



## Two-level model for the HS&B data

1. Description of the data .....	1
2. Construct the MDM file for HLM2 with SPSS file input .....	2
3. Description of the model .....	7
4. Creating the command file .....	8
5. Interpreting the output .....	14

### 1. Description of the data

**High School and Beyond (HS&B)** is a national longitudinal study originally funded by the United States Department of Education's National Center for Education Statistics (NCES) as a part of their longitudinal studies program. Its purpose was to document the “educational, vocational, and personal development of young people following them over time as they begin to take on adult roles and responsibilities”. Here a subset of the data representing 160 schools and a total of 7,185 students is used to illustrate the fitting and interpretation of a two-level hierarchical linear model.

This example is the first in a set of seven examples based on these data described on the HLM Support page.

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

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

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

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

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

**Figure 2.1** First ten cases in HSB1.SAV

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

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

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

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

**Figure 2.2** First ten cases in HSB2.SAV

## 2. Construct the MDM file for HLM2 with SPSS file input

The first task in using HLM2 is to construct the Multivariate Data Matrix (MDM) from raw data or from a statistical package. We generally work with two raw data files: a level-1 file and a

level-2 file. Both files must be sorted by the level-2 ID (It is possible, however, to build the MDM file from a single data file, though this option is not suggested when the level-1 file is very large. The level-1 file must again be sorted by level-2 ID.

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

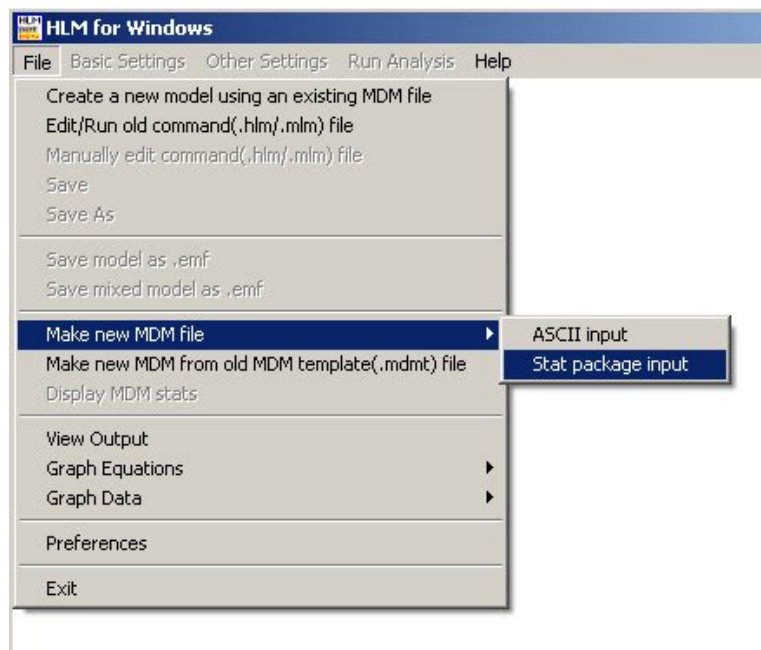
The procedure to create an MDM file consists of three major steps. The user needs to

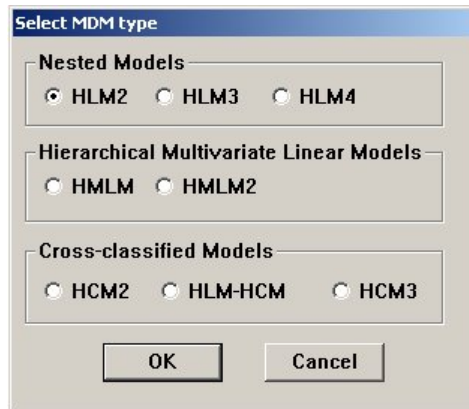
- Inform HLM of the input and MDM file type.
- Supply HLM with the appropriate information for the data, the command and the MDM files.
- Check if the data have been properly read into HLM.

We illustrate the use of SPSS file input.

### To inform HLM of the input and MDM file type

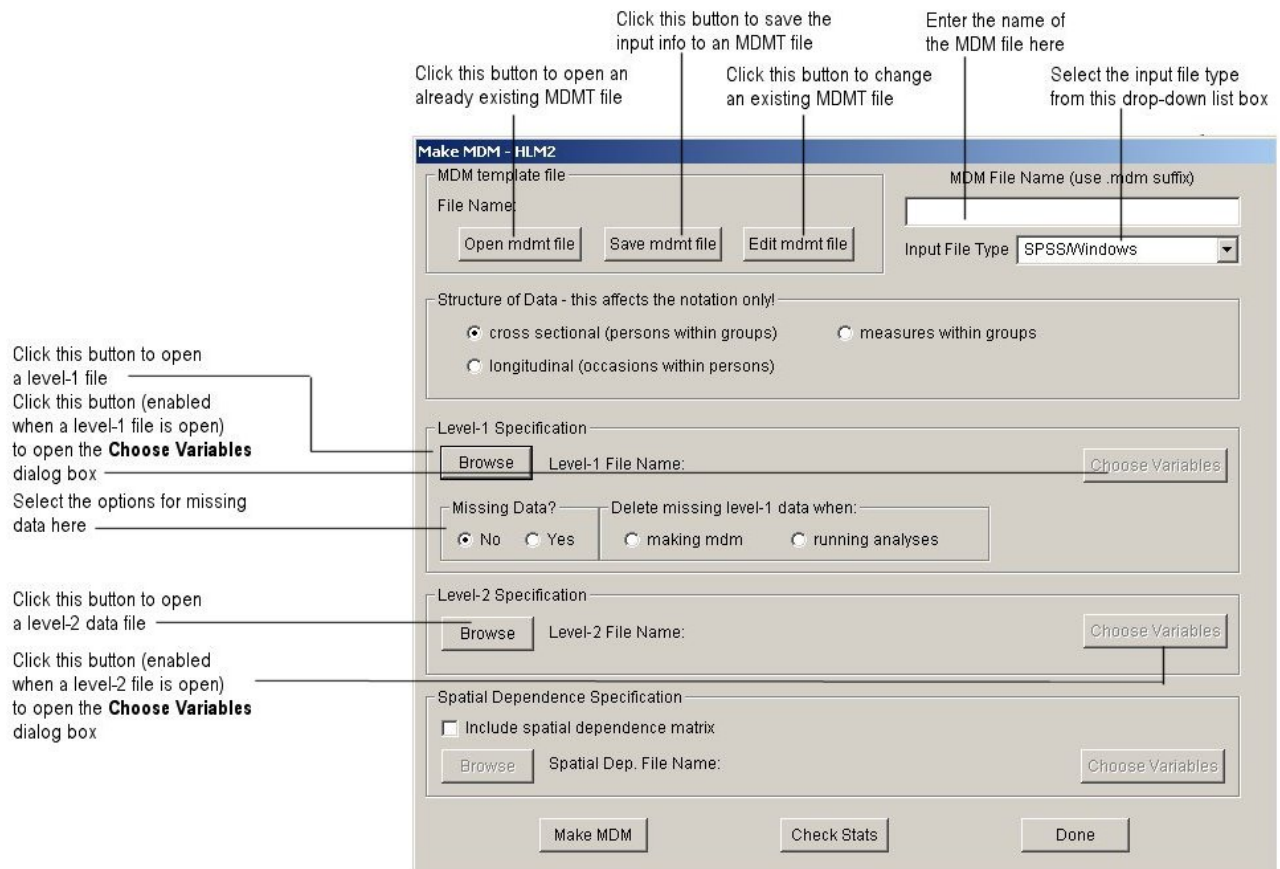
1. At the **WHLM** window, open the **File** menu.
2. Choose **Make new MDM file...Stat package input**. A **Select MDM type** dialog box opens.
3. Select **HLM2** and click **OK**. A **Make MDM - HLM2** dialog box will open.



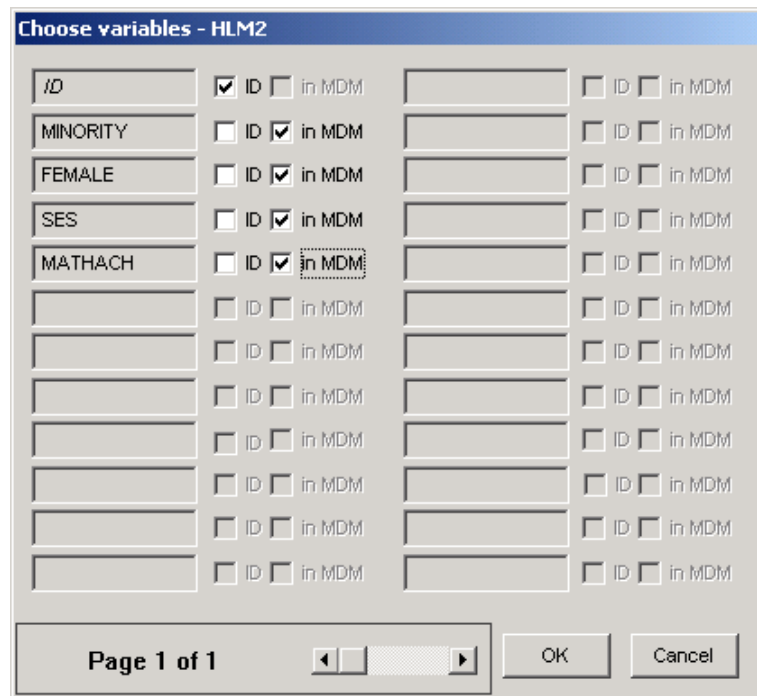


To supply HLM with appropriate information for the data, the command, and the MDM files:

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



- Click **Choose Variables** to open the **Choose Variables - HLM2** dialog box and choose the ID and variables by clicking the appropriate check boxes. To deselect, click the box again.



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

**Choose variables - HLM2**

ID	<input checked="" type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
SIZE	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
SECTOR	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
PRACAD	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
DISCLIM	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
HIMINTY	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
MEANSES	<input type="checkbox"/> ID	<input checked="" type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
	<input type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
	<input type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
	<input type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
	<input type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM
	<input type="checkbox"/> ID	<input type="checkbox"/> in MDM		<input type="checkbox"/> ID	<input type="checkbox"/> in MDM

Page 1 of 1

OK Cancel

13. Click **Save mdmt file** in the **MDM template file** section to open a **Save MDM template file** dialog box. Enter a name for the MDMT file (for example, HSBSPSS.MDMT). Click **Save** to save the file. The command file saves all the input information entered by the user. It can be re-opened by clicking the **Open mdmt file** button. To make changes to an existing MDMT file, click the **Edit mdmt file** button.
14. Note that HLM will also save the input information into another file called CREATMDM.MDMT when the MDM is created.
15. Click the **Make MDM** button. A screen displaying the prompts and responses for MDM creation will appear.
16. When the screen disappears, the level-1 and level-2 descriptive statistics will automatically be displayed. Pay particular attention to the N column. It is not an uncommon mistake to forget to sort by the ID variable, which can lead to a lot (or most) of the data not being processed. Close the Notepad window when done. Use the **Save As** option to give it a new name if later use of this file is anticipated. The file can also be opened by clicking on the **Display Stats** button.

The screenshot shows a Notepad window with the title 'HLM2MDM.STS'. It contains two tables of descriptive statistics. The first table, 'LEVEL-1 DESCRIPTIVE STATISTICS', shows data for variables MINORITY, FEMALE, SES, and MATHACH across 7185 cases. The second table, 'LEVEL-2 DESCRIPTIVE STATISTICS', shows data for variables SIZE, SECTOR, PRACAD, DISCLIM, HIMINTY, and MEANSES across 160 cases.

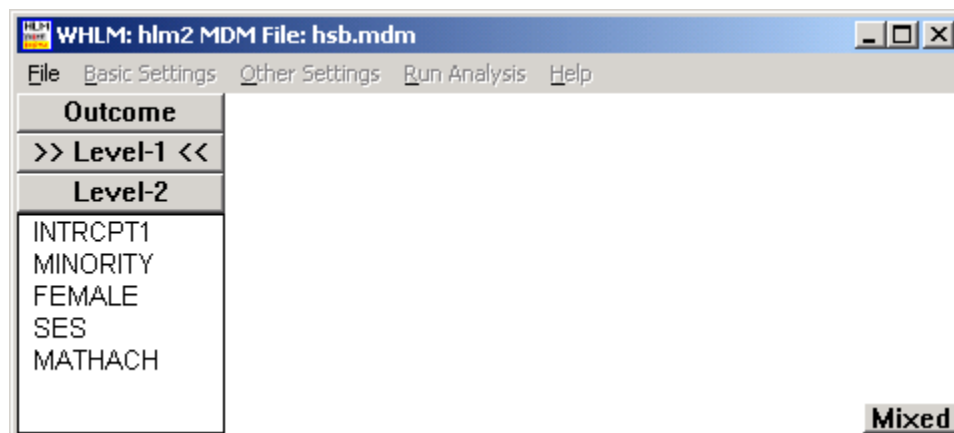
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
MINORITY	7185	0.27	0.45	0.00	1.00
FEMALE	7185	0.53	0.50	0.00	1.00
SES	7185	0.00	0.78	-3.76	2.69
MATHACH	7185	12.75	6.88	-2.83	24.99

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

To check whether the data have been properly read into HLM

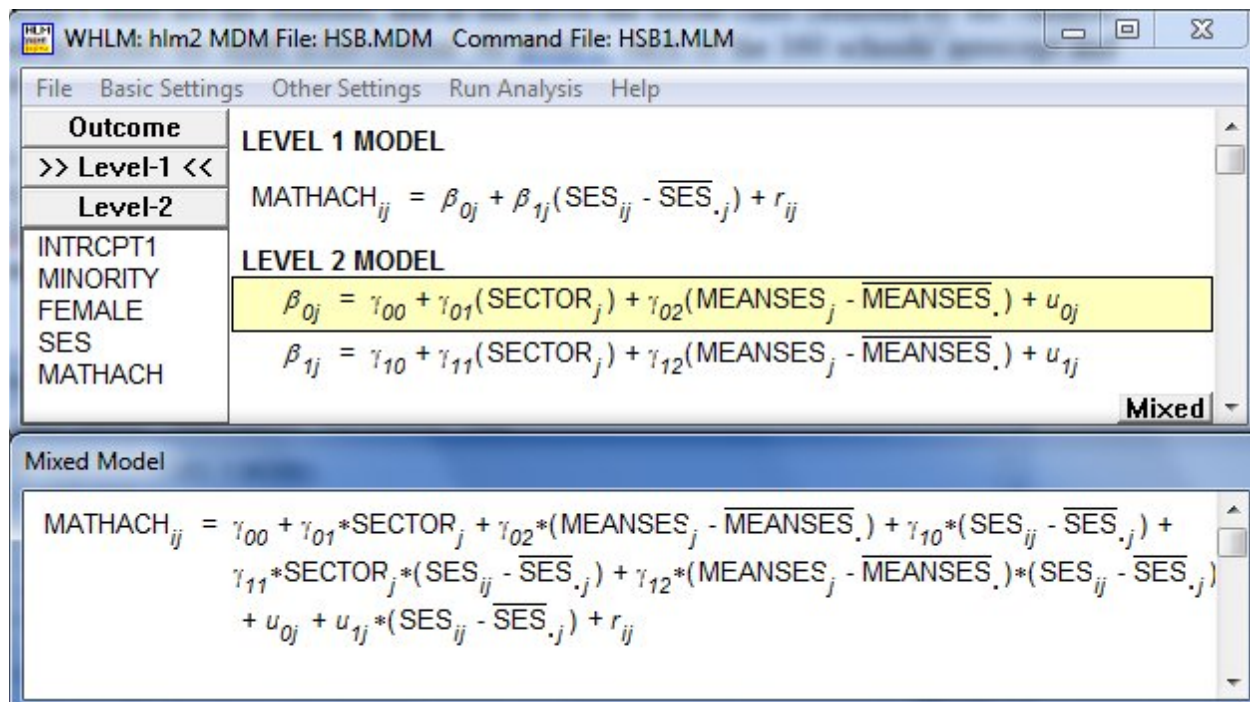
- Click **Done**. The WHLM window displays the type and name on its title bar (**hlm2** & **HSB.MDM**) and the level-1 variables on a drop-down menu.



### 3. Description of the model

Of interest in this example is the relationship between a student's mathematical achievement (MATHACH) and socio-economic status (SES). Socio-economic status was measured at both levels: at the student level, a measure of the individual's socio-economic status is represented by the variable SES. For each school, a measure of the mean SES for the school is represented by the variable MEANSES.

The 7185 level-1 units are the students, and at this level the social class (denoted by the variable SES) is used to model the math achievement. At level-2, each of the 160 schools' intercept and slope are predicted by school sector and school mean social class.



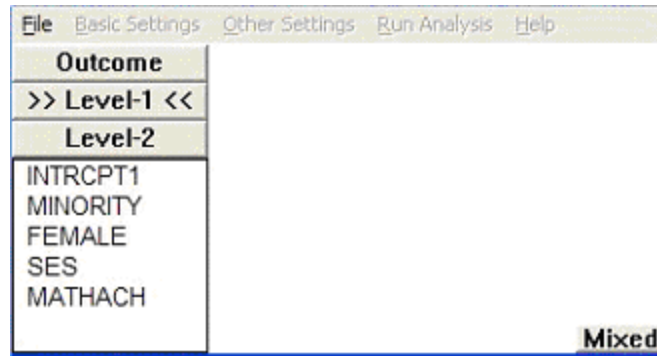
The information above implies that the level-1 model will have two coefficients for each student: the intercept ( $\beta_{0j}$ ) and the SES slope ( $\beta_{1j}$ ). At level-2, the intercept and SES slope are modeled as functions of sector and mean SES. Note that both the intercept and the slope are modeled as having randomly varying residuals. The assumption is that the intercept and slope vary not only as a function of the two predictors SECTOR and MEANSES, but also as a function of a unique school effect.

#### 4. Creating the command file

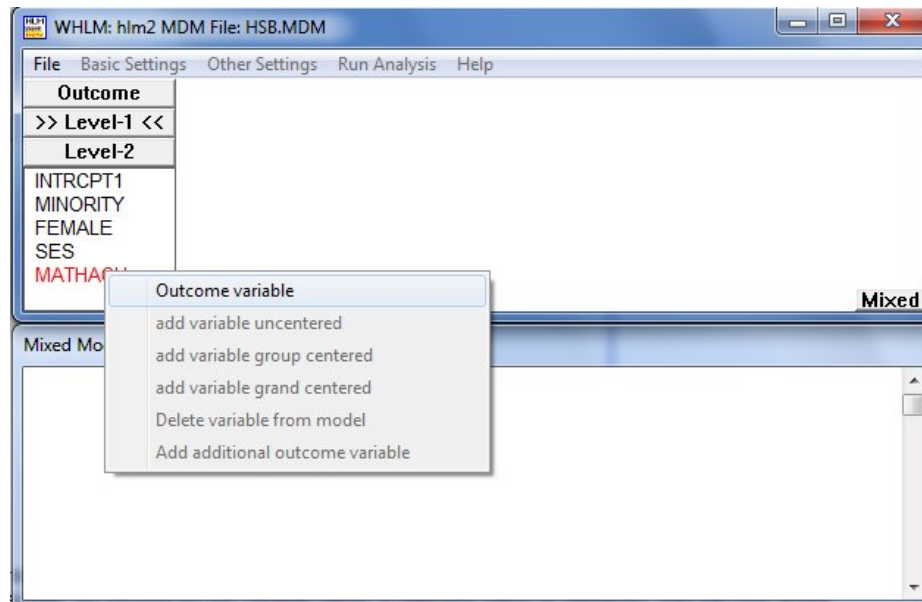
Before creating a new command file for the model given above, the relevant **MDM** file on which the analysis is to be based must be selected. In this case, we will simply continue using the **MDM** file constructed previously. However, to use another previously created **MDM** file, the **Create a new model using an existing MDM file** option should be selected from the **File** menu.

On the left of the main window, the level-1 variables for which information is available in the **MDM** file are given in a list box.

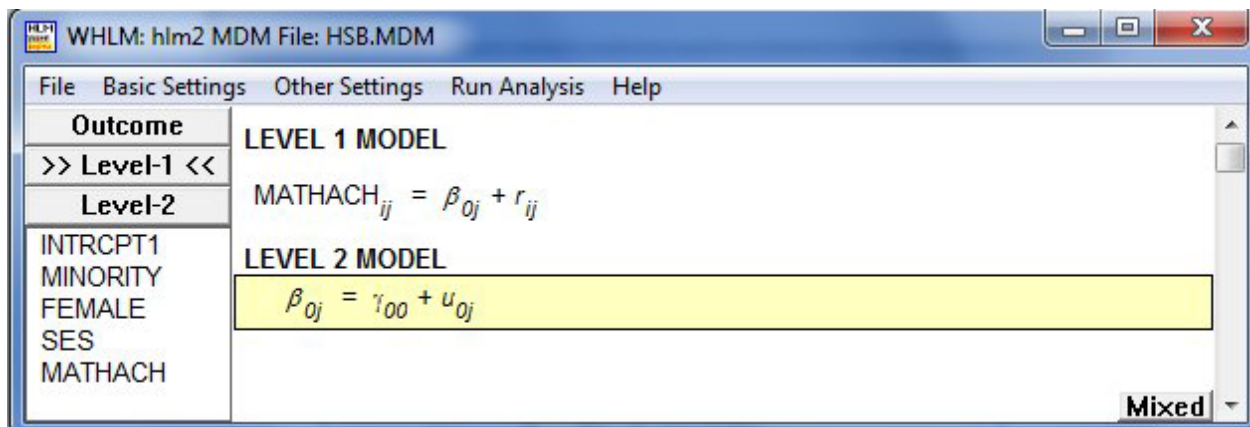




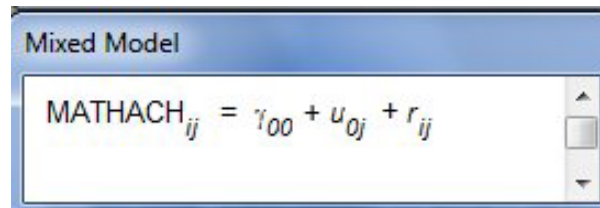
As a first step, the variable MATHACH, representing a student's math achievement, is selected as the outcome variable. By clicking on the variable name, a selection menu is displayed, with the only active option **Outcome variable**, as shown below:



Once the outcome variable has been selected, the following basic model is displayed in the **WHLM** window. This model is commonly referred to as the unconditional model.



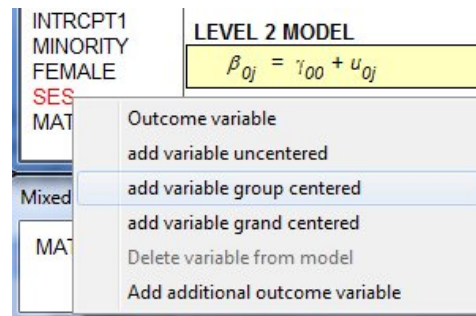
The unconditional model makes provision for a single common intercept. Variation in the outcome is parceled out between two variance components,  $u_{0j}$  representing the random variation in intercepts over the level-2 units (schools) and  $r_{ij}$  the residual variation at level-1.



Mixed Model

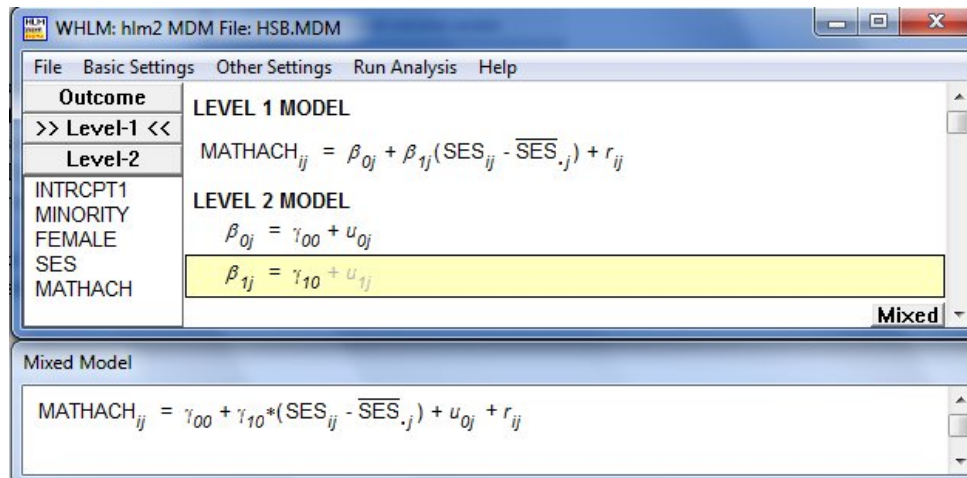
$$\text{MATHACH}_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

As the math achievement on level 1 of the model is to be modeled in terms of the student's social class, the variable SES is now added to the level-1 model. Clicking on the variable name SES leads to the display of a pop-up menu, on which a number of options are available.



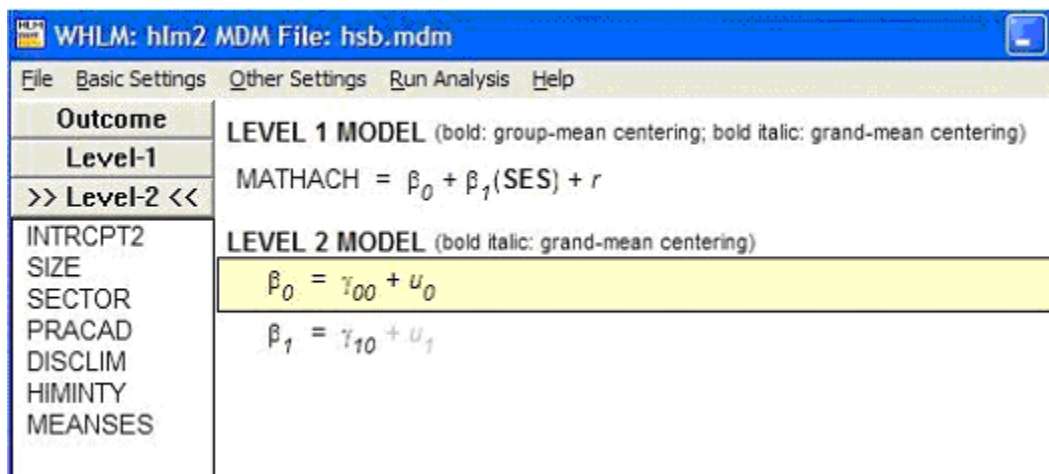
Select the option **add variable group centered**, indicating that the group mean is to be used as point of reference. Once this option has been selected, the additional variable is displayed in the **WHLM** window. Note that for each individual value of  $\text{SES}_{ij}$  the value of the mean socioeconomic status in unit  $i$  namely  $\text{SES}_{.j}$  is subtracted. The predictor now appears in the equation window and each regression coefficient associated with it has become an outcome in the level-2 model. By default, HLM only assumes a random intercept so only the coefficient  $u_0$  denoting random variation in the intercept over level-2 units is activated. The slope of the level-1 predictor(s) is assumed fixed. To fit a random SES slope as well, click on the slope equation ( $\beta_{1j}$ ) and then on the term  $u_{1j}$ .

This concludes the specification of the level-1 model.

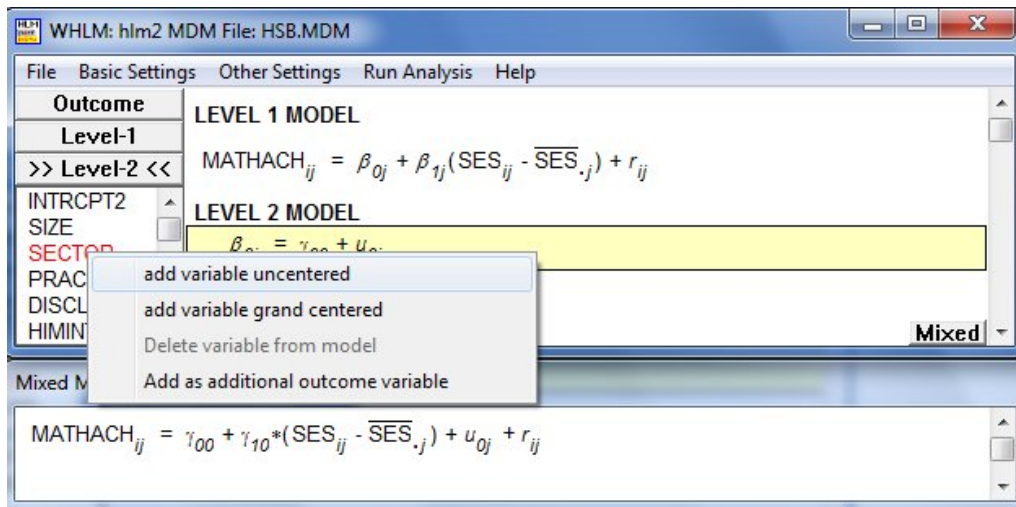


In the level-2 model, both the intercept and SES slope are to be modeled as dependent on the school's mean social class (MEANSES) and school sector (SECTOR).

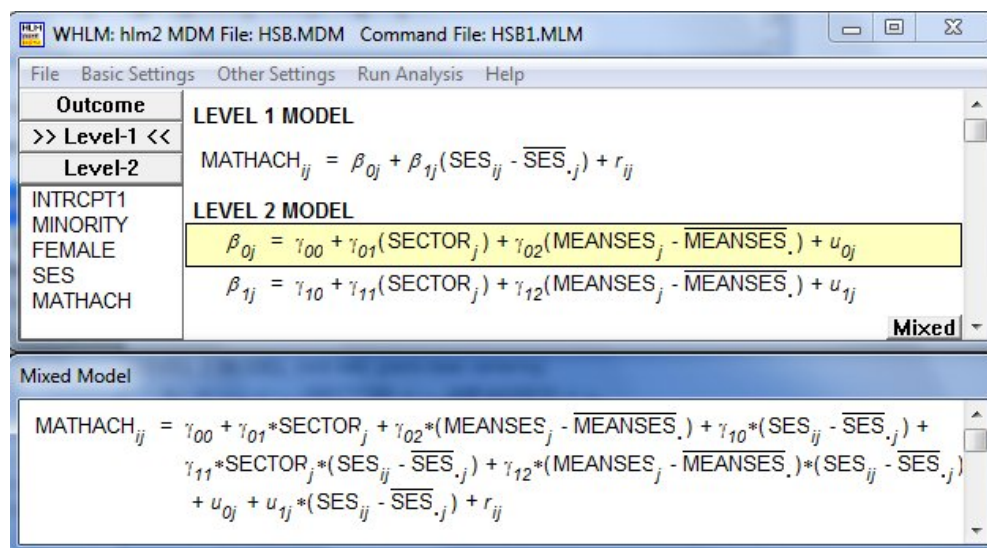
Start by clicking on the **Level-2** button on the left of the window. This leads to the display of the variables on which information at this level (>> **Level-2** <<)) is available in the **MDM** file. Select the equation for  $\beta_0$  by clicking on it, as predictors will be added to this equation first. The procedure to add variables to the equation for  $\beta_1$  is similar and is initiated by clicking on the equation for  $\beta_1$  before adding variables.



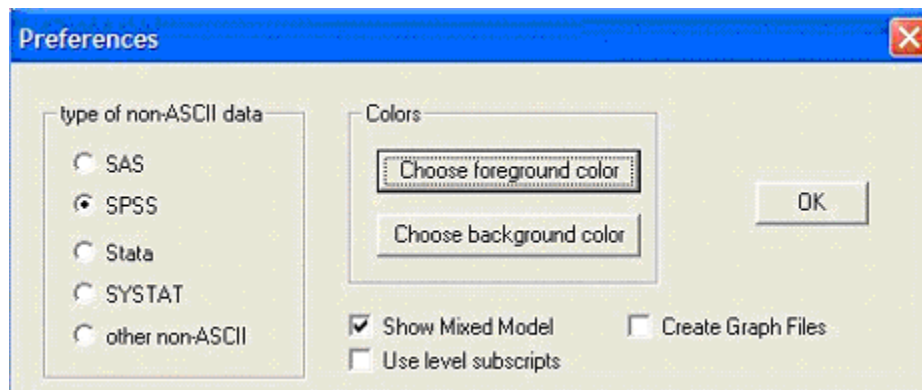
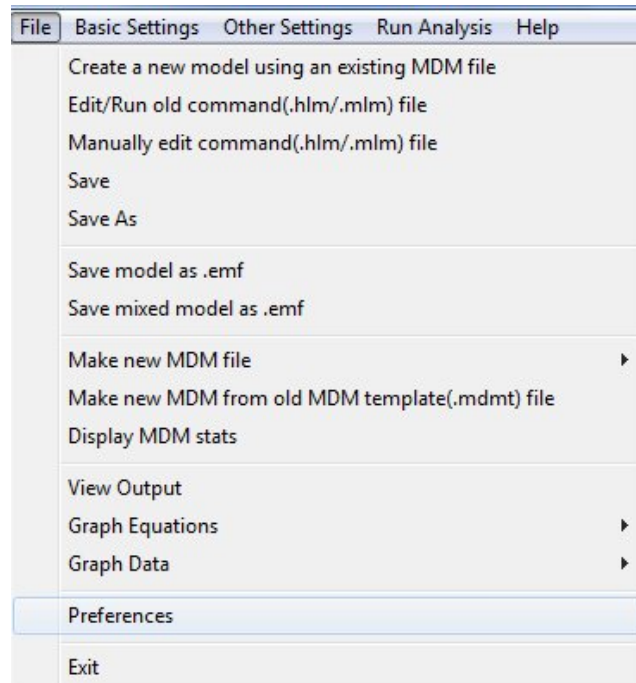
Select variables to be added to the level-2 equations as before. Click on a variable name to display a pop-up menu. The variables MEANSES is added **grand centered** and SECTOR is added **uncentered** to the model:



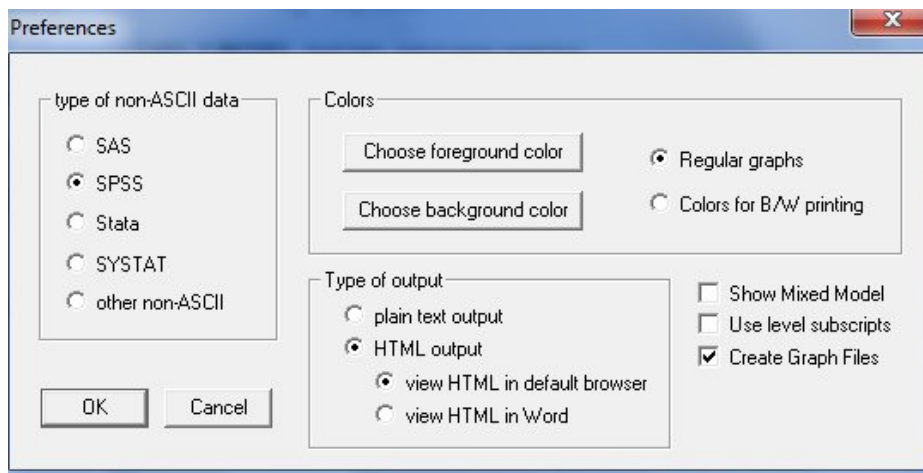
Repeat the variable selection procedure on the equation for  $\beta_1$  so that the variables MEANSES and SECTOR are also added **uncentered** to the equation. The  $u_1$  term at the end of the slope equation for  $\beta_1$  equation is grayed out. Enable this term as described previously and save the completed command file by using the **File, Save As** option. The final model is displayed below.



Should the user prefer the model displayed without subscripts or the mixed model display at the modeling window, the **Preferences** option from the **File** menu may be used to open the **Preferences** dialog box as shown below.

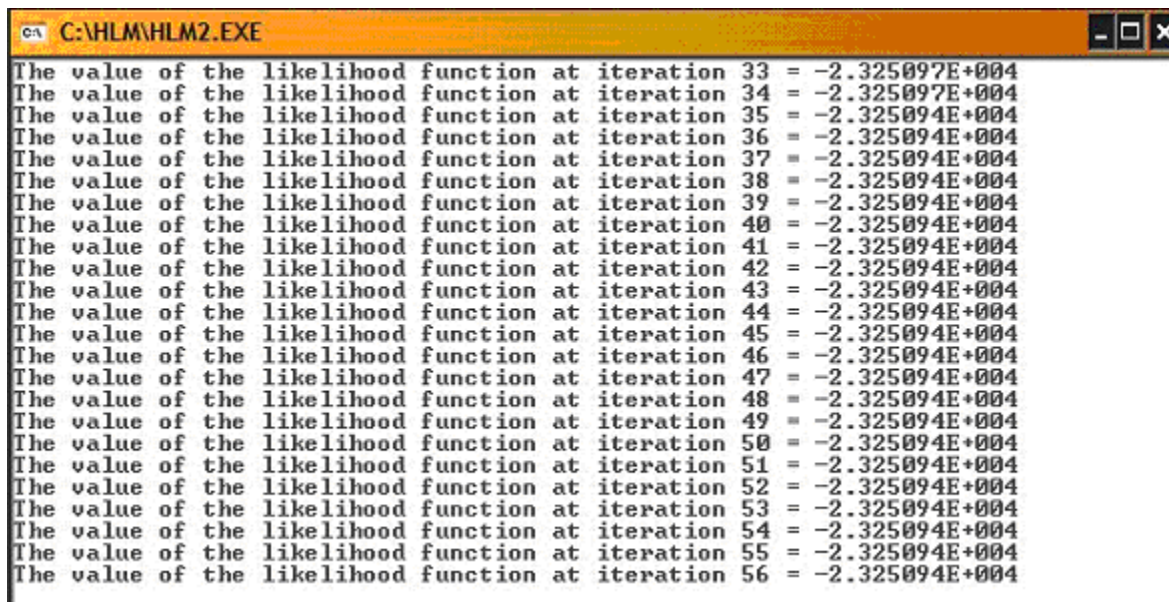


Uncheck the boxes for the request **Use level subscripts** and **Show Mixed Model**. Click the **OK** button.





After saving the command file, clicking the **Run Analysis** option at the top of the main window initiates the analysis. A screen showing details of the iterative procedure is shown. Note that the information on the screen is also given as part of the **WHLM** output file.



```
C:\HLM\HLM2.EXE
The value of the likelihood function at iteration 33 = -2.325097E+004
The value of the likelihood function at iteration 34 = -2.325097E+004
The value of the likelihood function at iteration 35 = -2.325094E+004
The value of the likelihood function at iteration 36 = -2.325094E+004
The value of the likelihood function at iteration 37 = -2.325094E+004
The value of the likelihood function at iteration 38 = -2.325094E+004
The value of the likelihood function at iteration 39 = -2.325094E+004
The value of the likelihood function at iteration 40 = -2.325094E+004
The value of the likelihood function at iteration 41 = -2.325094E+004
The value of the likelihood function at iteration 42 = -2.325094E+004
The value of the likelihood function at iteration 43 = -2.325094E+004
The value of the likelihood function at iteration 44 = -2.325094E+004
The value of the likelihood function at iteration 45 = -2.325094E+004
The value of the likelihood function at iteration 46 = -2.325094E+004
The value of the likelihood function at iteration 47 = -2.325094E+004
The value of the likelihood function at iteration 48 = -2.325094E+004
The value of the likelihood function at iteration 49 = -2.325094E+004
The value of the likelihood function at iteration 50 = -2.325094E+004
The value of the likelihood function at iteration 51 = -2.325094E+004
The value of the likelihood function at iteration 52 = -2.325094E+004
The value of the likelihood function at iteration 53 = -2.325094E+004
The value of the likelihood function at iteration 54 = -2.325094E+004
The value of the likelihood function at iteration 55 = -2.325094E+004
The value of the likelihood function at iteration 56 = -2.325094E+004
```

Once the iterative procedure has converged, the output will automatically be displayed in the format set on the **Preferences** dialog box.

## 5. Interpreting the output

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

### Specifications for this HLM2 run

Problem Title: Intercepts and Slopes-as-outcomes Model

The data source for this run = HSB.MDM

The command file for this run = HSB1.MLM

Output file name = hlm2.html

The maximum number of level-1 units = 7185

The maximum number of level-2 units = 160

The maximum number of iterations = 100

Method of estimation: restricted maximum likelihood

The outcome variable is MATHACH

## Summary of the model specified

### Level-1 Model

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

### Level-2 Model

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(\text{SECTOR}_j) + \gamma_{02}(\text{MEANSES}_j) + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}(\text{SECTOR}_j) + \gamma_{12}(\text{MEANSES}_j) + u_{1j}\end{aligned}$$

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

### Mixed Model

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

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

### Level-1 OLS Regressions

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

The **Output Settings – HLM2** dialog box (obtained by selecting the **Output Settings** option) in the **Other Settings** menu of **WHLM** may be used to change the number of units printed. When first analyzing a new data set, examining the OL equations for all the units may be helpful in identifying possible outlying cases and bad data.

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

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

### Least Squares Estimates

$$\sigma^2 = 39.03409$$

### Least-squares estimates of fixed effects

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, $\beta_0$ INTRCPT2,					
$\gamma_{00}$	12.083837	0.106889	113.050	7179	<0.001
SECTOR, $\gamma_{01}$ MEANSES,	1.280341	0.157845	8.111	7179	<0.001
$\gamma_{02}$	5.163791	0.190834	27.059	7179	<0.001
For SES slope, $\beta_1$ INTRCPT2,					
$\gamma_{10}$	2.935664	0.155268	18.907	7179	<0.001
SECTOR, $\gamma_{11}$ MEANSES,	-1.642102	0.240178	-6.837	7179	<0.001
$\gamma_{12}$	1.044120	0.299885	3.482	7179	<0.001

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

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, $\beta_0$ INTRCPT2,					
$\gamma_{00}$	12.083837	0.169507	71.288	7179	<0.001
SECTOR, $\gamma_{01}$ MEANSES,	1.280341	0.299077	4.281	7179	<0.001
$\gamma_{02}$	5.163791	0.334078	15.457	7179	<0.001
For SES slope, $\beta_1$ INTRCPT2,					
$\gamma_{10}$	2.935664	0.147576	19.893	7179	<0.001
SECTOR, $\gamma_{11}$ MEANSES,	-1.642102	0.237223	-6.922	7179	<0.001
$\gamma_{12}$	1.044120	0.332897	3.136	7179	0.002

The least-squares likelihood value = -2.336211E+004

Deviance = 46724.22267

Number of estimated parameters = 1

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



## Starting Values

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

$\tau_{(0)}$

INTRCPT1, $\beta_0$	2.56964	0.28026
SES, $\beta_1$	0.28026	-0.01614

New  $\tau_{(0)}$

INTRCPT1, $\beta_0$	2.56964	0.28026
SES, $\beta_1$	0.28026	-0.01614

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

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

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, $\beta_0$					
INTRCPT2,					
$\gamma_{00}$	12.094864	0.204326	59.194	157	<0.001
SECTOR, $\gamma_{01}$	1.226266	0.315204	3.890	157	<0.001
MEANSES,					
$\gamma_{02}$	5.335184	0.379879	14.044	157	<0.001
For SES slope, $\beta_1$					
INTRCPT2,					
$\gamma_{10}$	2.935219	0.168674	17.402	157	<0.001
SECTOR, $\gamma_{11}$	-1.634083	0.260672	-6.269	157	<0.001
MEANSES,					
$\gamma_{12}$	1.015061	0.323523	3.138	157	0.002

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

The value of the log-likelihood function at iteration 1 = -2.325199E+004

The value of the log-likelihood function at iteration 2 = -2.325182E+004

The value of the log-likelihood function at iteration 3 = -2.325174E+004

The value of the log-likelihood function at iteration 4 = -2.325169E+004

The value of the log-likelihood function at iteration 5 = -2.325154E+004

...

The value of the log-likelihood function at iteration 57 = -2.325094E+004

The value of the log-likelihood function at iteration 58 = -2.325094E+004

The value of the log-likelihood function at iteration 59 = -2.325094E+004

The value of the log-likelihood function at iteration 60 = -2.325094E+004

The third section of the output provides the final estimates of the fixed and random effects specified in the model.

\*\*\*\*\* ITERATION 61 \*\*\*\*\*

$\sigma^2 = 36.70313$

*Level-1 variance components*

$\tau$

*Level-2 variance-covariance components*

INTRCPT1, $\beta_0$	2.37996	0.19058
SES, $\beta_1$	0.19058	0.14892

*Level-2 variance-covariance components expressed as correlations*

$\tau$  (as correlations)

INTRCPT1, $\beta_0$	1.000	0.320
SES, $\beta_1$	0.320	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, $\beta_0$	0.733
SES, $\beta_1$	0.073

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

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

In the next few lines of the output, reliability estimates are given for the level-1 coefficients. This is an overall or average reliability for each level-1 coefficient across the set of 160 level-2 units (schools). These estimates are calculated according to Equation 3.59 in *Hierarchical Linear Models*, p. 50 and depend on two factors: the degree to which the true underlying parameters vary from school to school and the precision with which each school's regression equation is estimated.

The reliability of the mean of group  $j$  is defined as

$$\lambda_j = \frac{\tau_0^2}{\tau_0^2 + \sigma^2 / n_j}$$

where  $\sigma^2 / n_j$  is the variance of  $\bar{Y}_j$  as an estimator of  $\beta_{0j}$  while  $\tau_0$  is the variance of the true means,  $\beta_{0j}$ , around the grand mean  $\gamma_{00}$ . The error variance  $\sigma^2 / n_j$  is dependent on the sample size  $n_j$  and varies from level-2 unit to level-2 unit, while  $\tau_0$  is constant over level-2 units. Note that in the case of two groups with the first group having a group size twice that of a second group, the reliability of the mean for the first group would be twice that of the second group:

$$\lambda_{\text{group1}} / \lambda_{\text{group2}} = \frac{\tau_0^2 + \sigma^2 / 2n_2}{\tau_0^2} / \frac{\tau_0^2}{\tau_0^2 + \sigma^2 / n_2} = 2.$$

The reliability of aggregated variables increases as the number of level-1 units nested within a level-2 unit increases. If  $n_j$  is very large,  $\lambda_j$  will get closer to a value of 1. Similarly, when the group means vary substantially across the groups (holding constant the sample size per group),  $\lambda_j$  will tend to approach a value of 1.

The reliabilities of group-level parameter estimates in the output indicate the reliability with which one can discriminate among the groups using their OLS estimates of the random parameter(s). Note that low reliabilities do not invalidate the HLM analysis, but that very low reliabilities (e.g.  $< .10$ ) often indicate that a random coefficient might be considered fixed in subsequent analyses.

The precision of estimation of the intercept (which in this application is a school mean) depends on the sample size within each school. The precision of estimation of the slope depends both on the sample size and on the variability of SES within that school. Schools that are homogeneous with respect to SES will exhibit slope estimation with poor precision.

From the table, it can be seen that the slope estimates are far less reliable than the intercept estimates. The primary reason for the lack of reliability of the slopes is that the true slope variance across schools is much smaller than the variance of the true means. Also, the slopes are estimated with less precision than are the means because many schools are relatively homogeneous on SES.

The next three tables present the final estimates for: the fixed effects with GLS and robust standard errors, variance components at level-1 and level-2, and related test statistics. HLM produces two final tables of fixed effects: one with and one without robust standard errors. Robust standard errors are standard errors that are relatively insensitive to misspecification at the levels of the model and the distributional assumptions at each level. If the robust and model-based standard errors differ substantially, it is recommended that the user further investigate the tenability of the HLM assumptions of multivariate normality. Residual analysis based on the residual files produced by HLM2 and HLM3 may be used in this regard. The difference between the two sets of standard errors provided is in the use of a weight matrix in the case of the robust standard errors. Elements of this weight matrix are used to adjust the standard errors.

Note that the robust standard errors should be trusted only when the number of higher-level units is moderately large relative to the number of explanatory variables at a higher level. At some point, the number of higher-level units can become so small that these standard errors are not computable as the information matrix is uninvertible or not positive-definite. In such cases, a message that robust standard errors could not be computed for a model may be printed to the HLM output file.

The first table provides model-based estimates of the standard errors while the second table provides robust estimates of the standard errors. Note that the two sets of standard errors are similar, indicating that there is no reason to suspect a violation of underlying assumptions.

By looking at the estimates of the level-2 fixed effects we can answer the following questions:

- Do SECTOR and MEANSES significantly predict the intercept  $\beta_0$ ?
- Do SECTOR and MEANSES significantly predict the SES slope  $\beta_1$ ?

From the final estimates of the fixed effects, we see that all the predictors made a significant contribution to the explanation of the variation in the outcome variable MATHACH. The school sector had a positive influence on the intercept, but a negative influence on the slope of SES. As Catholic schools were coded 1, this implies that the intercept for students in Catholic schools was, on average, 1.226 units higher than for students in public schools. The SES slope for Catholic schools was 1.641 units lower than for public schools. The largest contribution to the intercept, however, is from the mean social class variable (MEANSES). The higher the mean SES of the group, the higher the intercept for a particular school.

**Final estimation of fixed effects:**

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, $\beta_0$					
INTRCPT2, $\gamma_{00}$	12.095006	0.198717	60.865	157	<0.001
SECTOR, $\gamma_{01}$	1.226384	0.306272	4.004	157	<0.001
MEANSES, $\gamma_{02}$	5.333056	0.369161	14.446	157	<0.001
For SES slope, $\beta_1$					
INTRCPT2, $\gamma_{10}$	2.937787	0.157119	18.698	157	<0.001
SECTOR, $\gamma_{11}$	-1.640954	0.242905	-6.756	157	<0.001
MEANSES, $\gamma_{12}$	1.034427	0.302566	3.419	157	<0.001

**Final estimation of fixed effects  
(with robust standard errors)**

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, $\beta_0$					
INTRCPT2, $\gamma_{00}$	12.095006	0.173688	69.637	157	<0.001
SECTOR, $\gamma_{01}$	1.226384	0.308484	3.976	157	<0.001
MEANSES, $\gamma_{02}$	5.333056	0.334600	15.939	157	<0.001
For SES slope, $\beta_1$					
INTRCPT2, $\gamma_{10}$	2.937787	0.147615	19.902	157	<0.001
SECTOR, $\gamma_{11}$	-1.640954	0.237401	-6.912	157	<0.001
MEANSES, $\gamma_{12}$	1.034427	0.332785	3.108	157	0.002

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

### Final estimation of variance components

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

From the final estimate of the level-1 variance and its standard error it can be seen that the variation over students within schools is quite small. The coefficient of 2.37996 (with corresponding chi-square value of 605) indicates significant variability among schools in terms of their average math achievement. The largest variance component is at level-1 of the model (36.70313), indicating that quite a lot of the variation in the outcome remains unexplained by this model.

There is not significant variability in terms of the SES slopes for the level-2 units, as indicated by the estimate of 0.14892 (with p-value of 0.369) for the level-2 component  $u_1$ . This insignificant variability is also shown above as the small value of percentage of level-2 variance that is in SES slope.

### Statistics for current covariance components model

Deviance = 46501.875643

Number of estimated parameters = 4

The final output given is the deviance statistic for this model, together with the degrees of freedom. This information may be used to compare the fit of two models, as can be seen in the section on hypothesis testing.

In the next example on this page modification of this model and the testing of hypotheses concerning both fixed effects and the homogeneity of the level-1 variance will be considered.