sigma^2 < 1.0$.

```
Do you want to do the Laplace-6 iterations? N
Do you want to do the Laplace-8 iterations? N
```

An answer of "Y" here allows us to obtain highly accurate Laplace approximation to maximum likelihood. See Sections 7.6.3 and 8.9.2. The user will be prompted to enter maximum number of Laplace macro iterations.

```
                   SPECIFYING A LEVEL-1 OUTCOME VARIABLE

Please specify a level-1 outcome variable
 The choices are:
 For     MALE enter  1    For     PPED enter  2    For     REP1 enter  3

What is the outcome variable: 3

Do you wish to:

   Examine means,variances,chi-squared, etc? Enter 1
   Specify an HLM model?                     Enter 2
   Define a new outcome variable?            Enter 3
   Exit?                                     Enter 4
What do you want to do? 2

                          SPECIFYING AN HLM MODEL
Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
 For     MALE enter  1    For     PPED enter  2
 level-1 predictor? (Enter 0 to end)  1
 level-1 predictor? (Enter 0 to end)  2
```

Thus, we have set up a level-1 model with repetition (REP1) as the outcome and with gender (MALE) and pre-primary experience (PPED) as predictors.

```
 Do you want to center any level-1 predictors? N

Do you want to set the level-1 intercept to zero in this analysis? N

Level-2 predictor variable specification

Which level-2 variables do you wish to use?

 The choices are:
 For    MSESC enter  1

 Which level-2 predictors to model INTRCPT1?
  Level-2 predictor? (Enter 0 to end)  1
 Which level-2 predictors to model     MALE slope?
```

```
  Level-2 predictor? (Enter 0 to end)  0
 Which level-2 predictors to model     PPED slope?
  Level-2 predictor? (Enter 0 to end)  0
```

Thus we have modeled the level-1 intercept as depending on the mean SES (MSESC) of the school. The coefficients associated with gender and pre-primary experience are fixed. Mean SES has been centered around its grand mean.

```
Do you want to constrain the variances in any of the level-2 random
 effects to zero? Y
 Do you want to fix INTRCPT1?  N
 Do you want to fix     MALE?  Y
 Do you want to fix     PPED?  Y

 Do you want to center any level-2 predictors? Y
 (Enter 0 for no centering, 2 for grand-mean)
 How do you want to center    MSESC? 2

                        ADDITIONAL PROGRAM FEATURES

Select the level-2 variables that you might consider for
inclusion as predictors in subsequent models.
 The choices are:
 For    MSESC enter  1

Which level-2 variables to model INTRCPT1?
 Level-2 variable? (Enter 0 to end)  0
Do you want to constrain any (more) of the gammas? N
Do you wish to use any of the optional hypothesis testing procedures? N
Do you want to do a latent variable regression? Y
Setting method of estimation to full.

Enter o for outcome, p for predictor, or i to ignore
How do you want to model INTRCPT1? P


                      OUTPUT SPECIFICATION

Do you want a level-1 residual file? Y

Enter additional variables to go in residual file
The choices are:
For    MALE enter  1   For PPED enter  2   For REP1 enter 3

Level-1 variable? (Enter 0 to end)  1
Level-1 variable? (Enter 0 to end)  2
Level-1 variable? (Enter 0 to end)  3

Enter additional variables to go in residual file
The choices are:
Level-1 variable? (Enter 0 to end)  1
For   MSESC enter 1

Level-2 variable? (Enter 0 to end)   1

Do you want a level-2 residual file? Y
```

```
Enter additional variables to go in residual file
The choices are:
Level-1 variable? (Enter 0 to end)  1
For   MSESC enter 1

Level-2 variable? (Enter 0 to end)   1

Enter type of stat package you will use:
   for SYSTAT       enter 1
   for SAS          enter 2
   for SPSS         enter 3
   for Stata        enter 4
   for Free Format  enter 5
Type? 3

Do you want to see OLS estimates for all of the level-2 units? N
 Enter a problem title: Bernoulli output, Thailand data
 Enter name of output file: THAIBERN.OUT

                        MACRO ITERATION 1

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -2.400265E+003
The value of the likelihood function at iteration 2 = -2.399651E+003
The value of the likelihood function at iteration 3 = -2.399620E+003
The value of the likelihood function at iteration 4 = -2.399614E+003
The value of the likelihood function at iteration 5 = -2.399612E+003
The value of the likelihood function at iteration 6 = -2.399612E+003
The value of the likelihood function at iteration 7 = -2.399612E+003
```

Macro iteration number 1 has converged after six micro iterations. This macro iteration actually computes the linear-model estimates (using the identity link function as if the level-1 errors were assumed normal).

These results are then transformed and input to start macro iteration 2, which is, in fact, the first non-linear iteration.

```
MACRO ITERATION 2

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.067218E+004
The value of the likelihood function at iteration 2 = -1.013726E+004
The value of the likelihood function at iteration 3 = -1.011008E+004
The value of the likelihood function at iteration 4 = -1.010428E+004
The value of the likelihood function at iteration 5 = -1.010265E+004
The value of the likelihood function at iteration 6 = -1.010193E+004
The value of the likelihood function at iteration 7 = -1.010188E+004
The value of the likelihood function at iteration 8 = -1.010188E+004
The value of the likelihood function at iteration 9 = -1.010187E+004
The value of the likelihood function at iteration 10 = -1.010187E+004
The value of the likelihood function at iteration 11 = -1.010187E+004
The value of the likelihood function at iteration 12 = -1.010187E+004
```

Macro interaction 2, the first non-linear macro iteration, converged after twelve micro iterations.

```
                         MACRO ITERATION 3

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -9.954836E+003
The value of the likelihood function at iteration 2 = -9.954596E+003
The value of the likelihood function at iteration 3 = -9.954567E+003
The value of the likelihood function at iteration 4 = -9.954558E+003
The value of the likelihood function at iteration 5 = -9.954555E+003
The value of the likelihood function at iteration 6 = -9.954554E+003
The value of the likelihood function at iteration 7 = -9.954553E+003

                         MACRO ITERATION 4

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000019E+004
The value of the likelihood function at iteration 2 = -1.000018E+004
The value of the likelihood function at iteration 3 = -1.000018E+004
The value of the likelihood function at iteration 4 = -1.000017E+004
The value of the likelihood function at iteration 5 = -1.000017E+004
The value of the likelihood function at iteration 6 = -1.000017E+004
The value of the likelihood function at iteration 7 = -1.000017E+004

                         MACRO ITERATION 5

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000347E+004
The value of the likelihood function at iteration 2 = -1.000347E+004
The value of the likelihood function at iteration 3 = -1.000347E+004

                         MACRO ITERATION 6

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000375E+004
The value of the likelihood function at iteration 2 = -1.000375E+004

                             MACRO ITERATION 7

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000375E+004
The value of the likelihood function at iteration 2 = -1.000375E+004
```

Note that macro iteration 7 converged with just 2 micro iterations. Also, the change in parameter estimates between macro iterations 6 and 7 was found negligible (less than the criterion for convergence) so that macro iteration 8 was the final "unit-specific" macro iteration. One final "population average" iteration is computed, and screen output for that is given below.

```
                         MACRO ITERATION 8

Starting values computed.  Iterations begun.
Should you wish to terminate the iterations prior to convergence, enter cntl-c
The value of the likelihood function at iteration 1 = -1.000374E+004
The value of the likelihood function at iteration 2 = -1.000374E+004
```

Thus concludes the interactive terminal session. See Section 8.2 for an annotated output for this run.

The interactive session annotated above produced the following command file (NEWCMD.HLM).

```
#This command file was run with thaiugr.mdm
STOPMICRO:0.0000010000
STOPMACRO:0.0001000000
MACROIT:25
MICROIT:20
NONLIN:BERNOULLI
LAPLACE:n,50
LAPLACE8:n,50
LEVEL1:REP1=INTRCPT1+MALE+PPED+RANDOM
LEVEL2:INTRCPT1=INTRCPT2+MSESC,2+RANDOM/
LEVEL2:MALE=INTRCPT2/
LEVEL2:PPED=INTRCPT2/
LEVEL1WEIGHT:NONE
LEVEL2WEIGHT:NONE
RESFILTYPE:SPSS
RESFIL1:Y/MALE,PPED,REP1/MSESC
RESFIL1NAME:resfil1.sav
RESFIL2:Y/MSESC
RESFIL2NAME:resfil2.sav
HETEROL1VAR:n
ACCEL:5
LVR:P
LEV1OLS:10
MLF:y
HYPOTH:n
FIXSIGMA2:1.000000
FIXTAU:3
CONSTRAIN:N
OUTPUT:n
FULLOUTPUT:Y
TITLE:Bernoulli output, Thailand data
```

If one types at the system prompt:

```
HLM2 THAIUGRP.MDM NEWCMD.HLM
```

the output above would be reproduced. It is a good idea to rename the NEWCMD.HLM file if it is to be edited and re-used. Each execution of the program will produce a NEWCMD.MLM file that will overwrite the old one.

Note that the "NEWCMD.HLM" file above is similar to the same file produced by a linear-model analysis, with the addition of the following lines:

```
STOPMICRO:0.000010                    (default convergence criterion for micro iterations)
STOPMACRO:0.000100                    (default convergence criterion for micro iterations)
MACROIT:25                                           (maximum number if macro iterations)
MICROIT:20                     (maximum number if micro iterations per macro iteration)
NONLIN:BERNOULLI                                                (type of non-linear model)
```

See Tables A.1 and B.1 for a description of the keywords and options.

E Using HMLM in Interactive and Batch Mode

This appendix describes prompts and commands for creating MDM files and executing analyses based on the MDM files. References are made to appropriate sections in the manual where the procedures are described in greater details. To start HMLM or HMLM2, type HMLM or HMLM2 at the system prompt.

E.1 Constructing an MDM file

The procedure for MDM creation is similar to the one for MDM for HLM2. The only difference is that the user will be prompted with questions regarding the number of occasions contained in the data and which the indicator variables. To create a MDM file using the NYS data sets described in Section 10.1.1, for example, HMLM will display the following prompts to request the needed information:

```
How many occasions are contained in the data? 5

Please select the 5 indicator variables:
Is    ATTIT     an indicator variable?  N
Is      AGE     an indicator variable?  N
Is    AGE11     an indicator variable?  N
Is    AGE13     an indicator variable?  N
Is   AGE11S     an indicator variable?  N
Is   AGE13S     an indicator variable?  N
Is     IND1     an indicator variable?  Y
Is     IND2     an indicator variable?  Y
Is     IND3     an indicator variable?  Y
Is     IND4     an indicator variable?  Y
Is     IND5     an indicator variable?  Y
```

E.2 Executing analyses based on MDM files

The procedure for executing analyses based on MDM files is similar to the one based on MDM files. A major difference is that only coefficients associate with variables that are invariant across all level-1 units, *i.e.*, their values do not vary across the units, can be specified as random. Otherwise, the coefficients will be automatically set as non-random by the program. The following displays prompts unique to HMLM and HMLM2 for the NYS example described in Section 10.2.

```
C:\HLM> HMLM NYS.MDM

Enter type HMLM analysis:
```

See Chapter 9 for details regarding type of HMLM analysis.

```
  1) Unrestricted
  2) Random effects model with homogeneous level-1 variance
  3) Random effects model with heterogeneous level-1 variance
  4) Random effects model with log-linear model for level-1 variance
  5) Random effects model with first-order autoregressive level-1 variance

type of analysis: 3
```

For choices 2 to 5, the user will be prompted.

```
Do you want to skip the unrestricted iterations? N
```

If "4"(log-linear model for level-1 variance) is chosen, HMLM will ask the user to enter variables to model $sigma^2$*, for example:*

```
Should VAR1 be in C?
```

An interactive session will output a command file NEWCMD.MLM. An example for one of the analyses discussed in Section 10.2 is given below.

```
#This command file was run with nys.mdm
LEVEL1:ATTIT=INTRCPT1+AGE13+AGE13S+RANDOM
LEVEL2:INTRCPT1=INTRCPT2+RANDOM
LEVEL2:AGE13=INTRCPT2+RANDOM
LEVEL2:AGE13S=INTRCPT2+RANDOM
NUMIT:50
STOPVAL:0.0000010000
FIXTAU:3
OUTPUT:nys1.out
FULLOUTPUT:Y
TITLE:HMLM OUPUT, NYS DATA
ACCEL:5
R_E_MODEL:UNRESTRICTED
LVR:N
```

If one types at the system prompt:

```
HMLM NYS.MDM NEWCMD.MLM
```

the result will be the output for a model with an unrestricted covariance structure given in Section 10.3. It is a good idea to rename the NEWCMD.MLM file if it is to be edited and re-used. Each execution of the program will produce a NEWCMD.MLM file that will overwrite the old one.

The following keywords have the same definitions and options in HMLM as in HLM2 (Table A.1)

```
ACCEL      DEVIANCE      DF           FIXTAU
LEVEL1     LEVEL2        GAMMA#       HYPOTH
NUMIT      OUTPUT        PRINTVARIANCE-COVARIANCE
STOPVAL TITLE            FULLOUTPUT   LVR
```

The following keywords are not available in HMLM:

```
LEV1OLS HOMVAR     HETERO1VAR MLF     LEVEL1DELETION     CONSTRAIN
FIXSIGMA2          HYPOTH             LAPLACE            MACROIT
MICROIT NONLIN     PLAUSVALS          RESFIL1            RESFIL1NAME
RESFIL2            RESFIL2NAME        RESFILTYPE         STOPMACRO
STOPMICRO
```

11 E.2.1 Table of keywords and options

Table E.1 Keywords and options unique to the HMLM command file

Keyword	Function	Option	Definition
R_E_MODEL	Choose type of model	UNRESTRICTED	Do only unrestricted iterations
		HOMOL1VAR	Do homogeneous model
		HETL1VAR	Do homogeneous and heterogeneous model
		AUTOREG	Do homogeneous and auto-regressive model
		LOGLIN/var	Do homogeneous and log-linear model
UNRESTRICTED	Possible suppression of unrestricted	Y	Do unrestricted iterations
		N	Don't do unrestricted iterations. Only possible it R_E_MODEL is not UNRESTRICTED

The following keywords have the same definitions and options in HMLM2 as in HLM3 (Table B.1)

ACCEL	DEVIANCE	DF	FIXTAU2	FIXTAU3
LEVEL1	LEVEL2	LEVEL3	GAMMA#	HYPOTH
NUMIT	OUTPUT	PRINTVARIANCE-COVARIANCE		
STOPVAL	TITLE	FULLOUTPUT		LVR
RESFIL3				

The following HLM3 keywords are not available in HMLM2:

LEV1OLS	HOMVAR	LEVEL1DELETION	CONSTRAIN	FIXSIGMA2
HYPOTH	LAPLACE	MACROIT	MICROIT	NONLIN
PLAUSVALS	RESFIL1	RESFIL1NAME	RESFIL2	RESFIL2NAME
RESFIL3	RESFIL3NAME	RESFILTYPE	STOPMACRO	STOPMICRO
LVR-BETA				

12 E.2.2 Table of HMLM2 keywords and options

Table E.1 Keywords and options unique to the HMLM2 command file

Keyword	Function	Option	Definition
R_E_MODEL	Choose type of model	UNRESTRICTED	Do only unrestricted iterations
		HOMOL1VAR	Do homogeneous model
		HETL1VAR	Do homogeneous and heterogeneous model
		AUTOREG	Do homogeneous and auto-regressive model
		LOGLIN/var	Do homogeneous and log-linear model
UNRESTRICTED	Possible suppression of unrestricted	Y	Do unrestricted iterations
		N	Don't do unrestricted iterations. Only possible it R_E_MODEL is not UNRESTRICTED

Note that HMLM and HMLM2 do not allow non-linear outcomes, use of plausible values and multiply-imputed values, constraints of gammas, and they do not write out any residual files.

F Using Special Features in Interactive and Batch Mode

This appendix describes and illustrates how to use the special features in interactive and batch mode to execute analyses. References are made to appropriate sections in the manual where the procedures are described in greater details.

F.1 Example: Latent variable analysis using the National Youth Study data sets

The following interactive session illustrates a latent variable analysis example using the National Youth Study (NYS) data sets. A description of the data files and the model specification can be found in Sections 10.1.1 and 11.1.1.

```
C:\HLM> HMLM NYS.MDM
Enter type HMLM analysis:

  1) Unrestricted
  2) Random effects model with homogeneous level-1 variance
  3) Random effects model with heterogeneous level-1 variance
  4) Random effects model with log-linear model for level-1 variance
  5) Random effects model with first-order autoregressive level-1 variance

type of analysis: 2
```

We select the homogeneous level-1 variance option for this model. Thus, using HLM2 will yield identical results in this case.

```
Do you want to skip the unrestricted iterations? Y

                   SPECIFYING A LEVEL-1 OUTCOME VARIABLE

Please specify a level-1 outcome variable
 The choices are:
 For    ATTIT enter  1    For      AGE enter  2    For    AGE11 enter  3
 For    AGE13 enter  4    For   AGE11S enter  5    For   AGE13S enter  6
 For     IND1 enter  7    For     IND2 enter  8    For     IND3 enter  9
 For     IND4 enter 10    For     IND5 enter 11
What is the outcome variable: 1
```

The outcome is tolerance towards deviant behavior.

```
                         SPECIFYING AN HMLM MODEL
Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
                          For      AGE enter  2    For    AGE11 enter  3
 For    AGE13 enter  4    For   AGE11S enter  5    For   AGE13S enter  6
 For     IND1 enter  7    For     IND2 enter  8    For     IND3 enter  9
 For     IND4 enter 10    For     IND5 enter 11

 level-1 predictor? (Enter 0 to end)  3
 level-1 predictor? (Enter 0 to end)  0
```

AGE11 is the age of participant at a specific time minus 11.

```
Do you want to center any level-1 predictors? N

Do you want to set the level-1 intercept to zero in this analysis? N

Level-2 predictor variable specification

Which level-2 variables do you wish to use?

 The choices are:
 For   FEMALE enter  1    For MINORITY enter  2    For   INCOME enter  3

 Which level-2 predictors to model INTRCPT1?
  Level-2 predictor? (Enter 0 to end)  1
  Level-2 predictor? (Enter 0 to end)  0

 Which level-2 predictors to model    AGE11 slope?
  Level-2 predictor? (Enter 0 to end)  1
  Level-2 predictor? (Enter 0 to end)  0

Do you want to constrain the variances in any of the level-2 random
 effects to zero? N

Do you want to center any level-2 predictors? N

                        ADDITIONAL PROGRAM FEATURES
Do you want to do a latent variable regression? Y

Enter o for outcome, p for predictor, or i to ignore
How do you want to handle INTRCPT1? P
How do you want to handle    AGE11? O
```

INTRCPT1, the level of tolerance at age 11, is used as a predictor to model the outcome, AGE11, the linear growth rate. Note that INTRCPT1 and AGE11 are latent variables, that is, they are free of measurement error.

```
Do you want to specify a multivariate hypothesis for the fixed effects? N

                           OUTPUT SPECIFICATION
How many iterations do you want to do? 50
Enter a problem title: Latent variable regression, NYS Data
Enter name of output file: NYS2.OUT

Computing . . ., please wait
```

Partial output for this analysis is given in Section 11.1.1.

F.2 A latent variable analysis to run regression with missing data

The following interactive session illustrates a latent variable analysis to run regression with missing data with an artificial data set. A description of the data files and the model specification can be found in Section 11.1.2.

```
C:\HLM> HMLM MISSING.MDM
Enter type HMLM analysis:

  1) Unrestricted
  2) Random effects model with homogeneous level-1 variance
```

```
  3) Random effects model with heterogeneous level-1 variance
  4) Random effects model with log-linear model for level-1 variance
  5) Random effects model with first-order autoregressive level-1 variance

type of analysis: 1

                   SPECIFYING A LEVEL-1 OUTCOME VARIABLE

Please specify a level-1 outcome variable

 The choices are:
 For MEASURES enter  1    For     IND1 enter  2    For     IND2 enter  3
 For     IND3 enter  4

What is the outcome variable: 1

                         SPECIFYING AN HMLM MODEL

Level-1 predictor variable specification

Which level-1 predictors do you wish to use?

 The choices are:
 For     IND1 enter  2    For     IND2 enter  3    For     IND3 enter  4

 level-1 predictor? (Enter 0 to end)  1
 That is the outcome variable!
 level-1 predictor? (Enter 0 to end)  2
 level-1 predictor? (Enter 0 to end)  3
 level-1 predictor? (Enter 0 to end)  4

Do you want to center any level-1 predictors? N

Do you want to set the level-1 intercept to zero in this analysis? Y
```

Note that a no-intercept model is formulated (see Section 2.9.6).

```
Level-2 predictor variable specification

Which level-2 variables do you wish to use?

 The choices are:
 For    DUMMY enter  1

 Which level-2 predictors to model     IND1 slope?
  Level-2 predictor? (Enter 0 to end)  0
 Which level-2 predictors to model     IND2 slope?
  Level-2 predictor? (Enter 0 to end)  0

 Which level-2 predictors to model     IND3 slope?
  Level-2 predictor? (Enter 0 to end)  0
```

IND2 and IND3 are selected to predict IND1.

```
Do you want to constrain the variances in any of the level-2 random
 effects to zero? N
```

```
ADDITIONAL PROGRAM FEATURES

Do you want to do a latent variable regression? Y

Enter o for outcome, p for predictor, or i to ignore
How do you want to handle     IND1? O
How do you want to handle     IND2? P
How do you want to handle     IND3? P

Do you want to specify a multivariate hypothesis for the fixed effects? N

                           OUTPUT SPECIFICATION
How many iterations do you want to do? 50
 Enter a problem title: Latent variable analysis, Missing data example
 Enter name of output file: MISSING1.OUT
```

Partial output for this analysis is given in Section 11.1.2.

F.3 Commands to apply HLM to multiply-imputed data

To analyze data with multiply-imputed values for the outcome variable or only one covariate, the user needs to manually add the following line into the command file:

```
PLAUSVALS: VARLIST
```

where VARLIST lists variables containing the multiply-imputed values.

To analyze data with multiply-imputed values for the outcome and/or covariates, the user needs to prepare multiple MDM files. After setting up the multiple MDM files, the user have to submit the command files to HLM2 and HLM3 as many times as the number of multiple MDM files with an extra flag, -MI#, where # is the sequence number, starting from 0. On the last run, you also need the -E flag, (E for estimate).

Suppose there are 4 sets of multiply-imputed data for a two-level model, called MDATA1.MDM, MDATA2.MDM, MDATA3.MDM, and MDATA4.MDM and the command file is ANALYSE.MLM; the following commands need to be typed in at the system prompt:

```
HLM2 -MI0 MDATA1.MDM ANALYSE.MLM
HLM2 -MI1 MDATA2.MDM ANALYSE.MLM
HLM2 -MI2 MDATA3.MDM ANALYSE.MLM
HLM2 -MI3 -E MDATA.4MDM ANALYSE.MLM
```

F.4 Commands to apply HLM2 to create a Spatial model

This option can only be invoked if the spatial dependence information was added when the mdm file was created. Then, add the following command line to the HLM file to accommodate spatial dependence:

```
dospatialcorrelation:y
```

G Using HCM2 in Interactive and Batch Mode

This appendix describes and illustrates how to use HCM2 in interactive construct MDM files, and in both interactive and batch mode to execute analyses based on the MDM file. It also lists and defines command keywords and options unique to HCM2. References are made to appropriate sections in the manual where the procedures are described in greater details. In the next Section, we show the construction of an MDM file using the educational attainment data as described in Chapter 13.

G.1 Using HCM2 in interactive mode

13 G.1.1 Example: constructing an MDM file for the educational attainment data using SPSS file input

```
C:\HLM>HCM2                         (type the program name at the system prompt to start)

Will you be starting with raw data? Y
Enter type of raw data:
   for ASCII input                  enter 1
   for SYSTAT .SYS file             enter 2
   for SAS V5 transport file        enter 3
   for SPSS file (UNIX or windows)  enter 4
   for STATA .dta file              enter 5
   for anything DBMSCOPY reads      enter 6
   for anything Stat/Transfer reads enter 7
Type? 4
```

The "anything DBMSCOPY reads" prompt is only present on PC versions.

```
Input name of level-1 file: ATTAINW.SAV
Input name of row file: ATTAINR.SAV
Input name of column file: ATTAINCO.SAV
```

See Section 13.1.2 for a description of variables in the data files.

```
 The available level-1 variables are:
 For  NEIGHID  enter  1    For    SCHID  enter  2    For   ATTAIN  enter  3
 For    P7VRQ  enter  4    For   P7READ  enter  5    For   DADOCC  enter  6
 For DADUNEMP  enter  7    For    DADED  enter  8    For    MOMED  enter  9
 For     MALE  enter 10
 What variable is the row ID? 1
 What variable is the column ID? 2
```

Note there are two linking ID's in the level-1 or within-cell file.

```
 Please specify level-1 variable # 1 (enter 0 to end): 3
 Please specify level-1 variable # 2 (enter 0 to end): 4
 Please specify level-1 variable # 3 (enter 0 to end): 5
 Please specify level-1 variable # 4 (enter 0 to end): 6
 Please specify level-1 variable # 5 (enter 0 to end): 7
 Please specify level-1 variable # 6 (enter 0 to end): 8
 Please specify level-1 variable # 7 (enter 0 to end): 9
```

```
Please specify level-1 variable # 8 (enter 0 to end): 10

The available row-level variables are:

 For  NEIGHID  enter  1    For  DEPRIVE  enter  2
 What variable is the row ID? 1
```

Note there is one row ID the level-2 row factor file.

```
 Please specify row-level variable # 1 (enter 0 to end): 2

 The available column-level variables are:
 For    SCHID  enter  1    For    DUMMY  enter  2
 What variable is the column ID? 1
```

Note there is one column ID the level-2 column factor file.

```
 Please specify column-level variable # 1 (enter 0 to end): 2

Are there missing data in the level-1 file? y

Enter name of MDM file: ATTAIN.MDM
```

HCM2 save send the descriptive statistics of variables for each file to the screen. It is important to examine these carefully to ensure that no errors were made. The program will save these statistics in a file name HCM2MDM.STS.

LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
ATTAIN	2310	0.09	1.00	-1.33	2.42
P7VRQ	2310	0.51	10.65	-27.03	42.97
P7READ	2310	-0.04	13.89	-31.87	28.13
DADOCC	2310	-0.46	11.78	-23.45	29.23
DADUNEMP	2310	0.11	0.31	0.00	1.00
DADED	2310	0.22	0.41	0.00	1.00
MOMED	2310	0.25	0.43	0.00	1.00
MALE	2310	0.48	0.50	0.00	1.00

ROW LEVEL DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
DEPRIVE	524	0.04	0.62	-1.08	2.96

COLUMN LEVEL DESCRIPTIVE STATISTICS

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
DUMMY	17	2.41	1.18	1.00	4.00

```
 2310 level-1 records have been processed
 524 row-level records have been processed
 17 column-level records have been processed
```

14 G.1.2 Example: Executing an unconditional model analysis using ATTAIN.MDM

```
C:\HLM> HCM2 ATTAIN.MDM

Do you want to do a non-linear analysis? N

                   SPECIFYING A LEVEL-1 OUTCOME VARIABLE

Please specify a level-1 outcome variable
 The choices are:
 For   ATTAIN enter  1    For    P7VRQ enter  2    For   P7READ enter  3
 For   DADOCC enter  4    For DADUNEMP enter  5    For    DADED enter  6
 For    MOMED enter  7    For     MALE enter  8
What is the outcome variable: 1
```

We shall model educational attainment with an unconditional model and specific the residual row, column, and cell-specific effects as random. See Section 13.2 .

```
                         SPECIFYING AN HCM2 MODEL

Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
                          For    P7VRQ enter  2    For   P7READ enter  3
 For   DADOCC enter  4    For DADUNEMP enter  5    For    DADED enter  6
 For    MOMED enter  7    For     MALE enter  8
level-1 predictor? (Enter 0 to end)  0

Do you want to set the level-1 intercept to zero in this analysis? N


Level-1/row predictor variable specification

Which row variables do you wish to use?


 The choices are:
 For  DEPRIVE enter  1

Which row-level predictor to model INTRCPT1, P0?
 Row-level predictor? (Enter 0 to end)  0


Column-level predictor variable specification

Which column-level variables do you wish to use?

 The choices are:
 For    DUMMY enter  1

Which column-level predictor to model INTRCPT1, P0?
 Column-level predictor? (Enter 0 to end)  0


Do you want to constrain the variances in any of the row-level random
 effect to zero? N


Do you want to constrain the variances in any of the column-level random
 effect to zero? N
```

```
Enter type of deflection:
 for independent(default) enter 1
 for cumulative           enter 2
Type? 1
```

Select 2 if to define a cumulative effect model.

```
                          OUTPUT SPECIFICATION

Do you want a row-level residual file? N
Do you want a column-level residual file? N
How many iterations do you want to do? 100
Enter a problem title: UNCONDITIONAL MODEL
Enter name of output file: ATTAIN1.TXT

Computing . . ., please wait
The value of the likelihood function at iteration 1 = -3.208601E+003
The value of the likelihood function at iteration 2 = -3.207263E+003
The value of the likelihood function at iteration 3 = -3.205187E+003
The value of the likelihood function at iteration 4 = -3.201693E+003
The value of the likelihood function at iteration 5 = -3.196031E+003
The value of the likelihood function at iteration 6 = -3.188714E+003
The value of the likelihood function at iteration 7 = -3.182922E+003
The value of the likelihood function at iteration 8 = -3.180439E+003
The value of the likelihood function at iteration 9 = -3.179676E+003
The value of the likelihood function at iteration 10 = -3.179379E+003
The value of the likelihood function at iteration 11 = -3.179212E+003
The value of the likelihood function at iteration 12 = -3.179104E+003
The value of the likelihood function at iteration 13 = -3.179032E+003
The value of the likelihood function at iteration 14 = -3.178984E+003
The value of the likelihood function at iteration 15 = -3.178881E+003
The value of the likelihood function at iteration 16 = -3.178879E+003
The value of the likelihood function at iteration 17 = -3.178878E+003
The value of the likelihood function at iteration 18 = -3.178877E+003
The value of the likelihood function at iteration 19 = -3.178876E+003
The value of the likelihood function at iteration 20 = -3.178874E+003
The value of the likelihood function at iteration 21 = -3.178874E+003
```

See Section 13.2 for a discussion of the results of this unconditional model.

15 G.1.3 Example: Executing a conditional model analysis using ATTAIN.MDM

```
C:\HLM> HCM2 ATTAIN.MDM

Do you want to do a non-linear analysis? N

                  SPECIFYING A LEVEL-1 OUTCOME VARIABLE

Please specify a level-1 outcome variable
 The choices are:
 For   ATTAIN enter  1    For    P7VRQ enter  2    For   P7READ enter  3
 For   DADOCC enter  4    For DADUNEMP enter  5    For    DADED enter  6
 For    MOMED enter  7    For     MALE enter  8
What is the outcome variable: 1
```

We shall model educational attainment with all the level-1 predictor variables. All the level-1 coefficients associated with the predictors are fixed. See Section 13.3.

```
SPECIFYING AN HCM2 MODEL

Level-1 predictor variable specification

Which level-1 predictors do you wish to use?
 The choices are:
                          For    P7VRQ enter  2    For   P7READ enter  3
 For   DADOCC enter  4    For DADUNEMP enter  5    For    DADED enter  6
 For    MOMED enter  7    For     MALE enter  8
level-1 predictor? (Enter 0 to end)  2
level-1 predictor? (Enter 0 to end)  3
level-1 predictor? (Enter 0 to end)  4
level-1 predictor? (Enter 0 to end)  5
level-1 predictor? (Enter 0 to end)  6
level-1 predictor? (Enter 0 to end)  7
level-1 predictor? (Enter 0 to end)  8
level-1 predictor? (Enter 0 to end)  0

Do you want to center any level-1 predictors? y
 Enter 0 for no centering, 2 for grand-mean
How do you want to center    P7VRQ? 2
How do you want to center   P7READ? 2
How do you want to center   DADOCC? 2
How do you want to center DADUNEMP? 2
How do you want to center    DADED? 2
How do you want to center    MOMED? 2
How do you want to center     MALE? 2

Do you want to set the level-1 intercept to zero in this analysis? N

Level-1/row predictor variable specification

Which row variables do you wish to use?

 The choices are:
 For  DEPRIVE enter  1
```

We shall use DEPRIVE to model the level-1 intercept.

```
Which row-level predictor to model INTRCPT1, P0?
 Row-level predictor? (Enter 0 to end)  1
Which row-level predictor to model    P7VRQ, P2 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model   P7READ, P3 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model   DADOCC, P4 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model DADUNEMP, P5 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model    DADED, P6 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model    MOMED, P7 slope?
 Row-level predictor? (Enter 0 to end)  0
Which row-level predictor to model     MALE, P8 slope?
 Row-level predictor? (Enter 0 to end)  0


Column-level predictor variable specification

Which column-level variables do you wish to use?
```

```
 The choices are:

 For    DUMMY enter  1

Which column-level predictor to model INTRCPT1, P0?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model    P7VRQ, P2 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model   P7READ, P3 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model   DADOCC, P4 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model DADUNEMP, P5 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model    DADED, P6 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model    MOMED, P7 slope?
 Column-level predictor? (Enter 0 to end)  0
Which column-level predictor to model     MALE, P8 slope?
 Column-level predictor? (Enter 0 to end)  0

Do you want to center any row-level predictors? y
 Enter 0 for no centering, 2 for grand-mean
How do you want to center  DEPRIVE? 2


Do you want to constrain the variances in any of the row-level random
 effect to zero? y
Do you want to fix INTRCPT1/ICPTROW? N
Do you want to fix P7VRQ/ICPTROW? Y
Do you want to fix P7READ/ICPTROW? Y
Do you want to fix DADOCC/ICPTROW? Y
Do you want to fix DADUNEMP/ICPTROW? Y
Do you want to fix DADED/ICPTROW? Y
Do you want to fix MOMED/ICPTROW? Y
Do you want to fix MALE/ICPTROW? Y


Do you want to constrain the variances in any of the column-level random
 effect to zero? Y
```

We shall treat the association between social deprivation and educational attainment as fixed across all schools. See Section 13.2.

```
Do you want to fix INTRCPT1/DEPRIVE? Y
Do you want to fix INTRCPT1/ICPTCOL? N
Do you want to fix P7VRQ/ICPTCOL? Y
Do you want to fix P7READ/ICPTCOL? Y
Do you want to fix DADOCC/ICPTCOL? Y
Do you want to fix DADUNEMP/ICPTCOL? Y
Do you want to fix DADED/ICPTCOL? Y
Do you want to fix MOMED/ICPTCOL? Y
Do you want to fix MALE/ICPTCOL? Y
 Enter type of deflection:
 for independent(default) enter 1
 for cumulative           enter 2
Type? 1
```

```
                              OUTPUT SPECIFICATION

Do you want a row-level residual file? N

Do you want a column-level residual file? N

How many iterations do you want to do? 100
Enter a problem title: CONDITIONAL MODEL WITH THE EFFECT ASSOCIATED WITH A ROW-SPECIFIC
PREDICTOR FIXED
Enter name of output file: ATTAIN2.TXT

Computing . . ., please wait

The value of the likelihood function at iteration 1 = -2.391226E+003
The value of the likelihood function at iteration 2 = -2.390450E+003
The value of the likelihood function at iteration 3 = -2.390158E+003
The value of the likelihood function at iteration 4 = -2.389892E+003
The value of the likelihood function at iteration 5 = -2.389646E+003
…
The value of the likelihood function at iteration 28 = -2.384804E+003
The value of the likelihood function at iteration 29 = -2.384803E+003
The value of the likelihood function at iteration 30 = -2.384803E+003
The value of the likelihood function at iteration 31 = -2.384802E+003
The value of the likelihood function at iteration 32 = -2.384802E+003
The value of the likelihood function at iteration 33 = -2.384802E+003
The value of the likelihood function at iteration 34 = -2.384802E+003
```

See Section 13.2 for a discussion of the results of this conditional model.

G.2 Using HCM2 in batch mode

The interactive session in G.1.1 produced the following command file, NEWCMD.HLM.

```
#WHLM CMD FILE FOR C:\HLM\ATTAIN.MDM
NUMIT:100
STOPVAL:0.0000010000
LEVEL1:ATTAIN=INTRCPT1+RANDOM
ROWCOL:INTRCPT1=THETA+RANDOMB+RANDOMC
FIXTAU:3
FIXDELTA:3
ACCEL:5
DEFLECTION:INDEPENDENT
TITLE:UNCONDITIONAL MODEL
OUTPUT:C:\HLM\ATTAIN1.TXT
FULLOUTPUT:N
```

If one types at the system prompt:

```
HCM2 ATTAIN.MDM NEWCMD.HLM
```

the result will be the output for the unconditional model. Note that each execution of the program will produce a NEWCMD.HLM file that will overwrite the old one.

For the conditional model, the command file is

```
#WHLM CMD FILE FOR C:\HLM\ATTAIN.MDM
NUMIT:100
STOPVAL:0.0000010000
LEVEL1:ATTAIN=INTRCPT1+P7VRQ,2+P7READ,2+DADOCC,2+DADUNEMP,2+DADED,2+MOMED,2+
```

```
MALE,2+RANDOM
ROWCOL:INTRCPT1=THETA+DEPRIVE(FIXED),2+RANDOMB+RANDOMC
ROWCOL:P7VRQ=THETA
ROWCOL:P7READ=THETA
ROWCOL:DADOCC=THETA
ROWCOL:DADUNEMP=THETA
ROWCOL:DADED=THETA
ROWCOL:MOMED=THETA
ROWCOL:MALE=THETA
FIXTAU:3
FIXDELTA:3
ACCEL:5
DEFLECTION:INDEPENDENT
TITLE:CONDITIONAL MODEL, WITH SOCIAL DEPRIVATION EFFECT FIXED
OUTPUT:C:\HLM\ATTAIN2.TXT
FULLOUTPUT:N
```

The following keywords in the above command files have the same definition and options in HCM2 as in the other modules (*e.g.* Tables A.1 and B.1)

```
ACCEL      FULLOUTPUT    FIXTAU     NONLIN     NUMIT      OUTPUT     STOPVAL    TITLE
FIXSIGMA2  STOPMICRO     STOPMACRO  DEVIANCE   DF         GAMMA
```

Had we requested residual level-1, and row and column files during the interaction session, the command files would contain the following additional command lines specifying the type (SPSS system file) and the names for each of the files (RESFIL1.SAV, RESROW.SAV, and RESCOL.SAV):

```
RESFILTYPE:SPSS
RESFIL1NAME:RESFIL1.SAV
RESFIL1:Y

RESROWNAME:RESROW.SAV
RESROW:Y
RESCOLNAME:RESCOL.SAV
RESCOL:Y
```

Table G.1 Keywords and options unique for HCM2 command file

Keyword	Function	Option	Definition
LEVEL1	Level-1 or within-cell model specification	INTRCPT1	Level-1 intercept
		+VARNAME	Level-1 predictor (no centering)
		+VARNAME,1	Level-1 predictor (group-mean centering)
		+VARNAME,2	Level-1 predictor (grand-mean centering)
ROWCOL: INTRCPT1 or L-1 VARNAME	Level-2 or between-cell model specification	THETA	Level-2 intercept
		+VARNAME(FIXED)	+ Level-2 predictor (fixed and grand-mean centering)
		+VARNAME(FIXED),2	+ Level-2 predictor (fixed and no centering)
		+VARNAME(RANDOM),2	+ Level-2 predictor (random and grand-mean centering)
		+VARNAME(RANDOM)	+ Level-2 predictor (random and no centering)
		+RANDOMB	+ Random main effect of the row factor
		+RANDOMC	+ Random main effect of the column factor
DEFLECTION	Define the use of a cumulative effect model	1	Independent
		2	Cumulative

H Using HCM3 in Batch Mode

Unlike the older modules (HLM2, HLM3, etc.), HCM3 does not have interactive modes to create the MDM or specify a model. If the windows interface is not available, these file must be created with an ASCII editor and submit them to obtain results.

H.1 Example: Creating an HCM3 MDM file from raw data

The first thing that needs creating is an mdm template file (usually suffixed with .mdmt), which tells HCM3 how to read the raw data. Here is the MDMT file from section 15.1.1:

```
#HCM3 MDM CREATION TEMPLATE
rawdattype:spss
l1fname:growth.sav
rowfname:student.sav
colfname:teacher.sav
clusfname:school.sav
l1missing:n
timeofdeletion:now
mdmname:growth.mdm
*begin l1vars
rowid:STUDID
colid:TCHRID
clusid:SCHLID
MATH
YEAR
G4D1
G4D21
G5D22
TWOWAY
*end l1vars
*begin rowvars
rowid:STUDID
DUMMY
*end rowvars
*begin colvars
colid:TCHRID
clusid:SCHLID
DUMMY
*end colvars
*begin clusvars
clusid:SCHLID
DUMMY
*end clusvars
```

The file is broken into two sections. The first is to declare the filenames of the raw data and other characteristics of the MDM file to be made, the second chooses the variables to be included at the various levels. Below is the first part with explanation in parentheses:

```
rawdattype:spss                    (This declares the type of input data. Possible values
are                                   spss, sas (version 5 transport file), stata, and ascii)
l1fname:growth.sav                 (The next four lines declare the names and locations of
                                     the four input files; level-1, row, column, and cluster,
                                                                              respectively.)
rowfname:student.sav
colfname:teacher.sav
clusfname:school.sav
l1missing:n                   (This declares whether or not there are missing data at level-
                               1. Possible values are n for not missing, or y for missing
                                                                            data present.)
```

```
timeofdeletion:now          (This may be n[ow] , where all level-1 cases with missing data
                                   on selected variables will be deleted, or a[nalysis] where
                                     the missing data will be left in and deleted at run-time
                                                                  based on the model specified.)

mdmname:growth.mdm                                          (Specifies the name of the mdm file.)
```

The second part of the mdmt file specifies which variables are ID variables, and which ones go into the mdm file as possible analysis variables. The structure looks like this:

```
*begin l1vars
rowid:STUDID
colid:TCHRID
clusid:SCHLID
MATH
YEAR
G4D1
G4D21
G5D22
TWOWAY
*end l1vars
*begin rowvars
rowid:STUDID
DUMMY
*end rowvars
*begin colvars
colid:TCHRID
clusid:SCHLID
DUMMY
*end colvars
*begin clusvars
clusid:SCHLID
DUMMY
*end clusvars
```

The ID s must be specified in the order shown, and must all be of the same type, either numeric (preferable) or alphanumeric(not advised). The level-1 file needs to be sorted primarily by row ID, secondarily by cluster ID, and thirdly and the column level. The row file should be sorted by row ID. The column file should be sorted by column ID within cluster ID, and the cluster file sorted by cluster ID.

Once the mdmt file is created, the file must be submitted to HCM3:

```
C:\HLM> HCM3 –r growth.mdmt
```

The results on the screen should then be examined to make sure the data were read correctly. These descriptive statistics will also be contained in a file named HCM3MDM.STS.

H.2 Example: Creating an HCM3 HLM file and running the model

The next step is to create a file that specifies the desired model. (This is usually suffixed with a .hlm) For example, we will use the model shown in section 15.2.

```
nonlin:n
numit:100
stopval:0.0000010000
level1:MATH=INTRCPT1+YEAR+G4D1+G4D21+G5D22+TWOWAY+RANDOM
rowcol:INTRCPT1=theta+randomb+randomc
clus:theta=ICPTCLUS+randomd
```

```
rowcol:YEAR=theta+randomb+randomc
clus:theta=ICPTCLUS+randomd
rowcol:G4D1=theta
clus:theta=ICPTCLUS
rowcol:G4D21=theta
clus:theta=ICPTCLUS
rowcol:G5D22=theta
clus:theta=ICPTCLUS
rowcol:TWOWAY=theta
clus:theta=ICPTCLUS
fixtau:3
fixdelta1:3
fixdelta2:3
accel:5
level1weight:none
rowweight:none
clusterweight:none
hypoth:n
resfiltype:spss
resfil1:n
resfil1fname:resfil1.sav
resrow:n
resrowfname:resrow.sav
rescol:n
rescolfname:rescol.sav
resclus:n
resclusfname:resclus.sav
deflection:cumulative
title:Unweighted model
output:docdef1.html
fulloutput:n
```

The above is very similar to an HCM2 model file, with the exception of the model specification at the top where an extra level is shown. Here is the model part that better demonstrates the nested nature of the model specification (the shown indentation will not run):

```
level1:MATH=INTRCPT1+YEAR+G4D1+G4D21+G5D22+TWOWAY+RANDOM
rowcol:INTRCPT1=theta+randomb+randomc
clus:theta=ICPTCLUS+randomd
rowcol:YEAR=theta+randomb+randomc
clus:theta=ICPTCLUS+randomd
rowcol:G4D1=theta
clus:theta=ICPTCLUS
rowcol:G4D21=theta
clus:theta=ICPTCLUS
rowcol:G5D22=theta
clus:theta=ICPTCLUS
rowcol:TWOWAY=theta
```

The rule here is that for every variable in the level1: line, there needs to be a rowcol: line in the same order as the variables are declared in the level1: line. For each variable in a rowcol: line, there must be clus: line. Also, note that instead of some form of INTRCPT, HCM3 uses the special name theta to denote the intercept in the rowcol: lines.

Note that a level-1 variable may vary at the row (randomb), column(randomc), or cluster(randomd) level. A row variable may vary at either the column or cluster levels. A column variable may vary at the row or cluster level, and cluster variable may vary at the row level. This can make for a very complicated model specification. For example, consider this skeleton section for just the level-1 intercept where rowvar, colvar and clusvar are arbitrary row, column, and cluster level variables:

```
level1:outcome=intrcpt1+random
```

```
rowcol:intrcpt1=intrcpt1+rowvar(random)+colvar(random)+rvar*colvar+randomb+randomc
  clus:intrcpt1=theta+clusvar(randomb)+randomd
  clus:rowvar=theta+clusvar(randomb)
  clus:colvar=theta+clusvar(fixed)+randomd
  clus:rowvar*colvar=icptclus+clusvar[norandom]+randomd
```

In the rowcol: line, there are four variables: the intercept, an arbitrary row variable (rowvar) an arbitrary column variable (colvar), and a row by column interaction term rowvar*colvar. The random in parentheses tells the program let the variables vary. If the variable should be fixed, substitute the word fixed instead. The interaction term cannot vary, so there is no way to specify this. Finally, the randomb and randomc at the end of the line tells the program to let the level-1 intercept vary across rows and columns respectively. Either +randomb and +randomc can be omitted if the level-1 variable should not be allowed to vary across rows or columns respectively.

The clus: lines all take on the same basic form. In this example, all the variables are modeled with a cluster intercept, which is random at level-3 except for the variable rowvar, where the +randomd is omitted. In the clus:colvar line, clusvar is fixed at the row level, where in the previous two lines it is allowed to vary. In the row/column interaction line, clusvar has no random/fixed declaration because this term cannot vary at any level.

Assuming that the above file is named growth.hlm, then the following command should be run:

```
C:\HLM> HCM3 GROWTH.MDM GROWTH.HLM
```

The following keywords in the above command files have the same definition and options in HCM2 as in the other modules (*e.g.* Tables A.1 and B.1)

```
ACCEL      FULLOUTPUT    FIXTAU     NONLIN     NUMIT      OUTPUT     STOPVAL    TITLE
FIXSIGMA2  STOPMICRO     STOPMACRO  DEVIANCE   DF         TITLE      GAMMA      RESFILTYPE
```

Table H.1 Keywords and options unique for HCM3 command file

Keyword	Function	Option	Definition
LEVEL1	Level-1 or within-cell model specification	INTRCPT1 +VARNAME +VARNAME,2 +VARNAME(FIXED),2	Level-1 intercept Level-1 predictor (no centering) Level-1 predictor (grand-mean centering) + Level-2 predictor (fixed and grand-mean centering)
ROWCOL: INTRCPT1 or ROW/COLUMN VARNAME	Level-2 or between-cell model specification	+VARNAME(FIXED) +VARNAME(RANDOM),2 +VARNAME(RANDOM) +RANDOMB +RANDOMC	+ Level-2 predictor (fixed and no centering) + Level-2 predictor (random and grand-mean centering) + Level-2 predictor (random and no centering) + Random main effect of the row factor + Random main effect of the column factor

Table H.1 Keywords and options unique for HCM3 command file (continued)

Keyword	Function	Option	Definition

`CLUS:`	Level-3 model specification	`THETA`
		`+VARNAME(FIXED)`
		`+VARNAME(RANDOM)`
		`+VARNAME(FIXED),2`
			Level-2 intercept
			+cluster-level predictor (fixed and no centering)
			+cluster-level predictor(random and no centering)
			+cluster-level predictor(fixed and grand-mean centering)

`DEFLECTION`	Define the use of a cumulative effect model	`independent` `cumulative`	Independent Cumulative

I Using HLMHCM in Batch Mode

Unlike the older modules (HLM2, HLM3, etc.), HLMHCM does not have interactive modes to create the MDM or specify a model. If the windows interface is not available, these file must be created with an ASCII editor and submit them to obtain results.

I.1 Example: Creating an HLMHCM MDM file from raw data

The first thing that needs creating is an mdm template file (usually suffixed with .mdmt), which tells HLMHCM how to read the raw data. Here is the MDMT file from section 17.1.1:

```
#HLMHCM MDM CREATION TEMPLATE
rawdattype:spss
l1fname:growth.sav
l2fname:student.sav
rowfname:school.sav
colfname:neigh.sav
l1missing:n
timeofdeletion:now
mdmname:growth.mdm
*begin l1vars
level2id:STUDID
AGE8
MATH
*end l1vars
*begin l2vars
level2id:STUDID
rowid:SCHID
colid:NEIGHID
FEMALE
BLACK
HISPANIC
*end l2vars
*begin rowvars
rowid:SCHID
SCHPOV
*end rowvars
*begin colvars
colid:NEIGHID
DISADV
*end colvars
```

The file is broken into two sections. The first is to declare the filenames of the raw data and other characteristics of the MDM file to be made, the second chooses the variables to be included at the various levels. Below is the first part with explanation in parentheses:

```
rawdattype:spss                 (This declares the type of input data. Possible values
                                 are spss, sas (version 5 transport file), stata, and
                                                                              ascii)
l1fname:growth.sav               (The next four lines declare the names and locations
                                   of the four input files; level-1, row, column, and
                                                               cluster, respectively.)
l2fname:student.sav
rowfname:school.sav
colfname:neigh.sav
l1missing:n                       (This declares whether or not there are missing data
                                 at level-1. Possible values are n for not missing,
                                                    or y for missing data present.)
timeofdeletion:now                  (This may be n[ow] , where all level-1 cases with
                                          missing data on selected variables will be
```

deleted, or a[nalysis] where the missing data will be left in and deleted at run-time based on the model specified.)

```
mdmname:growth.mdm                              (Specifies the name of the mdm file.)
```

The second part of the mdmt file specifies which variables are ID variables, and which ones go into the mdm file as possible analysis variables. The structure looks like this:

```
*begin l1vars
level2id:STUDID
(list of level-1 variables, one per line)
*end l1vars
*begin l2vars
level2id:STUDID
rowid:SCHID
colid:NEIGHID
(list of level-2 variables, one per line)
*end l2vars
*begin rowvars
rowid:SCHID
(list of row variables, one per line)
*end rowvars
*begin colvars
colid:NEIGHID
(list of column variables, one per line)
*end colvars
```

The IDs must be specified in the order shown, and must all be of the same type, either numeric (preferable) or alphanumeric(not advised).

Once the MDMT file is created, the file must be submitted to HLMHCM:

```
C:\HLM> HLMHCM –r growth.mdmt
```

The results on the screen should then be examined to make sure the data were read correctly. These descriptive statistics will also be contained in a file named HLMHCMMDM.STS.

I.2 Example: Creating an HLMHCM HLM file and running the model

The next step is to create a file that specifies the desired model. (This is usually suffixed with a .hlm) For example, we will use the model shown in section 15.2.

```
nonlin:n
numit:100000
stopval:0.0000010000
level1:MATH=INTRCPT1+AGE8+RANDOM
level2:INTRCPT1=INTRCPT2+BLACK+HISPANIC+random
rowcol:INTRCPT2=theta+DISADV(RANDOM)+randomb+randomc
rowcol:BLACK=theta
rowcol:HISPANIC=theta
level2:AGE8=INTRCPT2+BLACK+HISPANIC+random
rowcol:INTRCPT2=theta+DISADV(RANDOM)+randomb+randomc
rowcol:BLACK=theta
rowcol:HISPANIC=theta
fixtau:3
fixdelta:3
fixomega:3
accel:5
```

```
deviance:3800.651318
df:18
hypoth:n
resfiltype:spss
resfil1:n
resfil1fname:resfil1.sav
resfil2:n
resfil2fname:resfil2.sav
resrow:n
resrowfname:resrow.sav
rescol:n
rescolfname:rescol.sav
title:CONDITIONAL LINEAR GROWTH MODEL,WITH NEIGHBORHOOD DISADVANTAGE EFFECT RANDOM
output:growth3.html
fulloutput:n
```

The above is very similar to an HCM2 model file, with the exception of the model specification at the top where an extra level is shown. Here is the model part that better demonstrates the nested nature of the model specification (the shown indentation will not run):

```
level1:MATH=INTRCPT1+AGE8+RANDOM
    level2:INTRCPT1=INTRCPT2+BLACK+HISPANIC+random
        rowcol:INTRCPT2=theta+DISADV(RANDOM)+randomb+randomc
        rowcol:BLACK=theta
        rowcol:HISPANIC=theta
    level2:AGE8=INTRCPT2+BLACK+HISPANIC+random
        rowcol:INTRCPT2=theta+DISADV(RANDOM)+randomb+randomc
        rowcol:BLACK=theta
        rowcol:HISPANIC=theta
```

Assuming that the above file is named growth.hlm, then the following command should be run:

```
C:\HLM> HLMHCM GROWTH.MDM GROWTH.HLM
```

The following keywords in the above command files have the same definition and options in HCM2 as in the other modules (*e.g.* Tables A.1 and B.1).

```
ACCEL      FULLOUTPUT FIXTAU     NONLIN     NUMIT      OUTPUT     STOPVAL    TITLE
FIXSIGMA2  STOPMICRO  STOPMACRO  DEVIANCE   DF         TITLE      GAMMA      RESFILTYPE
```

References

Barnett, R. C., Brennan, R. T., Raudenbush, S. W., & Marshall, N. L. (1993). Gender and the relationship between marital role-quality and psychological distress: A study of dual-earner couples. *Journal of Personality and Social Psychology,* **64***,* 794-806.

Bloom, H. S., Raudenbush, S. W., Weiss, M. J., & Porter, K. (2017). Using multisite experiments to study cross-site variation in treatment effects: A hybrid approach with fixed intercepts and a random treatment coefficient. *Journal of Research on Educational Effectiveness*, **10**(**4**), 817-842.

Bock, R. (1975). *Multivariate Statistical Methods in Behavioral Research.* New York: McGraw-Hill.

Breslow, N., & Clayton, D. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association,* **88***,* 9-25.

Bryk, A., & Raudenbush, S. W. (1992). *Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods.* Newbury Park, CA: Sage.

Cheong, Y. F., Fotiu, R. P., & Raudenbush, S. W. (2001). Efficiency and robustness of alternative estimators for 2- and 3-level models: The case of NAEP. *Journal of Educational and Behavioral Statistics*, **26**, 411-429.

Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series* **B**(**39**), 1-8.

Elliot, D., Huizinga, D., & Menard, S. (1989). *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems.* New York: Springer-Verlag.

Garner, C., & Raudenbush, S. (1991). Neighborhood effects on educational attainment: A multi-level analysis of the influence of pupil ability, family, school, and neighborhood. *Sociology of Education*, **64**(**4**), 251-262.

Goldstein, H. (1991). Non-linear multilevel models with an application to discrete response data. *Biometrika,* **78***,* 45-51.

Hedeker, D., & Gibbons, R. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics,* **50**, pp. 993-944.

Hong, G., & Raudenbush, S. W. (2008). Causal inference for time-varying instructional treatments. *Journal of Educational and Behavioral Statistics,* **33**, 333-362.

Hough, H. J., Bryk, A., Pinnell, G. S., Kerbow, D., Fountas, I., & Scharer, P. L. (2008). The effects of school-based coaching: *Measuring change in the practice of teachers engaged in literacy collaborative professional development.* Unpublished manuscript, Stanford University.

Howard S. Bloom, Stephen W. Raudenbush, Michael J. Weiss & Kristin Porter (2017). Using Multisite Experiments to Study Cross-Site Variation in Treatment Effects: A Hybrid Approach with Fixed Intercepts and a Random Treatment Coefficient, *Journal of Research on Educational Effectiveness*, **10**(**4**). Pages 817-842. DOI: 10.1080/19345747.2016.1264518. Winner of the best article award by the Society for Research on Educational Effectiveness, 2017.

Huttenlocher, J.E., Haight, W., Bryk, A.S., & Seltzer, M. (1991). Early vocabulary growth: Relations to language input and gender. *Developmental Psychology*, **22**(**2**), 236-249.

Jennrich, R., & Schluchter, M. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics,* **42***,* 805-820.

Johnson, C. S. (2006). *Applications of Laplace approximation for hierarchical generalized linear models in educational research*. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Kang, S.J. (1992). *A mixed linear model for unbalanced two-way crossed multilevel data with estimation via the EM algorithm*. Unpublished doctoral dissertation, Michigan State University, East Lansing.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

Little, R., & Schenker, N. (1995). Missing data. In G. Arminger, C. C. Clogg & M. E. Sobel (Eds.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences* (pp. 39-76). New York: Plenum Press.

Longford, N. (1993). *Random Coefficient Models.* Oxford: Clarendon Press.

McCullagh, P., & Nelder, J. (1989). *Generalized Linear Models, 2nd Edition.* London: Chapman and Hill.

Pfefferman, D., Skinner, C.J., Homes, D.J., Goldstein, H., and Rasbash, J. (1998). Weighting for unequal selection models in multilevel models. *Journal of the Royal Statistical Society, Series* **B**, **60**, 1, 23-40.

Pinheiro, P. C., & Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, **4**, 12-15.

Ramirez, D., Yuen, S., Ramey, R., & Pasta, D. (1991). *The immersion study: Final Report*. Washington, DC: U.S. Office of Educational Research and Improvement.

Raudenbush, S.W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics*, **18**(**4**), 321-349.

Raudenbush, S. W. (1999). Hierarchical models. In S. Kotz (Ed.), *Encyclopedia of Statistical Sciences, Update Volume 3* (pp. 318-323). New York: John Wiley & Sons, Inc.

Raudenbush, S. W. (2001). Toward a coherent framework for comparing trajectories of individual change. Collins, L., & Sayer, A. (Eds.), *Best Methods for Studying Change* (pp. 33-64). Washington, DC: The American Psychological Association.

Raudenbush, S. W. (2009). Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. Education Finance and Policy, **4**(**4**), 468-491.

Raudenbush, S. W., & Bhumirat, C. (1992). The distribution of resources for primary education and its consequences for educational achievement in Thailand. *International Journal of Educational Research,* pp. 143-164.

Raudenbush, S. W., & Bloom, H. S. (2015). Learning about and from a distribution of program impacts using multisite trials. *American Journal of Evaluation*, **36**(**4**), 475-499.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition.* Newbury Park, CA: Sage.

Raudenbush, S. W., & Chan, W.S. (1993). Application of hierarchical linear models to study adolescent deviance in an overlapping cohort design. *Journal of Clinical and Consulting Psychology,* **61**(**6**), 941-951.

Raudenbush, S. W., & Sampson, R. (1999). Assessing direct and indirect associations in multilevel designs with latent variables *Sociological Methods and Research*, **28**(**2**), 123-153.

Raudenbush, S.W., Yang, Meng-Li, & Yosef, M. (2000). Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate Laplace approximation. *Journal of Computational and Graphical Statistics,* **9**, 141-157.

Rodriguez, G., & Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. *Journal of the royal Statistical Society,* **A**, **158**, 73-89.

Rogers, A., & *et. al*. (1992). *National Assessment of Educational Progress: 1990 Secondary-use Data Files User Guide.* Princeton, New Jersey: Educational Testing Service.

Rowan, B., Raudenbush, S., & Cheong, Y. (1993). Teaching as a non-routine task: Implications for the organizational design of schools. *Educational Administration Quarterly,* **29**(**4**), 479-500.

Rowan, R., Raudenbush, & Kang, S. (1991). Organizational design in high schools: A multilevel analysis. *American Journal of Education,* **99**(**2**), 238-266.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, **63**(**3**), 581-592.

Sampson, R., S. Raudenbush, S.W. & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, **277**, pp. 918-924.

Schafer, J. (1997). *Analysis of Incomplete Multivariate Data.* London: Chapman & Hall.

Schall, R. (1991). Estimation in generalized linear models with random effects. *Biometrika,* **40***,* 719-727.

Shin, Y. (2013). Efficient handling of predictors and outcomes having missing values. L., Rutkowski, M., von Davier, D. Rutkowski,(Eds.), A handbook of international large-scale assessment data analysis: Background, technical issues, and methods of data analysis, 451-479.

Shin, Y., & Raudenbush, S. W. (2007). Just-Identified Versus Overidentified Two-Level Hierarchical Linear Models with Missing Data. *Biometrics*, **63**(**4**), 1262-1268.

Shin, Y., & Raudenbush, S. W. (2011). The causal effect of class size on academic achievement: Multivariate instrumental variable estimators with data missing at random. *Journal of Educational and Behavioral Statistics*, **36**(**2**), 154-185.

Shin, Y., & Raudenbush, S. W. (2013). Efficient analysis of Q-level nested hierarchical general linear models given ignorable missing data. *The international journal of biostatistics*, **9**(**1**), 109-133.

Stiratelli, R., Laird, N., & Ware, J. (1984). Random effects models for serial observations with binary response. *Biometrics,* **40***,* 961-971.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., & Najarian, M. (2009*). Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Combined User's Manual for the ECLS-K Eighth-Grade and K-8 Full Sample Data Files and Electronic Codebooks*. NCES 2009-004. National Center for Education Statistics.

Verbitsky‐Savitz, N., & Raudenbush, S.W., (2009). Exploiting spatial dependence to improve measurement of neighborhood social processes. *Sociological Methodology*, **39**, pp 151‐183.

von Hippel, P. T. (2018). How many imputations do you need? A two-stage calculation using a quadratic rule. *Sociological Methods & Research*, 0049124117747303.

Weiss, M. J., Bloom, H. S., Verbitsky-Savitz, N., Gupta, H., Vigil, A. E., & Cullinan, D. N. (2017). How much do the effects of education and training programs vary across sites? Evidence from past multisite randomized trials. *Journal of Research on Educational Effectiveness*, **10**(**4**), 843-876. doi:10.1080/19345747.2017.1300719.

Wong, G., & Mason, W. (1985). The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association,* **80**(**391**), 513-524.

Yang, M. (1995). *A simulation study for the assessment of the non-linear hierarchical model estimation via approximate maximum likelihood*. Unpublished apprenticeship paper, College of Education, Michigan State University.

Yang, M.L. (1998). *Increasing the efficiency in estimating multilevel Bernoulli models* [Diss], East Lansing, MI: Michigan State University.

Yosef, M. (2001). *A comparison of alternative approximations to maximum likelihood estimation for hierarchical generalized linear models: The logistic-normal model case.* Unpublished doctoral dissertation, Michigan State University, East Lansing.

Zeger, S., Liang, K.Y., & Albert, P. (1988). Models for longitudinal data: A likelihood approach. *Biometrics,* **44***,* 1049-60.

Zeger, S., & Liang, L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika,* **73***,* 13-22.

Subject Index

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

Z