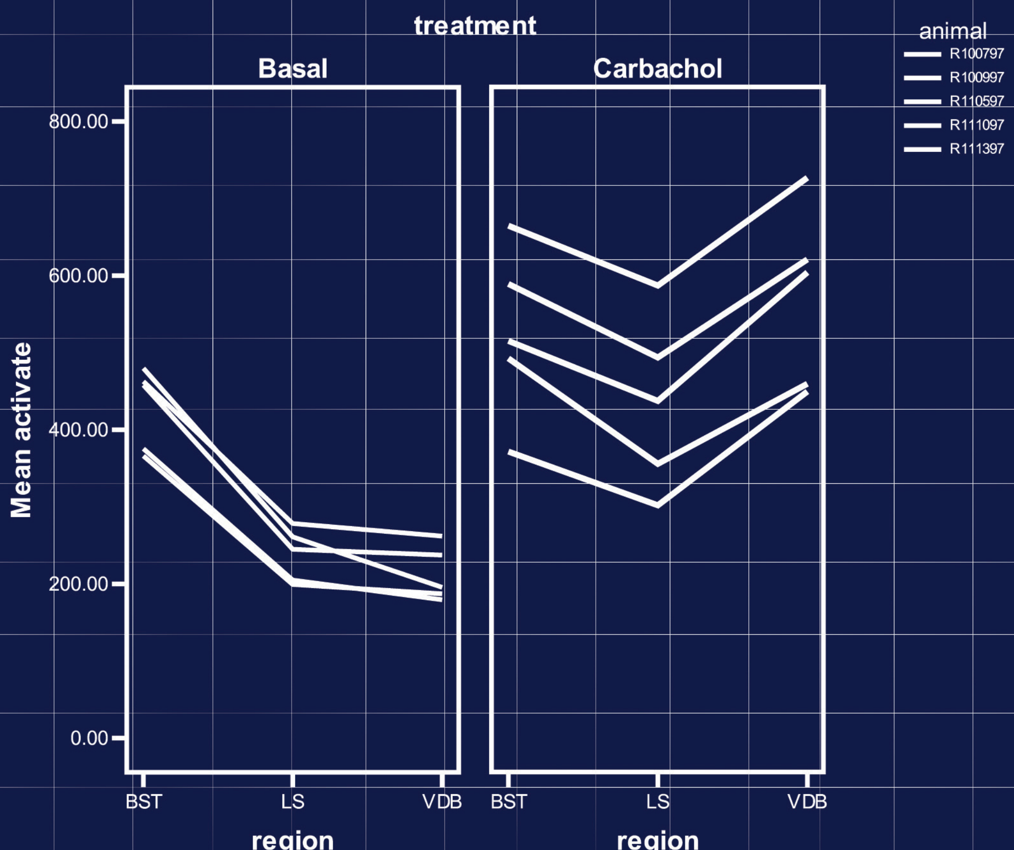
LINEAR MIXED MODELS

A Practical Guide Using Statistical Software

THIRD EDITION



Brady T. West
Kathleen B. Welch
Andrzej T. Gałecki
with contributions from Brenda W. Gillespie

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

4

Three-Level Models for Clustered Data: The Classroom Example

4.1 Introduction

In this chapter, we illustrate models for clustered study designs having three levels of data. In three-level clustered data sets, the units of analysis (Level 1) are nested within randomly sampled clusters (Level 2), which are in turn nested within other larger randomly sampled clusters (Level 3). Such study designs allow us to investigate whether covariates measured at each level of the data hierarchy have an impact on the dependent variable, which is always measured on the units of analysis at Level 1.

Designs that lead to three-level clustered data sets can arise in many fields of research, as illustrated in Table 4.1. For example, in the field of education research, as in the Classroom data set analyzed in this chapter, students' math achievement scores are studied by first randomly selecting a sample of schools, then sampling classrooms within each school, and finally sampling students from each selected classroom (Figure 4.1). In medical research, a study to evaluate treatments for blood pressure may be carried out in multiple clinics, with several doctors from each clinic selected to participate and multiple patients treated by each doctor participating in the study. In a laboratory research setting, a study of birth weights in rat pups similar to the Rat Pup study that we analyzed in Chapter 3 might have replicate experimental runs, with several litters of rats involved in each run, and several rat pups in each litter.

In Table 4.1, we provide possible examples of three-level clustered data sets in different research settings. In these studies, the dependent variable is measured on one occasion for each Level 1 unit of analysis, and covariates may be measured at each level of the data hierarchy. For example, in a study of student achievement, the dependent variable, math achievement score, is measured once for each student. In addition, student characteristics such as age, classroom characteristics such as class size, and school characteristics such as neighborhood poverty level are all measured. In the multilevel models that we fit to three-level data sets, we relate the effects of covariates measured at each level of the data to the dependent variable. Thus, in an analysis of the study of student achievement, we might use student characteristics to help explain between-student variability in math achievement scores, classroom characteristics to explain between-classroom variability in the classroom-specific average math scores, and school-level covariates to help explain between-school variation in school-specific mean math achievement scores.

Figure 4.1 illustrates the hierarchical structure of the data for an educational study similar to the Classroom study analyzed in this chapter. This figure depicts the data structure for a single (hypothetical) school with two randomly selected classrooms.

Models applicable for data from such sampling designs are known as **three-level hierarchical linear models (HLMs)** and are extensions of the two-level models introduced in Chapter 3. Two-level, three-level, and higher-level models are generally referred to as **multilevel models**. Although multilevel models in general may have random effects associated

DOI: 10.1201/9781003181064-4

TABLE 4.1: Examples of Three-Level Data in Different Research Settings

Level of Data		Research Setting		
		Education	**Medicine**	**Biology**
Level 3	**Cluster of clusters (random factor)**	School	Clinic	Replicate
	Covariates	School size, poverty level of neighborhood surrounding the school	Number of doctors in the clinic, clinic type (public or private)	Experimental Run, Instrument calibration, ambient temperature, run order
Level 2	**Cluster of units (random factor)**	Classroom	Doctor	Litter
	Covariates	Class size, years of experience of teacher	Specialty, years of experience	Litter size, weight of mother rat
Level 1	**Unit of analysis**	Student	Patient	Rat Pup
	Dependent variable	Test scores	Blood pressure	Birth weight
	Covariates	Sex, age	Age, illness severity	Sex

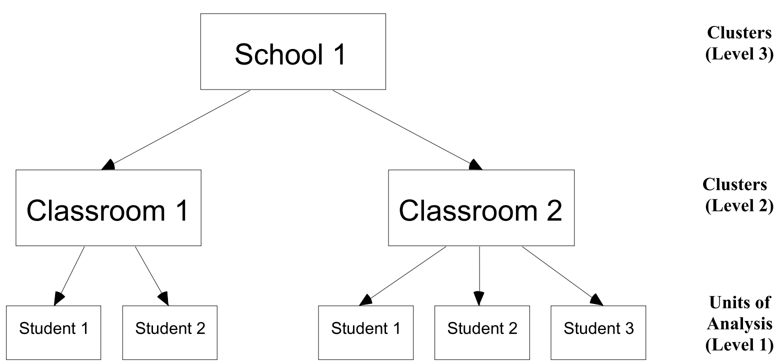


FIGURE 4.1: Nesting structure of a clustered three-level data set in an educational setting.

with both the intercept and with other covariates, in this chapter we restrict our discussion to **random intercept models**, in which random effects at each level of clustering are associated with the intercept. The random effects associated with the clusters at Level 2 of a three-level data set are often referred to as **nested random effects**, because the Level 2 clusters are nested within the clusters at Level 3. In the absence of fixed effects associated with covariates, a three-level HLM is also known as a **variance components model**. The HLM software (Raudenbush et al., 2005; Raudenbush & Bryk, 2002), which we highlight in this chapter, was developed specifically for analyses involving these and related types of models. We specifically focus on Version 8.1.4.10 of the HLM software in this third edition of the book.

4.2 The Classroom Study

4.2.1 Study Description

The Study of Instructional Improvement[1], or SII (Hill, Rowan, and Ball, 2005), was carried out by researchers at the University of Michigan to study the math achievement scores of first- and third-grade students in randomly selected classrooms from a national U.S. sample of elementary schools. In this example, we analyze data for 1,190 first-grade students sampled from 312 classrooms in 107 schools. The dependent variable, MATHGAIN, measures change in student math achievement scores from the spring of kindergarten to the spring of first grade.

The SII study design resulted in a three-level data set, in which students (Level 1) are nested within classrooms (Level 2), and classrooms are nested within schools (Level 3).

We examine the contribution of selected student-level, classroom-level and school-level covariates to the variation in students' math achievement gain. Although one of the original study objectives was to compare math achievement gain scores in schools participating in comprehensive school reforms (CSRs) to the gain scores from a set of matched comparison schools not participating in the CSRs, this comparison is not considered here.

A sample of the Classroom data is shown in Table 4.2. The first two rows in the table are included to distinguish between the different types of variables; in the actual electronic data set, the variable names are defined in the first row. Each row of data in Table 4.2 contains the school ID, classroom ID, student ID, the value of the dependent variable, and values of three selected covariates, one at each level of the data.

The layout of the data in the table reflects the hierarchical nature of the data set. For instance, the value of MATHPREP, a classroom-level covariate, is the same for all students in the same classroom, and the value of HOUSEPOV, a school-level covariate, is the same for all students within a given school. Values of student-level variables (e.g., the dependent variable, MATHGAIN, and the covariate, SEX) vary from student to student (row to row) in the data.

The following variables are considered in the analysis of the Classroom data:

[1]Work on this study was supported by grants from the U.S. Department of Education to the Consortium for Policy Research in Education (CPRE) at the University of Pennsylvania (Grant # OERI-R308A60003), the National Science Foundation's Interagency Educational Research Initiative to the University of Michigan (Grant #s REC-9979863 & REC-0129421), the William and Flora Hewlett Foundation, and the Atlantic Philanthropies. Opinions expressed in this book are those of the authors and do not reflect the views of the U.S. Department of Education, the National Science Foundation, the William and Flora Hewlett Foundation, or the Atlantic Philanthropies.

TABLE 4.2: Sample of the Classroom Data Set

School (Level 3)		Classroom (Level 2)		Student (Level 1)		
Cluster ID	Covariate	Cluster ID	Covariate	Unit ID	Dependent Variable	Covariate
SCHOOLID	HOUSEPOV	CLASSID	MATHPREP	CHILDID	MATHGAIN	SEX
1	0.0817	160	2.00	1	32	1
1	0.0817	160	2.00	2	109	0
1	0.0817	160	2.00	3	56	1
1	0.0817	217	3.25	4	83	0
1	0.0817	217	3.25	5	53	0
1	0.0817	217	3.25	6	65	1
1	0.0817	217	3.25	7	51	0
1	0.0817	217	3.25	8	66	0
1	0.0817	217	3.25	9	88	1
1	0.0817	217	3.25	10	7	0
1	0.0817	217	3.25	11	60	0
2	0.0823	197	2.50	12	2	1
2	0.0823	197	2.50	13	101	0
2	0.0823	211	2.33	14	30	0
2	0.0823	211	2.33	15	65	0
...						

Note: "..." indicates portion of the data not displayed.

School (Level 3) Variables

- **SCHOOLID** = School ID number[a]
- **HOUSEPOV** = Proportion of households in the neighborhood of the school below the poverty level

Classroom (Level 2) Variables

- **CLASSID** = Classroom ID number[a]
- **YEARSTEA**[b] = First-grade teacher's years of teaching experience
- **MATHPREP** = First-grade teacher's mathematics preparation: number of mathematics content and methods courses
- **MATHKNOW**[b] = First-grade teacher's mathematics content knowledge; based on a scale based composed of 30 items (higher values indicate higher content knowledge)

Student (Level 1) Variables

- **CHILDID** = Student ID number[a]
- **MATHGAIN** = Student's gain in math achievement score from the spring of kindergarten to the spring of first grade (the dependent variable)

- **MATHKIND**[b] = Student's math score in the spring of their kindergarten year
- **SEX** = Indicator variable (0 = boy, 1 = girl)
- **MINORITY**[b] = Indicator variable (0 = non-minority student, 1 = minority student)
- **SES**[b] = Student socioeconomic status

[a]The original ID numbers in the study were randomly reassigned for the data set used in this example.
[b]Not shown in Table 4.2.

4.2.2 Data Summary

The descriptive statistics and plots for the Classroom data presented in this section are obtained using the HLM software (Version 8.1.4.10). Syntax to carry out these descriptive analyses using the other software packages is included on the book's website (see Appendix A).

4.2.2.1 Data Set Preparation

To perform the analyses for the Classroom data set using the HLM software, we need to prepare three separate data sets:

1. The *Level 1 (student-level) Data Set* has one record per student, and contains variables measured at the student level, including the dependent variable, MATHGAIN, the student-level covariates, and the identifying variables: SCHOOLID, CLASSID, and CHILDID. The Level 1 data set is sorted in ascending order by SCHOOLID, CLASSID, and CHILDID.
2. The *Level 2 (classroom-level) Data Set* has one record per classroom, and contains classroom-level covariates, such as YEARSTEA and MATHKNOW, that are measured for the teacher. Each record contains the identifying variables SCHOOLID and CLASSID, and the records are sorted in ascending order by these variables.
3. The *Level 3 (school-level) Data Set* has one record per school, and contains school-level covariates, such as HOUSEPOV, and the identifying variable, SCHOOLID. The data set is sorted in ascending order by SCHOOLID.

The Level 1, Level 2, and Level 3 data sets can easily be derived from a single data set having the "long" structure illustrated in Table 4.2. All three data sets should be stored in a format readable by HLM, such as ASCII (i.e., raw data in text files), or a data set specific to a statistical software package. For ease of presentation, we assume that the three data sets for this analysis have been set up in SPSS format (Version 28). These three SPSS data files can be downloaded from the book's website.

4.2.2.2 Preparing the Multivariate Data Matrix (MDM) File

We prepare two MDM files for this initial data summary. One includes only variables that do not have any missing values (i.e., it excludes MATHPREP and MATHKNOW), and consequently, all students ($n = 1190$) are included. The second MDM file contains all variables, including MATHPREP and MATHKNOW, and thus includes complete cases ($n = 1081$) only.

We create the first **Multivariate Data Matrix (MDM)** file using the Level 1, Level 2, and Level 3 data sets defined earlier. After starting HLM, locate the main menu, and click on **File**, **Make new MDM file**, and then **Stat package input**. In the dialog box that opens, select **HLM3** to fit a three-level hierarchical linear model with nested random effects, and click **OK**. In the next window, choose the **Input File Type** as SPSS/Windows.

In the **Level-1 Specification** area of this window, we select the Level 1 data set. **Browse** to the location of the student-level file. Click **Choose Variables**, and select the following variables: CLASSID (check "L2id," because this variable identifies clusters of students at Level 2 of the data), SCHOOLID (check "L3id," because this variable identifies clusters of classrooms at Level 3 of the data), MATHGAIN (check "in MDM," because this is the dependent variable), and the student-level covariates MATHKIND, SEX, MINORITY, and SES (check "in MDM" for all of these variables). Click **OK** when finished selecting these variables.

In the **Missing Data?** box, select "Yes," because some students may have missing values on some of the variables. Then choose **Delete missing level-1 data when: running analyses**, so that observations with missing data will be deleted from individual analyses and not from the MDM file.

In the **Level-2 Specification** area of this window, select the Level 2 (classroom-level) data set defined above. **Browse** to the location of the classroom-level data set. In the **Choose Variables** dialog box, choose the CLASSID variable (check "L2id") and the SCHOOLID variable (check "L3id"). In addition, select the classroom-level covariate that has complete data for all classrooms[2], which is YEARSTEA (check "in MDM"). Click **OK** when finished selecting these variables.

In the **Level-3 Specification** area of this window, select the Level 3 (school-level) data set defined earlier. **Browse** to the location of the school-level data set, and choose the SCHOOLID variable (click on "L3id") and the HOUSEPOV variable, which has complete data for all schools (check "in MDM"). Click **OK** when finished.

Once all three data sets have been identified and the variables have been selected, go to the **MDM template file** portion of the window, where you will see a white box for an **MDM File Name** (with an .mdm suffix). Enter a name for the MDM file (such as classroom.mdm), including the .mdm suffix, and then click **Save mdmt file** to save an MDM Template file that can be used later when creating the second MDM file. Finally, click on **Make MDM** to create the MDM file using the three input files.

After HLM finishes processing the MDM file, click **Check Stats** to view descriptive statistics for the selected variables at each level of the data, as shown in the following HLM output:

[2]We do not select the classroom-level variables MATHPREP and MATHKNOW when setting up the first MDM file, because information on these two variables is missing for some classrooms, which would result in students from these classrooms being omitted from the initial data summary.

```
                    LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
 SEX              1190       0.51        0.50       0.00          1.00
 MINORITY         1190       0.68        0.47       0.00          1.00
 MATHKIND         1190     466.66       41.46     290.00        629.00
 MATHGAIN         1190      57.57       34.61    -110.00        253.00
 SES              1190      -0.01        0.74      -1.61          3.21

                   LEVEL-2 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
YEARSTEA           312      12.28        9.65       0.00         40.00

                   LEVEL-3 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
HOUSEPOV           107       0.19        0.14       0.01          0.56
```

Because none of the selected variables in the MDM file have any missing data, this is a descriptive summary for all 1190 students, within the 312 classrooms and 107 schools. We note that 51% of the 1190 students are female and that 68% of them are of minority status. The average number of years teaching for the 312 teachers is 12.28, and the mean proportion of households in poverty in the neighborhoods of the 107 schools is 0.19 (19%).

We now construct the second MDM file. After closing the window containing the descriptive statistics, click **Choose Variables** in the **Level-2 Specification** area. At this point, we add the MATHKNOW and MATHPREP variables to the MDM file by checking "in MDM" for each of these variables; click **OK** after selecting them. Then, save the .mdm file and the .mdmt template file under different names and click **Make MDM** to generate a new MDM file containing these additional Level 2 variables. After clicking **Check Stats**, the following output is displayed:

```
                   LEVEL-1 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
      SEX         1081       0.50        0.50          0.00       1.00
 MINORITY         1081       0.67        0.47          0.00       1.00
 MATHKIND         1081     467.15       42.00        290.00     629.00
 MATHGAIN         1081      57.84       34.70        -84.00     253.00
      SES         1081      -0.01        0.75         -1.61       3.21

                   LEVEL-2 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
 YEARSTEA          285      12.28        9.80          0.00      40.00
 MATHKNOW          285      -0.08        1.00         -2.50       2.61
 MATHPREP          285       2.58        0.96          1.00       6.00

                   LEVEL-3 DESCRIPTIVE STATISTICS

VARIABLE NAME       N       MEAN          SD        MINIMUM     MAXIMUM
HOUSEPOV           105       0.20        0.14          0.01       0.56
```

Note that 27 classrooms have been dropped from the Level 2 file because of missing data on the additional Level 2 variables, MATHKNOW and MATHPREP. Consequently, there

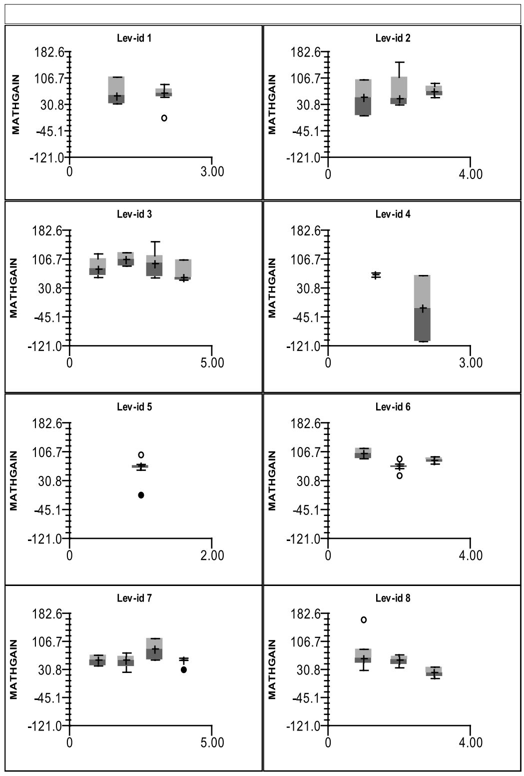


FIGURE 4.2: Boxplots of the MATHGAIN responses for students in the selected classrooms in the first eight schools for the Classroom data set.

were two schools (SCHOOLID = 48 and 58) omitted from the **Check Stats** summary, because none of the classrooms in these two schools had information on MATHKNOW and MATHPREP. This resulted in only 1081 students being analyzed in the data summary based on the second MDM file, which is 109 fewer than were included in the first summary. We close the window containing the descriptive statistics and click **Done** to proceed to the HLM model-building window.

In the model-building window we continue the data summary by examining boxplots of the observed MATHGAIN responses for students in the selected classrooms from the first eight schools in the first (complete) MDM file. (Note: before generating plots, you may need to fit a simple intercept-only model in HLM. To do this in the model building window, select MATHGAIN as the outcome variable, and then click Run Analysis, Run the model shown. This will create a valid HLM graphing file if one does not already exist for your .mdm file.) Select **File**, **Graph Data**, and **box-whisker plots**. To generate these boxplots, select MATHGAIN as the **Y-axis** variable and select **Group at level-3** to generate a separate panel of boxplots for each school. For **Number of groups**, we select **First ten groups** (corresponding to the first 10 schools). Finally, click **OK**. HLM produces boxplots for the first 10 schools, and we display the plots for the first eight schools in Figure 4.2.

In Figure 4.2, we note evidence of both between-classroom and between-school variability in the MATHGAIN responses. We also see differences in the within-classroom variability, which may be explained by student-level covariates. By clicking **Graph Settings** in the HLM graph window, one can select additional Level 2 (e.g., YEARSTEA) or Level 3 (e.g., HOUSEPOV) variables as **Z-focus** variables that color-code boxplots, based on values of the classroom- and school-level covariates. Readers should refer to the HLM manual for more information on additional graphing features.

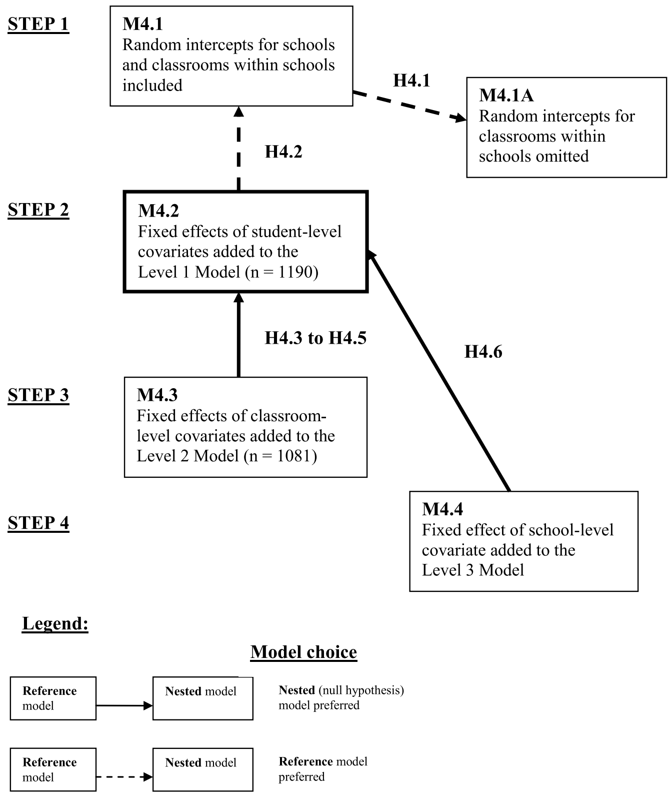


FIGURE 4.3: Model selection and related hypotheses for the Classroom data analysis.

4.3 Overview of the Classroom Data Analysis

For the analysis of the Classroom data, we follow the “step-up” modeling strategy outlined in Subsection 2.7.2. This approach differs from the strategy used in Chapter 3, in that it starts with a simple model, containing only a single fixed effect (the overall intercept), random effects associated with the intercept for classrooms and schools, and residual errors, and then builds the model by adding fixed effects of covariates measured at the various levels.

In Subsection 4.3.1 we outline the analysis steps, and informally introduce related models and hypotheses to be tested. In Subsection 4.3.2 we present the specification of selected models that will be fitted to the Classroom data, and in Subsection 4.3.3 we detail the hypotheses tested in the analysis. To follow the analysis steps outlined in this section, we refer readers to the schematic diagram presented in Figure 4.3.

4.3.1 Analysis Steps

Step 1: Fit the initial “unconditional” (variance components) model (Model 4.1).

Fit a three-level model with a fixed intercept, and random effects associated with the intercept for classrooms (Level 2) and schools (Level 3), and decide whether to keep the random intercepts for classrooms (Model 4.1 vs. Model 4.1A).

Because Model 4.1 does not include fixed effects associated with any covariates, it is referred to as a "means-only" or "unconditional" model in the HLM literature and is known as a variance components model in the classical ANOVA context. We include the term "unconditional" in quotes because it indicates a model that is not conditioned on any fixed effects other than the intercept, although it is still conditional on the random effects. The model includes a fixed overall intercept, random effects associated with the intercept for classrooms within schools, and random effects associated with the intercept for schools.

After fitting Model 4.1, we obtain estimates of the initial variance components, i.e., the variances of the random effects at the school level and the classroom level, and the residual error variance at the student level. We use the variance component estimates from Model 4.1 fit to estimate intraclass correlation coefficients (ICCs) of MATHGAIN responses at the school level and at the classroom level (see Section 4.8).

We also test Hypothesis 4.1 to decide whether the random effects associated with the intercepts for classrooms nested within schools can be omitted from Model 4.1. We fit a model without the random classroom effects (Model 4.1A) and perform a likelihood ratio test. We decide to retain the random intercepts associated with classrooms nested within schools, and we also retain the random intercepts associated with schools in all subsequent models, to preserve the hierarchical structure of the data. Model 4.1 is preferred at this stage of the analysis.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

Add fixed effects associated with covariates measured on the students to the Level 1 Model to obtain Model 4.2, and evaluate the reduction in the residual error variance.

In this step, we add fixed effects associated with the four student-level covariates (MATHKIND, SEX, MINORITY, and SES) to Model 4.1 and obtain Model 4.2. We informally assess the related reduction in the between-student variance (i.e., the residual error variance).

We also test Hypothesis 4.2, using a likelihood ratio test, to decide whether we should add the fixed effects associated with all of the student-level covariates to Model 4.1. We decide to add these fixed effects and choose Model 4.2 as our preferred model at this stage of the analysis.

Step 3: Build the Level 2 Model by adding Level 2 covariates (Model 4.3).

Add fixed effects associated with the covariates measured on the Level 2 clusters (classrooms) to the Level 2 model to create Model 4.3, and decide whether to retain the effects of the Level 2 covariates in the model.

We add fixed effects associated with the three classroom-level covariates (YEARSTEA, MATHPREP, and MATHKNOW) to Model 4.2 to obtain Model 4.3. At this point, we would like to assess whether the Level 2 component of variance (i.e., the variance of the nested random classroom effects) is reduced when we include the effects of these Level 2 covariates in the model. However, Model 4.2 and Model 4.3 are fitted using different sets of observations, owing to the missing values on the MATHPREP and MATHKNOW covariates, and a simple comparison of the classroom-level variance components obtained from these two models is therefore not appropriate.

We also test Hypotheses 4.3, 4.4, and 4.5, to decide whether we should keep the fixed effects associated with YEARSTEA, MATHPREP, and MATHKNOW in Model 4.3, using individual t-tests for each hypothesis. Based on the results of the t-tests, we decide that none of the fixed effects associated with these Level 2 covariates should be retained and choose Model 4.2 as our preferred model at this stage. We do not use a likelihood ratio test for the fixed effects of all Level 2 covariates at once, as we did for all Level 1 covariates in Step 2, because different sets of observations were used to fit Model 4.2 and Model 4.3.

A likelihood ratio test is only possible if both models were fitted using the same set of observations. In Subsection 4.11.4, we illustrate syntax that could be used to construct a "complete case" data set in each software package.

Step 4: Build the Level 3 Model by adding the Level 3 covariate (Model 4.4).

Add a fixed effect associated with the covariate measured on the Level 3 clusters (schools) to the Level 3 model to create Model 4.4, and evaluate the reduction in the variance component associated with the Level 3 clusters.

In this last step, we add a fixed effect associated with the only school-level covariate, HOUSEPOV, to Model 4.2 and obtain Model 4.4. We assess whether the variance component at the school level (i.e., the variance of the random school effects) is reduced when we include this fixed effect at Level 3 of the model. Because the same set of observations was used to fit Model 4.2 and Model 4.4, we informally assess the relative reduction in the between-school variance component in this step.

We also test Hypothesis 4.6 to decide whether we should add the fixed effect associated with the school-level covariate to Model 4.2. Based on the result of a *t*-test for the fixed effect of HOUSEPOV in Model 4.4, we decide not to add this fixed effect, and choose Model 4.2 as our final model. We consider diagnostics for Model 4.2 in Section 4.9, using residual files generated by the HLM software.

Figure 4.3 provides a schematic guide to the model selection process and hypotheses considered in the analysis of the Classroom data. See Subsection 3.3.1 for a detailed interpretation of the elements in the figure. Table 4.3 provides a summary of the various models considered in the Classroom data analyses.

4.3.2 Model Specification

4.3.2.1 General Model Specification

We specify Model 4.3 in (4.1), because it is the model with the most fixed effects that we consider in the analysis of the Classroom data. Models 4.1, 4.1A, and 4.2 are simplifications of this more general model. Selected models are summarized in Table 4.3.

The general specification for Model 4.3 is:

$$\begin{aligned}\text{MATHGAIN}_{ijk} = \; & \left.\begin{aligned}&\beta_0 + \beta_1 \times \text{MATHKIND}_{ijk} + \beta_2 \times \text{SEX}_{ijk}\\ &+ \beta_3 \times \text{MINORITY}_{ijk} + \beta_4 \times \text{SES}_{ijk} + \beta_5 \times \text{YEARSTEA}_{jk}\\ &+ \beta_6 \times \text{MATHPREP}_{jk} + \beta_7 \times \text{MATHKNOW}_{jk}\end{aligned}\right\}\text{fixed}\\ & \qquad\qquad\qquad\qquad + u_k + u_{j|k} + \varepsilon_{ijk} \qquad \}\ \text{random}\end{aligned} \tag{4.1}$$

In this specification, MATHGAIN_{ijk} represents the value of the dependent variable for student i in classroom j nested within school k; β_0 through β_7 represent the fixed intercept and the fixed effects of the covariates (e.g., MATHKIND,..., MATHKNOW); u_k is the random effect associated with the intercept for school k; $u_{j|k}$ is the random effect associated with the intercept for classroom j within school k; and ε_{ijk} represents the residual error. To obtain Model 4.4, we add the fixed effect (β_8) of the school-level covariate HOUSEPOV and omit the fixed effects of the classroom-level covariates (β_5 through β_7).

The distribution of the random effects associated with the schools in Model 4.3 is written as

$$u_k \sim \mathcal{N}(0, \sigma^2_{int:school}),$$

where $\sigma^2_{int:school}$ represents the variance of the school-specific random intercepts.

The distribution of the random effects associated with classrooms nested within a given school is

$$u_{j|k} \sim \mathcal{N}(0, \sigma^2_{int:classroom}),$$

where $\sigma^2_{int:classroom}$ represents the variance of the random classroom-specific intercepts at any given school. This between-classroom variance is assumed to be constant for all schools.

The distribution of the residual errors associated with the student-level observations is

$$\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2),$$

where σ^2 represents the residual error variance.

We assume that the random effects, u_k, associated with schools, the random effects, $u_{j|k}$, associated with classrooms nested within schools, and the residual errors, ε_{ijk}, are all mutually independent.

The general specification of Model 4.3 corresponds closely to the syntax that is used to fit the model in SAS, SPSS, Stata, and R. In Subsection 4.3.2.2 we provide the hierarchical specification that more closely corresponds to the HLM setup of Model 4.3.

4.3.2.2 Hierarchical Model Specification

We now present a hierarchical specification of Model 4.3. The following model is in the form used in the HLM software, but employs the notation used in the general specification of Model 4.3 in (4.1), rather than the HLM notation. The correspondence between the notation used in this section and that used in the HLM software is shown in Table 4.3.

The hierarchical model has three components, reflecting contributions from the three levels of data shown in Table 4.2. First, we write the **Level 1** component as

Level 1 Model (Student)

$$\begin{aligned} \text{MATHGAIN}_{ijk} = \quad & b_{0j|k} + \beta_1 \times \text{MATHKIND}_{ijk} + \beta_2 \times \text{SEX}_{ijk} \\ & + \beta_3 \times \text{MINORITY}_{ijk} + \beta_4 \times \text{SES}_{ijk} + \varepsilon_{ijk} \end{aligned} \tag{4.2}$$

where

$$\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma^2).$$

The **Level 1** model (4.2) shows that at the student level of the data, we have a set of simple classroom-specific linear regressions of MATHGAIN on the student-level covariates. The unobserved classroom-specific intercepts, $b_{0j|k}$, are related to several other fixed and random effects at the classroom level, and are defined in the following **Level 2** model.

The **Level 2** model for the classroom-specific intercepts can be written as

Level 2 Model (Classroom)

$$\begin{aligned} b_{0j|k} = \quad & b_{0k} + \beta_5 \times \text{YEARSTEA}_{jk} + \beta_6 \times \text{MATHPREP}_{jk} \\ & + \beta_7 \times \text{MATHKNOW}_{jk} + u_{j|k} \end{aligned} \tag{4.3}$$

where

$$u_{j|k} \sim \mathcal{N}(0, \sigma^2_{int:classroom}).$$

The **Level 2** model (4.3) assumes that the intercept, $b_{0j|k}$, for classroom j nested within school k, depends on the unobserved intercept specific to the k-th school, b_{0k}, the

TABLE 4.3: Summary of Selected Models Considered for the Classroom Data

		Term/Variable	General Notation	HLM Notation	Model 4.1	Model 4.2	Model 4.3	Model 4.4
Fixed effects		**Intercept**	β_0	γ_{000}	√	√	√	√
		MATHKIND	β_1	γ_{300}		√	√	√
		SEX	β_2	γ_{100}		√	√	√
		MINORITY	β_3	γ_{200}		√	√	√
		SES	β_4	γ_{400}		√	√	√
		YEARSTEA	β_5	γ_{010}			√	
		MATHPREP	β_6	γ_{030}			√	
		MATHKNOW	β_7	γ_{020}			√	
		HOUSEPOV	β_8	γ_{001}				√
Random effects	**Classroom** (j)	**Intercept**	$u_{j\|k}$	r_{jk}	√	√	√	√
	School (k)	**Intercept**	u_k	u_{00k}	√	√	√	√
Residuals	**Student** (i)		ε_{ijk}	e_{ijk}	√	√	√	√
Covariance parameters ($\boldsymbol{\theta}_D$) **for** $\boldsymbol{D}$ **matrix**	**Classroom level**	**Variance of intercepts**	$\sigma^2_{int:class}$	τ_π	√	√	√	√
	School level	**Variance of intercepts**	$\sigma^2_{int:school}$	τ_β	√	√	√	√
Covariance parameters (θ_R) **for** R_i **matrix**	**Student level**	**Residual variance**	σ^2	σ^2	√	√	√	√

classroom-specific covariates associated with the teacher for that classroom (YEARSTEA, MATHPREP, and MATHKNOW), and a random effect, $u_{j|k}$, associated with classroom j within school k.

The **Level 3** model for the school-specific intercepts in Model 4.3 is:

Level 3 Model (School)

$$b_{0k} = \beta_0 + u_k \tag{4.4}$$

where

$$u_k \sim \mathcal{N}(0, \sigma^2_{int:school}).$$

The **Level 3** model shows that the school-specific intercept in Model 4.3 depends on the overall fixed intercept, β_0, and the random effect, u_k, associated with the intercept for school k.

By substituting the expression for b_{0k} from the Level 3 model into the Level 2 model, and then substituting the resulting expression for $b_{0j|k}$ from the Level 2 model into the Level 1 model, we recover Model 4.3 as it was specified in (4.1).

We specify Model 4.4 by omitting the fixed effects of the classroom-level covariates from Model 4.3, and adding the fixed effect of the school-level covariate, HOUSEPOV, to the **Level 3** model.

4.3.3 Hypothesis Tests

Hypothesis tests considered in the analysis of the Classroom data are summarized in Table 4.4.

Hypothesis 4.1: The random effects associated with the intercepts for classrooms nested within schools can be omitted from Model 4.1.

We do not directly test the significance of the random classroom-specific intercepts, but rather test null and alternative hypotheses about the *variance* of the classroom-specific intercepts. The null and alternative hypotheses are:

$$H_0 : \sigma^2_{int:classroom} = 0$$
$$H_A : \sigma^2_{int:classroom} > 0$$

We use a REML-based likelihood ratio test for Hypothesis 4.1 in SAS, SPSS, R, and Stata. The test statistic is calculated by subtracting the –2 REML log-likelihood for Model 4.1 (the reference model, including the nested random classroom effects) from the corresponding value for Model 4.1A (the nested model). To obtain a p-value for this test statistic, we refer it to a mixture of χ^2 distributions, with 0 and 1 degrees of freedom and equal weight 0.5.

The HLM3 procedure does not use REML estimation and is not able to fit a model without any random effects at a given level of the model, such as Model 4.1A. Therefore, an LRT cannot be performed, so we consider an alternative chi-square test for Hypothesis 4.1 provided by HLM in Subsection 4.7.2.

We decide that the random effects associated with the intercepts for classrooms nested within schools should be retained in Model 4.1. We do not explicitly test the variance of the random school-specific intercepts, but we retain them in Model 4.1 and all subsequent models to reflect the hierarchical structure of the data.

Hypothesis 4.2: The fixed effects associated with the four student-level covariates should be added to Model 4.1.

TABLE 4.4: Summary of Hypotheses Tested in the Classroom Analysis

Hypothesis Specification			Hypothesis Test				
				Models Compared			
Label	**Null (H_0)**	**Alternative (H_A)**	**Test**	**Nested Model (H_0)**	**Ref. Model (H_A)**	**Est. Method**	**Test Stat. Dist. under H_0**
4.1	Drop $u_{j\|k}$ ($\sigma^2_{int:classroom} = 0$)	Retain $u_{j\|k}$ ($\sigma^2_{int:classroom} > 0$)	LRT	Model 4.1A	Model 4.1	REML	$0.5\chi^2_0 + 0.5\chi^2_1$
4.2	Fixed effects of student-level covariates are all zero ($\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$)	At least one fixed effect at the student level is different from zero	LRT	Model 4.1	Model 4.2	ML	χ^2_4
4.3	Fixed effect of YEARSTEA is zero $\beta_5 = 0$	$\beta_5 \neq 0$	t-test	N/A	Model 4.3	REML/ ML	t^a_{177}
4.4	Fixed effect of MATHPREP is zero (β_6=0)	$\beta_6 \neq 0$	t-test	N/A	Model 4.3	REML/ ML	t^a_{177}
4.5	Fixed effect of MATHKNOW is zero ($\beta_7 = 0$)	$\beta_7 \neq 0$	t-test	N/A	Model 4.3	REML/ ML	t^a_{177}
4.6	Fixed effect of HOUSEPOV is zero ($\beta_8 = 0$)	$\beta_8 \neq 0$	t-test	N/A	Model 4.4	REML/ ML	t^a_{105}

[a] Degrees of freedom for the t-statistics are those reported by the HLM3 procedure.

The null and alternative hypotheses for the fixed effects associated with the student-level covariates, MATHKIND, SEX, MINORITY, and SES, are:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$
$$H_{\mathrm{A}}: \text{At least one fixed is not equal to zero.}$$

We test Hypothesis 4.2 using a likelihood ratio test, based on maximum likelihood (ML) estimation. The test statistic is calculated by subtracting the –2 ML log-likelihood for Model 4.2 (the reference model with the fixed effects of all student-level covariates included) from that for Model 4.1 (the nested model). Under the null hypothesis, the distribution of this test statistic is asymptotically a χ^2 with 4 degrees of freedom.

We decide that the fixed effects associated with the Level 1 covariates should be added and select Model 4.2 as our preferred model at this stage of the analysis.

Hypotheses 4.3, 4.4, and **4.5:** The fixed effects associated with the classroom-level covariates should be retained in Model 4.3.

The null and alternative hypotheses for the fixed effects associated with the classroom-level covariates YEARSTEA, MATHPREP, and MATHKNOW are written as follows:

Hypothesis 4.3 for the fixed effect associated with YEARSTEA:

$$H_0 : \beta_5 = 0$$
$$H_A : \beta_5 \neq 0$$

Hypothesis 4.4 for the fixed effect associated with MATHPREP:

$$H_0 : \beta_6 = 0$$
$$H_A : \beta_6 \neq 0$$

Hypothesis 4.5 for the fixed effect associated with MATHKNOW:

$$H_0 : \beta_7 = 0$$
$$H_A : \beta_7 \neq 0$$

We test each of these hypotheses using t-tests based on the fit of Model 4.3. We decide that none of the fixed effects associated with the classroom-level covariates should be retained, and keep Model 4.2 as our preferred model at this stage of the analysis.

Because there are classrooms with missing values for the classroom-level covariate MATHKNOW, there are different sets of observations used to fit Model 4.2 and Model 4.3. In this case, we cannot use a likelihood ratio test to decide if we should keep the fixed effects of all covariates at Level 2 of the model, as we did for the fixed effects of the covariates at Level 1 of the model when we tested Hypothesis 4.2.

To perform a likelihood ratio test for the effects of all of the classroom-level covariates, we would need to fit both Model 4.2 and Model 4.3 using the same cases (i.e., those cases with no missing data on all covariates would need to be considered for both models). See Subsection 4.11.4 for a discussion of how to set up the necessary data set, with complete data for all covariates, in each of the software procedures.

Hypothesis 4.6: The fixed effect associated with the school-level covariate HOUSEPOV should be added to Model 4.2.

The null and alternative hypotheses for the fixed effect associated with HOUSEPOV are specified as follows:

$$H_0 : \beta_8 = 0$$
$$H_A : \beta_8 \neq 0$$

We test Hypothesis 4.6 using a t-test for the significance of the fixed effect of HOUSEPOV in Model 4.4. We decide that the fixed effect of this school-level covariate should not be added to the model, and choose Model 4.2 as our final model. For the results of these hypothesis tests, see Section 4.5.

4.4 Analysis Steps in the Software Procedures

We compare results for selected models across the software procedures in Section 4.6.

4.4.1 SAS

We begin by reading the comma-delimited data file, classroom.csv (assumed to have the "long" data structure displayed in Table 4.2, and located in the `C:\temp` directory) into a temporary SAS data set named `classroom`:

```
proc import out = WORK.classroom
   datafile = "C:\temp\classroom.csv";
   dbms = csv replace;
   getnames = YES;
   datarow = 2;
   guessingrows = 20;
run;
```

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

Prior to fitting the models in SAS, we sort the data set by SCHOOLID and CLASSID. Although not necessary for fitting the models, this sorting makes it easier to interpret elements in the marginal covariance and correlation matrices that we display later.

```
proc sort data = classroom;
   by schoolid classid;
run;
```

The SAS `proc glimmix` code for fitting Model 4.1 using REML estimation is shown below. We do not need to explicitly specify the estimation method because `proc glimmix` assumes by default that the conditional distribution of the outcome is normal, and that RSPL estimation (equivalent to REML for a normal outcome) should be used.

The `model` statement has no covariates on the right-hand side of the equal sign because the only fixed effect in Model 4.1 is the intercept, which is included by default in the fixed effects for the model.

The two `random` statements specify that we wish the model to include a random intercept for each school and a random intercept for each classroom nested within a school `classid(schoolid)`.

```
ods exclude solutionr;
ods output solutionr = eblupdat;
title "Model 4.1";
proc glimmix data = classroom noclprint;
   class classid schoolid;
   model mathgain =   / solution;
```

```
    random intercept / subject = schoolid solution v=69 vcorr=69;
    random intercept / subject = classid(schoolid) solution;
    covtest general 0 1 0/ est;
run;
```

We specify the noclprint option in the proc glimmix statement. This option suppresses listing of levels of the class variables in the output to save space; we recommend using this option only after an initial examination of the levels of the class variables to be sure the data set is being read correctly.

The solution option in the model statement causes SAS to print estimates of the fixed-effect parameters included in the model, which for Model 4.1 is simply the overall intercept.

The first random statement in the proc glimmix syntax for Model 4.1:

```
random intercept / subject = schoolid solution v=69 vcorr=69;
```

indicates that the model should include a random intercept for each school, and that SCHOOLID should be considered to be the overall subject in the model. The solution option requests that the EBLUPs for each school be displayed in the output.

The second random statement:

```
random intercept / subject = classid(schoolid) solution;
```

indicates that the model should include a random intercept for each classroom nested within a school. The solution option tells SAS to display the EBLUPs for the classrooms in the output. Because we included the CLASSID variable in the class statement, SAS sets up the same number of columns for CLASSID in the $\boldsymbol{Z}$ matrix for each subject (i.e., school), regardless of the number of classrooms in the school, as we will see below.

We use the v=69 option in the first random statement to request that a single block (corresponding to the sixty-ninth school) of the estimated $\boldsymbol{V}$ matrix of the marginal model implied by Model 4.1 be displayed in the output (see Subsection 2.2.2 for more details on the $\boldsymbol{V}$ matrix in SAS). In addition, we specify the vcorr=69 option to request that the corresponding marginal correlation matrix be displayed. The v and vcorr options can be added to either of the random statements with the same results. If the v and vcorr options are specified without a school being listed after the equal sign, the output will simply display these matrices corresponding to the first school. We display and discuss the estimated marginal covariance matrix and the corresponding correlation matrix for the observations in school 69 based on the fit of Model 4.1, in Section 4.8.

The Dimensions table, based on the fit of Model 4.1 using the two random statements shown above, indicates that there are 107 Subjects, which correspond to the 107 schools in the data set. There are 2 G-side covariance parameters, corresponding to the variance of the random intercept for school and the variance of the random intercept for classroom nested within school. There is 1 R-side covariance parameter, corresponding to the residual error variance. The $\boldsymbol{X}$ matrix has 1 column, corresponding to the fixed effect for the intercept. There are 10 columns in the $\boldsymbol{Z}$ matrix per subject (i.e., school), corresponding to the maximum number of classrooms in any given school (9), plus a single column for the random intercept for the school. The maximum number of students in any subject (i.e., school) is 31.

```
                    Dimensions

G-side Cov. Parameters           2
R-side Cov. Parameters           1
Columns in X                     1
Columns in Z per Subject        10
Subjects (Blocks in V)         107
Max Obs per Subject             31
```

This syntax for Model 4.1 allows SAS to fit the model efficiently, by taking blocks of the marginal $\boldsymbol{V}$ matrix for each subject (i.e., school) into account. However, it also means that if we specify the `solution` option in the corresponding `random` statements in order to obtain the EBLUPs for the random school effects and random classroom effects nested within school, SAS will display output for 9 classrooms for each school, even for schools that have fewer than 9 classrooms. This results in extra unwanted output being generated. A portion of the output generated when the `solution` option is specified in both `random` statements in Model 4.1 is shown below. The EBLUPs for SCHOOLID 1 and SCHOOLID 2 (with two and three classrooms, respectively) are depicted.

```
                          Solution for Random Effects

                                                        Std Err
Effect       Subject                     Estimate          Pred      DF    t Value     Pr > |t|

Intercept    schoolid 1                    0.9416        7.1675     878       0.13       0.8955
Intercept    classid(schoolid) 160 1       1.6385        8.9204     878       0.18       0.8543
Intercept    classid(schoolid) 217 1      -0.4329        8.1163     878      -0.05       0.9575
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept    schoolid 2                    2.5481        7.0641     878       0.36       0.7184
Intercept    classid(schoolid) 197 2      -1.6947        9.1921     878      -0.18       0.8538
Intercept    classid(schoolid) 211 2       2.5135        8.6895     878       0.29       0.7725
Intercept    classid(schoolid) 307 2       2.4439        8.6895     878       0.28       0.7786
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
Intercept                                       0        9.9613     878       0.00       1.0000
```

Note that the standard error, along with a t-test, is displayed for each EBLUP. The test statistic is calculated by dividing the predicted EBLUP by its estimated standard error, with the degrees of freedom being the same as indicated by the `ddfm =` option. In this case, we did not specify this option, so SAS uses the default "containment" method to determine the df for each of these tests (see Subsection 3.11.6 for a discussion of different methods of calculating denominator degrees of freedom in SAS). Although it may not be of particular interest to test whether the predicted random effect for a given school or classroom is equal to zero, large values of the t-statistic may indicate an outlying value for a given school or classroom that analysts may want to investigate.

To suppress the display of the EBLUPs in the output, but still obtain them in a SAS data set, we use the following ODS statements prior to invoking `proc glimmix`:

```
ods exclude solutionr;
ods output solutionr = eblupdat;
```

The first `ods` statement prevents the EBLUPs from being displayed in the output. The second `ods` statement requests that the EBLUPs be placed in a new SAS data set named `eblupdat`. The distributions of the EBLUPs contained in the `eblupdat` data set can be investigated graphically to check for possible outliers (see Subsection 4.10.1 for diagnostic plots of the EBLUPs associated with Model 4.2). Note that the `eblupdat` data set will only be created if the `solution` option is added to one or both of the `random` statements.

Alternative syntax for Model 4.1 may be specified by omitting the `subject =` option in the `random` statement for SCHOOLID, and retaining the same `random` statement for CLASSID nested within SCHOOLID, as shown below:

```
random schoolid / solution;
random intercept / subject = classid(schoolid) solution;
```

By including the `solution` option in the first and second `random` statements shown above, the Solution for Random Effects table will display the EBLUPs for all 107 schools, followed by the EBLUPs for the 312 classrooms, with no extra output being produced.

A portion of the resulting output for the EBLUPs follows:

```
                         Solution for Random Effects

                                                         Std Err
Effect      schoolid      Subject                Estimate     Pred     DF  t Value  Pr > |t|

schoolid           1                               0.9416   7.1675    878     0.13    0.8955
schoolid           2                               2.5481   7.0641    878     0.36    0.7184
schoolid           3                              13.6776   6.6404    878     2.06    0.0397
schoolid           4                              -6.4638   7.6028    878    -0.85    0.3954
schoolid           5                               0.6099   7.7682    878     0.08    0.9374
 . . .
Intercept                 classid(schoolid) 160 1   1.6385   8.9204    878     0.18    0.8543
Intercept                 classid(schoolid) 217 1  -0.4329   8.1163    878    -0.05    0.9575
Intercept                 classid(schoolid) 197 2  -1.6947   9.1921    878    -0.18    0.8538
Intercept                 classid(schoolid) 211 2   2.5135   8.6895    878     0.29    0.7725
Intercept                 classid(schoolid) 307 2   2.4439   8.6895    878     0.28    0.7786
Intercept                 classid(schoolid) 11  3   3.8699   8.6621    878     0.45    0.6552
Intercept                 classid(schoolid) 137 3   5.5645   9.1834    878     0.61    0.5447
Intercept                 classid(schoolid) 145 3   8.1027   8.4635    878     0.96    0.3386
Intercept                 classid(schoolid) 228 3  -0.02350  8.8986    878    -0.00    0.9979
```

The resulting "Dimensions" output for this alternative model specification is as follows:

```
          Dimensions

G-side Cov. Parameters           2
R-side Cov. Parameters           1
Columns in X                     1
Columns in Z                   419
Subjects (Blocks in V)           1
Max Obs per Subject           1190
```

Importantly, SAS now considers there to be only one subject in this model, even though we specified a `subject =` option in the second `random` statement. In this case, the $\boldsymbol{Z}$ matrix has 419 columns, which correspond to the total number of schools (107) plus classrooms (312). The one subject now has 1190 observations (the total sample size).

Note that we did not include the `v` or `vcorr` options for this alternative specification of the `random` statements, because this would result in SAS displaying matrices that are of the same dimension as the total number of observations in the entire data set. In the case of the Classroom data, this would result in a 1190×1190 matrix being displayed for both the marginal covariance and marginal correlation matrices. We recommend that readers be cautious about including options, such as `v` and `vcorr`, for displaying the marginal covariance and correlation matrices when using random statements with no `subject =` option, as shown in the first `random` statement in the alternative syntax for Model 4.1. We also advise users to check the Dimensions table, to verify the number of subjects that are included in the model.

We carry out a likelihood ratio test of Hypothesis 4.1 (whether we can delete the random effects associated with classrooms nested within schools from the model) by using the `covtest` statement in the `proc glimmix` code for Model 4.1.

```
covtest general 0 1 0/ est;
```

The `covtest` statement makes it possible to carry out a LR test of the covariance parameters without the necessity to fit Model 4.1A (the nested model). In this case, we use the `covtest` statement with the `general` option, which requires that we specify the coefficients for a linear combination of the covariance parameters that will be set equal to zero under the null hypothesis. The linear combination for this hypothesis test specifies that the variance for the random classroom effects (the second covariance parameter) is being tested, while the variance of the random intercept for schools (the first covariance parameter) and the variance of the residual errors (the third covariance parameter) are assigned coefficients of zero, which indicates that they are not involved in Hypothesis 4.1 and are both allowed to vary freely. Results of the REML-based likelihood ratio test for Hypothesis 4.1 are discussed in detail in Subsection 4.5.1. The output for this hypothesis test is shown below. Note that in the portion of the output labeled "Estimates H0" (the estimated values of the parameters under the null hypothesis), the value of the variance of the random classroom intercepts is set to zero. We also see that SAS has used a mixture of chi-squares (denoted `MI`) to compute the p-value for this test, as desired.

```
                    Tests of Covariance Parameters
                   Based on the Restricted Likelihood

     Label          DF     -2 Res Log Like       ChiSq    Pr > ChiSq

     General         1               11777        7.90        0.0025

                    Tests of Covariance Parameters
                   Based on the Restricted Likelihood

                         ----------Estimates H0----------
         Label         Est1          Est2          Est3          Note

         General        109.24              0      1094.00       MI

             MI: P-value based on a mixture of chi-squares.
```

We decide to retain the nested random classroom effects in Model 4.1 based on the significant ($p = 0.0025$) result of this test. We also keep the random school-specific intercepts in the model to reflect the hierarchical structure of the data.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

To fit Model 4.2, again using REML estimation, we modify the `model` statement to add the fixed effects of the student-level covariates MATHKIND, SEX, MINORITY, and SES: We retain the same `random` statements that we used in the original specification of Model 4.1.

```
title "Model 4.2";
proc glimmix data = classroom noclprint;
   class classid schoolid;
   model mathgain = mathkind sex minority ses / solution;
   random intercept / subject = schoolid;
   random intercept / subject = classid(schoolid);
run;
```

Because the covariates SEX and MINORITY are indicator variables, having values of 0 and 1, they do not need to be identified as categorical variables in the `class` statement.

We formally test Hypothesis 4.2 (whether any of the fixed effects associated with the student-level covariates are different from zero) by performing an ML-based likelihood ratio test, subtracting the –2 ML log-likelihood of Model 4.2 (the reference model) from that of Model 4.1 (the nested model, excluding the four student-level fixed effects being tested), and referring the difference to a χ^2 distribution with 4 degrees of freedom. Note that the `noreml` option is used to request that maximum likelihood estimation (rather than the default REML estimation) be used for both models (see Subsection 2.6.2.1 for a discussion of likelihood ratio tests for fixed effects):

```
title "Model 4.1: ML Estimation";
proc glimmix data = classroom noclprint noreml;
   class classid schoolid;
   model mathgain =   / solution;
   random intercept / subject = schoolid;
   random intercept / subject = classid(schoolid);
run;

title "Model 4.2: ML Estimation";
proc glimmix data = classroom noclprint noreml;
   class classid schoolid;
   model mathgain = mathkind sex minority ses / solution;
   random intercept / subject = schoolid;
   random intercept / subject = classid(schoolid);
run;
```

We calculate the likelihood ratio test statistic for Hypothesis 4.2 by subtracting the value of the –2 log likelihood of Model 4.2 (the reference model, including the four fixed effects at the student level) from that of Model 4.1 (the nested model) and referring the result to a χ^2 distribution with 4 degrees of freedom to obtain the p-value. The syntax to derive the likelihood ratio test statistic and the p-value of the test is shown below. The `noobs` option is used in the `proc print` statement to prevent SAS from displaying the observation number in the output.

```
title "P-value for Hypothesis 4.2";
data test;
lrstat = 11771.33 - 11390.96;
df = 4;
pvalue = 1-probchi(lrstat,df);
format pvalue 10.8;
run;
proc print data=test noobs;
run;
```

We reject the null hypothesis that all fixed effects associated with the student-level covariates are equal to zero, based on the result of this test ($p < 0.001$), and choose Model 4.2 as our preferred model at this stage of the analysis.

The significant result for the test of Hypothesis 4.2 suggests that the fixed effects of the student-level covariates explain at least some of the variation at the student level of the data. Informal examination of the estimated residual error variance in Model 4.2 suggests that the Level 1 (student-level) residual error variance is substantially reduced compared to that for Model 4.1 (see Section 4.6).

Step 3: Build the Level 2 Model by adding Level 2 Covariates (Model 4.3).

To fit Model 4.3, we add the fixed effects of the classroom-level covariates YEARSTEA, MATHPREP, and MATHKNOW to the `model` statement that we specify for Model 4.2:

```
title "Model 4.3";
proc glimmix data = classroom noclprint;
   class classid schoolid;
   model mathgain = mathkind sex minority ses
                    yearstea mathprep mathknow / solution;
   random intercept / subject = schoolid;
   random intercept / subject = classid(schoolid);
run;
```

We consider t-tests of Hypotheses 4.3, 4.4, and 4.5, for the individual fixed effects of the classroom-level covariates added to the model in this step. The results of these t-tests indicate that none of the fixed effects associated with the classroom-level covariates are significant, so we keep Model 4.2 as the preferred model at this stage of the analysis. Note that results of the t-tests are based on an analysis with 109 observations omitted due to missing data on the MATHKNOW variable.

Step 4: Build the Level 3 Model by adding the Level 3 Covariate (Model 4.4).

To fit Model 4.4, we add the fixed effect of the school-level covariate, HOUSEPOV, to the `model` statement for Model 4.2:

```
title "Model 4.4";
proc glimmix data = classroom noclprint;
   class classid schoolid;
   model mathgain = mathkind sex minority ses housepov / solution;
   random intercept / subject = schoolid;
   random intercept / subject = classid(schoolid);
run;
```

To test Hypothesis 4.6, we carry out a t-test for the fixed effect of HOUSEPOV, based on the REML fit of Model 4.4. The result of the t-test indicates that the fixed effect for this school-level covariate is not significant ($p = 0.25$). We also note that the estimated variance of the random school effects in Model 4.4 has not been reduced compared to that of Model 4.2. We therefore do not retain the fixed effect of the school-level covariate and choose Model 4.2 as our final model.

We now refit our final model, using REML estimation:

```
ods exclude solutionr;
ods output solutionr = eblupdat;
title "Model 4.2 (Final)";
proc glimmix data = classroom noclprint
              plots =(residualpanel(blup)
                      studentpanel(blup)
                      boxplot(student blup));
   class classid schoolid;
   model mathgain = mathkind sex minority ses / solution;
   output out = pdat
                  pred(blup) = p_cond
                  resid(blup) = r_cond
                  student(blup) = s_cond;
   random intercept / subject = schoolid solution;
   random intercept / subject = classid(schoolid) solution;
run;
```

We request diagnostic plots for this model by using the `plots =` option in the `proc glimmix` statement. In addition, we capture the data set `pdat` (requested with the `ods output` statement). This data set contains the conditional predicted value, P_COND, for each student, based on the estimated fixed effects in the model and the EBLUPs for the random school and classroom effects. We also request the conditional residuals for each observation, R_COND, and the studentized conditional residuals, S_COND. This data set can be used to further assess model diagnostics, if desired (see Subsection 4.10.2).

4.4.2 SPSS

We first import the comma-delimited data file named classroom.csv, which has the "long" format displayed in Table 4.2, using the following SPSS (Version 28) syntax:

```
GET DATA /TYPE = TXT
 /FILE = "E:\LMM Data\classroom.csv"
 /DELCASE = LINE
 /DELIMITERS = ","
 /ARRANGEMENT = DELIMITED
 /FIRSTCASE = 2
 /IMPORTCASE = ALL
 /VARIABLES =
   sex F1.0
   minority F1.0
   mathkind F3.2
   mathgain F4.2
   ses F5.2
   yearstea F5.2
   mathknow F5.2
   housepov F5.2
   mathprep F4.2
   classid F3.2
   schoolid F1.0
   childid F2.1
```

```
.
CACHE.
EXECUTE.
```

Now that we have a data set in SPSS in the "long" form appropriate for fitting linear mixed models (LMMs), we proceed with the analysis.

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

There are two ways to set up an LMM with nested random effects using SPSS syntax. We begin by illustrating how to set up Model 4.1 without specifying any "subjects" in the RANDOM subcommands:

```
* Model 4.1 (less efficient syntax).
MIXED
mathgain BY classid schoolid
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
   SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
   LCONVERGE (0, ABSOLUTE) DFMETHOD(KENWARDROGER)
   PCONVERGE(0.000001, ABSOLUTE)
 /FIXED = | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM classid(schoolid) | COVTYPE(VC)
 /RANDOM schoolid | COVTYPE(VC) .
```

The first variable listed after invoking the MIXED command is the dependent variable, MATHGAIN. The two random factors (CLASSID and SCHOOLID) have been declared as categorical factors by specifying them after the BY keyword. The CRITERIA subcommand specifies the estimation criteria to be used when fitting the model, which are the defaults, along with the approximate degrees of freedom calculation method to be used for testing the fixed effects (we use the Kenward–Roger method in this case).

The FIXED subcommand has no variables on the right-hand side of the equal sign before the vertical bar (|), which indicates that an intercept-only model is requested (i.e., the only fixed effect in the model is the overall intercept). The SSTYPE(3) option after the vertical bar specifies that SPSS should use the default "Type 3" analysis, in which the tests for the fixed effects are adjusted for all other fixed effects in the model. This option is not critical for Model 4.1, because the only fixed effect in this model is the intercept.

We use the default REML estimation method by specifying REML in the METHOD subcommand.

The PRINT subcommand specifies that the printed output should include the estimated fixed effects (SOLUTION).

Note that there are two RANDOM subcommands: the first identifies CLASSID nested within SCHOOLID as a random factor, and the second specifies SCHOOLID as a random factor. No "subject" variables are specified in either of these RANDOM subcommands. The COVTYPE(VC) option specified after the vertical bar indicates that a variance components (VC) covariance structure for the random effects is desired.

In the following syntax, we show an alternative specification of Model 4.1 that is more efficient computationally for larger data sets. We specify INTERCEPT before the vertical bar and a SUBJECT variable after the vertical bar for each RANDOM subcommand. This syntax means that for each level of the SUBJECT variables, we add a random effect to the model associated with the INTERCEPT:

```
* Model 4.1 (more efficient syntax).
MIXED
mathgain
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  DFMETHOD(KENWARDROGER)
 /FIXED = | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

There is no `BY` keyword in this syntax, meaning that the random factors in this model are not identified initially as categorical factors.

The first `RANDOM` subcommand declares that there is a random effect associated with the `INTERCEPT` for each subject identified by a combination of the classroom ID and school ID variables (CLASSID*SCHOOLID). Even though this syntax appears to be setting up a crossed effect for classroom by school, it is equivalent to the nested specification for these two random factors that was seen in the previous syntax (`CLASSID(SCHOOLID)`). The asterisk is necessary because nested `SUBJECT` variables cannot currently be specified when using the `MIXED` command. The second `RANDOM` subcommand specifies that there is a random effect in the model associated with the `INTERCEPT` for each subject identified by the SCHOOLID variable. The `SOLUTION` option computes and displays the EBLUPs of the random classroom and school effects at each level in the SPSS output.

> **Software Note:** Fitting Model 4.1 using the less efficient SPSS syntax takes a relatively long time compared to the time required to fit the equivalent model using the other packages; and it also takes longer compared to the more efficient version of the SPSS syntax. We use the more computationally efficient syntax for the remainder of the analysis in SPSS.

We now carry out a likelihood ratio test of Hypothesis 4.1 to decide if we wish to omit the nested random effects associated with classrooms from Model 4.1. To do this, we fit Model 4.1A by removing the first `RANDOM` subcommand from the more efficient syntax for Model 4.1:

```
* Model 4.1A .
MIXED
mathgain
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  DFMETHOD(KENWARDROGER)
 /FIXED = | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

To calculate the test statistic for Hypothesis 4.1, we subtract the –2 REML log-likelihood value for the reference model, Model 4.1, from the corresponding value for Model 4.1A (the nested model). We refer the resulting test statistic to a mixture of χ^2 distributions, with 0 and 1 degrees of freedom, and equal weight of 0.5 (see Subsection 4.5.1 for a discussion of this test). Based on the significant result of this test ($p = 0.002$), we decide to retain the nested random classroom effects in Model 4.1 and all future models. We also retain the random effects associated with schools in all models without testing their significance, to reflect the hierarchical structure of the data in the model specification.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

Model 4.2 adds the fixed effects of the student-level covariates MATHKIND, SEX, MINORITY, and SES to Model 4.1, using the following syntax:

```
* Model 4.2 .
MIXED
mathgain WITH mathkind sex minority ses
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  DFMETHOD(KENWARDROGER)
 /FIXED = mathkind sex minority ses | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

Note that the four student-level covariates have been identified as continuous by listing them after the `WITH` keyword. This is an acceptable approach for SEX and MINORITY, even though they are categorical, because they are both indicator variables having values 0 and 1. All four student-level covariates are also listed in the `FIXED =` subcommand, so that fixed effect parameters associated with each covariate will be added to the model.

Software Note: If we had included SEX and MINORITY as categorical factors by listing them after the `BY` keyword, SPSS would have used the *highest* levels of each of these variables as the reference categories (i.e., SEX = 1, girls; and MINORITY = 1, minority students). The resulting parameter estimates would have given us the estimated fixed effect of being a boy, and of being a non-minority student, respectively, on math achievement score. These parameter estimates would have had the opposite signs of the estimates resulting from the syntax that we used, in which MINORITY and SEX were listed after the `WITH` keyword.

We test Hypothesis 4.2 to decide whether all fixed effects associated with the Level 1 (student-level) covariates are equal to zero, by performing a likelihood ratio test based on ML estimation (see Subsection 2.6.2.1 for a discussion of likelihood ratio tests for fixed effects). The test statistic is calculated by subtracting the –2 ML log-likelihood for Model 4.2 (the reference model) from the corresponding value for Model 4.1 (the nested model, excluding all fixed effects associated with the Level 1 covariates). We use the `/METHOD = ML` subcommand to request Maximum Likelihood estimation of both models, and remove the Kenward–Roger approximation, which is only possible with REML estimation:

```
* Model 4.1 (ML Estimation).
MIXED
mathgain
```

```
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED = | SSTYPE(3)
 /METHOD = ML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .

* Model 4.2 (ML Estimation).
MIXED
mathgain WITH mathkind sex minority ses
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
 /FIXED = mathkind sex minority ses | SSTYPE(3)
 /METHOD = ML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

We reject the null hypothesis ($p < 0.001$) and decide to retain all fixed effects associated with the Level 1 covariates. This result suggests that the Level 1 fixed effects explain at least some of the variation at Level 1 (the student level) of the data, which we had previously attributed to residual error variance in Model 4.1.

Step 3: Build the Level 2 Model by adding Level 2 Covariates (Model 4.3).

We fit Model 4.3 using the default REML estimation method by adding the Level 2 (classroom-level) covariates YEARSTEA, MATHPREP, and MATHKNOW to the SPSS syntax. We add these covariates to the `WITH` subcommand, so that they will be treated as continuous covariates, and we also add them to the `FIXED` subcommand:

```
* Model 4.3 .
MIXED
mathgain WITH mathkind sex minority ses yearstea mathprep mathknow
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  DFMETHOD(KENWARDROGER)
 /FIXED = mathkind sex minority ses yearstea
  mathprep mathknow | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

We cannot perform a likelihood ratio test of Hypotheses 4.3 through 4.5, owing to the presence of missing data on some of the classroom-level covariates; we instead use t-tests for the fixed effects of each of the Level 2 covariates added at this step. The results of these t-tests are reported in the SPSS output for Model 4.3. Because none of these t-tests are significant, we do not add these fixed effects to the model, and select Model 4.2 as our preferred model at this stage of the analysis.

Step 4: Build the Level 3 Model by adding the Level 3 Covariate (Model 4.4).

We fit Model 4.4 using REML estimation, by updating the syntax used to fit Model 4.2. We add the Level 3 (school-level) covariate, HOUSEPOV, to the `WITH` subcommand and to the `FIXED` subcommand, as shown in the following syntax:

```
* Model 4.4.
MIXED
mathgain WITH mathkind sex minority ses housepov
 /CRITERIA = CIN(95) MXITER(100) MXSTEP(5) SCORING(1)
  SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
  LCONVERGE (0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
  DFMETHOD(KENWARDROGER)
 /FIXED = mathkind sex minority ses housepov | SSTYPE(3)
 /METHOD = REML
 /PRINT = SOLUTION
 /RANDOM INTERCEPT | SUBJECT(classid*schoolid) COVTYPE(VC) SOLUTION
 /RANDOM INTERCEPT | SUBJECT(schoolid) COVTYPE(VC) SOLUTION .
```

We use a *t*-test for Hypothesis 4.6, to decide whether we wish to add the fixed effect of HOUSEPOV to the model. The result of this *t*-test is reported in the SPSS output, and indicates that the fixed effect of HOUSEPOV is not significant ($p = 0.25$). We therefore choose Model 4.2 as our final model.

4.4.3 R

We begin by reading the comma-delimited raw data file, having the structure described in Table 4.2 and with variable names in the first row, into a data frame object named `class`. The `h = T` option instructs R to read a header record containing variable names from the first row of the raw data:

```
> class <- read.csv("E:\\LMM Data\\classroom.csv", h = T)
```

4.4.3.1 Analysis Using the `lme()` Function

We first load the `nlme` package, so that the `lme()` function can be used in the analysis:

```
> library(nlme)
```

We now proceed with the analysis steps.

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

We fit Model 4.1 to the Classroom data using the `lme()` function as follows:

```
> # Model 4.1.
> model4.1.fit <- lme(mathgain ~ 1, random = ~ 1 | schoolid/classid,
    class, method = "REML")
```

We describe the syntax used in the `lme()` function below:

- `model4.1.fit` is the name of the object that will contain the results of the fitted model.

- The first argument of the function, `mathgain ~ 1`, defines the response variable, MATHGAIN, and the single fixed effect in this model, which is associated with the intercept (denoted by a 1 after the ~).

- The second argument, `random = ~ 1 | schoolid/classid`, indicates the nesting structure of the random effects in the model. The `random = ~ 1` portion of the argument indicates that the random effects are to be associated with the intercept. The first variable listed after the vertical bar (|) is the random factor at the highest level (Level 3) of the data (i.e., SCHOOLID). The next variable after a forward slash (/) indicates the random factor (i.e., CLASSID) with levels nested within levels of the first random factor. This notation for the nesting structure is known as the **Wilkinson–Rogers notation**.

- The third argument, `class`, indicates the name of the data frame object to be used in the analysis.

- The final argument of the function, `method = "REML"`, tells R that REML estimation should be used for the variance components in the model. REML is the default estimation method in the `lme()` function.

Estimates saved in the model fit object can be obtained by applying the `summary()` function:

```
> summary(model4.1.fit)
```

Software Note: The `getVarCov()` function, which can be used to display blocks of the estimated marginal $\boldsymbol{V}$ matrix for a two-level model, can also be used to display blocks of the estimated $\boldsymbol{V}$ matrix for the models considered in this example, but some additional coding is needed. We've included this code on the Chapter 4 page of the book's website.

The EBLUPs of the random school effects and the nested random classroom effects in the model can be obtained by using the `random.effects()` function:

```
> random.effects(model4.1.fit)
```

At this point we perform a likelihood ratio test of Hypothesis 4.1, to decide if we need the nested random classroom effects in the model. We first fit a nested model, Model 4.1A, omitting the random effects associated with the classrooms. We do this by excluding the CLASSID variable from the nesting structure for the random effects in the `lme()` function:

```
> # Model 4.1A.
> model4.1A.fit <- lme(mathgain ~ 1, random = ~1 | schoolid,
  data = class, method = "REML")
```

The `anova()` function can now be used to carry out a likelihood ratio test for Hypothesis 4.1, to decide if we wish to retain the nested random effects associated with classrooms.

```
> anova(model4.1.fit, model4.1A.fit)
```

The `anova()` function subtracts the –2 REML log-likelihood value for Model 4.1 (the reference model) from that for Model 4.1A (the nested model), and refers the resulting test statistic to a χ^2 distribution with 1 degree of freedom. However, because the appropriate null distribution for the likelihood ratio test statistic for Hypothesis 4.1 is a mixture of two

χ^2 distributions, with 0 and 1 degrees of freedom and equal weights of 0.5, we multiply the p-value provided by the `anova()` function by 0.5 to obtain the correct p-value. Based on the significant result of this test ($p < 0.01$), we retain the nested random classroom effects in Model 4.1 and in all future models. We also retain the random school effects as well, to reflect the hierarchical structure of the data in the model specification.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

After obtaining the estimates of the fixed intercept and the variance components in Model 4.1, we modify the syntax to fit Model 4.2, which includes the fixed effects of the four Level 1 (student-level) covariates MATHKIND, SEX, MINORITY, and SES. Note that these covariates are added on the right-hand side of the ~ in the first argument of the `lme()` function:

```
> # Model 4.2.
> model4.2.fit <- lme(mathgain ~ mathkind + sex + minority + ses,
  random = ~1 | schoolid/classid, class,
  na.action = "na.omit", method = "REML")
```

Because some of the students might have missing data on these covariates (which is actually not the case for the Level 1 covariates in the Classroom data set), we include the argument `na.action = "na.omit"`, to tell the two functions to drop cases with missing data from the analysis.

Software Note: Without the `na.action = "na.omit"` specification, the `lme()` function will not run if there are missing data on any of the variables input to the function.

The `1` that was used to identify the intercept in the fixed part of Model 4.1 does not need to be specified in the syntax for Model 4.2, because the intercept is automatically included in any model with at least one fixed effect.

We assess the results of fitting Model 4.2 using the `summary()` function:

```
> summary(model4.2.fit)
```

We now test Hypothesis 4.2, to decide whether the fixed effects associated with all Level 1 (student-level) covariates in Model 4.2 are equal to zero, by carrying out a likelihood ratio test using the `anova()` function. To do this we refit the nested model, Model 4.1, and the reference model, Model 4.2, using ML estimation. The test statistic is calculated by the `anova()` function by subtracting the –2 ML log-likelihood for Model 4.2 (the reference model) from that for Model 4.1 (the nested model), and referring the test statistic to a χ^2 distribution with 4 degrees of freedom.

```
> # Model 4.1: ML estimation with lme().
> model4.1.ml.fit <- lme(mathgain ~ 1,
   random = ~1 | schoolid/classid, class, method = "ML")
> # Model 4.2: ML estimation with lme().
> model4.2.ml.fit <- lme(mathgain ~ mathkind + sex + minority + ses,
   random = ~1 | schoolid/classid, class,
   na.action = "na.omit", method = "ML")
> anova(model4.1.ml.fit, model4.2.ml.fit)
```

We see that at least one of the fixed effects associated with the Level 1 covariates is significant, based on the result of this test ($p < 0.001$); Subsection 4.5.2 presents details on testing Hypothesis 4.2. We therefore proceed with Model 4.2 as our preferred model.

Step 3: Build the Level 2 Model by adding Level 2 Covariates (Model 4.3).

We fit Model 4.3 by adding the fixed effects of the Level 2 (classroom-level) covariates, YEARSTEA, MATHPREP, and MATHKNOW, to Model 4.2:

```
> # Model 4.3.
> model4.3.fit <- update(model4.2.fit,
fixed = ~ mathkind + sex + minority + ses + yearstea + mathprep + mathknow)
```

We investigate the resulting parameter estimates and standard errors for the estimated fixed effects by applying the `summary()` function to the model fit object:

```
> summary(model4.3.fit)
```

We cannot consider a likelihood ratio test for the fixed effects added to Model 4.2, because some classrooms have missing data on the MATHKNOW variable, and Model 4.2 and Model 4.3 are fitted using different observations as a result. Instead, we test the fixed effects associated with the classroom-level covariates (Hypothesis 4.3 through Hypothesis 4.5) using t-tests. None of these fixed effects are significant based on the results of these t-tests (provided by the `summary()` function), so we choose Model 4.2 as the preferred model at this stage of the analysis.

Step 4: Build the Level 3 Model by adding the Level 3 Covariate (Model 4.4).

Model 4.4 can be fitted by adding the Level 3 (school-level) covariate to the formula for the fixed-effects portion of the model in the `lme()` function. We add the fixed effect of the HOUSEPOV covariate to the model by updating the `fixed =` argument for Model 4.2:

```
> # Model 4.4.
> model4.4.fit <- update(model4.2.fit,
    fixed = ~ mathkind + sex + minority + ses + housepov)
```

We apply the `summary()` function to the model fit object to obtain the resulting parameter estimates and t-tests for the fixed effects (in the case of `model4.4.fit`):

```
> summary(model4.4.fit)
```

The t-test for the fixed effect of HOUSEPOV is not significant, so we choose Model 4.2 as our final model for the Classroom data set.

4.4.3.2 Analysis Using the `lmer()` Function

We begin by loading the `lme4` package, so that the `lmer()` function can be used in the analysis:

```
> library(lme4)
```

We now proceed with the analysis steps.

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

We fit Model 4.1 to the Classroom data using the `lmer()` function as follows:

```
> # Model 4.1.
> model4.1.fit.lmer <- lmer(mathgain ~ 1 + (1|schoolid) + (1|classid),
   class, REML = T)
```

We describe the syntax used in the `lmer()` function below:

- `model4.1.fit.lmer` is the name of the object that will contain the results of the fitted model.
- Like the `lme()` function, the first argument of the function, `mathgain ~ 1`, defines the response variable, MATHGAIN, and the single fixed effect in this model, which is associated with the intercept (denoted by a 1 after the ~).
- Next, random intercepts associated with each level of SCHOOLID and CLASSID are added to the model formula (using `+` notation), using the syntax `(1|schoolid)` and `(1|classid)`. Note that a specific nesting structure does not need to be indicated here.
- The third argument, `class`, once again indicates the name of the data frame object to be used in the analysis.
- The final argument of the function, `REML = T`, tells R that REML estimation should be used for the variance components in the model. REML is also the default estimation method in the `lmer()` function.

Estimates from the model fit and 95% confidence intervals for the parameters being estimated can then be obtained using the `summary()` and `confint()` functions:

```
> summary(model4.1.fit.lmer)
> confint(model4.1.fit.lmer)
```

The EBLUPs of the random school effects and the nested random classroom effects in the model can be obtained using the `ranef()` function:

```
> ranef(model4.1.fit.lmer)
```

We then compute measures of uncertainty associated with these predictions of the random effects at each level of the Classroom data using the `merTools` package within R. After loading this package, we use the `REsim()` and `plotREsim()` functions from this package to display and then plot estimated standard deviations of these predictions:

```
> library(merTools)
> REsim(model4.1.fit.lmer)
> plotREsim(REsim(model4.1.fit.lmer))
```

This code generates the paneled plot shown in Figure 4.4. The function conveniently generates plots for both the school and classroom levels, enabling visualization of the distributions of the predicted random effects at each level. We don't observe any extreme outliers at either level in these plots, given that the intervals defined by the estimated standard deviations of the predictions all cover the horizontal 0 line. But these plots still do a good job of conveying the variance in the random effects at each level of the data (along with their uncertainty).

At this point we perform a likelihood ratio test of Hypothesis 4.1, to decide if we need the nested random classroom effects in the model. We first fit a nested model, Model 4.1A, omitting the random effects associated with the classrooms. We do this by excluding `(1|CLASSID)` from the model formula for the `lmer()` function:

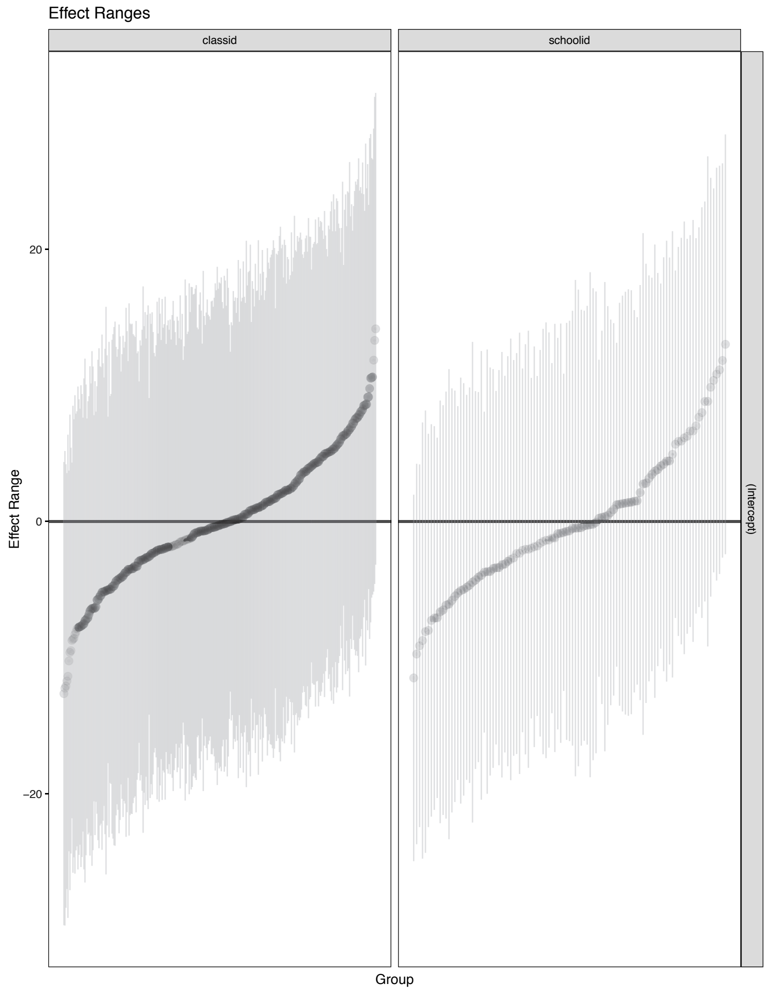


FIGURE 4.4: Plots of predicted EBLUPs at the classroom and school levels in R, including estimated standard deviations of the predictions.

```
> # Model 4.1A.
> model4.1A.fit.lmer <- lmer(mathgain ~ 1 + (1|schoolid),
   class, REML = T)
```

The anova() function can now be used to carry out a likelihood ratio test for Hypothesis 4.1, to decide if we wish to retain the nested random effects associated with classrooms.

```
> anova(model4.1.fit.lmer, model4.1A.fit.lmer)
```

The anova() function subtracts the −2 REML log-likelihood value for Model 4.1 (the reference model) from that for Model 4.1A (the nested model), and refers the resulting test statistic to a χ^2 distribution with 1 degree of freedom. However, because the appropriate null distribution for the likelihood ratio test statistic for Hypothesis 4.1 is a mixture of two χ^2 distributions, with 0 and 1 degrees of freedom and equal weights of 0.5, we multiply the p-value provided by the anova() function by 0.5 to obtain the correct p-value. Based on the significant result of this test ($p < 0.01$), we retain the nested random classroom effects in Model 4.1 and in all future models. We also retain the random school effects as well, to reflect the hierarchical structure of the data in the model specification.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

After obtaining the estimates of the fixed intercept and the variance components in Model 4.1, we modify the syntax to fit Model 4.2, which includes the fixed effects of the four Level 1 (student-level) covariates MATHKIND, SEX, MINORITY, and SES. Note that these covariates are added on the right-hand side of the ~ in the first argument of the `lmer()` function:

```
> # Model 4.2.
> model4.2.fit.lmer <- lmer(mathgain ~ mathkind + sex + minority + ses
   + (1|schoolid) + (1|classid),
class, na.action = "na.omit", REML = T)
```

Because some of the students might have missing data on these covariates (which is actually not the case for the Level 1 covariates in the Classroom data set), we include the argument `na.action = "na.omit"`, to tell the two functions to drop cases with missing data from the analysis.

Software Note: Without the `na.action = "na.omit"` specification, the `lmer()` function will not run if there are missing data on any of the variables input to the function.

Software Note: The `lmer()` function in the `lme4` package does not automatically compute p-values for the t-statistics that are generated by dividing the fixed-effect parameter estimates by their standard errors (for testing the hypothesis that a given fixed effect parameter is equal to zero). This is primarily due to the fact that these test statistics are approximate, and several methods exist for approximating the degrees of freedom. Instead, approximate tests available in the `lmerTest` package can be used; see Chapter 3 for an example using two-level models. When possible, we use likelihood ratio tests for the fixed-effect parameters in this chapter.

The `1` that was used to identify the intercept in the fixed part of Model 4.1 does not need to be specified in the syntax for Model 4.2, because the intercept is automatically included by the two functions in any model with at least one fixed effect.

We assess the results of fitting Model 4.2 using the `summary()` function:

```
> summary(model4.2.fit.lmer)
```

We now test Hypothesis 4.2, to decide whether the fixed effects associated with all Level 1 (student-level) covariates in Model 4.2 are equal to zero, by carrying out a likelihood ratio test using the `anova()` function. To do this, we refit the nested model, Model 4.1, and the reference model, Model 4.2, using ML estimation (note the `REML = F` arguments below). The test statistic is calculated by the `anova()` function by subtracting the –2 ML log-likelihood for Model 4.2 (the reference model) from that for Model 4.1 (the nested model), and referring the test statistic to a χ^2 distribution with 4 degrees of freedom.

```
> # Model 4.1: ML estimation with lmer().
```

```
> model4.1.lmer.ml.fit <- lmer(mathgain ~ 1 + (1|schoolid) + (1|classid),
   class, REML = F)
> # Model 4.2: ML estimation with lmer().
> model4.2.lmer.ml.fit <- lmer(mathgain ~ mathkind + sex + minority + ses
   + (1|schoolid) + (1|classid),
   class, REML = F)
> anova(model4.1.lmer.ml.fit, model4.2.lmer.ml.fit)
```

We see that at least one of the fixed effects associated with the Level 1 covariates is significant, based on the result of this test ($p < 0.001$); Subsection 4.5.2 presents details on testing Hypothesis 4.2. We therefore proceed with Model 4.2 as our preferred model.

Step 3: Build the Level 2 Model by adding Level 2 Covariates (Model 4.3).

We fit Model 4.3 by adding the fixed effects of the Level 2 (classroom-level) covariates, YEARSTEA, MATHPREP, and MATHKNOW, to Model 4.2:

```
> # Model 4.3.
> model4.3.fit.lmer <- lmer(mathgain ~ mathkind + sex + minority + ses
   + yearstea + mathprep + mathknow
   + (1|schoolid) + (1|classid),
   class, na.action = "na.omit", REML = T)
```

We investigate the resulting parameter estimates and standard errors for the estimated fixed effects by applying the `summary()` function to the model fit object:

```
> summary(model4.3.fit.lmer)
```

We cannot consider a likelihood ratio test for the fixed effects added to Model 4.2, because some classrooms have missing data on the MATHKNOW variable, and Model 4.2 and Model 4.3 are fitted using different observations as a result. Instead, we can refer the test statistics provided by the `summary()` function to standard normal distributions to make approximate inferences about the importance of the effects (where a test statistic larger than 1.96 in absolute value would suggest a significant fixed effect at the 0.05 significance level, under asymptotic assumptions). These tests would suggest that none of these fixed effects are significant, so we would not retain them in this model.

Step 4: Build the Level 3 Model by adding the Level 3 Covariate (Model 4.4).

Model 4.4 can be fitted by adding the Level 3 (school-level) covariate HOUSEPOV to the formula for the fixed-effects portion of Model 4.2:

```
> # Model 4.4.
> model4.4.fit.lmer <- lmer(mathgain ~ mathkind + sex + minority + ses
   + housepov + (1|schoolid) + (1|classid),
   class, na.action = "na.omit", REML = T)
```

We apply the `summary()` function to the model fit object to obtain the resulting parameter estimates and standard errors:

```
> summary(model4.4.fit.lmer)
```

Based on the test statistic for the fixed effect of HOUSEPOV (-1.151), we once again do not have enough evidence to say that this effect is different from 0 (at the 0.05 level), so we choose Model 4.2 as our final model for the Classroom data set.

4.4.4 Stata

We start by using web-aware Stata (Version 17) to import the Classroom data directly from the book's website:

```
. insheet using http://www-personal.umich.edu/~bwest/classroom.csv, clear
```

The `mixed` command can then be used to fit three-level hierarchical models with nested random effects.

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

We first specify the `mixed` syntax to fit Model 4.1, including the random effects of schools and of classrooms nested within schools:

```
. * Model 4.1.
. mixed mathgain || schoolid: || classid:, variance reml dfmethod(kroger)
```

The first variable listed after invoking `mixed` is the continuous dependent variable, MATHGAIN. No covariates are specified after the dependent variable, because the only fixed effect in Model 4.1 is the intercept, which is included by default.

After the first clustering indicator (`||`), we list the random factor identifying clusters at Level 3 of the data set, SCHOOLID, followed by a colon (`:`). We then list the nested random factor, CLASSID, after a second clustering indicator. This factor identifies clusters at Level 2 of the data set, and is again followed by a colon.

> **Software Note:** If a multilevel data set is organized by a series of nested groups, such as classrooms nested within schools as in this example, the random effects structure of the mixed model is specified in `mixed` by listing the random factors defining the structure, separated by two vertical bars (`||`). The nesting structure reads left to right; e.g., SCHOOLID is the highest level of clustering, with levels of CLASSID nested within each school.
>
> If no variables are specified after the colon at a given level of the nesting structure, the model will only include a single random effect (associated with the intercept) for each level of the random factor. Additional covariates with random effects at a given level of the nesting structure can be specified after the colon.

Finally, the `variance` and `reml` options specified after a comma request that the estimated variances of the random school and classroom effects, rather than their estimated standard deviations, should be displayed in the output, and that REML estimation should be used to fit this model (ML estimation is the default in Stata). The `dfmethod(kroger)` option requests that the Kenward–Roger method be used to approximate the degrees of freedom for the t-statistics used to test hypotheses about the fixed effect parameters (see Chapter 2 for details).

Information criteria, including the REML log-likelihood, can be obtained by using the `estat ic` command after submitting the `mixed` command:

```
. estat ic
```

In the output associated with the fit of Model 4.1, Stata automatically reports a likelihood ratio test, calculated by subtracting the –2 REML log-likelihood of Model 4.1 (including the random school effects and nested random classroom effects) from the –2 REML log-likelihood of a simple linear regression model without the random effects. Stata reports the following note along with the test:

```
Note: LR test is conservative and provided only for reference
```

Stata performs a classical likelihood ratio test here, where the distribution of the test statistic (under the null hypothesis that both variance components are equal to zero) is asymptotically χ^2_2 (where the 2 degrees of freedom correspond to the two variance components in Model 4.1). Appropriate theory for testing a model with multiple random effects (e.g., Model 4.1) vs. a model without any random effects has yet to be developed, and Stata discusses this issue in detail if users click on the `LR test is conservative` note. The p-value for this test statistic is known to be larger than it should be (making it conservative).

We recommend testing the need for the random effects by using individual likelihood ratio tests, based on REML estimation of nested models. To test Hypothesis 4.1, and decide whether we want to retain the nested random effects associated with classrooms in Model 4.1, we fit a nested model, Model 4.1A, again using REML estimation:

```
. * Model 4.1A.
. mixed mathgain || schoolid:, variance reml dfmethod(kroger)
```

The test statistic for Hypothesis 4.1 can be calculated by subtracting the –2 REML log-likelihood for Model 4.1 (the reference model) from that of Model 4.1A (the nested model). The p-value for the test statistic (7.9) is based on a mixture of χ^2 distributions with 0 and 1 degrees of freedom, and equal weight 0.5. Because of the significant result of this test ($p = 0.002$), we retain the nested random classroom effects in Model 4.1 and in all future models (see Subsection 4.5.1 for a discussion of this test). We also retain the random effects associated with schools, to reflect the hierarchical structure of the data set in the model.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

We fit Model 4.2 by adding the fixed effects of the four student-level covariates, MATHKIND, SEX, MINORITY, and SES, using the following syntax:

```
. * Model 4.2.
. mixed mathgain mathkind sex minority ses || schoolid: || classid:, ///
   variance reml dfmethod(kroger)
```

Information criteria associated with the fit of this model can be obtained by using the `estat ic` command after the `mixed` command has finished running.

We test Hypothesis 4.2 to decide whether the fixed effects that were added to Model 4.1 to form Model 4.2 are all equal to zero, using a likelihood ratio test. We first refit the nested model, Model 4.1, and the reference model, Model 4.2, using ML estimation. We specify the `mle` option to request maximum likelihood estimation for each model. The `est store` command is then used to store the results of each model fit in new objects.

```
. * Model 4.1: ML Estimation.
. mixed mathgain || schoolid: || classid:, variance mle
. est store model4_1_ml_fit

. * Model 4.2: ML Estimation.
. mixed mathgain mathkind sex minority ses || schoolid: || classid:, ///
   variance mle
. est store model4_2_ml_fit
```

We use the `lrtest` command to perform the likelihood ratio test. The likelihood ratio test statistic is calculated by subtracting the –2 ML log-likelihood for Model 4.2 from that for Model 4.1, and referring the difference to a χ^2 distribution with 4 degrees of freedom. The likelihood ratio test requires that both models are fitted using the same cases.

```
. lrtest model4_1_ml_fit model4_2_ml_fit
```

Based on the significant result ($p < 0.001$) of this test, we choose Model 4.2 as our preferred model at this stage of the analysis. We discuss the likelihood ratio test for Hypothesis 4.2 in more detail in Subsection 4.5.2.

Step 3: Build the Level 2 Model by adding Level 2 Covariates (Model 4.3).

To fit Model 4.3, we modify the `mixed` command used to fit Model 4.2 by adding the fixed effects of the classroom-level covariates, YEARSTEA, MATHPREP, and MATHKNOW, to the fixed portion of the command. We again use REML estimation for this model, and obtain the model information criteria by using the post-estimation command `estat ic`.

```
. * Model 4.3.
. mixed mathgain mathkind sex minority ses yearstea mathprep mathknow ///
   || schoolid: || classid:, variance reml dfmethod(kroger)
. estat ic
```

We do not consider a likelihood ratio test for the fixed effects added to Model 4.2 to form Model 4.3, because Model 4.3 was fitted using different cases, owing to the presence of missing data on some of the classroom-level covariates. Instead, we consider the t-tests reported by Stata for Hypotheses 4.3 through 4.5. None of the t-tests reported for the fixed effects of the classroom-level covariates are significant; recall that we used the Kenward–Roger method to approximate the degrees of freedom for the t-statistics, which is more appropriate when there are smaller sample sizes at some levels of the data hierarchy. Therefore, we do not retain these fixed effects in Model 4.3, and choose Model 4.2 as our preferred model at this stage of the analysis.

Step 4: Build the Level 3 Model by adding the Level 3 Covariate (Model 4.4).

To fit Model 4.4, we add the fixed effect of the school-level covariate, HOUSEPOV, to the model, by updating the `mixed` command that was used to fit Model 4.2. We again use the default REML estimation method, and use the `estat ic` post-estimation command to obtain information criteria for Model 4.4.

```
. * Model 4.4.
. mixed mathgain mathkind sex minority ses housepov ///
   || schoolid: || classid:, variance reml dfmethod(kroger)
. estat ic
```

To test Hypothesis 4.6, we use the t-test reported by the `mixed` command for the fixed effect of HOUSEPOV. Because of the nonsignificant test result ($p = 0.25$), we do not retain this fixed effect, and choose Model 4.2 as our final model for the analysis of the Classroom data.

At this point, we can visualize marginal predictions of MATHGAIN as a function of MATHKIND and MINORITY (two of the most important predictors) after refitting Model 4.2:

```
. * Model 4.2 (final model), with visualization of predicted values.
. mixed mathgain mathkind sex minority ses housepov ///
   || schoolid: || classid:, variance reml dfmethod(kroger)
. estat ic
. margins, at(mathkind=(300(50)600)) by(minority)
. marginsplot
```

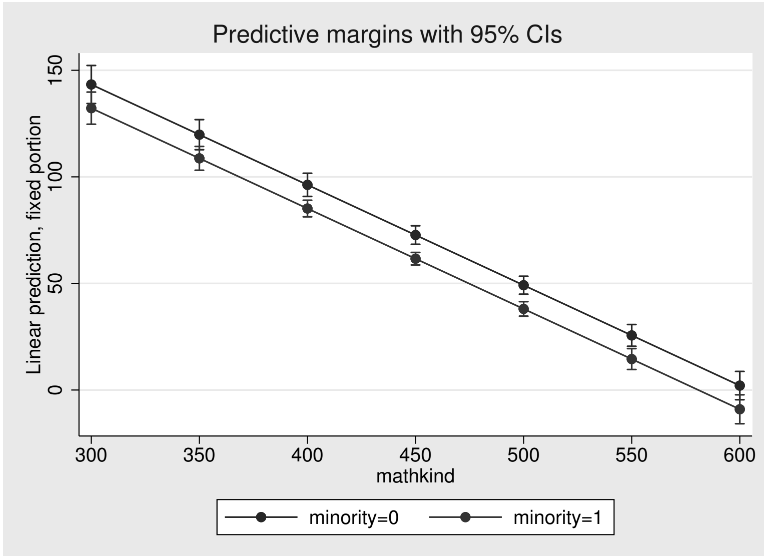


FIGURE 4.5: Plot of predicted marginal means for MATHGAIN in Stata, with 95% confidence intervals.

We note that the `at()` option allows for specifying a series of values on a predictor at which one wishes to see marginal predictions of the dependent variable. In this case, we say that we want values of MATHKIND ranging from 300 to 600, increasing by units of 50. We also request that predictions be displayed by values of the MINORITY variable. These commands produce the plot shown in Figure 4.5, which clearly shows the strong relationships of these variables with MATHGAIN.

4.4.5 HLM

We assume that the first MDM file discussed in the initial data summary (Subsection 4.2.2) has been generated using HLM3, and proceed to the model-building window.

Step 1: Fit the initial "unconditional" (variance components) model (Model 4.1), and decide whether to omit the random classroom effects (Model 4.1 vs. Model 4.1A).

We begin by specifying the Level 1 (student-level) model. In the model-building window, click on MATHGAIN, and identify it as the **Outcome variable**. Go to the **Basic Settings** menu and identify the outcome variable as a **Normal** (Continuous) variable (this is the default). Choose a title for this analysis (such as "Classroom Data: Model 4.1"), and choose a location and name for the output (.html) file that will contain the results of the model fit. Click **OK** to return to the model-building window. Under the **File** menu, click **Preferences**, and then click **Use level subscripts** to display subscripts in the model-building window.

Three models will now be displayed. The **Level 1 model** describes the "means-only" model at the student level. We show the Level 1 model below as it is displayed in the HLM model-building window:

Model 4.1: Level 1 Model

$$\text{MATHGAIN}_{ijk} = \pi_{0jk} + e_{ijk}$$

The value of MATHGAIN for an individual student i, within classroom j nested in school k, depends on the intercept for classroom j within school k, π_{0jk}, plus a residual error, e_{ijk}, associated with the student.

The **Level 2 model** describes the classroom-specific intercept in **Model 4.1** at the classroom level of the data set:

Model 4.1: Level 2 Model

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

The classroom-specific intercept, π_{0jk}, depends on the school-specific intercept, β_{00k}, and a random effect, r_{0jk} associated with the j-th classroom within school k.

The **Level 3 model** describes the school-specific intercept in Model 4.1:

Model 4.1: Level 3 Model

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

The school-specific intercept, β_{00k}, depends on the overall (grand) mean, γ_{000}, plus a random effect, u_{00k} associated with the school.

The overall "means-only" mixed model derived from the preceding Level 1, Level 2, and Level 3 models can be displayed by clicking on the **Mixed** button:

Model 4.1: Overall Mixed Model

$$\text{MATHGAIN}_{ijk} = \gamma_{000} + r_{0jk} + u_{00k} + e_{ijk}$$

An individual student's MATHGAIN depends on an overall fixed intercept, γ_{000} (which represents the overall mean of MATHGAIN across all students), a random effect associated with the student's classroom, r_{0jk}, a random effect associated with the student's school, u_{00k}, and a residual error, e_{ijk}. Table 4.3 shows the correspondence of this HLM notation with the general notation used in (4.1).

To fit Model 4.1, click **Run Analysis**, and select **Save as and Run** to save the .hlm command file. You will be prompted to supply a name and location for this .hlm file. After the estimation has finished, click on **File**, and select **View Output** to see the resulting parameter estimates and fit statistics.

At this point, we test the significance of the random effects associated with classrooms nested within schools (Hypothesis 4.1). However, because the HLM3 procedure does not allow users to remove all random effects from a given level of a hierarchical model (in this example, the classroom level, or the school level), we cannot perform a likelihood ratio test of Hypothesis 4.1, as was done in the other software procedures. Instead, HLM provides chi-square tests that are calculated using methodology described in Raudenbush & Bryk (2002). The following output is generated by HLM3 after fitting Model 4.1:

```
Final estimation of level-1 and level-2 variance components:
--------------------------------------------------------------------
Random Effect              Standard     Variance     df     Chi-     P-value
                           Deviation    Component          square
--------------------------------------------------------------------
INTRCPT1,          r0      10.02212     100.44281   205   301.95331  <0.001
level-1,           e       32.05828    1027.73315
```

```
Final estimation of level-3 variance components:
--------------------------------------------------------------------
Random Effect         Standard    Variance    df    Chi-     P-value
                      Deviation   Component         square
--------------------------------------------------------------------
INTRCPT1/      u00    8.66240     75.03712    106  165.74813  <0.001
   INTRCPT2,
```

The chi-square test statistic for the variance of the nested random classroom effects (301.95) is significant ($p < 0.001$), so we reject the null hypothesis for Hypothesis 4.1 and retain the random effects associated with both classrooms and schools in Model 4.1 and all future models. We now proceed to fit Model 4.2.

Step 2: Build the Level 1 Model by adding Level 1 Covariates (Model 4.1 vs. Model 4.2).

We specify the **Level 1 Model** for Model 4.2 by clicking on **Level 1** to add fixed effects associated with the student-level covariates to the model. We first select the variable MATHKIND, choose **add variable uncentered**, and then repeat this process for the variables SEX, MINORITY, and SES. Notice that as each covariate is added to the Level 1 model, the Level 2 and Level 3 models are also updated. The new Level 1 model is as follows:

Model 4.2: Level 1 Model

$$\begin{aligned}\text{MATHGAIN}_{ijk} = \quad & \pi_{0jk} + \pi_{1jk}(\text{SEX}_{ijk}) + \pi_{2jk}(\text{MINORITY}_{ijk}) \\ & + \pi_{3jk}(\text{MATHKIND}_{ijk}) + \pi_{4jk}(\text{SES}_{ijk}) + e_{ijk}\end{aligned}$$

This updated Level 1 model shows that a student's MATHGAIN now depends on the intercept specific to classroom j, π_{0jk}, the classroom-specific effects ($\pi_{1jk}, \pi_{2jk}, \pi_{3jk}$, and π_{4jk}) of each of the student-level covariates, and a residual error, e_{ijk}.

The **Level 2** portion of the model-building window displays the classroom-level equations for the student-level intercept (π_{0jk}) and for each of the student-level effects (π_{1jk} through π_{4jk}) defined in this model. The equation for each effect from HLM is as follows:

Model 4.2: Level 2 Model

$$\begin{aligned}\pi_{0jk} &= \beta_{00k} + r_{0jk} \\ \pi_{1jk} &= \beta_{10k} \\ \pi_{2jk} &= \beta_{20k} \\ \pi_{3jk} &= \beta_{30k} \\ \pi_{4jk} &= \beta_{40k}\end{aligned}$$

The equation for the student-level intercept (π_{0jk}) has the same form as in Model 4.1. It includes an intercept specific to school k, β_{00k}, plus a random effect, r_{0jk}, associated with each classroom in school k. Thus, the student-level intercepts are allowed to vary randomly from classroom to classroom within the same school.

The equations for each of the effects associated with the four student-level covariates (π_{1jk} through π_{4jk}) are all constant at the classroom level. This means that the effects of being female, being a minority student, kindergarten math achievement, and student-level SES are assumed to be the same for students within all classrooms (i.e., these coefficients do not vary across classrooms within a given school).

The **Level 3** portion of the model-building window shows the school-level equations for the school-specific intercept, β_{00k}, and for each of the school-specific effects in the classroom-level model, β_{10k} through β_{40k}:

Model 4.2: Level 3 Model

$$\begin{aligned}
\beta_{00k} &= \gamma_{000} + u_{00k} \\
\beta_{10k} &= \gamma_{100} \\
\beta_{20k} &= \gamma_{200} \\
\beta_{30k} &= \gamma_{300} \\
\beta_{40k} &= \gamma_{400}
\end{aligned}$$

The equation for the school-specific intercept includes a parameter for an overall fixed intercept, γ_{000}, plus a random effect, u_{00k}, associated with the school. Thus, the intercepts are allowed to vary randomly from school to school, as in Model 4.1. However, the effects (β_{10k} through β_{40k}) associated with each of the covariates measured at the student level are not allowed to vary from school to school. This means that the effects of being female, being a minority student, of kindergarten math achievement, and of student-level SES are assumed to be the same across all schools.

Click the **Mixed** button to view the overall LMM specified for Model 4.2:

Model 4.2: Overall Mixed Model

$$\begin{aligned}
\text{MATHGAIN}_{ijk} = \; & \gamma_{000} + \gamma_{100} * \text{SEX}_{ijk} + \gamma_{200} * \text{MINORITY}_{ijk} \\
& + \gamma_{300} * \text{MATHKIND}_{ijk} + \gamma_{400} * \text{SES}_{ijk} + r_{0jk} + u_{00k} + e_{ijk}
\end{aligned}$$

The HLM specification of the model at each level results in the same overall LMM (Model 4.2) that is fitted in the other software procedures. Table 4.3 shows the correspondence of the HLM notation with the general model notation used in (4.1).

At this point we wish to test Hypothesis 4.2, to decide whether the fixed effects associated with the Level 1 (student-level) covariates should be added to Model 4.1. We set up the likelihood ratio test for Hypothesis 4.2 in HLM before running the analysis for Model 4.2.

To set up a likelihood ratio test of Hypothesis 4.2, click on **Other Settings** and select **Hypothesis Testing**. Enter the **Deviance** (or –2 ML log-likelihood)[3] displayed in the output for Model 4.1 (deviance = 11771.33) and the **Number of Parameters** from Model 4.1 (number of parameters = 4: the fixed intercept, and the three variance components) in the Hypothesis Testing window. After fitting Model 4.2, HLM calculates the appropriate likelihood ratio test statistic and corresponding p-value for Hypothesis 4.2 by subtracting the deviance statistic for Model 4.2 (the reference model) from that for Model 4.1 (the nested model).

After setting up the analysis for Model 4.2, click **Basic Settings**, and enter a new title for this analysis, in addition to a new file name for the saved output. Finally, click **Run Analysis**, and choose **Save as and Run** to save a new .hlm command file for this model. After the analysis has finished running, click **File** and **View Output** to see the results.

Based on the significant ($p < 0.001$) result of the likelihood ratio test for the student-level fixed effects, we reject the null for Hypothesis 4.2 and conclude that the fixed effects at Level 1 should be retained in the model. The results of the test of Hypothesis 4.2 are discussed in more detail in Subsection 4.5.2.

The significant test result for Hypothesis 4.2 also indicates that the fixed effects at Level 1 help to explain variance in the residual errors at the student level of the data. A comparison

[3]HLM reports the value of the –2 ML log-likelihood for a given model as the model deviance.

of the estimated residual error variance for Model 4.2 vs. that for Model 4.1, both calculated using ML estimation in HLM3, provides evidence that the residual error variance at Level 1 is in fact substantially reduced in Model 4.2 (as discussed in Subsection 4.7.2). We retain the fixed effects of the Level 1 covariates in Model 4.2 and proceed to consider Model 4.3.

Step 3: Build the Level 2 Model by adding Level 2 covariates (Model 4.3).

Before fitting Model 4.3, we need to add the MATHPREP and MATHKNOW variables to the MDM file (as discussed in Subsection 4.2.2). We then need to recreate Model 4.2 in the model-building window.

We obtain Model 4.3 by adding the Level 2 (classroom-level) covariates to Model 4.2. To do this, first click on **Level 2**, then click on the Level 2 model for the intercept term (π_{0jk}); include the nested random classroom effects, r_{0jk}, and add the uncentered versions of the classroom-level variables, YEARSTEA, MATHPREP, and MATHKNOW, to the Level 2 model for the intercept. This results in the following Level 2 model for the classroom-specific intercepts:

Model 4.3: Level 2 Model for Classroom-Specific Intercepts

$$\begin{aligned}\pi_{0jk} = \; & \beta_{00k} + \beta_{01k}(\text{YEARSTEA}_{jk}) + \beta_{02k}(\text{MATHKNOW}_{jk}) \\ & + \beta_{03k}(\text{MATHPREP}_{jk}) + r_{0jk}\end{aligned}$$

We see that adding the classroom-level covariates to the model implies that the randomly varying intercepts at Level 1 (the values of π_{0jk}) depend on the school-specific intercept (β_{00k}), the classroom-level covariates, and the random effect associated with each classroom (i.e., the value of r_{0jk}).

The effects of the student-level covariates (π_{1jk} through π_{4jk}) have the same expressions as in Model 4.2 (they are again assumed to remain constant from classroom to classroom).

Adding the classroom-level covariates to the Level 2 model for the intercept causes HLM to include additional Level 3 equations for the effects of the classroom-level covariates in the model-building window, as follows:

Model 4.3: Level 3 Model (Additional Equations)

$$\begin{aligned}\beta_{01k} &= \gamma_{010} \\ \beta_{02k} &= \gamma_{020} \\ \beta_{03k} &= \gamma_{030}\end{aligned}$$

These equations show that the effects of the Level 2 (classroom-level) covariates are constant at the school level. That is, the classroom-level covariates are not allowed to have effects that vary randomly at the school level, although we could set up the model to allow this.

Click the **Mixed** button in the HLM model-building window to view the overall mixed model for Model 4.3:

Model 4.3: Overall Mixed Model

$$\begin{aligned}\text{MATHGAIN}_{ijk} = \; & \gamma_{000} + \gamma_{010} * \text{YEARSTEA}_{jk} + \gamma_{020} * \text{MATHKNOW}_{jk} \\ & + \gamma_{030} * \text{MATHPREP}_{jk} + \gamma_{100} * \text{SEX}_{ijk} + \gamma_{200} * \text{MINORITY}_{ijk} \\ & + \gamma_{300} * \text{MATHKIND}_{ijk} + \gamma_{400} * \text{SES}_{ijk} + r_{0jk} + u_{00k} + e_{ijk}\end{aligned}$$

We see that the LMM specified here is the same model that is being fit using the other software procedures. Table 4.3 shows the correspondence of the HLM model parameters with the parameters that we use in (4.1).

After setting up Model 4.3, click **Basic Settings** to enter a new name for this analysis and a new name for the .html output file. Click **OK**, and then click **Run Analysis**, and choose **Save as and Run** to save a new .hlm command file for this model before fitting the model. After the analysis has finished running, click **File** and **View Output** to see the results.

We use t-tests for Hypotheses 4.3 through 4.5 to decide if we want to keep the fixed effects associated with the Level 2 covariates in Model 4.3 (a likelihood ratio test based on the deviance statistics for Model 4.2 and Model 4.3 is not appropriate, due to the missing data on the classroom-level covariates). Based on the nonsignificant t-tests for each of the classroom-level fixed effects displayed in the HLM output, we choose Model 4.2 as our preferred model at this stage of the analysis.

Step 4: Build the Level 3 Model by adding the Level 3 covariate (Model 4.4).

In this step, we add the school-level covariate to Model 4.2 to obtain Model 4.4. We first open the .hlm file corresponding to Model 4.2 from the model-building window by clicking **File**, and then **Edit/Run old command file**. After locating the .hlm file saved for Model 4.2, open the file, and click the **Level 3** button. Click on the first Level 3 equation for the intercept that includes the random school effects (u_{00k}). Add the uncentered version of the school-level covariate, HOUSEPOV, to this model for the intercept. The resulting Level 3 model is as follows:

Model 4.4: Level 3 Model for School-Specific Intercepts

$$\beta_{00k} = \gamma_{000} + \gamma_{001}(\text{HOUSEPOV}_k) + u_{00k}$$

The school specific intercepts, β_{00k}, in this model now depend on the overall fixed intercept, γ_{000}, the fixed effect, γ_{001}, of HOUSEPOV, and the random effect, u_{00k}, associated with school k.

After setting up Model 4.4, click **Basic Settings** to enter a new name for this analysis and a new name for the .html output file. Click **OK**, and then click **Run Analysis**, and choose **Save as and Run** to save a new .hlm command file before fitting the model. After the analysis has finished running, click **File** and **View Output** to see the results.

We test Hypothesis 4.6 using a t-test for the fixed effect associated with HOUSEPOV in Model 4.4. Based on the nonsignificant result of this t-test ($p = 0.25$), we do not retain the fixed effect of HOUSEPOV, and choose Model 4.2 as our final model in the analysis of the Classroom data set.

We now generate residual files to be used in checking model diagnostics (discussed in Section 4.10) for Model 4.2. First, open the .hlm file for Model 4.2, and click **Basic Settings**. In this window, specify names and file types (we choose to save SPSS-format data files in this example) for the Level 1, Level 2, and Level 3 "Residual" files (click on the buttons for each of the three files). The Level 1 file will contain the Level 1 residuals in a variable named **l1resid**, and the conditional predicted values of the dependent variable in a variable named **fitval**. The Level 2 residual file will include a variable named **ebintrcp**, and the Level 3 residual file will include a variable named **eb00**; these variables will contain the Empirical Bayes (EB) predicted values (i.e., the EBLUPs) of the random classroom and school effects, respectively. These three files can be used for exploration of the distributions of the EBLUPs and the Level 1 residuals.

Covariates measured at the three levels of the Classroom data set can also be included in the three files, although we do not use that option here. Rerun the analysis for Model 4.2 to generate the residual files, which will be saved in the same folder where the .html output file was saved. We apply SPSS syntax to the resulting residual files in Section 4.10, to check the diagnostics for Model 4.2.

4.5 Results of Hypothesis Tests

4.5.1 Likelihood Ratio Tests for Random Effects

When the "step-up" approach to model building is used for three-level random intercept models, as for the Classroom data, random effects are usually retained in the model, regardless of the results of significance tests for the associated covariance parameters. However, when tests of significance for random effects are desired, we recommend using likelihood ratio tests, which require fitting a nested model (in which the random effects in question are omitted) and a reference model (in which the random effects are included). Both the nested and reference models should be fitted using REML estimation. The exception to this is SAS `proc glimmix` which allows the user to specify an appropriate likelihood ratio test for covariance parameters by fitting the reference model only, and then specifying an appropriate test using the `covtest` statement. This makes it unnecessary to fit a nested model in order to carry out a likelihood ratio test for covariance parameters when fitting LMMs using SAS `proc glimmix`.

TABLE 4.5: Summary of Hypothesis Test Results for the Classroom Analysis

Hypothesis Label	Test	Models Compared (Nested vs. Reference)[a]	Estimation Method[b]	Test Statistic Values (Calculation)	p-Value
4.1	LRT	4.1A vs. 4.1	REML	$\chi^2(0:1) = 7.9$ (11776.7–11768.8)	< 0.01
4.2	LRT	4.1 vs. 4.2	ML	$\chi^2(4) = 380.4$ (11771.3–11390.9)	< 0.01
4.3	t-test	4.3	REML ML	$t(792) = 0.34$ $t(177) = 0.35$	0.73 0.72
4.4	t-test	4.3	REML ML	$t(792) = 0.97$ $t(177) = 0.97$	0.34 0.34
4.5	t-test	4.3	REML ML	$t(792) = 1.67$ $t(177) = 1.67$	0.10 0.10
4.6	t-test	4.4	REML ML	$t(873) = -1.15$ $t(105) = -1.15$	0.25 0.25

Note: See Table 4.4 for null and alternative hypotheses, and distributions of test statistics under H_0.

[a] Nested models are not necessary for the t-tests of Hypothesis 4.4 through Hypothesis 4.6.

[b] The HLM3 procedure uses ML estimation only; we also report results based on REML estimation from SAS `proc glimmix`.

Likelihood ratio tests for the random effects in a three-level random intercept model are not possible when using the HLM3 procedure, because (1) HLM3 uses ML rather than REML estimation, and (2) HLM in general will not allow models to be specified that do not include random effects at each level of the data. Instead, HLM implements alternative chi-square tests for the variance of random effects, which are discussed in more detail in Raudenbush & Bryk (2002).

In this section, we present the results of the likelihood ratio test for the random effects in Model 4.1, based on fitting Model 4.1, and using the `covtest` statement to specify the hypothesis that we wish to test. (See Subsection 4.4.1 for the SAS code to carry out this test.)

Hypothesis 4.1: The random effects associated with classrooms nested within schools can be omitted from Model 4.1.

We calculate the likelihood ratio test statistic for Hypothesis 4.1 by subtracting the value of the −2 REML log-likelihood for Model 4.1 (the reference model) from the value for Model 4.1A (the nested model excluding the random classroom effects). The resulting test statistic is equal to 7.9 (see Table 4.5). Because a variance cannot be less than zero, the null hypothesis value of $\sigma^2_{int:classroom} = 0$ is at the boundary of the parameter space, and the null distribution of the likelihood ratio test statistic is a mixture of χ^2_0 and χ^2_1 distributions, each having equal weight 0.5 (Verbeke & Molenberghs, 2000). The calculation of the p-value for the likelihood ratio test statistic is as follows:

$$p\text{-}value = 0.5 \times p(\chi^2_0 > 7.9) + 0.5 \times p(\chi^2_1 > 7.9) < 0.01$$

Based on the result of this test, we conclude that there is significant variance in the MATHGAIN means between classrooms nested within schools, and we retain the random effects associated with classrooms in Model 4.1 and in all subsequent models. We also retain the random school effects, without testing them, to reflect the hierarchical structure of the data in the model specification.

4.5.2 Likelihood Ratio Tests and *t*-Tests for Fixed Effects

Hypothesis 4.2: The fixed effects, β_1, β_2, β_3, and β_4, associated with the four student-level covariates, MATHKIND, SEX, MINORITY, and SES, should be added to Model 4.1.

We test Hypothesis 4.2 using a likelihood ratio test, based on ML estimation. We calculate the likelihood ratio test statistic by subtracting the −2 ML log-likelihood for Model 4.2 (the reference model including the four student-level fixed effects) from the corresponding value for Model 4.1 (the nested model excluding the student-level fixed effects). The distribution of the test statistic, under the null hypothesis that the four fixed effect parameters are all equal to zero, is asymptotically a χ^2 with 4 degrees of freedom. Because the p-value is significant ($p < 0.001$), we add the fixed effects associated with the four student-level covariates to the model and choose Model 4.2 as our preferred model at this stage of the analysis.

Recall that likelihood ratio tests are only valid if both the reference and nested models are fitted using the same observations, and the fits of these two models are based on all 1190 cases in the data set.

Hypotheses 4.3, 4.4, and 4.5: The fixed effects, β_5, β_6, and β_7, associated with the classroom-level covariates, YEARSTEA, MATHKNOW, and MATHPREP, should be retained in Model 4.3.

We are unable to use a likelihood ratio test for the fixed effects of all the Level 2 (classroom-level) covariates, because cases are lost due to missing data on the MATHKNOW variable. Instead, we consider individual t-tests for the fixed effects of the classroom-level covariates in Model 4.3.

To illustrate testing Hypothesis 4.3, we consider the t-test reported by HLM3 for the fixed effect, β_5, of YEARSTEA in Model 4.3. Note that HLM used ML estimation for all models fitted in this chapter, so the estimates of the three variance components will be biased and, consequently, the t-tests calculated by HLM will also be biased (see Subsection 2.4.1). However, under the null hypothesis that $\beta_5 = 0$, the test statistic reported by HLM approximately follows a t-distribution with 177 degrees of freedom (see Subsection 4.11.3 for a discussion of the calculation of degrees of freedom in HLM3). Because the t-test for Hypothesis 4.3 is not significant ($p = 0.724$), we decide not to include the fixed effect associated with YEARSTEA in the model and conclude that there is not a relationship between the MATHGAIN score of the student and the years of experience of their teacher.

Similarly, we use t-statistics to test Hypotheses 4.4 and 4.5. Because neither of these tests is significant (see Table 4.5), we conclude that there is not a relationship between the MATHGAIN score of the student and the math knowledge or math preparation of their teacher, as measured for this study. Because the results of hypothesis tests 4.3 through 4.5 were not significant, we do not add the fixed effects associated with any classroom-level covariates to the model, and proceed with Model 4.2 as our preferred model at this stage of the analysis.

Hypothesis 4.6: The fixed effect, β_8, associated with the school-level covariate, HOUSEPOV, should be retained in Model 4.4.

We consider a t-test for Hypothesis 4.6. Under the null hypothesis, the t-statistic reported by HLM3 for the fixed effect of HOUSEPOV in Model 4.4 approximately follows a t-distribution with 105 degrees of freedom (see Table 4.5). Because this test is not significant ($p = 0.253$), we do not add the fixed effect associated with HOUSEPOV to the model, and conclude that the MATHGAIN score of a student is not related to the poverty level of the households in the neighborhood of their school. We choose Model 4.2 as our final model for the Classroom data analysis.

4.6 Comparing Results Across the Software Procedures

In Table 4.6 to Table 4.9, we present comparisons of selected results generated by the five software procedures after fitting Models 4.1, 4.2, 4.3, and 4.4, respectively, to the Classroom data.

4.6.1 Comparing Model 4.1 Results

The initial model fitted to the Classroom data, Model 4.1, is variously described as an unconditional, variance components, or "means-only" model. It has a single fixed-effect parameter, the intercept, which represents the mean value of MATHGAIN for all students. Despite the fact that HLM3 uses ML estimation, and the other five software procedures use REML estimation for this model, all six procedures produce the same estimates for the intercept and its standard error.

The REML estimates of the variance components and their standard errors are very similar across the procedures in SAS, SPSS, R, and Stata, whereas the ML estimates from HLM are somewhat different. Looking at the REML estimates, the estimated variance of the random school effects ($\sigma^2_{int:school}$) is 77.5, the estimated variance of the nested random classroom effects ($\sigma^2_{int:classroom}$) is 99.2, and the estimated residual error variance (σ^2) is approximately 1028.2; the largest estimated variance component is the residual error variance.

Table 4.6 also shows that the –2 REML log-likelihood values calculated for Model 4.1 agree across the procedures in SAS, SPSS, R, and Stata. The AIC and BIC information criteria based on the –2 REML log-likelihood values disagree across the procedures that compute them, owing to the different formulas that are used to calculate them (as discussed in Subsection 3.6.1). The HLM3 procedure does not calculate these information criteria.

TABLE 4.6: Comparison of Results for Model 4.1

	SAS: `proc glimmix`	**SPSS:** `MIXED`	**R:** `lme()` **function**	**R:**`lmer()` **function**	**Stata:** `mixed`	**HLM3**
Estimation method	**REML**	**REML**	**REML**	**REML**	**REML**	**ML**
Fixed-Effect Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*
β_0(intercept)	57.43(1.44)	57.43(1.44)	57.43(1.44)	57.43(1.44)	57.43(1.44)	57.43(1.44)[b]
Covariance Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*[c]	*Estimate (n.c.)*	*Estimate (SE)*	*Estimate (SE)*
$\sigma^2_{int:school}$	77.49(32.63)	77.49(32.62)	77.49[d]	77.49	77.50(32.62)	75.04(31.70)
$\sigma^2_{int:classroom}$	99.23(41.81)	99.23(41.81)	99.23	99.23	99.23(41.81)	100.44(38.45)
σ^2(residual variance)	1028.23(49.04)	1028.23(49.04)	1028.23	1028.23	1028.23(49.04)	1027.73(48.06)
Model Information Criteria						
–2 RE/ML log-likelihood	11768.8	11768.8	11768.8	11768	11768.8	11771.3
AIC	11774.8	11774.8	11776.8	11777	11776.8	n.c.
BIC	11782.8	11790.0	11797.1	11797	11797.1	n.c.

Note: (n.c.) = not computed.

Note: 1190 Students at Level 1; 312 Classrooms at Level 2; 107 Schools at Level 3.

[b] Model-based standard errors are presented for the fixed-effect parameter estimates in HLM; robust (sandwich-type) standard errors are also produced in HLM by default.

[c] Standard errors for the estimated covariance parameters are not reported in the output generated by the `summary()` function in R; 95% confidence intervals for the parameter estimates can be generated by applying the `intervals()` function in the `nlme` package to the object containing the results of an `lme()` fit (e.g., `intervals(model4.1.fit)`).

[d] These are squared values of the estimated standard deviations reported by the `nlme` version of the `lme()` function in R.

4.6.2 Comparing Model 4.2 Results

Model 4.2 includes four additional parameters, representing the fixed effects of the four student-level covariates. As Table 4.7 shows, the estimates of these fixed-effect parameters and their standard errors are very similar across the six procedures. The estimates produced using ML estimation in HLM3 are only slightly different from the estimates produced by the other four procedures.

The estimated variance components in Table 4.7 are all smaller than the estimates in Table 4.6, across the six procedures, owing to the inclusion of the fixed effects of the Level 1 (student-level) covariates in Model 4.2. The estimate of the variance between schools was the least affected, while the residual error variance was the most affected (as expected).

Table 4.7 also shows that the –2 REML log-likelihood values agree across the procedures that use REML estimation, as was noted for Model 4.1.

4.6.3 Comparing Model 4.3 Results

Table 4.8 shows that the estimates of the fixed-effect parameters in Model 4.3 (and their standard errors) are once again nearly identical across the five procedures that use REML estimation of the variance components (SAS, SPSS, R, and Stata). The parameter estimates are slightly different when using HLM3, due to the use of ML estimation (rather than REML) by this procedure.

The five procedures that use REML estimation agree quite well (with small differences likely due to rounding error) on the values of the estimated variance components. The variance component estimates from HLM3 are somewhat smaller.

The –2 REML log-likelihood values agree across the procedures in SAS, SPSS, R, and Stata. The AIC and BIC model fit criteria calculated using the –2 REML log-likelihood values for each program differ, due to the different calculation formulas used for the information criteria across the software procedures.

We have included the *t*-tests reported by five of the six procedures for the fixed-effect parameters associated with the classroom-level covariates in Table 4.8, primarily to illustrate the differences in the degrees of freedom computed by the different procedures for the approximate *t*-statistics. Despite the different methods used to calculate the approximate degrees of freedom for the *t*-tests (see Subsection 3.11.6 or 4.11.3), the results are nearly identical across the procedures. The output of the `mixed` command in Stata does not explicitly report the degrees of freedom, but simply indicates that the Kenward–Roger method was used for this computation. We remind readers that *p*-values for the *t*-statistics are not computed by default when using the `lme4` package version of the `lmer()` function in R to fit the model, because of the different approaches that exist for defining the reference distribution. Approximate *t*-statistics can be computed when using `lmer()` in combination with the `lmerTest` package.

4.6.4 Comparing Model 4.4 Results

The comparison of the results produced by the software procedures in Table 4.9 is similar to the comparisons in the other three tables. Test statistics calculated for the fixed effect of HOUSEPOV again show that the procedures that compute the test statistics agree in terms of the results of the tests, despite the different degrees of freedom calculated for the approximate *t*-statistics.

TABLE 4.7: Comparison of Results for Model 4.2

	SAS: `proc glimmix`	**SPSS:** `MIXED`	**R:** `lme()` **function**	**R:** `lmer()` **function**	**Stata:** `mixed`	**HLM3**
Estimation method	**REML**	**REML**	**REML**	**REML**	**REML**	**ML**
Fixed-Effect Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*
β_0(Intercept)	282.79(10.85)	282.79(10.85)	282.79(10.85)	282.79(10.85)	282.79(10.85)	282.73(10.83)
β_1(MATHKIND)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)
β_2(SEX)	−1.25(1.66)	−1.25(1.66)	−1.25(1.66)	−1.25(1.66)	−1.25(1.66)	−1.25(1.65)
β_3(MINORITY)	−8.26(2.34)	−8.26(2.34)	−8.26(2.34)	−8.26(2.34)	−8.25(2.34)	−8.25(2.33)
β_4(SES)	5.35(1.24)	5.35(1.24)	5.35(1.24)	5.35(1.24)	5.35(1.24)	5.35(1.24)
Covariance Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (n.c.)*	*Estimate (n.c.)*	*Estimate (SE)*	*Estimate (SE)*
$\sigma^2_{int:school}$	75.21(25.92)	75.20(25.92)	75.20	75.20	75.20(25.92)	72.88(26.10)
$\sigma^2_{int:classroom}$	83.28(29.38)	83.28(29.38)	83.28	83.29	83.28(29.38)	82.98(28.82)
σ^2(residual variance)	734.57(34.70)	734.57(34.70)	734.57	734.57	734.57(34.70)	732.22(34.30)
Model Information Criteria						
–2 RE/ML log-likelihood	11385.8	11385.8	11385.8	11386	11385.8	11390.9
AIC	11391.8	11391.8	11401.8	11402	11401.8	n.c.
BIC	11399.8	11407.0	11442.4	11442	11442.5	n.c.

Note: (n.c.) = not computed.
Note: 1190 Students at Level 1; 312 Classrooms at Level 2; 107 Schools at Level 3.

TABLE 4.8: Comparison of Results for Model 4.3

	SAS: `proc glimmix`	**SPSS:** `MIXED`	**R:** `lme()` **function**	**R:** `lmer()` **function**	**Stata:** `mixed`	**HLM3**
Estimation method	**REML**	**REML**	**REML**	**REML**	**REML**	**ML**
Fixed-Effect Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*
β_0(Intercept)	282.02(11.70)	282.02(11.70)	282.02(11.70)	282.02(11.70)	282.02(11.70)	281.90(11.65)
β_1(MATHKIND)	−0.48(0.02)	−0.48(0.02)	−0.48(0.02)	−0.48(0.02)	−0.48(0.02)	−0.47(0.02)
β_2(SEX)	−1.34(1.72)	−1.34(1.72)	−1.34(1.72)	−1.34(1.72)	−1.34(1.72)	−1.34(1.71)
β_3(MINORITY)	−7.87(2.42)	−7.87(2.42)	−7.87(2.42)	−7.87(2.42)	−7.83(2.42)	−7.83(2.40)
β_4(SES)	5.42(1.28)	5.42(1.28)	5.42(1.28)	5.42(1.28)	5.43(1.28)	5.43(1.27)
β_5(YEARSTEA)	0.04(0.12)	0.04(0.12)	0.04(0.12)	0.04(0.12)	0.04(0.12)	0.04(0.12)
β_6(MATHPREP)	1.09(1.15)	1.09(1.15)	1.09(1.15)	1.09(1.15)	1.09(1.15)	1.10(1.14)
β_7(MATHKNOW)	1.91(1.15)	1.91(1.15)	1.91(1.15)	1.91(1.15)	1.91(1.15)	1.89(1.14)
Covariance Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (n.c.)*	*Estimate (n.c.)*	*Estimate (SE)*	*Estimate (SE)*
$\sigma^2_{int:school}$	75.19(27.35)	75.19(27.35)	75.19	75.19	75.19(27.35)	72.16(27.44)
$\sigma^2_{int:classroom}$	86.68(31.43)	86.68(31.43)	86.68	86.68	86.68(31.43)	82.69(30.32)
σ^2(residual variance)	713.83(35.47)	713.83(35.47)	713.83	713.83	713.83(35.47)	711.50(35.00)
Model Information Criteria						
–2 RE/ML log-likelihood	10313.0	10313.0	10313.0	10312.0	10313.0	10320.1
AIC	10319.0	10319.0	10335.0	10335.0	10335.0	n.c.
BIC	10327.0	10333.9	10389.8	10390.0	10389.8	n.c.
Tests for Fixed Effects	*t-tests*	*t-tests*	*t-tests*	N/A	*t-tests*	*t-tests*
β_5(YEARSTEA)	$t(792.0) = 0.34$, $p = 0.73$	$t(227.7) = 0.34$, $p = 0.74$	$t(177.0) = 0.34$, $p = 0.73$		$t = 0.34$, $p = 0.74$	$t(177.0) = 0.35$, $p = 0.72$

(Cont.)

TABLE 4.8: Comparison of Results for Model 4.3 *(Cont.)*

	SAS: `proc glimmix`	**SPSS:** `MIXED`	**R:** `lme()` **function**	**R:** `lmer()` **function**	**Stata:** `mixed`	**HLM3**
Estimation method	**REML**	**REML**	**REML**	**REML**	**REML**	**ML**
β_6(MATHPREP)	$t(792.0) = 0.95$, $p = 0.34$	$t(206.2) = 0.95$, $p = 0.34$	$t(177.0) = 0.95$, $p = 0.34$		$t = 0.95$, $p = 0.34$	$t(177.0) = 0.97$, $p = 0.34$
β_7(MATHKNOW)	$t(792.0) = 1.67$, $p = 0.10$	$t(232.3) = 1.67$, $p = 0.10$	$t(177.0) = 1.67$, $p = 0.10$		$t = 1.66$, $p = 0.10$	$t(177.0) = 1.67$, $p = 0.10$

Note: (n.c.) = not computed.
Note: 1081 Students at Level 1; 285 Classrooms at Level 2; 105 Schools at Level 3.

TABLE 4.9: Comparison of Results for Model 4.4

	SAS: `proc glimmix`	**SPSS:** `MIXED`	**R:** `lme()` **function**	**R:** `lmer()` **function**	**Stata:** `mixed`	**HLM3**
Estimation method	**REML**	**REML**	**REML**	**REML**	**REML**	**ML**
Fixed-Effect Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*	*Estimate (SE)*
β_0(Intercept)	285.06(11.02)	285.06(11.02)	285.06(11.02)	285.06(11.02)	285.06(11.02)	284.92(10.99)
β_1(MATHKIND)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)	−0.47(0.02)
β_2(SEX)	−1.23(1.66)	−1.23(1.66)	−1.23(1.66)	−1.23(1.66)	−1.23(1.66)	−1.23(1.65)
β_3(MINORITY)	−7.76(2.39)	−7.76(2.38)	−7.76(2.38)	−7.76(2.38)	−7.76(2.38)	−7.74(2.37)
β_4(SES)	5.24(1.25)	5.24(1.25)	5.24(1.25)	5.24(1.25)	5.24(1.25)	5.24(1.24)
β_8(HOUSEPOV)	−11.44(9.94)	−11.44(9.94)	−11.44(9.94)	−11.44(9.94)	−11.44(9.94)	−11.30(9.83)
Covariance Parameter	*Estimate (SE)*	*Estimate (SE)*	*Estimate (n.c.)*	*Estimate (n.c.)*	*Estimate (SE)*	*Estimate (SE)*
$\sigma^2_{int:school}$	77.76(25.99)	77.76(25.99)	77.76	77.76	77.76(25.99)	74.14(26.16)
$\sigma^2_{int:classroom}$	81.55(29.07)	81.56(29.07)	81.56	81.56	81.56(29.07)	80.96(28.61)
σ^2(residual variance)	734.42(34.67)	734.42(34.67)	734.42	734.42	734.42(34.67)	732.08(34.29)
Model Information Criteria						
–2 RE/ML log-likelihood	11378.1	11378.1	11378.1	11378	11378.1	11389.6
AIC	11384.1	11384.1	11396.1	11396	11396.1	n.c.
BIC	11392.1	11399.3	11441.8	11442.0	11441.8	n.c.
Tests for Fixed Effects	*t-tests*	*t-tests*	*t-tests*	N/A	*t-tests*	*t-tests*
β_8(HOUSEPOV)	$t(873.0) = -1.15$, $p = 0.25$	$t(119.5) = -1.15$, $p = 0.25$	$t(105.0) = -1.15$, $p = 0.25$		$t = -1.15$, $p = 0.25$	$t(105.0) = -1.15$, $p = 0.25$

Note: (n.c.) = not computed.
Note: 1190 Students at Level 1; 312 Classrooms at Level 2; 107 Schools at Level 3.

4.7 Interpreting Parameter Estimates in the Final Model

We consider results generated by the HLM3 procedure in this section.

4.7.1 Fixed-Effect Parameter Estimates

Based on the results from Model 4.2, we see that gain in math score in the spring of first grade (MATHGAIN) is significantly related to math achievement score in the spring of kindergarten (MATHKIND), minority status (MINORITY), and student socioeconomic status (SES). The portion of the HLM3 output for Model 4.2 presented below shows that the individual tests for each of these fixed-effect parameters are significant ($p < 0.05$). The estimated fixed effect of SEX (females relative to males) is the only nonsignificant fixed effect in Model 4.2 ($p = 0.45$).

```
The outcome variable is MATHGAIN
Final estimation of fixed effects:
----------------------------------------------------------------------------
                                            Standard             Approx.
Fixed Effect               Coefficient     Error        T-ratio  d.f.    P-value
----------------------------------------------------------------------------
For INTRCPT1, P0
For INTRCPT2, B00
INTRCPT3, G000                282.726785   10.828453    26.110    106     <0.001
For SEX slope, P1
For INTRCPT2, B10
INTRCPT3, G100                 -1.251422    1.654663    -0.756    767      0.450
For MINORITY slope, P2
For INTRCPT2, B20
INTRCPT3, G200                 -8.253782    2.331248    -3.540    767      <0.001
For MATHKIND slope, P3
For INTRCPT2, B30
INTRCPT3, G300                 -0.469668    0.022216   -21.141    767      <0.001
For SES slope, P4
For INTRCPT2, B40
INTRCPT3, G400                  5.348526    1.238400     4.319    767      <0.001
----------------------------------------------------------------------------
```

The Greek letters for the fixed-effect parameters in the HLM version of Model 4.2 (see Table 4.3 and Subsection 4.4.5) are shown in the left-most column of the output, in their Latin form, along with the name of the variable whose fixed effect is included in the table. For example, G100 represents the overall fixed effect of SEX (γ_{100} in HLM notation). This fixed effect is actually the intercept in the Level 3 equation for the school-specific effect of SEX (hence, the INTRCPT3 notation). The column labeled "Coefficient" contains the fixed-effect parameter estimate for each of these covariates. The standard errors of the parameter estimates are also provided, along with the T-ratios (t-test statistics), approximate degrees of freedom (d.f.) for the T-ratios, and the p-value. We describe the HLM calculation of degrees of freedom for these approximate t-tests in Subsection 4.11.3.

The estimated fixed effect of kindergarten math score, MATHKIND, on math achievement score in first grade, MATHGAIN, is negative (-0.47), suggesting that students with higher math scores in the spring of their kindergarten year have a lower predicted gain in math achievement in the spring of first grade, after adjusting for the effects of other covariates (i.e., SEX, MINORITY, and SES). That is, students doing well in math in kindergarten will not improve as much over the next year as students doing poorly in kindergarten.

Minority students are predicted to have a mean MATHGAIN score that is 8.25 units lower than their non-minority counterparts, after adjusting for the effects of other covariates. In addition, students with higher SES are predicted to have higher math achievement gain than students with lower SES, adjusting for the effects of the other covariates in the model.

4.7.2 Covariance Parameter Estimates

The HLM output below presents the estimated variance components for Model 4.2, based on the HLM3 fit of this model.

```
Final estimation of level-1 and level-2 variance components:
--------------------------------------------------------------------
Random Effect       Standard    Variance   df   Chi-square  P-value
                    Deviation   Component
--------------------------------------------------------------------
INTRCPT1,      r0    9.10959     82.98470  205   298.96800   <0.001
Level-1,       e    27.05951    732.21715


Final estimation of level-3 variance components
---------------------------------------------------------------------------
Random Effect           Standard    Variance   df     Chi-square   P-value
                        Deviation   Component
---------------------------------------------------------------------------
INTRCPT1/INTRCPT2, u00  8.53721     72.88397   106    183.59757   <0.001
---------------------------------------------------------------------------
```

The variance components in this three-level model are reported in two blocks of output. The first block of output contains the estimated standard deviation of the nested random effects associated with classrooms (labeled r0, and equal to 9.11), and the corresponding estimated variance component (equal to 82.98). In addition, a chi-square test (discussed in the following text) is reported for the significance of this variance component. The first block of output also contains the estimated standard deviation of the residual errors (labeled e, and equal to 27.06), and the corresponding estimated variance component (equal to 732.22). No test of significance is reported for the residual error variance.

The second block of output above contains the estimated standard deviation of the random effects associated with schools (labeled u00), and the corresponding estimated variance component (equal to 72.88). HLM also reports a chi-square test of significance for the variance component at the school level.

The addition of the fixed effects of the student-level covariates to Model 4.1 (to produce Model 4.2) reduced the estimated residual error variance by roughly 29% (estimated residual error variance = 1027.73 in Model 4.1, vs. 732.22 in Model 4.2). The estimates of the classroom- and school-level variance components were also reduced by the addition of the fixed effects associated with the student-level covariates, although not substantially (the estimated classroom-level variance was reduced by roughly 17.4%, and the estimated school-level variance was reduced by about 2.9%). This suggests that the four student-level covariates are effectively explaining some of the random variation in the response values at the different levels of the data set, especially at the student level (as expected).

The magnitude of the variance components in Model 4.2 (and the significant chi-square tests reported for the variance components by HLM3) suggests that there is still unexplained random variation in the response values at all three levels of this data set.

We see in the output above that HLM3 produces chi-square tests for the variance components in the output (see Raudenbush & Bryk (2002) for details on these tests). These tests suggest that the variances of the random effects at the school level (u00) and the classroom level (r0) in Model 4.2 are both significantly greater than zero, even after the inclusion of the fixed effects of the student-level covariates. These test results indicate that a significant amount of random variation in the response values at all three levels of this data set remains unexplained. At this point, fixed effects associated with additional covariates could be added to the model, to see if they help to explain random variation at the different levels of the data.

4.8 Estimating the Intraclass Correlation Coefficients (ICCs)

In the context of a three-level hierarchical model with random intercepts, the **intraclass correlation coefficient** (ICC) is a measure describing the similarity (or homogeneity) of observed responses within a given cluster. For each level of clustering (e.g., classroom or school), an ICC can be defined as a function of the variance components. For brevity in this section, we represent the variance of the random effects associated with schools as σ_s^2 (instead of $\sigma^2_{int:school}$), and the variance of the random effects associated with classrooms nested within schools as σ_c^2 (instead of $\sigma^2_{int:classroom}$).

The school-level ICC is defined as the proportion of the total random variation in the observed responses (the denominator in (4.5)) due to the variance of the random school effects (the numerator in (4.5)):

$$\text{ICC}_{school} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_c^2 + \sigma^2} \tag{4.5}$$

The value of ICC_{school} is high if the total random variation is dominated by the variance of the random school effects. In other words, the ICC_{school} is high if the MATHGAIN scores of students in the same school are relatively homogeneous, but the MATHGAIN scores across schools tend to vary widely.

Similarly, the classroom-level ICC is defined as the proportion of the total random variation (the denominator in (4.6)) due to random between-school and between-classroom variation (the numerator in (4.6)):

$$\text{ICC}_{classroom} = \frac{\sigma_s^2 + \sigma_c^2}{\sigma_s^2 + \sigma_c^2 + \sigma^2} \tag{4.6}$$

This ICC is high if there is little variation in the responses of students within the same classroom (σ^2 is low) compared to the total random variation.

The ICCs for classrooms and for schools are estimated by substituting the estimated variance components from a random intercept model into the preceding formulas. Because variance components are positive or zero by definition, the resulting ICCs are also positive or zero.

The software procedures discussed in this chapter provide clearly labeled variance component estimates in the computer output when fitting a random intercepts model, allowing for easy calculation of estimates of these ICCs. We can use the estimated variance components from Model 4.1 to compute estimates of the intraclass correlation coefficients (ICCs) defined in (4.5) and (4.6). We estimate the ICC of observations on students within the same school to be 77.5 / (77.5 + 99.2 + 1028.2) = 0.064, and we estimate the ICC of observations on students within the same classroom nested within a school to be (77.5 + 99.2) / (77.5 + 99.2 + 1028.2) = 0.147. Observations on students in the same school are modestly correlated, while observations on students within the same classroom have a somewhat higher correlation.

To further illustrate ICC calculations, we consider the marginal covariance matrix $\boldsymbol{V}_k$ implied by Model 4.1 for a hypothetical school, k, having two classrooms, with the first classroom having two students, and the second having three students. The first two rows and columns of this matrix correspond to observations on the two students from the first classroom, and the last three rows and columns correspond to observations on the three students from the second classroom:

$$V_k = \begin{pmatrix} \begin{pmatrix} \sigma_s^2+\sigma_c^2+\sigma^2 & \sigma_s^2+\sigma_c^2 \\ \sigma_s^2+\sigma_c^2 & \sigma_s^2+\sigma_c^2+\sigma^2 \end{pmatrix} & \begin{matrix} \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \sigma_s^2 \end{matrix} \\ \begin{matrix} \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 \end{matrix} & \begin{pmatrix} \sigma_s^2+\sigma_c^2+\sigma^2 & \sigma_s^2+\sigma_c^2 & \sigma_s^2+\sigma_c^2 \\ \sigma_s^2+\sigma_c^2 & \sigma_s^2+\sigma_c^2+\sigma^2 & \sigma_s^2+\sigma_c^2 \\ \sigma_s^2+\sigma_c^2 & \sigma_s^2+\sigma_c^2 & \sigma_s^2+\sigma_c^2+\sigma^2 \end{pmatrix} \end{pmatrix}$$

The corresponding marginal correlation matrix for these observations can be calculated by dividing all elements in the matrix above by the total variance of a given observation, $[var(y_{ijk}) = \sigma_s^2 + \sigma_c^2 + \sigma^2]$, as shown below. The ICCs defined in (4.5) and (4.6) can easily be identified in this implied correlation matrix:

$$V_k(corr) = \begin{pmatrix} \begin{pmatrix} 1 & \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & 1 \end{pmatrix} & \begin{matrix} \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \end{matrix} \\ \begin{matrix} \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \end{matrix} & \begin{pmatrix} 1 & \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & 1 & \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} \\ \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & \frac{\sigma_s^2+\sigma_c^2}{\sigma_s^2+\sigma_c^2+\sigma^2} & 1 \end{pmatrix} \end{pmatrix}$$

We obtain estimates of the ICCs from the marginal covariance matrix for the MATHGAIN observations implied by Model 4.1 by using the `v` option in the `random` statement in SAS `proc glimmix`. The estimated 5×5 $\boldsymbol{V}_{69}$ matrix for the observations on the 5 students from school 69 is displayed as follows:

Estimated V Matrix for schoolid 69

Row	Col1	Col2	Col3	Col4	Col5
1	1204.95	176.72	77.4938	77.4938	77.4938
2	176.72	1204.95	77.4938	77.4938	77.4938
3	77.4938	77.4938	1204.95	176.72	176.72
4	77.4938	77.4938	176.72	1204.95	176.72
5	77.4938	77.4938	176.72	176.72	1204.95

The 2×2 submatrix in the upper-left corner of this matrix corresponds to the marginal variances and covariances of the observations for the two students in the first classroom in school 69, and the 3×3 submatrix in the lower-right corner represents the corresponding values for the three students from the second classroom in this school.

We note that the estimated covariance of observations collected on students in the same classroom is 176.72. This is the sum of the estimated variance of the nested random classroom effects, 99.23, and the estimated variance of the random school effects, 77.49. Observations collected on students attending the same school but having different teachers are estimated to have a common covariance of 77.49, which is the variance of the random school effects. Finally, all observations have a common estimated variance, 1204.95, which is equal to the sum of the three estimated variance components in the model (99.23 + 77.49 + 1028.23 = 1204.95), and is the value along the diagonal of this matrix.

The marginal covariance matrices for observations on students within any given school would have the same structure, but would be of different dimensions, depending on the number of students within the school. Observations on students in different schools will have zero covariance, because they are assumed to be independent of each other.

The estimated marginal correlations of observations for students within school 69 (or any other school) implied by Model 4.1 can be derived by using the `vcorr` option in the `random` statement in SAS `proc glimmix`. Note in the corresponding SAS output below that observations on different students within the same classroom in this school have an estimated marginal correlation of 0.1467, and observations on students in different classrooms within this school have an estimated correlation of 0.06431. These results match our initial ICC calculations based on the estimated variance components.

Covariates are not considered in the classical definitions of the ICC, either based on the random intercept model or the marginal model; however, covariates can easily be accommodated in the mixed model framework in either model setting. The ICC may be calculated from a model without fixed effects of other covariates (e.g., Model 4.1) or for a model including these fixed effects (e.g., Model 4.2 or Model 4.3). In either case, we can obtain the ICCs from the labeled variance component estimates or from the estimated marginal correlation matrix, as described earlier.

```
              Estimated V Correlation Matrix for schoolid 69

        Row         Col1          Col2          Col3          Col4          Col5

          1       1.0000        0.1467       0.06431       0.06431       0.06431
          2       0.1467        1.0000       0.06431       0.06431       0.06431
          3      0.06431       0.06431        1.0000        0.1467        0.1467
          4      0.06431       0.06431        0.1467        1.0000        0.1467
          5      0.06431       0.06431        0.1467        0.1467        1.0000
```

4.9 Calculating Predicted Values

4.9.1 Conditional and Marginal Predicted Values

In this section, we use the estimated fixed effects in Model 4.2, generated by the HLM3 procedure, to write formulas for calculating predicted values of MATHGAIN. Recall that three different sets of predicted values can be generated: conditional predicted values including the EBLUPs of the random school and classroom effects, and marginal predicted values based only on the estimated fixed effects. For example, considering the estimates of the fixed effects in Model 4.2, we can write a formula for the **conditional predicted values** of MATHGAIN for a student in a given classroom:

$$\begin{aligned} \widehat{\text{MATHGAIN}}_{ijk} = {} & 282.73 - 0.47 \times \text{MATHKIND}_{ijk} - 1.25 \times \text{SEX}_{ijk} \\ & - 8.25 \times \text{MINORITY}_{ijk} + 5.35 \times \text{SES}_{ijk} + \widehat{u}_k + \widehat{u}_{j|k} \end{aligned} \quad (4.7)$$

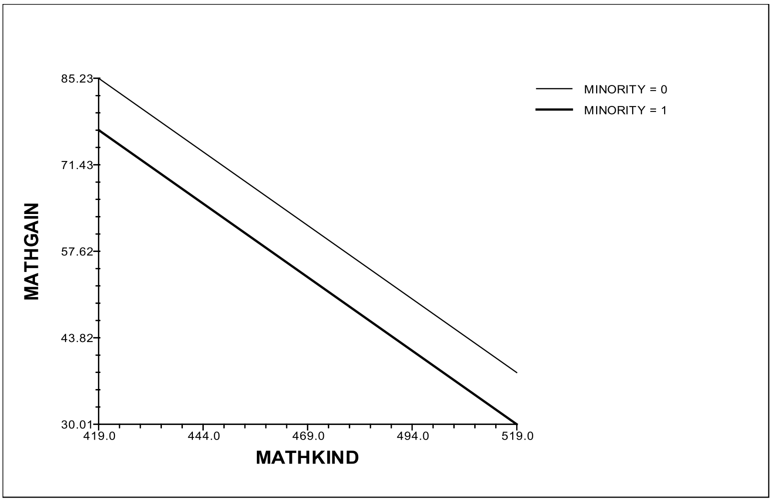


FIGURE 4.6: Marginal predicted values of MATHGAIN as a function of MATHKIND and MINORITY, based on the fit of Model 4.2 in HLM3.

This formula includes the EBLUPs of the random effect for this student's school, u_k, and the random classroom effect for this student, $u_{j|k}$. Residuals calculated based on these conditional predicted values should be used to assess assumptions of normality and constant variance for the residual errors (see Subsection 4.10.2). A formula similar to (4.7) that omits the EBLUPs of the random classroom effects ($u_{j|k}$) could be written for calculating a second set of conditional predicted values specific to schools:

$$\begin{aligned}\widehat{\text{MATHGAIN}}_{ijk} = \; & 282.73 - 0.47 \times \text{MATHKIND}_{ijk} - 1.25 \times \text{SEX}_{ijk} \\ & - 8.25 \times \text{MINORITY}_{ijk} + 5.35 \times \text{SES}_{ijk} + \hat{u}_k \end{aligned} \tag{4.8}$$

A third set of **marginal predicted values**, based on the marginal distribution of MATHGAIN responses implied by Model 4.2, can be calculated based only on the estimated fixed effects:

$$\begin{aligned}\widehat{\text{MATHGAIN}}_{ijk} = \; & 282.73 - 0.47 \times \text{MATHKIND}_{ijk} - 1.25 \times \text{SEX}_{ijk} \\ & - 8.25 \times \text{MINORITY}_{ijk} + 5.35 \times \text{SES}_{ijk} \end{aligned} \tag{4.9}$$

These predicted values represent average values of the MATHGAIN response (across schools and classrooms) for all students having given values on the covariates.

We discuss how to obtain both conditional and marginal predicted values based on the observed data using SAS, SPSS, R, and Stata in Chapters 3, 5, 6, and 7 respectively. Readers can refer to Subsection 4.4.5 for details on obtaining conditional predicted values in HLM.

4.9.2 Plotting Predicted Values Using HLM

The HLM software has several convenient graphical features that can be used to visualize the fit of an LMM. For example, after fitting Model 4.2 in HLM, we can plot the marginal predicted values of MATHGAIN as a function of MATHKIND for each level of MINORITY, based on the estimated fixed effects in Model 4.2. In the model-building window of HLM, click **File**, **Graph Equations**, and then **Model graphs**. In the Equation Graphing window, we set the parameters of the plot. First, set the Level-1 **X focus** to be MATHKIND, which will set the horizontal axis of the graph. Next, set the first Level-1 **Z focus** to be MINORITY. Finally, click on **OK** in the main Equation Graphing window to generate the graph in Figure 4.6.

We can see the significant negative effect of MATHKIND on MATHGAIN in Figure 4.6, along with the gap in predicted MATHGAIN for students with different minority status. The fitted lines are parallel because we did not include an interaction between MATHKIND and MINORITY in Model 4.2. We also note that the values of SES and SEX are held fixed at their mean when calculating the marginal predicted values in Figure 4.6.

We can also generate a graph displaying the fitted conditional MATHGAIN values as a function of MATHKIND for a sample of individual schools, based on both the estimated fixed effects and the predicted random school effects (i.e., EBLUPs) resulting from the fit of Model 4.2. In the HLM model-building window, click **File**, **Graph Equations**, and then **Level-1 equation graphing**. First, choose MATHKIND as the Level-1 **X focus**. For **Number of groups** (Level 2 units or Level 3 units), select **First ten groups**. Finally, set **Grouping** to be **Group at level 3**, and click OK. This plots the conditional predicted values of MATHGAIN as a function of MATHKIND for the first ten schools in the data set, in separate panels (not displayed here).

4.10 Diagnostics for the Final Model

In this section we consider diagnostics for our final model, Model 4.2, fitted using ML estimation in HLM.

4.10.1 Plots of the EBLUPs

Plots of the EBLUPs for the random classroom and school effects from Model 4.2 were generated by first saving the EBLUPs from the HLM3 procedure in SPSS data files (see Subsection 4.4.5), and then generating the plots in SPSS. Figure 4.7 below presents a normal Q–Q plot of the EBLUPs for the random classroom effects. This plot was created using the EBINTRCP variable saved in the Level 2 residual file by the HLM3 procedure:

```
* For Figure 4.7.
PPLOT
/VARIABLES=ebintrcp
/NOLOG
/NOSTANDARDIZE
/TYPE=Q-Q
/FRACTION=BLOM
/TIES=MEAN
/DIST=NORMAL.
```

We do not see evidence of any outliers in the random classroom effects, and the distribution of the EBLUPs for the random classroom effects is approximately normal. In the next plot (Figure 4.8), we investigate the distribution of the EBLUPs for the random school effects, using the EB00 variable saved in the Level 3 residual file by the HLM3 procedure:

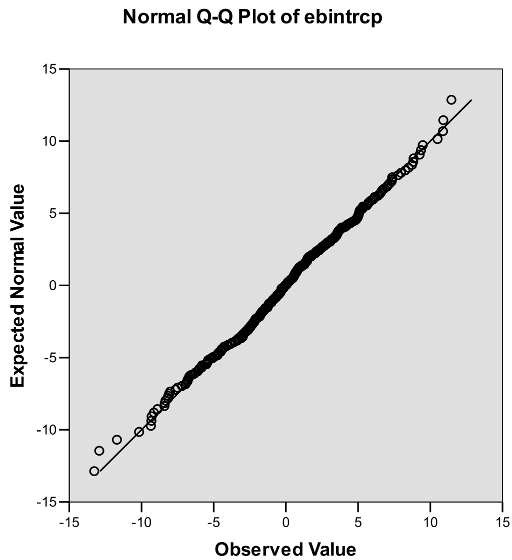


FIGURE 4.7: EBLUPs of the random classroom effects from Model 4.2, plotted using SPSS.

```
* For Figure 4.8.
PPLOT
/VARIABLES=eb00
/NOLOG
/NOSTANDARDIZE
/TYPE=Q-Q
/FRACTION=BLOM
/TIES=MEAN
/DIST=NORMAL.
```

We do not see any evidence of a deviation from a normal distribution for the EBLUPs of the random school effects, and more importantly, we do not see any extreme outliers. Plots such as these can be used to identify EBLUPs that are potential outliers, and further investigate the clusters (e.g., schools or classrooms) associated with the extreme EBLUPs. Note that evidence of a normal distribution in these plots does not always imply that the distribution of the random effects is in fact normal (see Subsection 2.8.3).

4.10.2 Residual Diagnostics

In this section, we investigate the assumptions of normality and constant variance for the residual errors, based on the fit of Model 4.2. These plots were created in SPSS, using the Level 1 residual file generated by the HLM3 procedure. We first investigate a normal Q–Q plot for the residuals:

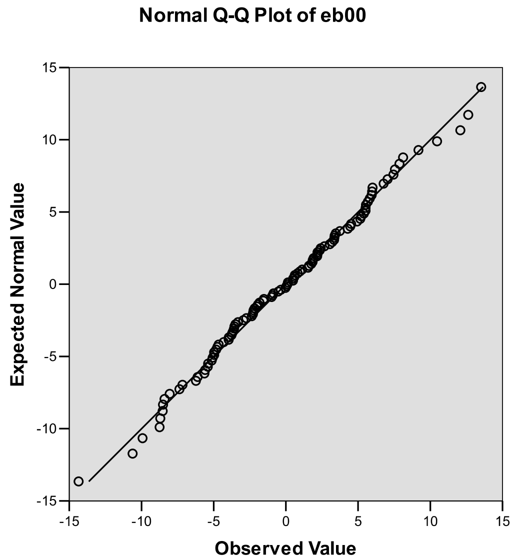


FIGURE 4.8: EBLUPs of the random school effects from Model 4.2, plotted using SPSS.

```
* For Figure 4.9.
PPLOT
/VARIABLES=l1resid
/NOLOG
/NOSTANDARDIZE
/TYPE=Q-Q
/FRACTION=BLOM
/TIES=MEAN
/DIST=NORMAL.
```

If the residual errors based on Model 4.2 followed an approximately normal distribution, all of the points in Figure 4.9 would lie on or near the straight line included in the figure. We see a deviation from this line at the tails of the distribution, which suggests a long-tailed distribution of the residual errors (since only the points at the ends of the distribution deviate from normality). There appear to be small sets of extreme negative and positive residuals that may warrant further investigation. Transformations of the response variable (MATHGAIN) could also be performed, but the scale of the MATHGAIN variable (where some values are negative) needs to be considered; for example, a log transformation of the response would not be possible without first adding a constant to each response to produce a positive value.

Next, we investigate a scatter plot of the conditional residuals vs. the fitted MATHGAIN values, which include the EBLUPs of the random school effects and the nested random classroom effects. These fitted values are saved by the HLM3 procedure in a variable named FITVAL in the Level 1 residual file. We investigate this plot to get a visual sense of whether or not the residual errors have constant variance:

```
* For Figure 4.10.
GRAPH
/SCATTERPLOT(BIVAR) = fitval WITH llresid
/MISSING = LISTWISE .
```

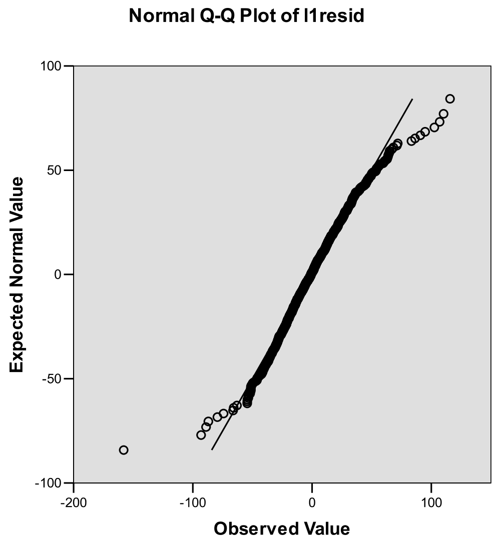


FIGURE 4.9: Normal quantile–quantile (Q–Q) plot of the residuals from Model 4.2, plotted using SPSS.

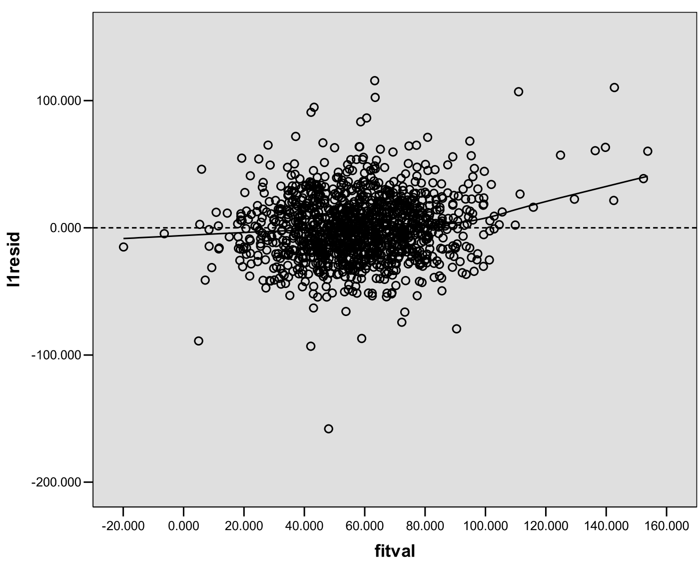


FIGURE 4.10: Residual vs. fitted plot from SPSS.

We have edited the scatter plot in SPSS (Figure 4.10) to include the fit of a smooth Loess curve, indicating the relationship of the fitted values with the residuals, in addition to a dashed reference line set at zero.

We see evidence of non-constant variance in the residual errors in Figure 4.10. We would expect there to be no relationship between the fitted values and the residuals (a line fitted to the points in this plot should look like the reference line, representing the zero mean of the residuals), but the Loess smoother shows that the residuals tend to get larger for larger predicted values of MATHGAIN.

This problem suggests that the model may be misspecified; there may be omitted covariates that would explain the large positive values and the low negative values of MATHGAIN that are not being well fitted. Scatter plots of the residuals against other covariates would be useful to investigate at this point, as there might be nonlinear relationships of the covariates with the MATHGAIN response that are not being captured by the strictly linear fixed effects in Model 4.2.

4.11 Software Notes

4.11.1 REML vs. ML Estimation

The procedures in SAS, SPSS, R, and Stata use restricted maximum likelihood (REML) estimation as the default estimation method for fitting LMMs with nested random effects to three-level data sets. These procedures estimate the variance and covariance parameters using REML (where ML estimation is also an option), and then use the estimated marginal $\boldsymbol{V}$ matrix to estimate the fixed-effect parameters in the models using generalized least squares (GLS). The procedure available in HLM (HLM3) utilizes ML estimation when fitting three-level models with nested random effects.

4.11.2 Setting up Three-Level Models in HLM

In the following text, we note some important differences in setting up three-level models using the HLM software as opposed to the other four packages:

- Three data sets, corresponding to the three levels of the data, are required to fit an LMM to a three-level data set. The other procedures require that all variables for each level of the data be included in a single data set, and that the data be arranged in the "long" format displayed in Table 4.2.

- Models in HLM are specified in multiple parts. For a three-level data set, Level 1, Level 2, and Level 3 models are identified. The Level 2 models are for the effects of covariates measured on the Level 1 units and specified in the Level 1 model; and the Level 3 models are for the effects of covariates measured on the Level 2 units and specified in the Level 2 models.

- In models for three-level data sets, the effects of any of the Level 1 predictors (including the intercept) are allowed to vary randomly across Level 2 and Level 3 units. Similarly, the effects of Level 2 predictors (including the intercept) are allowed to vary randomly across Level 3 units. In the models fitted in this chapter, we have allowed only the intercepts to vary randomly at different levels of the data.

4.11.3 Calculation of Degrees of Freedom for t-Tests in HLM

The degrees of freedom for the approximate t-statistics calculated by the HLM3 procedure, and reported for Hypothesis 4.3 through Hypothesis 4.6, are described in this subsection.

Level 1 Fixed Effects: df = number of Level 1 observations (i.e., number of students) – number of random effects at Level 2 – number of random effects at Level 3 – number of fixed effect parameters associated with the covariates at Level 1.

For example, for the t-tests for the fixed effects associated with the Level 1 (student-level) covariates in Model 4.2, we have df = 1190 – 312 – 107 – 4 = 767.

Level 2 Fixed Effects: df = number of random effects at Level 2 – number of random effects at Level 3 – number of fixed effects at Level 2.

For example, in Model 4.3, we have df = 285 – 105 – 3 = 177 for the t-tests for the fixed effects associated with the Level 2 (classroom-level) covariates, as shown in Table 4.5. Note that there are three fixed effects at Level 2 (the classroom level) in Model 4.3.

Level 3 Fixed Effects: df = number of random effects at Level 3 – number of fixed effects at Level 3.

Therefore, in Model 4.4, we have df = 107 – 2 = 105 for the t-test for the fixed effect associated with the Level 3 (school-level) covariate, as shown in Table 4.5. The fixed intercept is considered to be a Level 3 fixed effect, and there is one additional fixed effect at Level 3.

We note that the HLM output in Version 8.1.4.10 of HLM includes additional tables of fixed effect estimates with **robust standard errors**. These standard errors will allow for inferences for the fixed effects that are robust to possible misspecification of the random effects structure (which was a distinct possibility in this chapter, given that the models only included random intercepts). This output should be consulted to determine whether inferences related to the fixed effects are sensitive to this possible misspecification. If this is the case, we would recommend reporting the results based on the robust standard errors.

4.11.4 Analyzing Cases with Complete Data

We mention in the analysis of the Classroom data that likelihood ratio tests are not possible for Hypotheses 4.3 through 4.5, due to the presence of missing data for some of the Level 2 covariates.

An alternative way to approach the analyses in this chapter would be to begin with a data set having cases with complete data for all covariates. This would make either likelihood ratio tests or alternative tests (e.g., t-tests) appropriate for any of the hypotheses that we test. In the Classroom data set, MATHKNOW is the only classroom-level covariate with missing data. Taking that into consideration, we include the following syntax for each of the software packages that could be used to derive a data set where only cases with complete data on all covariates are included.

In SAS, the following data step could be used to create a new SAS data set, `classroom_nomiss`, which contains only observations with complete data:

```
data classroom_nomiss;
   set classroom;
   if mathknow ne .;
run;
```

In SPSS, the following syntax can be used to select cases that do not have missing data on MATHKNOW (the resulting data set should be saved under a different name):

```
FILTER OFF.
USE ALL.
SELECT IF (not MISSING (mathknow)).
EXECUTE.
```

In R, we could create a new data frame object excluding those cases with missing data on MATHKNOW:

```
> class.nomiss <- subset(class, !is.na(mathknow))
```

In Stata, the following command could be used to delete cases with missing data on MATHKNOW:

```
. keep if mathknow != .
```

Finally, in HLM, this can be accomplished by selecting **Delete data when** ... **making MDM** (rather than when running the analysis) when setting up the MDM file.

4.11.5 Miscellaneous Differences

Less critical differences between the five software procedures in terms of fitting three-level models are highlighted in the following text:

- Procedures in the HLM software package automatically generate both model-based standard errors and robust (or sandwich-type) standard errors for estimated fixed effects. The two different sets of standard errors are clearly distinguished in the HLM output. As we noted above, the robust standard errors are useful to report if one is unsure about whether the marginal covariance matrix for the data has been correctly specified; if the robust standard errors differ substantially from the model-based standard errors, we would recommend reporting the robust standard errors (for more details see Raudenbush & Bryk (2002)). Robust standard errors can be obtained in SAS by using the `empirical` option when invoking `proc glimmix`.
- Fitting three-level random intercept models using the `MIXED` command in SPSS tends to be computationally intensive, and can take longer than in the other software procedures. This is generally an issue when computing EBLUPs of the random effects at multiple levels, at least in Version 28 of SPSS.

4.12 Recommendations

Three-level models for cross-sectional data introduce the possibility of extremely complex random-effects structures. In the example analyses presented in this chapter, we only considered models with random intercepts; we could have allowed the relationships of selected covariates at Level 1 (students) of the Classroom data to randomly vary across classrooms and schools, and the relationships of selected covariates at Level 2 (classrooms) to vary across schools. The decision to include many additional random effects in a three-level model will result in a much more complex implied covariance structure for the dependent variable, including several covariance parameters (especially if an unstructured $\boldsymbol{D}$ matrix is used, which is the default random-effects covariance structure in HLM and the two R functions). This may result in estimation difficulties, or the software appearing to "hang" or "freeze" when attempting to estimate a model. For this reason, we only recommend including a large number of random effects (above and beyond random intercepts) at higher levels if there is *explicit research interest* in empirically describing (and possibly attempting to explain) the variance in the relationships of selected covariates with the dependent variable across higher level units. Including random intercepts at Level 2 and Level 3 of a given three-level data set will typically result in a reasonable implied covariance structure for a given continuous dependent variable in a cross-sectional three-level data set.

Because three-level models do introduce the possibility of allowing many relationships to vary across higher levels of the data hierarchy (e.g., the relationship of student-level SES with mathematical performance varying across schools), the ability to graphically explore variance in both the means of a dependent variable and the relationships of key independent variables with the dependent variable across higher-level units becomes very important when analyzing three-level data. For this reason, having good graphical tools "built in" to a given software package becomes very important. We find that the HLM software provides users with a useful set of "point-and-click" graphing procedures for exploring random coefficients without too much additional work (see Subsection 4.2.2.2). Creating similar graphs and figures in the other software tends to take more work and some additional programming, but is still possible.