Introductory Guide to HLM With HLM 7 Software

3

G. David Garson

HLM software has been one of the leading statistical packages for hierarchical linear modeling due to the pioneering work of Stephen Raudenbush and Anthony Bryk, who created the software and authored the leading text on hierarchical linear and nonlinear modeling (Bryk & Raudenbush, 1992; Raudenbush & Bryk, 2002). Though differences among software packages' capabilities have diminished over time, HLM 7 offers a number of appealing advantages and capabilities. Among these are what many consider to be a more intuitive model specification environment, greater ease in creating three- and four-level models, its wide choice of estimation options, integrated likelihood ratio hypothesis testing, graphics options, and the ability easily to handle heterogeneous hierarchical linear models (where the dependent is thought to have different error variances for different levels of some grouping variable such as Agency).

HLM SOFTWARE

Scientific Software International (SSI) distributes HLM 7. A free student edition of HLM 7 is available.[1] The student edition is full-featured, including examples, but is limited in the size and complexity of models (though it will work with all example files provided with the software). HLM 7 software operates through several modules, each designed for a different type of HLM model, only some of which can be illustrated here due to space constraints:

HLM2. For two-level linear and nonlinear models with one dependent variable.

HLM3 and HLM4. For three-level and four-level models with one dependent variable.

HGLM. For generalized linear models for distributions other than normal and link functions other than identity, handling binary, count, multinomial, and ordinal outcome variables in Bernoulli, binomial, Poisson, multinomial, and ordinal models.

HMLM. For multivariate normal models with more than one outcome variable, including when the level 1 covariance structure is homogenous, heterogeneous, loglinear, or AR(1) (first-order autoregressive).

HMLM2. For two-level HMLM models where level 1 is nested within level 2.

HCM2. For models where level 1 units are cross-classified by two level 2 units.

HCM3. For three-level cross-classified models.

HLMHCM. For two- and three-level hierarchical linear models with cross-classified random effects (ex., repeated test scores nested within students who are cross-classified by schools and neighborhoods).

In summary, HLM 7 is a versatile and full-featured environment for many linear and generalized linear mixed models.

ENTERING DATA INTO HLM 7

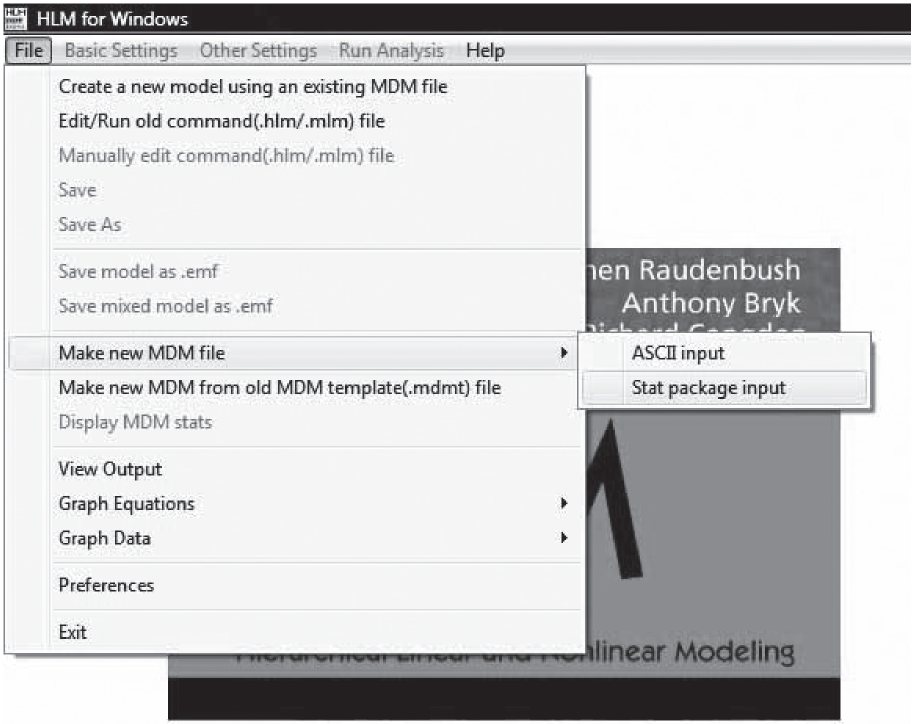
HLM software stores data in its own multivariate data matrix (MDM) format, which may be created from raw data or from data files imported from SPSS, SAS, Stata, SYSTAT, or other packages. MDM format files come in flavors keyed to the several types of HLM modules noted above. File creation options are accessed from the HLM File menu, illustrated in Figure 3.1 below. The example below illustrates data entry from an SPSS .sav file for models of type HLM2, but similar procedures are followed for other model types.

"Stat package input," depicted above, is the most common method of creating .mdm data files. Further, not only are data commonly prepared using statistical or data packages outside HLM 7, but as an additional preprocessing step, the researcher also should rule out multicollinearity among the level 2 (or higher) predictors. Having done this, there are two methods of importing files into HLM 7 from other statistical packages.

Input Method 1: Separate Files for Each Level

This method results in faster processing but requires more time to set up the data. It requires that separate files be created outside of HLM 7 for each level of HLM analysis. For SPSS, these are .sav files. For SAS, these are SAS 5 transport files. Separate SYSTAT and Stata files are also acceptable. For instance, HLM 7 software comes with example files from the Singer (1998) "High School and Beyond" study. The SPSS files for this example include HSB1.SAV, which contains the level 2 link field (ID is school ID) and any student-level variables. There are multiple rows per school, one row per student. It is critical that the level 1 file is sorted such that all students for a given school ID are adjacent.

Figure 3.1 HLM 7 file menu



Likewise, the school-level (level 2) file, HSB2.SAV, contains the same level 2 link field and any school-level variables.

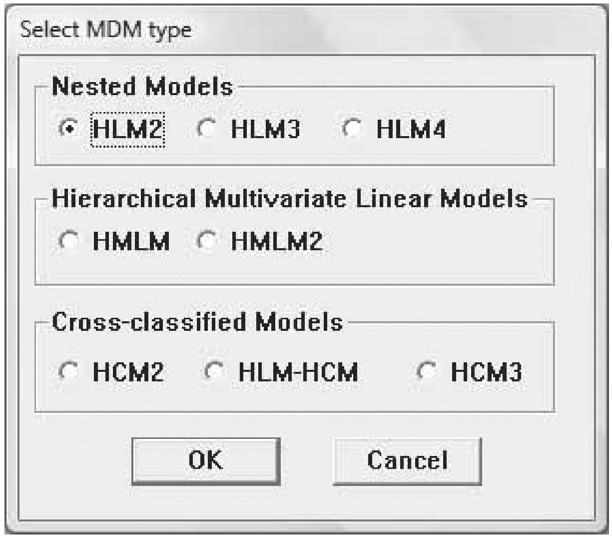
Input Method 2: Using a Single Statistics Program Data File

This method[2] is easier in terms of data management and is the one illustrated in this chapter. The same statistics package file formats as for Method 1 may be used. For the example, the single data file must be sorted such that all students for a given school ID are adjacent.

Making the MDM File

The next step is to create the .mdm file, which is HLM software's native data format. After it is created, the input data files are not needed. After creating the

Figure 3.2 HLM 7 select MDM type window



input data file in SPSS, SAS, or another package, HLM 7 is run and "Stat package input" is selected. This causes the "Select MDM type" window illustrated above to appear.

The researcher chooses the HLM model type wanted. For instance, for a simple two-level hierarchical linear model, the selection would be HLM2.

After selecting HLM2, the "Make MDM - HLM2" dialog box appears, illustrated in Figure 3.3.

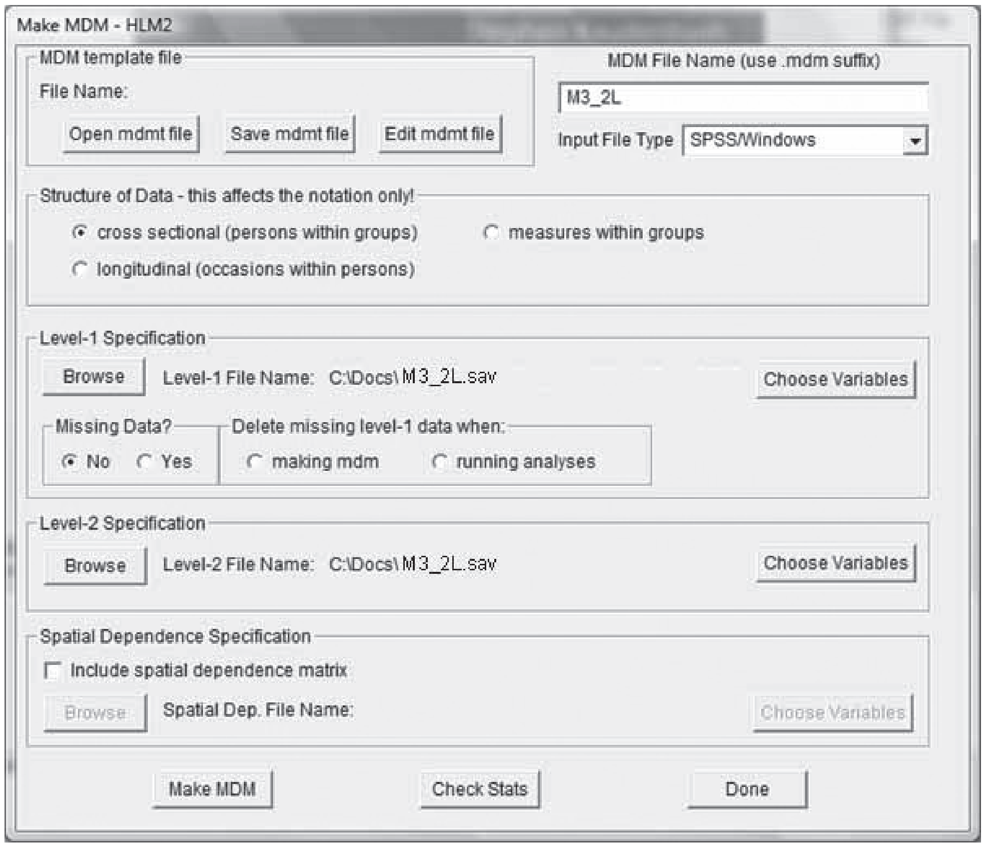
Here, the following steps are necessary:

Set the "Input File Type" to "SPSS/Windows" (or another statistical package format).

In the level 1 specification area, click the "Browse" button and browse to the input file for level 1. Then, as illustrated at the top of Figure 3.4, click the "Choose variables" button, click the checkbox indicating the level 2 link variable (id in the example), and click the checkboxes of any other level 1 variables in the analysis.

In the level 2 specification area, click the "Browse" button and browse to the input file for level 2. This may be the same file as for level 1 (following Method 2 above). Again click the "Choose variables" button, click the checkbox indicating the level 2 link variable (agency), and click the checkboxes of any other level 2 variables in the analysis, as indicated in the lower half of Figure 3.4.

Figure 3.3 Make MDM - HLM 2 window in HLM 7



Save the MDM template file by clicking the “Save mdmt file” button, making sure the file location window points to the desired folder and giving a filename (add the .mdmt extension), then clicking the “Save” button.

To complete the process, the researcher clicks the “Make MDM” button, giving a filename (here, M3_L2.mdm, standing for mixed linear model Chapter 3, 2-level). The .mdm file is created, and the descriptive statistics module runs. Alternatively, one may click the “Check Stats” button. This output, shown in Figure 3.5, should be examined to verify the results. For instance, it is prudent to examine the reported sample size, which, if low, flags that the researcher has not sorted the Level 1 file to assure that individual rows for the same level 2 ID (AGENCY in this example) are adjacent.

Figure 3.4 Choose variables windows in HLM 7

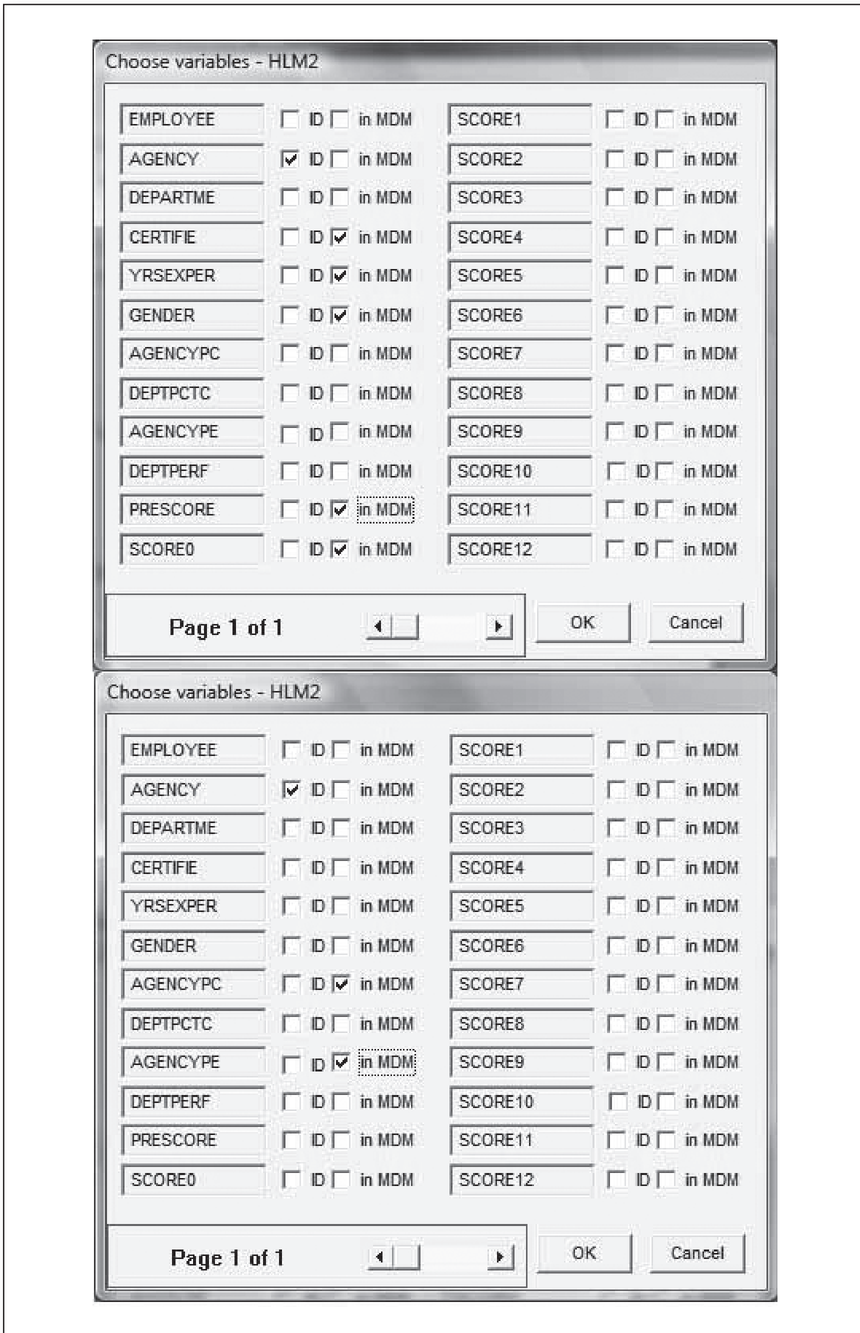
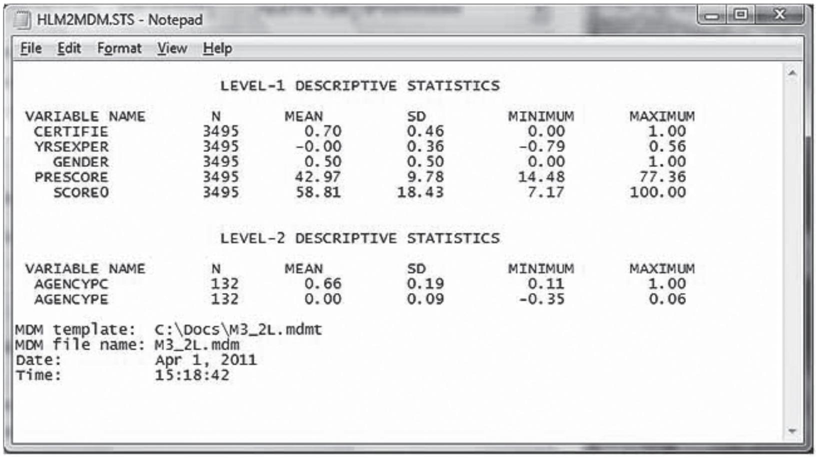


Figure 3.5 "Check Stats" button output in HLM 7



```
HLM2MDM.STS - Notepad
File Edit Format View Help

                     LEVEL-1 DESCRIPTIVE STATISTICS

 VARIABLE NAME        N      MEAN        SD       MINIMUM     MAXIMUM
  CERTIFIE          3495     0.70      0.46        0.00        1.00
  YRSEXPER          3495    -0.00      0.36       -0.79        0.56
    GENDER          3495     0.50      0.50        0.00        1.00
  PRESCORE          3495    42.97      9.78       14.48       77.36
    SCORE0          3495    58.81     18.43        7.17      100.00

                     LEVEL-2 DESCRIPTIVE STATISTICS

 VARIABLE NAME        N      MEAN        SD       MINIMUM     MAXIMUM
  AGENCYPC           132     0.66      0.19        0.11        1.00
  AGENCYPE           132     0.00      0.09       -0.35        0.06

MDM template:  C:\Docs\M3_2L.mdmt
MDM file name: M3_2L.mdm
Date:          Apr 1, 2011
Time:          15:18:42
```

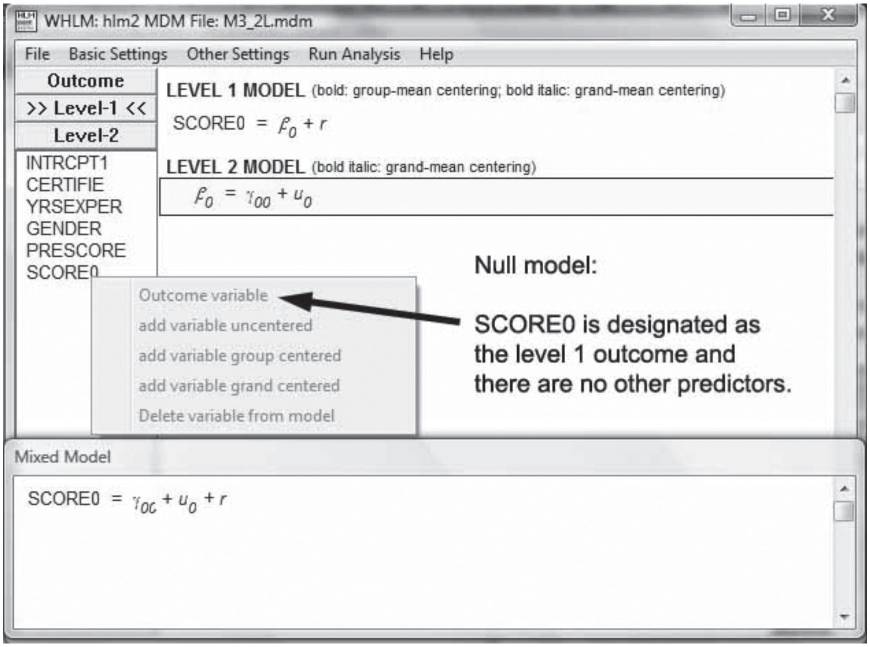
Click the "Done" button to exit to the WHLM model construction screen discussed below. At this point, the researcher will have saved three files to the disk: the newly created HLM-compatible data file, H3_L2.mdm in this example; the default template creatmdm.mdmt (the researcher may override the default name); and the output file above, HLM2MDM.STS (if desired, use File, Save As, to save output under a different name, as this default file may get reused with new content if there are multiple runs).

THE NULL MODEL IN HLM 7

After data are entered, the next step is to create the model. Typically, the first model created is the null model. The null model serves two purposes: (1) It is the basis for calculating the intraclass correlation coefficient (ICC), which is the usual test of whether multilevel modeling is needed; and (2) it outputs the deviance statistic (–2LL) and other coefficients used as a baseline for comparing later, more complex models. For the current example, the null model addresses the question, "Is there a (level 2) agency effect on the (level 1) intercept of performance score, which represents the mean score?" If there is an agency effect, then ordinary regression methods will suffer from correlated error, and some form of linear mixed modeling is required.

The null model, like all two-level hierarchical models in HLM 7, is created in the WHLM modeling dialog, illustrated in Figure 3.6. This dialog is reached either on clicking "Done" in the "MAKE MDM" dialog or, if the MDM file was

Figure 3.6 WHLM modeling window in HLM 7: Null model



previously saved, from the HLM menu by selecting File, "Create a new model using an existing MDM file," and then opening the appropriate .mdm file.

In the WHLM modeling dialog illustrated in Figure 3.6, the employee level (level 1) dependent variable performance score (SCORE0) is designated as the outcome variable. No other predictors are added. HLM 7 already knows "Agency" is the level 2 grouping variable and automatically assumes it is a predictor of the level 1 intercept of SCORE0.

Table 3.1 Summary of the Null Model

Level 1 Model

$SCORE0_{ij} = \beta_{0j} + r_{ij}$

Level 2 Model

$\beta_{0j} = \gamma_{00} + u_{0j}$

Mixed Model

$SCORE0_{ij} = \gamma_{00} + u_{0j} + r_{ij}$

When SCORE0 is designated as the outcome variable, HLM 7 constructs and displays the model, in this case the null model (also called the intercept-only model or the one-way ANOVA model with random effects). The null model is shown in Table 3.1 below. Clicking the "Mixed" button at the bottom of the WHLM dialog creates the combined HLM equation shown at the bottom of the figure: The two separate equations shown in the upper main window are mathematically equivalent to the single combined mixed model equation. For learning purposes, it is easier to examine the equations at each level. At level 1, SCORE0 is predicted by an intercept term and a random term. The symbol for the intercept term varies depending on the distribution specified for the outcome variable (this is done in the “Basic Settings” window, described below) and is expressed equivalently but differently in output.

The level 1 intercept term, expressed as β_{0j} in output, is a function of a random intercept term at level 2 (γ_{00}) and a level 1 residual error term (r_{ij}). The level 1 intercept, in turn, is a function of the grand mean (γ_{00}) across level 2 units, which are agencies in this example, plus a random error term (u_{0j}), signifying the intercept is modeled as a random effect. Substituting the right-hand side of the level 2 equation into the level 1 equation gives the mixed model equation for the null random intercept model. HLM 7 will create one level 1 regression for each agency, and then will utilize the variance in these intercepts when estimating parameters and standard errors at level 1. This is what makes the process different from ordinary regression, where a single overall intercept is estimated.

Before calculating estimates, the researcher may specify the distribution of the outcome variable by selecting “Basic Settings” from the main menu bar, yielding the window shown in Figure 3.7. The normal distribution, used in this example, is the default. Other available specifications support Bernoulli, Poisson, multinomial, and ordinal distributions. Selecting “Bernoulli” for a binary outcome variable applies a logistic link function, and in the ensuing multilevel logistic regression, interpretations are in terms of the log odds of the outcome rather than in terms of the raw outcome itself. Selecting “Multinomial” creates a multilevel multinomial regression using a logit link. “Ordinal” supports multilevel ordinal regression models. Multilevel Poisson regression models employ a Poisson log link and require an exposure variable (time, for example). In this window, one may also specify the name and location of the output statistics file and the output graphics file.

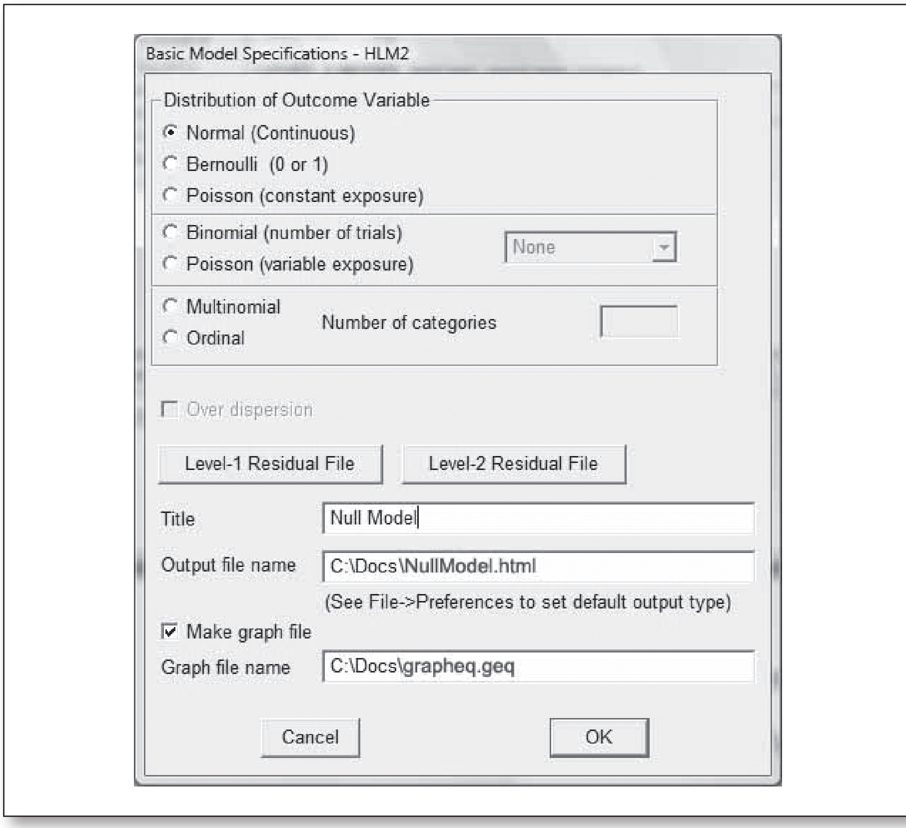
It is also possible to modify model estimation settings prior to running the model by selecting “Other Settings” from the main menu bar, then “Estimation Settings,” as illustrated in Figure 3.8. Estimation settings were discussed in Chapter 2. For the null model, we use the default setting, restricted maximum likelihood estimation. There is also an “Iterations Settings” window, also from the “Other Settings” menu. Though not illustrated here, it provides options discussed in Chapter 2 with regard to estimation settings.

One may also select “Other Settings” from the main menu bar, then “Output Settings” to obtain the window shown in Figure 3.9. For this model, one may choose to print out variance–covariance matrices or to restrict output to the main results. The default is restricted output and no matrices.

To run the null model, the researcher simply selects “Run Analysis” from the main menu bar. Output is sent to the file location and a name is specified in the “Basic Model Specifications” window (Figure 3.7). To view the output, select File, View Output, from the main menu bar. For this example, the critical output of the null model looks as shown in Table 3.2.

The phrase “Number of estimated parameters = 2” refers to the fact that in a null model, estimates are made for the level 1 intercept and the level 2 intercept. In the

Figure 3.7 Basic Model Specifications window from Basic Settings menu in HLM 7



final variance components table, the fact that the component for the intercept (161.94, which HLM labels tau, τ) is significant means that the intercept of the outcome variable, SCORE0, is significantly affected by its predictors, which in this example is the level 2 effect of agency. A non-significant intercept in the variance components table term (not the case here) would mean that after other variables in the model are controlled, there would be no residual between-groups variance in the level 1 dependent variable (Score0). The agency effect is smaller than the residual variance component (212.69, which HLM also labels sigma-squared, σ^2), indicating that there is still considerable residual variation in Score0 yet to be explained and that a model with additional predictors may be needed.

The fact that the intercept component is significant means that the intraclass correlation coefficient, ICC, is also significant, indicating that a multilevel model is appropriate and needed. ICC varies from +1.0 when group means differ but within any group there is no variation, to $-1/(n-1)$ when group means are all the same but within-group variation is very large. At the extreme, when ICC

Figure 3.8 Estimation Settings window from Other Settings menu in HLM 7

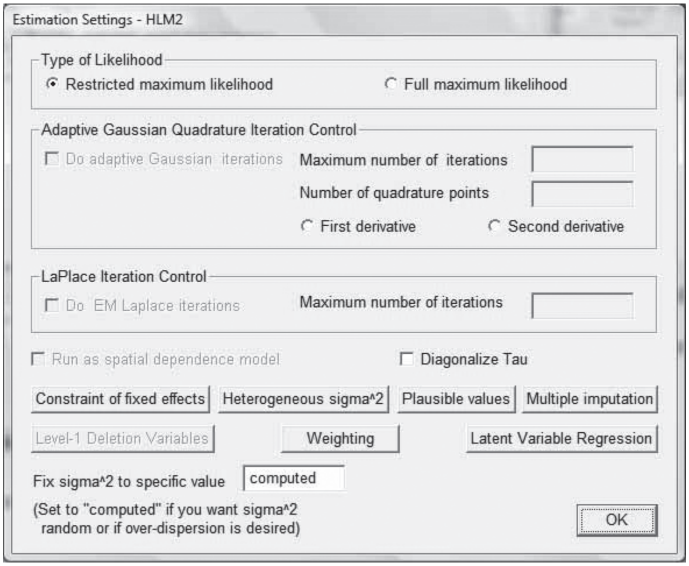


Figure 3.9 Output Settings window for HLM2

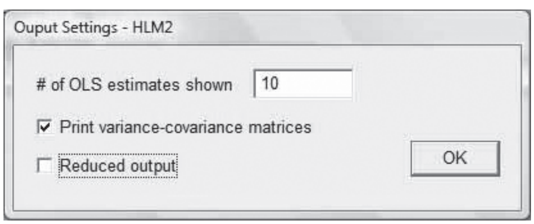


Table 3.2 Final Estimation of Variance Components for the Two-Level Null Model

Random effect	*Standard deviation*	*Variance component*	*d.f.*	χ^2	*p-value*
INTRCPT1, u_0	12.72558	161.94029	131	2391.61810	<0.001
level-1, r	14.58375	212.68575			
Statistics for current covariance components model					
Deviance = 29028.420032					
Number of estimated parameters = 2					

approaches 0 or is negative, hierarchical modeling is not appropriate. For this example, the magnitude of ICC may be calculated as the intercept variance component in the null model divided by the total of variance components. That is, ICC = 161.94/(161.94 + 212.69) = .43.

The fixed effect tables are of lesser interest in a null model but are presented in Table 3.3. Mean performance score (the intercept at level 1) is estimated to be 55.60 for this example, when the level 2 grouping variable, agency, is the only effect modeled. Confidence limits around the mean, of course, are approximately plus or minus two standard errors. The lower table, "with robust standard errors," produces the same estimate but has a slightly different standard error. Robust standard errors are recommended when it is possible the researcher has specified the wrong distribution of the dependent variable. Significant differences between the ordinary and robust estimates of the standard error may flag a problem with the distribution specified by the researcher. This is not the case in this example, which specified a normal distribution (which is the default).

The "Deviance" value of 29028.42 in Table 3.2 is the basis of model fit measures. While not used at this point, for the null model it is the baseline model fit. More complex models are assessed in part by how greatly they reduce deviance (which is also called –2 log likelihood, –2LL, and model chi-square). These tests of the difference in deviance values between models are likelihood ratio tests, requested in HLM 7 by selecting "Other Settings" from the main menu bar, then "Hypothesis Testing," as discussed later in this chapter. In summary, at the end of analysis of the null model we have demonstrated that there is a significant agency effect on employee performance scores; that therefore multilevel modeling is

Table 3.3 Fixed Effects Tables for the Null Model

Final estimation of fixed effects

Fixed effect	*Coefficient*	*Standard error*	t-*ratio*	*Approx.*d.f.	p-*value*
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	55.598248	1.145778	48.524	131	<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}	55.598248	1.141423	48.710	131	<0.001

needed; and that additional, more complex models with more predictors should reduce significantly the baseline deviance value of 29028.42.

A RANDOM COEFFICIENTS REGRESSION MODEL IN HLM 7

Given level 1 representing employees, with performance score as an outcome (dependent) variable, and level 2 representing agencies, a random coefficients regression model is one with one or more level 1 predictors such as gender, years of experience, or a binary indicator for whether the employee is certified or not. The level 2 grouping variable (Agency) remains a random factor, but there are no other level 2 predictors. The "coefficients" term in the label means that the agency effect is used not only to model the level 1 intercept of SCORE0 as an outcome, but also to model the regression coefficients of the level 1 predictors.

As an example of random coefficients (RC) regression, employee performance score (score0) at level 1 is predicted from the level 1 covariates years of experience (YrsExper) and sex (Gender, where 0 = male, 1 = female). Note that HLM 7 enters binary variables like Gender as covariates by default. There are no predictors at level 2, but Agency is the subjects variable under which employees are grouped. The intercept of score0 at level 1 and the b coefficient of YrsExper at level 1 are both modeled as random effects of Agency. Gender is treated as a simple level 1 fixed effect. This model explores whether the Agency effect discovered in the null model may be attributed in part to some agencies having more experienced employees than others. The model also explores whether the demographic variable, Gender, modifies the relationship of years of experience to performance score.

Figure 3.10 illustrates this RC regression model. An often-cited advantage of HLM software is how its user interface clearly separates regression models at different levels. Here, at level 1, score0 is predicted from YrsExper and Gender, plus an intercept term β_{0j} and an error term r_{ij}:

$$\text{SCORE0}_{ij} = \beta_{0j} + \beta_{1j}{*}(\text{YRSEXPER}_{ij}) + \beta_{2j}{*}(\text{GENDER}_{ij}) + r_{ij}$$

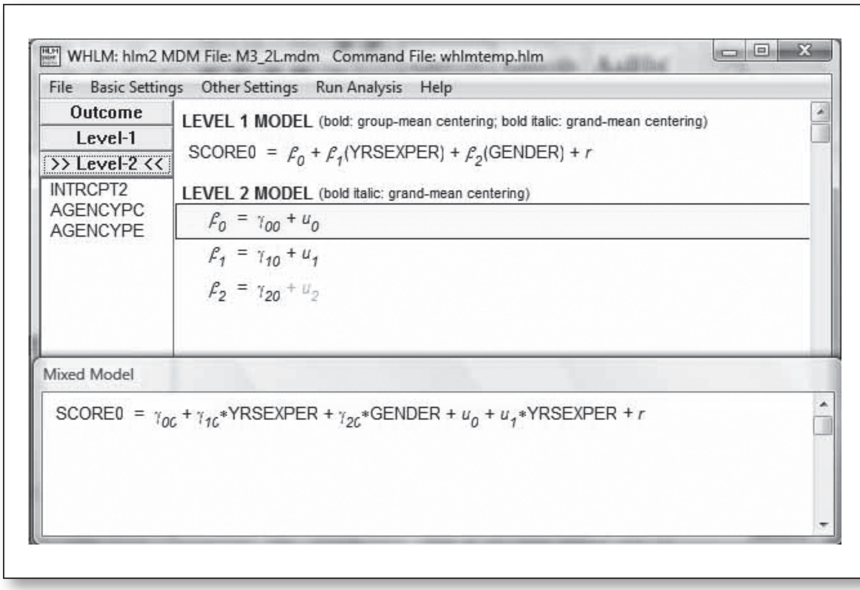
At level 2, there are no predictors. However, the level 1 intercept is predicted by the level 2 mean (γ_{00}) of score0 plus a level 2 error term (u_{0j}). The level 2 error term represents the random effect of agency on score0 at level 1. Also, the level 1 regression coefficient (slope) of YrsExper (β_{1j}) is predicted by the mean of agency regression coefficients where this mean is based on the 132 agencies in the sample, plus a level 2 error term (u_{1j}) representing the random effect of Agency on the level 1 regression of score0 on YrsExper. HLM 7 output represents the level 2 equations as below:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20}$$

The level 1 and level 2 equations can be combined, through substitution, into the single mixed model equation below:

$$\text{SCORE0}_{ij} = \gamma_{00} + \gamma_{10}\text{*YRSEXPER}_{ij} + \gamma_{20}\text{*GENDER}_{ij} + u_{0j} + u_{1j}\text{*YRSEXPER}_{ij} + r_{ij}$$

Figure 3.10 A two-level random coefficients regression model in HLM 7



As explained in Chapter 1, by designating a level 2 subjects variable, Agency, one is requesting that regressions be created separately for each agency so that the variance in intercepts and coefficients can be calculated and used in subsequent estimates of fixed and random effects. If full rather than reduced output is requested from the Other Settings, Output Settings menu selection, HLM 7 will print the OLS regression coefficients for the first 10 agencies (the default of 10 can be overridden by the researcher to get all coefficients), as shown in Table 3.4. Using these intercepts and slopes, one can create a plot of multiple different regression lines across agencies, graphically illustrating the nature of random intercepts and random slopes for the given data.

The likelihood ratio test, discussed in Chapter 2, can be used as an overall test of whether the RC regression model with predictors is a significantly better fit than the intercept-only (null) model without predictors. In HLM 7, this is done from the main menu by selecting Other Settings, Hypothesis Testing, leading to the "Hypothesis Testing" window shown in Figure 3.11. In this window, one

Table 3.4 OLS Coefficients for the First 10 Agencies in the RC Regression Model

Level-1 OLS Regressions

Level 2 Unit	*INTRCPT1*	*YRSEXPER slope*
1	26.91987	–8.59458
2	33.31829	2.13703
3	–190.35526	–308.72487
4	32.08813	–3.68785
5	31.26302	–4.19888
6	29.55992	–1.72763
7	31.67209	1.19479
8	36.08018	5.47836
9	48.24087	27.07489
10	37.73941	2.49587

The average OLS level 1 coefficient for INTRCPT1 = 56.98519

The average OLS level 1 coefficient for YRSEXPER = 11.88690

Figure 3.11 Hypothesis Testing window in HLM 7

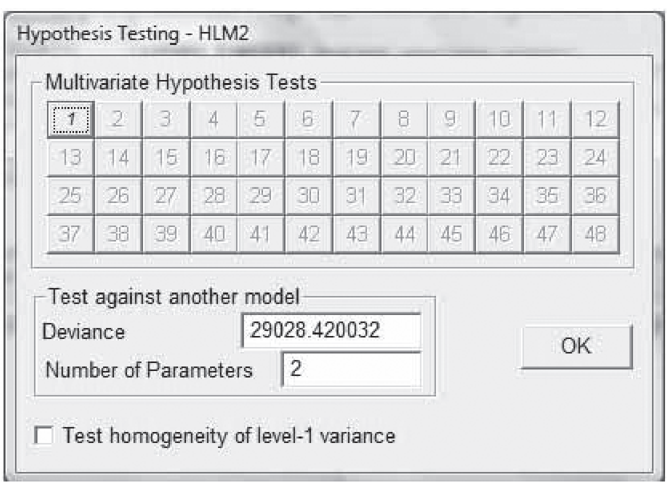


Table 3.5 HLM 7 Likelihood Ratio Test for the RC Regression Model Compared to the Null Model

Deviance

Deviance = 28201.160055

Number of estimated parameters = 4

Variance-Covariance components test

χ^2 statistic = 827.25998

Degrees of freedom = 2

p-value = <0.001

manually enters the deviance value and number of estimated parameters from the null model. The greater the drop in the deviance (–2LL), the more likely the fit is to be significantly better. For this example, deviance dropped from 29028.42 in the null model to 28201.16 in the RC regression model. The likelihood ratio test, shown in HLM 7 output in Table 3.5, shows this difference to be significant at better than the .001 level. Later, the likelihood ratio test can be used to compare any two models if one is nested within the other, as the null model is nested within the RC regression model.[3]

Another way to assess improvement in model fit is to examine the residual variance component in the "Final Estimation of Variance Components" table (Table 3.6). The residual variance component is variance associated with the within-agency variation in score0 not accounted for by the random effects of Agency on the intercept of score0 and on the slope of YrsExper. As the random effects explain more, the residual component will drop. HLM 7 lists the residual variance component as "level-1, *r*." For these data, the residual component drops from 212.69 in the null model to 163.77 in the RC regression model. Since both were models in which the off-diagonal covariances were constrained to zero,[4] we can calculate that residual variance was 163.774/(163.774 + 130.569 + 63.299) = 46% of total variance in the RC regression model compared by similar calculation to 57% in the null model.

In the same manner, in Table 3.6 we can calculate that the random effect of Agency on mean performance scores by agency (the "INTRCPT1, u_0" effect) is 37% of total effects. This means that there is a significant, moderately strong

Table 3.6 Variance Components Table for the RC Regression Model

Final estimation of variance components

Random effect	*Standard deviation*	*Variance component*	*d.f.*	χ^2	p-*value*
INTRCPT1, u_0	11.42668	130.56897	131	1910.92848	<0.001
YRSEXPER slope, u_1	7.95609	63.29930	131	299.94754	<0.001
level-1, *r*	12.79742	163.77400			

tendency for some agencies to have higher mean scores than others. The random effect of Agency on slopes of YrsExper (the "YRSEXPER slope, u_1" effect) is 18% of total effects. This means that there is a significant but weaker tendency for YrsExper to have a stronger effect on performance score in some agencies than others.

Table 3.7 shows the fixed effects table for the RC regression model. HLM 7 prints two fixed effects tables, one using robust standard errors and one not. Robust standard errors should be used when it is possible that the distribution of the dependent variable was misspecified, though for this example the coefficients are identical. For these data, score0 is normally distributed, as specified (for example, skew and kurtosis were both within +/– 1.0), and coefficients and the probability levels are identical between the two versions of the fixed effects table. From Table 3.7, we conclude that both YrsExper and Gender are significant predictors of employee performance scores (score0). YrsExper was centered (mean = 0) and for Gender "male" was coded 0. Therefore, a male with mean years of experience could be expected to score 59.4 points on the performance test. The fact that the slope of Gender was negative means that being female (Gender = 1) was associated with scoring 6.5 points less, controlling for other variables in the model. The t-ratios are the regression coefficients divided by their standard errors. The fact that the absolute t-ratios for YrsExper and Gender are similar indicates similarity in effect size. (Note that a 1-year increase in years of experience does *not* predict an increase of score of 16.25 points because YrsExper was not only centered but was also transformed to decimal form, ranging from –.79 to +.56: That is, the units were no longer raw years.)

Table 3.7 The Fixed Effects Table for the RC Regression Model

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}	59.390114	1.035557	57.351	131	<0.001
For YRSEXPER slope, β_1					
INTRCPT2, γ_{10}	16.252170	1.058662	15.352	131	<0.001
For GENDER slope, β_2					
INTRCPT2, γ_{20}	–6.482774	0.434132	–14.933	3230	<0.001

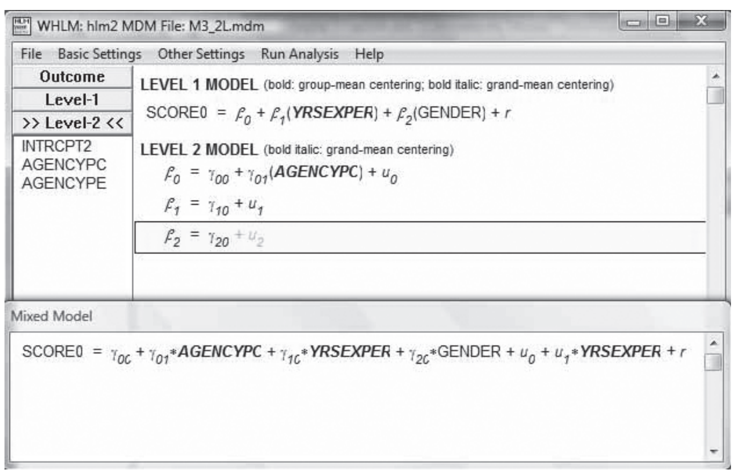
HOMOGENOUS AND HETEROGENEOUS FULL RANDOM COEFFICIENTS MODELS

Also called an "intercepts-and-slopes-as-outcomes" model, the full random coefficients model is a type of hierarchical linear model in which, for two levels, there are predictors at both levels, and both the level 1 intercept and the level 1 slopes are predicted as random effects. As shown in Figure 3.12 below, the level 1 model for this example remains similar to the previous model.

The essential features of the model are these:

- At level 1, Score0 is predicted by YrsExper and Gender.
- Agency remains the grouping variable defining level 2, which means that as many level 1 regressions are run as there are agencies in the sample, yielding an estimate of the variability of level 1 slopes and intercepts.
- The intercept at level 1 is predicted as a random effect of Agency and of the level 2 predictor AgencyPC, which is a newly added covariate measuring percent of employees certified in an agency.
- The slope of YrsExper at level 1 is predicted as a random effect of Agency. In a subsequent example later in this section, a different model will illustrate modeling the slope of YrsExper as a random effect of both Agency and AgencyPC at level 2. The current model tests the proposition that the mean (intercept) for Score0 is a random effect of Agency and of AgencyPC at level 2, and that the strength of relationship (slope) of YrsExper to Score0 is a function of Agency but not of AgencyPC.
- The slope of Gender is not predicted as a random effect, as signified by its random error term (u2) being grayed out in Figure 3.12.

Figure 3.12 A two-level random coefficients model in HLM 7



Though not illustrated, we click "Other Settings" to obtain a window similar to that shown in Figure 3.11 above. We enter the deviance (28201.16) and number of parameters (4) from the previous RC regression model without the level 2 agency percent certified variable. This requests a likelihood ratio test of the model fit difference, to be discussed below, between the current model and the previous one. In a second run, a likelihood test was also requested comparing the current model with the null model. In addition, the "Test level-1 homogeneity of variance" checkbox, also illustrated in Figure 3.11 above, was checked. (Figure 3.11 also shows the HLM 7 option for "Multivariate Hypotheses Tests," as illustrated, for example, in Chapter 12.)

The overall test of the model that is reflected in the likelihood ratio test of the difference in deviance (–2LL) between models is shown in Table 3.8. While deviance declined another 40.39 points compared to the RC regression model and reflects significantly better fit than the null model, the difference between the current model and the RC regression model is not significant (*p*-value = >.500). That is, agency percent certified as a level 2 covariate in the current model reduced deviance by a non-significant amount. On parsimony grounds, the researcher would prefer the RC regression model. Nonetheless, for instructional purposes, the remaining HLM 7 output is examined below.

Table 3.8 HLM 7 Likelihood Ratio Test for the Two-Level Homogenous Full RC Model

Statistics for current covariance components model

Deviance = 28160.766143

Number of estimated parameters = 4

Variance-Covariance components test
(compared to RC regression model)

χ^2 statistic = 40.39391

Degrees of freedom = 0

p-value = >.500

The "Test level-1 homogeneity of variance" output, not previously discussed, is shown in Table 3.9. Although not invoking this test is the default in HLM 7, it is sufficiently critical that Raudenbush and Bryk (2002) state, "investigators generally will wish to begin with this assumption" (p. 263). This test refers to the assumption that when the model is run for each of the 132 agencies in the current example, the residual variances are homogenous. Optionally, the researcher may override the default and test this assumption. For the example data, the test p-value is 0.000. This finding of significance means that residual variances do differ significantly for these data across agencies. Raudenbush and Bryk (p. 263) note that heterogeneity of error variance is a serious problem if variances are not random but are a function of

Table 3.9 Test of Homogeneity of Level 1 Variance

Test of homogeneity of level 1 variance

χ^2 statistic = 271.67767

Degrees of freedom = 131

p-value = 0.000

level 1 or level 2 predictors. Heterogeneity may indicate one of four problems in the research design:

1. Model misspecification. One or more level 1 predictors have been omitted from the model, where the variables in question are distributed with unequal variance across groups. This is the most likely cause of heterogeneity, making this test a form of screening for model misspecification.

2. A level 1 predictor has been modeled as a fixed effect when in fact it is a random effect.

3. One or more level 1 predictors are non-normal (for example, kurtotic with heavy tails), causing the significance test statistic for homogeneity of residual variance, which assumes normality, to report lack of homogeneity. Transformation of the predictor, as in OLS regression, may mitigate non-normality. Likewise, use of a link function other than identity in a generalized linear mixed model may also moderate the effects of non-normality.

4. Coding or other data entry errors, or presence of outliers, may cause heterogeneous error variance in some groups.

It is possible to visually inspect residual variance by level 2 groups (agencies for this example) in a variety of ways (see Raudenbush & Bryk, 2002, pp. 263–267; Raudenbush et al., 2011, pp. 274–278). For instance, within HLM 7 one may select File, Graph Equations, "Level 1 residual versus predicted value" to obtain a plot where points represent agencies. A more complex graphical method is to use a combination of HLM 7, SPSS, and Excel, yielding a plot of residual variances by agency, as shown in Figure 3.13. The process is described in the endnotes to this chapter.[5] The steeper the trend line and the more outliers in this chart, the less homogenous residual variance is. Sometimes it may be possible to obtain a non-significant test of homogeneity by identifying outliers, such as Agency 48 in Figure 3.13, and removing them from the sample for separate analysis. However, if the trend line is steep, the researcher may need to create a heterogeneous model. Endnote 5 describes how to obtain a table of the correlations of Agency residual variance with the predictor variables, showing for these data that such variance is correlated with years of experience though not with gender.

To deal with the problem of heterogeneous residual variance, the four possibilities above need to be explored first. Ultimately, however, it may be necessary to create a heterogeneous variance model. This is done in HLM 7 by selecting Other Settings, Estimation Settings, giving the window shown in Figure 3.8 above. Clicking the "Heterogeneous sigma^2" button leads to the dialog screen of the same name, shown in Figure 3.14. Here, one must enter a level 1 variable, which may account for heterogeneous residual variance. The variable may be one not otherwise used in the model.

Figure 3.13 Residual variance by agency for the two-level random coefficients model

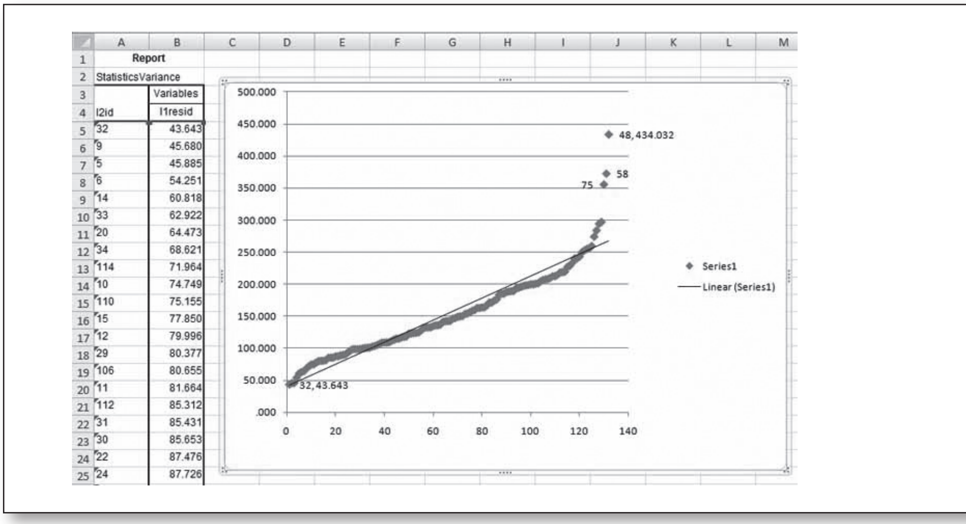
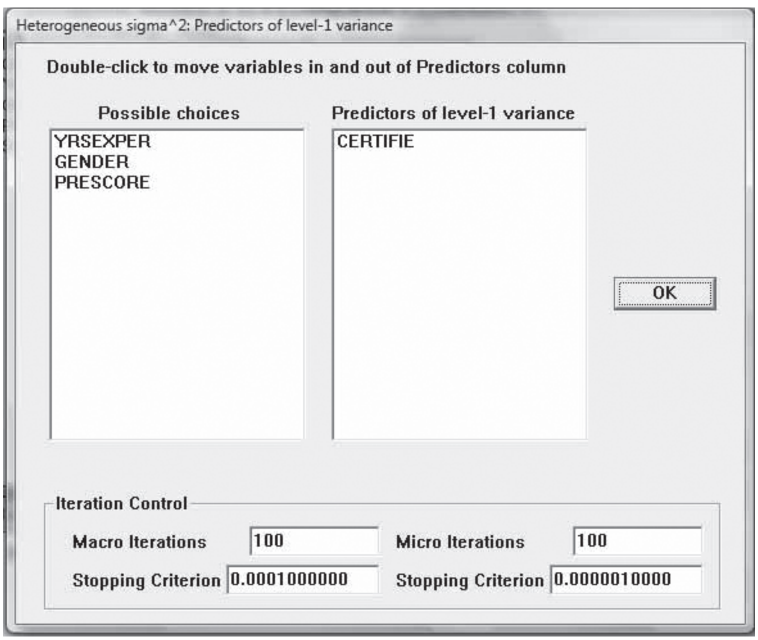


Figure 3.14 Modeling heterogeneity in HLM 7



Determining which variable to use as the predictor of level 1 residual variance can be explored using a statistical package to view the variances of the OLS residuals by candidate variables. In this example, in SPSS, select Analyze, Regression, Linear; set the dependent to Score0. Set the independents to YrsExper and Gender, as in the current model. Set the Selection variable to be Certified (coded 0 = not certified, 1 = certified). Set the Rule to be Certified = 0 in a first run and Certified = 1 in a second run. In the ANOVA table output, residual sum of squares for not certified is 57,488.9 and for certified is 668,407.4. The larger the difference, the better the selection variable is as a candidate predictor to be specified in Figure 3.14 when creating a heterogeneous variance model.

Table 3.10 Likelihood Ratio Test and Homogeneity of Residual Variance Test for the Heterogeneous Full Random Coefficients Model

Statistics for the current model

Deviance = 27686.815423

Number of estimated parameters = 9

Model comparison test

χ^2 statistic = 514.34458

Degrees of freedom = 5

p-value = <0.001

Test of homogeneity of level 1 variance

χ^2 statistic = 136.59649

Degrees of freedom = 131

p-value = 0.351

Having determined that the level 1 variable Certified is a good candidate to predict residual variance, Certified could be incorporated in the model in one of three ways: (1) It could be added as a level 1 variable not modeled as a random effect of level 2; (2) it could be added and also modeled as a random effect; or (3) it could be used as the predictor of level 1 residual variance by using a heterogeneous residual variance model. For these data, option (1) would yield results that fail the test of level 1 homogeneity of variance and that fail to show significantly better fit than the previous RC regression model. Option (2) would drop the deviance value enough to show significantly better fit but would also fail the test of level 1 homogeneity of residual variance. As shown in Table 3.10, option 3, the heterogeneous model, yields results that pass the homogeneity test (the homogeneity p-value is non-significant) and that also show significantly better fit than the RC regression model discussed earlier (the model comparison likelihood ratio test p-value is significant).

As shown in Figure 3.15, the heterogeneous model is identical to that shown in Figure 3.12 except that at level 1, terms are added that model residual variance as a function of the predictor variable Certified.

Fixed effects of the heterogeneous model are shown in Table 3.11. HLM prints out two fixed effects tables, one with robust standard errors and one with ordinary standard errors. Robust standard errors are advisable when there is misspecification of the distribution of the dependent variable. Therefore, significant differences between the ordinary and robust estimates of the standard error may flag a problem with the distribution specified by the researcher (normal is default). This does not appear to be a problem for this example, and

Figure 3.15 The heterogeneous random coefficients model in HLM 7

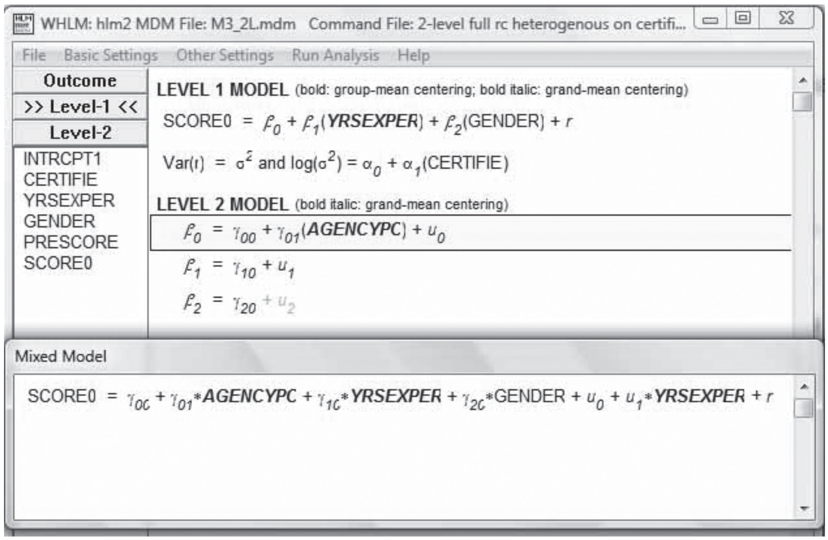


Table 3.11 Fixed Effects for the Heterogeneous Full Random Coefficients Model

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}	56.525630	0.820256	68.912	130	<0.001
AGENCYPC, γ_{01}	19.826460	2.563049	7.735	130	<0.001
For YRSEXPER slope, β_1					
INTRCPT2, γ_{10}	17.868835	0.817503	21.858	131	<0.001
For GENDER slope, β_2					
INTRCPT2, γ_{20}	–4.786541	0.426161	–11.232	3230	<0.001

one arrives at the same statistical inferences either way. For this example, all fixed effects are significant:

- The level 1 intercept of 56.5 gives the mean performance score across all agencies when other variables in the model are controlled at zero. Since YrsExper and

AgencyPC were centered, controlling means when both are at their mean values. Since Gender was coded such that 0 = male, controlling means "for men." That is, men with mean years of experience working in agencies with a mean percent of certified employees are predicted to have performance scores, on the average, of 56.5.
- The level 2 predictor AgencyPC (percent certified for a given agency) is significant. This variable, which ranged from .11 to 1.0, was centered when used in the HLM model. Its coefficient is positive, meaning that the higher the percent certified in an agency, the higher the intercept (hence the higher the mean performance score) for employees of the agency. Its slope, 19.8, means that when the percent certified goes up by 1 percent (which is .01 units on a scale from .11 to 1.0), mean performance scores are expected to rise by .198 units.
- The slope of the level 1 predictor variable Gender is significant and negative. Since Gender was coded 0 = male, 1 = female, a 1-unit increase means being female. That is, being female lowered the expected performance score by 4.79 points, on the average, controlling for other variables in the model.
- The slope of the level 1 predictor YrsExper was significant and positive. This variable was centered on its mean and expressed in standardized units ranging from −.79 to +.56 (not in raw years). The fact that it is positive means the more years experience, the higher the predicted performance score, controlling other variables in the model. That its t-ratio in absolute terms is about twice that for Gender means YrsExper has a greater effect on performance score than does Gender.

Table 3.12 shows the random effects components in the heterogeneous random coefficients model, of which there are two. Because residual variance is being modeled in a heterogeneous model, there is no residual variance component as, for instance, shown in Table 3.6 for a related homogenous model. The Intrcpt1 component is the effect on mean performance scores at level 1 due to Agency as a random factor at level 2. The YrsExper slope component is the effect on the slope of YrsExper and level 1 due to Agency as a random factor at level 2. Both effects are significant. Thus, both mean performance score and the regression coefficients for YrsExper vary significantly by Agency, confirming that multilevel analysis is required to properly model employee performance score at level 1. Coefficients in the fixed effects table discussed above will be more reliable than coefficients from the equivalent model in OLS regression due to the existence of these significant random effects. That the intercept component

Table 3.12 Random Effects for the Heterogeneous Random Coefficients Model

Final estimation of variance components

Random effect	*Standard deviation*	*Variance component*	*d.f.*	χ^2	p-*value*
INTRCPT1, u_0	8.89447	79.11164	130	1283.08288	<0.001
YRSEXPER slope, u_1	6.64525	44.15935	131	330.23181	<0.001

is significant even controlling for other variables in the model means that there remains significant variation, which might be explained by adding additional predictors to the model.

As a final example in this section, Figure 3.16 illustrates a full random coefficients model in HLM otherwise paralleling the foregoing one, but now with the level 2 predictor, centered agency percent certified (AgencyPC), used to model the level 1 slope of years of experience (YrsExper) as well as the level 1 intercept. As can be seen in Figure 3.16, this is simply a matter of adding AgencyPC to the level 2 equation that models the level 1 slope of YrsExper. In the level 2 equation, $\beta_{1j} = \gamma_{10} + \gamma_{11}*(AGENCYPC_j) + u_{1j}$, the u_{1j} term is the Agency effect on the slope of YrsExper and the $\gamma_{11}*(AGENCYPC_j)$ term is the AgencyPC effect on the slope.

Table 3.13 shows the deviance model fit statistic for this model. It is very close to the homogenous RC model in which AgencyPC modeled only the level 1 intercept and not the slope (Table 3.8). This strongly hints at the finding shown below in the fixed effects table, which shows the AgencyPC effect on the slope of YrsExper to be non-significant (Table 3.14). We conclude that agency percent certified at level 2 does not account for a significant portion of the variance in the strengths (slopes) of the relation of YrsExper with performance scores across

Figure 3.16 A full random coefficients model with a level 2 covariate modeling level 1 slope and intercept

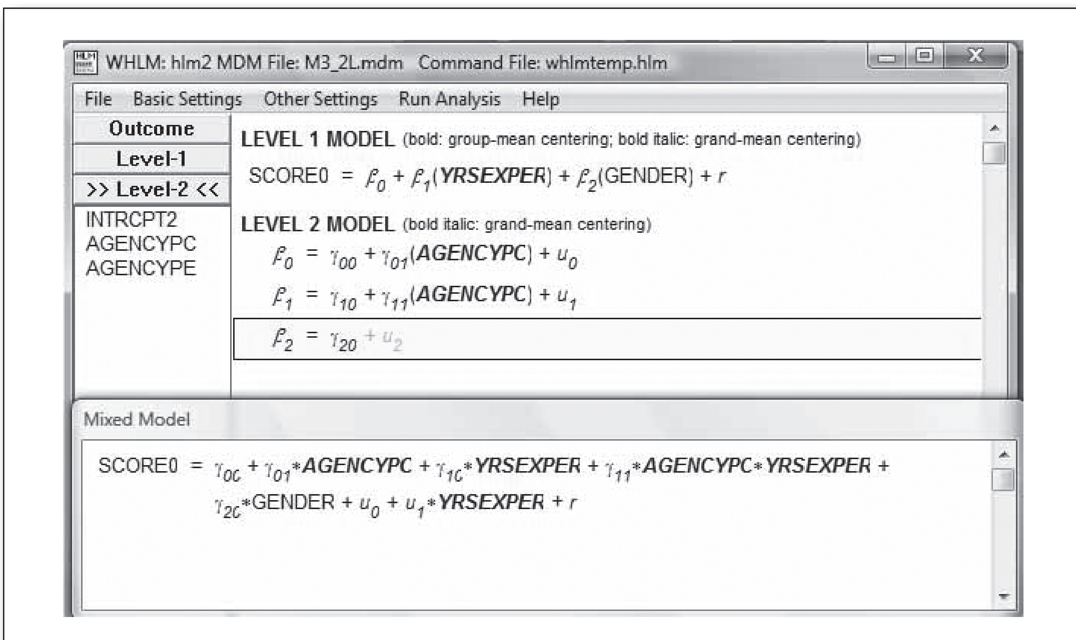


Table 3.13 Deviance and Likelihood Ratio Test of the Homogenous RC Model With Agency Modeling Both Slope and Intercept, Compared to the RC Regression Model

Statistics for current covariance components model
Deviance = 28156.789317
Number of estimated parameters = 4

Variance-Covariance components test
χ^2 statistic = 44.37074
Degrees of freedom = 0
p-value = >.500

agencies. Neither of these homogenous models has as good a fit as did the heterogeneous model just discussed. The likelihood ratio tests compare the full homogenous RC model (Table 3.8), the full heterogeneous RC model (Table 3.10), and the full homogenous RC model with agency percent certified modeling slope as well as intercept. (Table 3.13 shows only the heterogeneous model to be better fitting than the RC regression model with no level 2 predictors.)

Because the interpretation parallels that of full RC models discussed above, the variance components table for the homogenous RC model with agency percent certified modeling slope as well as intercept is not presented. Also, though not presented in table form, it may be noted that this homogenous full RC model just discussed fails the test of homogeneity of level 1 variance, as did the earlier homogenous full RC model. However, for pedagogic reasons and because SAS and SPSS cannot easily compute models for heterogeneous error variance, this is the model reproduced in SAS 9.2 in Chapter 4 and in SPSS 19 in Chapter 5. It should be noted that such testing for homogeneity of error variance routinely returns a finding of heterogeneity, yet it is common practice to execute homogenous models anyway, in part because the same substantive

Table 3.14 Fixed Effects for the Homogenous RC Model With AGENCYPC Modeling Both Slope and Intercept

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}	59.360637	0.952706	62.307	130	< 0.001
AGENCYPC, γ_{01}	30.975071	5.264133	5.884	130	< 0.001
For YRSEXPER slope, β_1					
INTRCPT2, γ_{10}	15.505130	1.077111	14.395	130	< 0.001
AGENCYPC, γ_{11}	–3.500075	6.720186	–0.521	130	0.603
For GENDER slope, β_2					
INTRCPT2, γ_{20}	–6.491174	0.433456	–14.975	3230	< 0.001

conclusions are often arrived at. Also, even if error variance is heterogeneous, it may not be correlated with predictor variables (though it is correlated with YrsExper for these data). How to check is discussed in the endnotes.[6]

Figure 3.17 Correlations of residual error variance with level 1 predictors

Correlations

		ResidVar	yrsexper_mean_1	gender_mean
ResidVar	Pearson Correlation	1	.462**	-.043
	Sig. (2-tailed)		.000	.626
	N	132	132	132
yrsexper_mean_1	Pearson Correlation	.462**	1	-.074
	Sig. (2-tailed)	.000		.401
	N	132	132	132
gender_mean	Pearson Correlation	-.043	-.074	1
	Sig. (2-tailed)	.626	.401	
	N	132	132	132

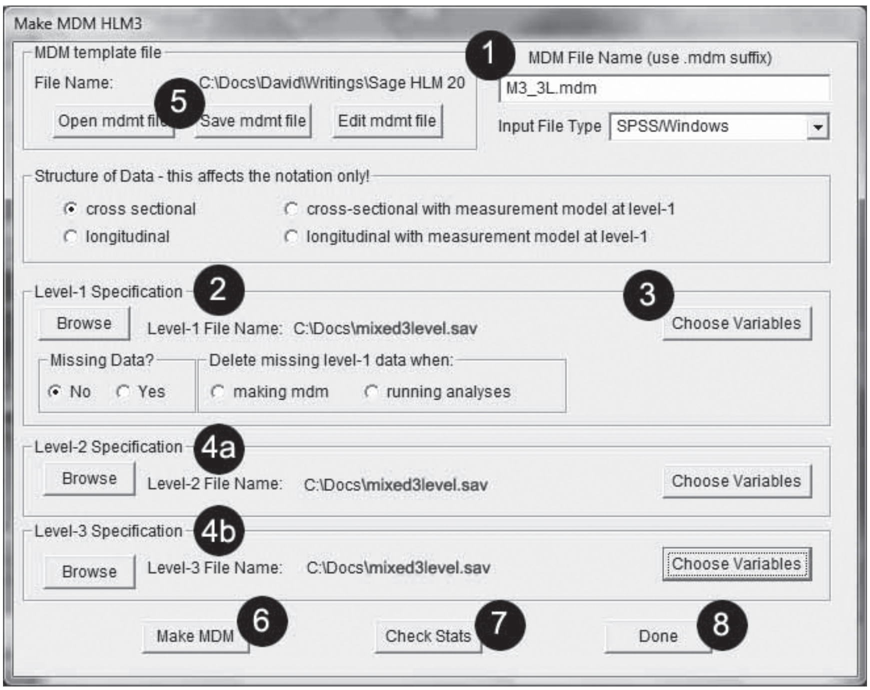
**. Correlation is significant at the 0.01 level (2-tailed).

THREE-LEVEL HIERARCHICAL LINEAR MODELS

Three-level problems occur frequently in real-world data: cross-sectional studies of students nested within classrooms nested within schools, for instance, or longitudinal studies of yearly tests nested within students nested within schools. To create a basic, cross-sectional three-level hierarchical linear model, we will use the same SPSS data file, mixed3level.sav, which had within it a previously undiscussed field for the third level: Department (Employees are nested within Agencies, and Agencies within Departments, with Department being the department ID variable) and two department-level covariates to serve as potential level 3 predictors. As HLM 7 has only minimal data management capabilities, like most HLM users, we initially create the dataset in another statistical package, in this case SPSS 19. SAS may also be used for this purpose, as HLM 7 reads both SPSS and SAS formats directly. While separate data files might be created for levels 1, 2, and 3, this is not necessary and was not done for the example data.

The first step is to import the data from SPSS .sav format and to create the HLM 7 multivariate data matrix file (.mdm format) and data template file (.mdmt format). This assumes that the data are previously sorted in a nested order: employees within agencies within departments. The .mdm file will be incorrect if data are not sorted properly in SPSS or other statistical packages used for data management. Three-level cross-sectional models are created in the HLM3 module. When HLM 7 opens, select File, Make new MDM file, Stat package input. A “Select MDM Type” window opens. Check the HLM3 radio button and click OK to arrive at the “Make MDM HLM3” window shown in Figure 3.18.

Figure 3.18 HLM 7 setup for three-level cross-sectional models (HLM3)



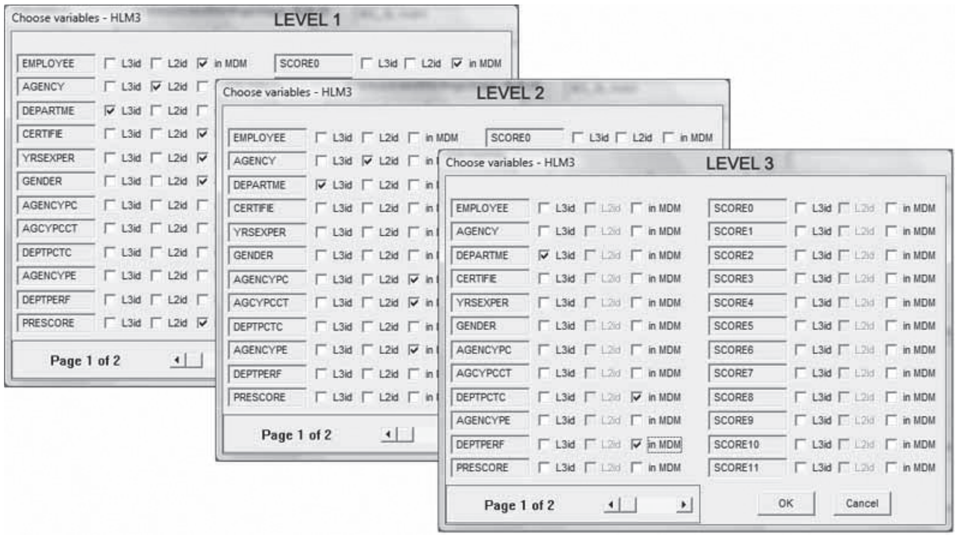
1. In the “Make MDM HLM3” window, follow these steps:

2. In the upper right, first enter the desired filename for the data matrix file. Here, it is M3_3L.mdm.

3. In the “Level-1 Specification” area, click the Browse button and browse to the SPSS .sav file and enter it. Also click the appropriate radio buttons regarding how to handle missing data, if any.

4a. Click the Choose Variables button to enter level 1 variables as shown in Figure 3.19. For level 1, enter all three ID variables (Department, Agency, and Employee)

Figure 3.19 Variable selection for three-level models in HLM 7



as shown, as well as entering any level 1 covariates and the level 1 dependent variable in “in MDM” checkboxes. For level 2, enter just Department and Agency, plus any level 2 predictors. For level 3, enter just Department, plus any level 3 predictors.

4b. Repeat for levels 2 and 3.

5. Save the template file.

6. Make (and save) the data matrix file.

7. Click “Check Statistics” to get a window showing the count of cases and descriptive statistics (mean, standard deviation, minimum, maximum) at each level. While this is an optional step, it is wise to check to see if data are being imported as expected.

8. Click “Done” to finish, and exit to the “WHLM: hlm3File” window, from which one may select options and run a three-level model. Alternatively, one may close and later run HLM 7, then select File, Create a new model from an existing .mdm file.

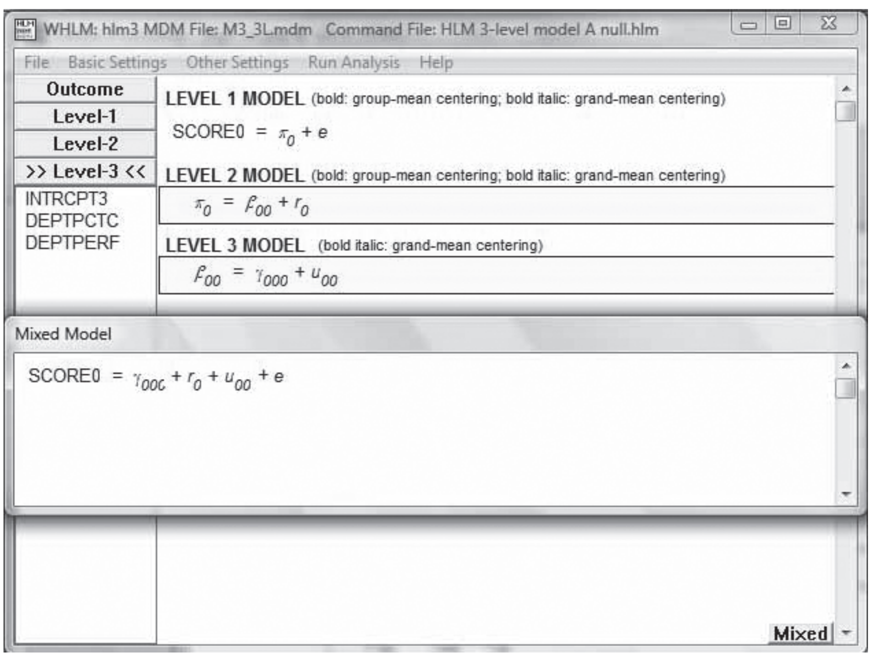
After the .mdm file is created, any of a variety of different cross-sectional three-level models may be created. With any model, the first analysis step is to specify the distribution of the dependent variable. For the example data, the dependent is employee performance score (Score0), which is normally distributed. This is the default selection, shown earlier in Figure 3.7. In the same “Basic

Model Specifications" window, the researcher should also specify the file location for the output and the title for the given run of the model.

Model A

Figure 3.20 shows the three-level null model in HLM 7 (compare to the two-level null model in Figure 3.6). For the null model, there are no predictors at any level. However, the intercept of performance score at level 1 is modeled for the Agency grouping effect at level 2 (signified by the r_0 random effects term in the level 2 equation). There is also a Department grouping effect at level 3 (signified by the u_{00} random effects term). The key element of the output is that the deviance is 28880.49, a value that can be compared with later models with predictors. Also, both the level 2 and level 3 intercepts were significant, confirming the existence of both a level 2 Agency effect and a level 3 Department effect. Finally, the residual within-group variance component was 212.40, a baseline value that will decrease as predictors are added to the model.

Figure 3.20 The three-level null model in HLM 7 (Model A)



Model B

Figure 3.21 displays the three-level random intercepts model, but one that adds a Department effect at level 3 to further model the level 1 intercept of performance score. At level 1, performance score is still predicted from years of experience and gender. The intercept of performance score is also still modeled as a random effect of Agency at level 2 and of Department at level 3. This is done by modeling the intercept of the level 2 Agency effect on the level 1 intercept, as a function of Department at level 3, as shown in the figure below.

Table 3.15 presents model fit and random effects output for this model. The deviance is now 28102.81, some 777.67 points lower in the direction of better model fit. The likelihood ratio (model comparison) test comparing the null model confirms the current model is a significantly better fit. This is also reflected in the fact that the within-Agencies residual variance component has dropped from 212.40 in the null model to 170.99 in the random intercepts model (Model B), since its effects explain some of the previously unexplained within-groups variance.

The variance components shown in the random effects tables have to do with partitioning the variance in performance scores at level 1. These components are

Figure 3.21 A three-level random intercepts model in HLM 7 (Model B)

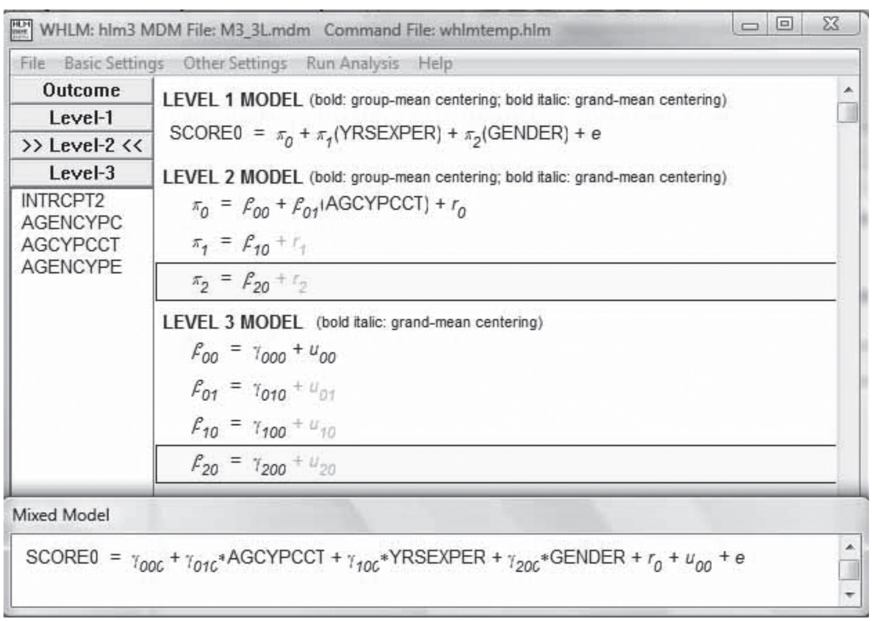


Table 3.15 Random Effects and Model Fit for the Three-Level Random Intercepts Model (Model B)

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	1.42285	2.02449	71	91.84951	0.049
level-1, *e*	13.07647	170.99418			

Final estimation of level 3 variance components

Random Effect	*Standard deviation*	*Variance component*	*d.f.*	χ^2	p-*value*
INTRCPT1/INTRCPT2,u_{00}	8.95949	80.27238	61	1325.17420	<0.001

Statistics for the current model

Deviance = 28102.809960

Number of estimated parameters = 7

Model comparison test

χ^2 statistic = 777.67997

Degrees of freedom = 3

p-value = <0.001

not to be interpreted in the fashion discussed below of estimates of fixed effects, which have an interpretation similar to regression. The Department effect on the level 1 intercept (which reflects mean performance score) is large (80.27) and highly significant (p = <.001). The Agency effect on the intercept, however, is small (2.02) and only barely significant (p = 0.049). Though we conclude there is both an Agency and a Department effect on performance score, the latter dominates the former. That the within-group residual variance is large (170.99) means that there is considerable within-Agency variation in performance scores as yet unexplained by modeled effects, suggesting the need to add additional effects and predictors to the model.

Table 3.16 shows the fixed effects in the three-level random intercepts model. These effects are interpreted similarly to regression. There are four fixed effects, all significant. The uppermost one (61.74, $p < 0.001$) is the estimate of the intercept, which is the mean performance score at level 1, controlling for other variables in the model. In descending order, the next (7.89, $p = 0.043$) is the effect of the level 2 predictor, centered agency percent certified (AGCYPCCT), on the level 1 intercept. This variable ranged from −.59 to +.30. A 0.1 increase in centered agency percent certified increases mean performance score by 0.79 points.

Table 3.16 Fixed Effects for the Three-Level Random Intercepts Model (Model B)

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}	61.740246	1.143603	53.987	61	<0.001
For AGCYPCCT, β_{01}					
INTRCPT3, γ_{010}	7.893035	3.822309	2.065	71	0.043
For YRSEXPER slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	16.316538	1.630712	10.006	3297	<0.001
For GENDER slope, π_2					
For INTRCPT2, β_{20}					
INTRCPT3, γ_{200}	–6.635115	0.403094	–16.460	3297	<0.001

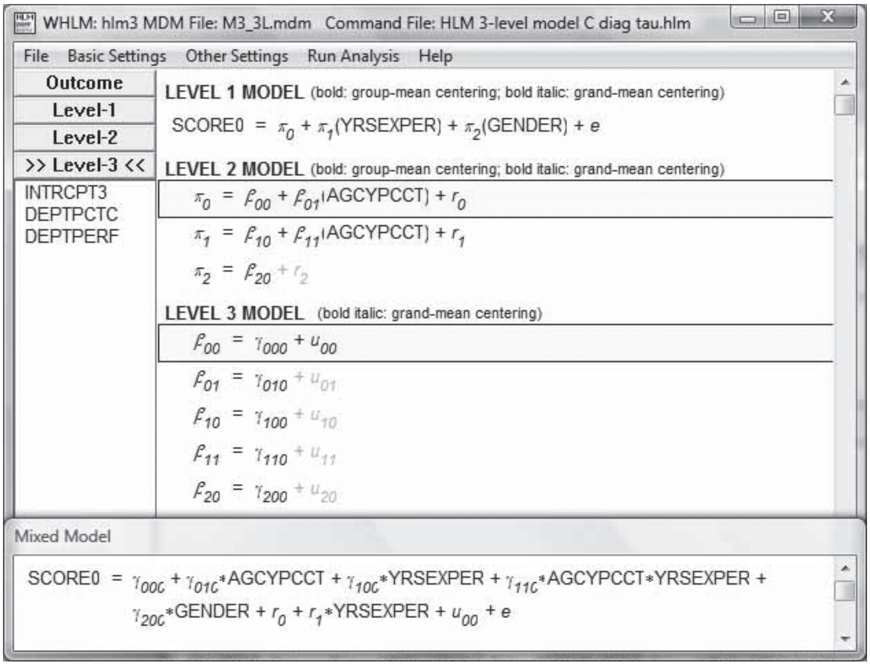
The next coefficient is the estimate of the slope of years of experience at level 1 (16.32, $p < 0.001$). This centered variable ranged from –.79 to + .56. A 0.1 increase in years of experience as coded increases mean performance score by 1.63 points. The t-ratios suggest that the level 1 effect of years of experience is stronger than that of agency percent certified at level 2. Finally, the Gender effect at level 1 is –6.35. Gender was coded 0 = male, 1 = female. A 1-unit increase in Gender (in other words, being female) is associated with a 6.35 drop in performance score below the mean of 61.74, controlling for other variables in the model. The t-ratios suggest that the Gender effect is somewhat stronger than the YrsExper effect, controlling for other variables in the model.

Model C

As a final three-level model illustration, we now create a random coefficients model, shown in Figure 3.22, where centered agency percent certified at level 2 is used as a random factor with respect to a random effect not only on the level 1 intercept of performance score, but also on the slope of years of experience at level 1 (compare to the two-level model in Figure 3.16, as well as the previous three-level random intercepts model in Figure 3.21).

The effect added in Model C is reflected in the second of the level 2 equations shown in Figure 3.22. Each equation is a potential random effect and hence associated with a variance component, but Model C invokes only three of the eight possible random effects. The variance components (random effects) table will

Figure 3.22 A three-level random coefficients model in HLM 7 (Model C)



have four estimates: the three shown as active (by virtue of their random error terms being toggled on and not grayed out) plus the within-groups residual effect. At level 2 (the Agency level), Model C invokes the first two random effects, which are the Agency effect on the level 1 intercept (mean performance score) and the Agency effect on the level 1 slope of the predictor YrsExper. The third potential level 1 random effect, the Agency effect on the slope of Gender, is not activated. At level 3, only the first potential random effect is activated: the Department effect on the level 1 intercept (created by way of an effect on the level 2 intercept for the equation representing the Agency effect on the level 1 intercept of performance score). The remaining four level 3 potential effects, not activated, are in descending order. They are all Department effects on (a) the slope of AGCYPCCT as a level 2 predictor of the level 1 intercept of performance score; (b) on the slope of YrsExper, by way of the level 2 equation for the Agency effect on this slope; (c) on the slope of AGCYPCCT as a level 2 predictor of the slope of YrsExper at level 1; and (d) on the slope of Gender, by way of the level 2 equation for the Agency effect on this slope. This recitation highlights the obvious fact that mixed models can quickly become very complex as levels and effects are added. It also highlights the usefulness of the HLM 7 user interface in tracking just what is being modeled more easily than other packages.

While this three-level model may seem only slightly different from the two-level model, adding Department effects, there is now a very important difference: There will now be two random effects (the random effects on the intercept and on the slope of YrsExper). Because there is now more than one between-groups effect (not counting the residual within-groups effect), the researcher must confront the issue of covariance of random effects. As discussed in Chapter 2, the default in HLM 7 approaches this issue quite differently from the default in SPSS or SAS, with the consequence that estimated coefficients may be significantly different. By default, HLM 7 estimates each coefficient in the entire covariance matrix, whereas by default SPSS and SAS impose a variance components (VC) constraint, which requires zero covariance among random effects. The advantage of the unstructured approach of HLM is that it may be more realistic to expect and allow non-zero covariances on the off-diagonal and heterogeneous variances on the diagonal (VC models constrain solutions to have zeros on the off-diagonal and homogenous variances on the diagonal.). The advantage of the VC approach is that it may result in equally good model fit even though a simpler model, and that the model may converge more quickly and with fewer problems. Refer back to Chapter 2 for a discussion of selecting an appropriate covariance type.

Different packages offer different covariance type models, which is to say, different types of constraints on the variance–covariance matrix. For two- and three-level cross-sectional models of the type discussed in this chapter, HLM 7 offers only the default unstructured type and a diagonal type (invoked from the "Estimation" window shown in Figure 3.8 by checking the "Diagonalize tau" radio button). The diagonal model constrains the solution to one with covariances of 0 (like VC) but allows heterogeneous variances (unlike VC). SPSS and SAS also support the DIAG covariance type, though the default VC type results are often very similar or identical. When there is the possibility of covarying random effects, HLM 7 results often will be closest to those in SAS and SPSS if the comparison is between (a) an HLM 7 model with diagonalized tau vs. an SPSS or SAS model with DIAG or VC covariance type, or (b) a default HLM 7 model versus an SPSS or SAS model with UN (unstructured) type. For reasons given in Chapter 2, such as differences in estimation algorithms, estimates may still vary between packages.

Not shown here, the model in Figure 3.21 was run on the default (unstructured) basis in HLM 7 and on a diagonalized basis. A likelihood ratio test was run to compare the two models. While the unstructured approach did yield a lower deviance (lower is better model fit), it was not enough lower to be significant. Therefore, the simpler, diagonalized model is discussed below. The SAS and SPSS parallel examples in Chapters 4 and 5, respectively, illustrate the default variance components model, but for these data, the VC and DIAG models generate very similar estimates.

Table 3.17 shows model fit and random effects for the final three-level random coefficients model, with the slope of years of experience modeled as a random

Table 3.17 Random Effects and Model Fit for the Final Three-Level Random Coefficients Model (Model C)

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	1.58976	2.52733	70	125.63509	<0.001
YRSEXPER slope, r_1	7.84488	61.54207	131	301.72408	<0.001
level 1, *e*	12.80223	163.89697			

Note. The chi-square statistics reported above are based on only 133 of 134 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*
INTRCPT1/INTRCPT2, u_{00}	9.34926	87.40861	61	1258.08679	<0.001

Statistics for the current model

Deviance = 28055.839714

Number of estimated parameters = 10

Model comparison test

χ^2 statistic = 47.05989

Degrees of freedom = 3

p-value = <0.001

effect of Agency, and AGCYPCCT. The likelihood ratio test for model comparison with the previous model (Model B) shows a reduction in deviance of 47.06, an amount significant at $p < 0.001$. Likewise, we see that the within-Agency residual variance component has dropped further, from 170.99 in the previous model (B) compared to 163.89 in this model (C). This suggests that the changes made in Model C compared to Model B were significant.

There are three other variance components beyond the residual, which HLM 7 labels "level-1, *e*" and which reflects as-yet unexplained within-agency variation in performance scores. All other components reflect between-group (in this example, between-agency and between-department) variance. In the "Final estimation of level 1 and level 2 variance components" section of Table 3.17, the "INTRCPT1" term is variance associated with the Agency effect. Agency as the level 2 subjects (grouping) variable has a small (2.53) but significant ($p < 0.001$) effect on mean performance score (the intercept). The "YRSEXPER slope" term is variance associated with random effects on the slope of years of experience as a predictor of level 1 performance scores. This random effect, which was modeled as a function of both Agency and AGCYPCCT at level 2, is moderately large (61.54) and significant ($p < 0.001$). In the lower, "Final estimation of level-3 variance components" portion of Table 3.17, the single intercept term reflects variance associated with the Department effect on the intercept of performance score. It is large (87.41) and significant ($p < 0.001$).

Table 3.18 displays the fixed effects for Model C. There are now five fixed effects: all those for the previous model (see Table 3.16) plus a new one for the random effect of agency percent certified centered (AGCYPCCT) at level 2 on the slope of YrsExper at level 1. This added effect has the lowest t-ratio and is the least significant, though it is significant by the usual .05 alpha criterion.

In Table 3.18, the uppermost coefficient is the estimate (61.74) of the intercept or mean value of performance score, controlling for other variables in the model.

Table 3.18 Fixed Effects for the Final Three-Level Random Coefficients Model (Model C)

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}	61.742258	1.241603	49.728	61	<0.001
For AGCYPCCT, β_{01}					
INTRCPT3, γ_{010}	10.517761	4.065604	2.587	69	0.012
For YRSEXPER slope, π_1					
For INTRCPT2, β_{10}					
INTRCPT3, γ_{100}	15.571292	1.360052	11.449	69	<0.001
For AGCYPCCT, β_{11}					
INTRCPT3, γ_{110}	2.641930	9.479845	0.279	69	0.781
For GENDER slope, π_2					
For INTRCPT2, β_{20}					
INTRCPT3, γ_{200}	−6.579907	0.409705	−16.060	3164	<0.001

It is virtually identical to that in Model B (Table 3.16). In descending order, the next coefficient (10.52) is the effect of agency percent certified centered (AGCYPCCT) on the level 1 intercept, now estimated to be larger than in Model B but interpreted similarly. Next is the estimate of the slope of years of experience at level 1 (15.57), quite close to that in Model B and also interpreted similarly. The fourth coefficient down in Table 3.18 is the AGCYPCCT effect on the slope of YrsExper as a level 1 predictor of performance score. This effect is non-significant. Finally, the Gender effect at level 1 is −6.58, little changed from the previous model.

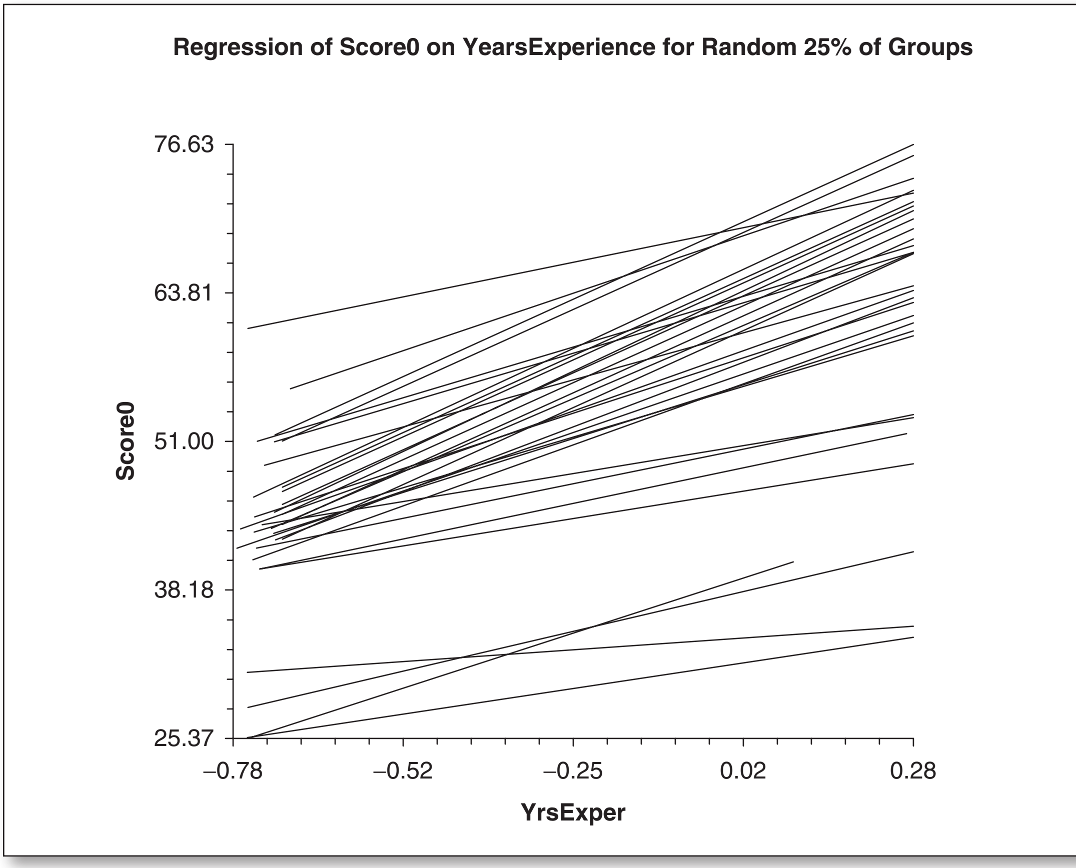
We conclude that there is both a Department and an Agency effect in explaining variance in mean performance score, with the Department effect being much greater than the Agency effect. There is also an Agency-level effect on the variance in the strength of relationship (slope) of YrsExper as a predictor of performance scores. At level 1, YrsExper and Gender are both significant predictors of performance score. We also conclude that while agency percent certified has a significant effect on mean employee performance score, it does not have a significant effect on the strength of relationship between years of experience and score. The model could be revised by dropping this non-significant effect. The remaining large residual variance component also indicates that additional effects and variables may be needed to fully explain the variance in employee-level performance scores.

GRAPHICS IN HLM 7

Before closing this chapter, it should be noted that HLM 7 also offers a variety of graphics options useful for analysis and diagnostics. These options are described in Chapter 18 of the HLM 7 manual by Raudenbush et al. (2011). Options are also illustrated in the present volume, in Chapters 8 and 10. To illustrate briefly here, Figure 3.23 below shows level 1 equation graphing for the two-level full random coefficients model described earlier in this chapter. Recall that linear mixed modeling computes separate regressions for each level 2 group (agencies for this example). Figure 3.23 shows the regression lines for 25% of the 132 agencies, predicting Score0 from YrsExper. Although it is true for all agencies that scores increase as experience increases, Figure 3.23 shows graphically that (a) intercepts differ by group, with some agencies scoring higher than other agencies at all levels of experience; and (b) slopes differ by group, with the steepness of the regression lines showing that within some agencies, experience is more related to score than for other agencies.

Figure 3.23 was obtained from the HLM 7 menu by selecting File, Graph equations, Level 1 equation graphing; then, in the "Level 1 Equation Graphing" dialog, the x-focus was made to be YrsExper and the number of groups set to "random Sample" with a probability of .25; and in the same dialog, under "Range/Titles/Color," the graph title was set.

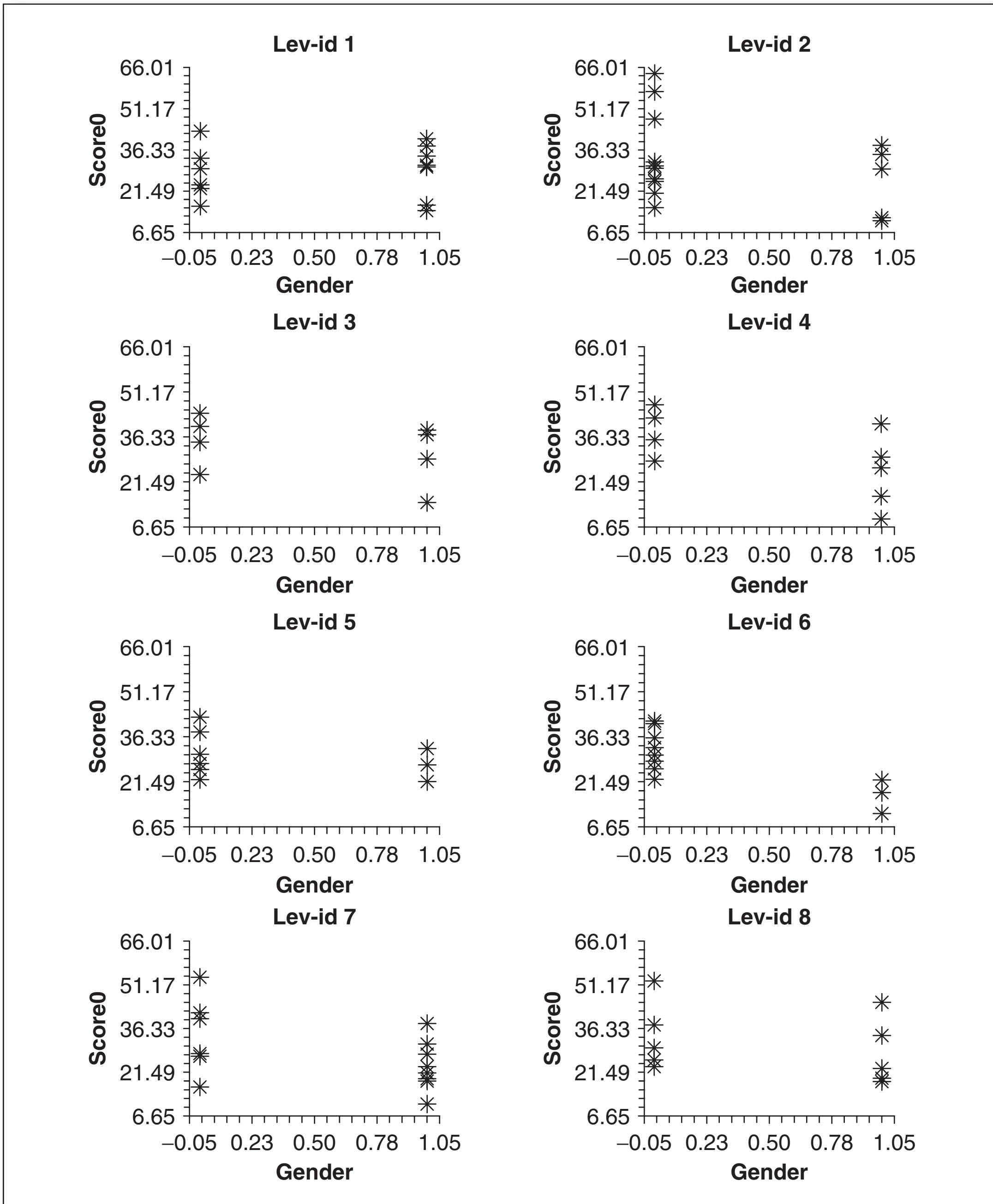
Figure 3.23 Score0 by YrsExper for 25% of the groups in the two-level full random coefficients model



As a second illustration, Figure 3.24 displays the relationship of Gender to Score0 for the first 10 groups in the same study. With men coded 0 and women coded 1, the graphs show that men tend to score higher than women, though this tendency is more pronounced in some groups than others. This figure was obtained from the HLM 7 menu by selecting File, Graph Data, "Line plots, scatter plots"; then, the number of groups was set to "First 10 groups."

Other HLM 7 graphing options support plots of residuals versus predicted values, residual box-whisker plots, coefficient confidence intervals, and more. For all figures, in the figure display window one may select File, Save As, and save the figure in either extended metafile format (.emf, compatible with Microsoft Word's Insert, Picture command and with Microsoft PowerPoint) or Windows metafile format (.wmf, compatible with Microsoft Office and with Adobe Illustrator).

Figure 3.24 Score0 by Gender for first 10 groups in the two-level full random coefficients model



SUMMARY

As this is intentionally an introductory guide and due to space limitations, this chapter has sought to present only some of the most important types of linear mixed models that may be implemented with HLM 7 software. Multilevel modeling supports a far, far richer array of models and variations, some of which are treated in subsequent application chapters in this volume. These include cohort models, longitudinal and growth models, and cross-classified models not discussed in this chapter, to name a few. Moreover, generalized linear mixed modeling software extends many of these models to ones using nonlinear link functions where the outcome variable is not a normally distributed continuous variable (ex., to binomial, multinomial, ordinal, or Poisson distributions, implemented in GLMM in SPSS, GLIMMIX in SAS, and built into HLM 7). Additional types of models are presented in comprehensive texts on multilevel modeling, which this work is intended to introduce and for which it may serve as a supplement. While only the surface of what is possible with linear mixed modeling is presented in this volume, the basic principles concerning types of effects, model assumptions, and inference logic apply to all types and hopefully serve as a stepping stone toward more advanced work.

NOTES

1. At this writing, the student version is available at http://www.ssicentral.com/hlm/downloads/HLM7StudentSetup.exe.

2. Method 2 is supported for HLM 6 and higher.

3. For pedagogical reasons, the default REML output has been illustrated. However, keep in mind the admonitions in Chapter 2 that model fit comparisons using likelihood ratio tests or information criteria measures when models differ in fixed effects require ML estimation.

4. For reasons of comparability to parallel SPSS and SAS output, output for this and subsequent examples (except the heterogeneous model) is obtained based on a diagonalized tau constraint, explained later in this chapter.

5. To create Figure 3.14, in HLM 7, select Basic Settings, which opens the "Basic Model Specifications" window. Click the "Level 1 Residual File" to save residuals. (The default filename will be resfil1.sav if SPSS format is desired. SAS, Stata, and ASCII alternatives are available.) In the .sav file, the agency id field is saved as I2id, signifying it is the variable designating level 2 units. The residual values for each employee are saved as I1resid, signifying level 1 residuals. In SPSS, to aggregate the residuals by agency and compute the variance, select Analyze, Compare Means, and in the "Means" window, let I1resid be the dependent and I2id the independent; and under the Options button, let variance be the only statistic. In SPSS syntax, the command is as follows: MEANS TABLES=l1resid BY l2id / CELLS VAR. The resulting report lists residual variances by agency id. This is copied into Excel, where the table is sorted in ascending order by I1resid. Block out the sorted column and select Insert, and in the Chart Layouts block, select Scatter as the chart type. Click on a point to select all points, and then right-click and select "Add trend line." To label outlier points, block both the I2id and I1resid columns; double-click on a point to select it, and then

right click and choose "Add data label. Then re-select the point, right click, select "Format data label," and set the contents field and label alignment.

6. In SPSS, with the residual file () discussed previously in this chapter as the active dataset, select Data, Aggregate, and let Agency be the "Break Variable"; move the predictor variables and the residual variable (I1resid by default) into the "Aggregated Variables" area, and then click on the "Functions" button to let the statistic for the residual variable be the standard deviation (variance is not an option). The other predictor variable statistics will default to means. Choose to save the aggregated file under a different filename. After running the Aggregate function, open the file (aggr.sav by default) and use Transform, Compute Variable, to create a new residual variance variable based on the residual standard deviation. Then select Analyze, Correlate, to obtain the correlations of residual variance by Agency with the predictor variables. For the example data, output is shown in Figure 3.17.

REFERENCES

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods* (1st ed.). Newbury Park, CA: Sage.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, A. S., Fai, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*(4), 323–355.