



Checking homogeneity of level-1 variance using the HSB data

1. Description of the data	1
2. Creating the command file	3
3. Interpreting the output.....	4

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 second in a set of seven examples based on these data described on the HLM Support page and used the MDM file created in the first example (**HSB.MDM**).

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. 1.1.

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 1.1 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)

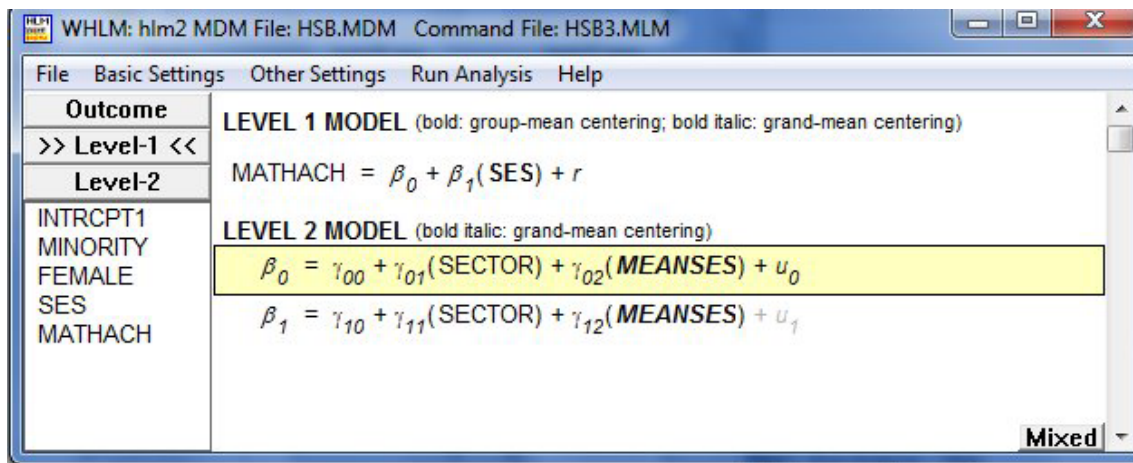
The data for the first ten schools are displayed in Fig 1.2.

	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

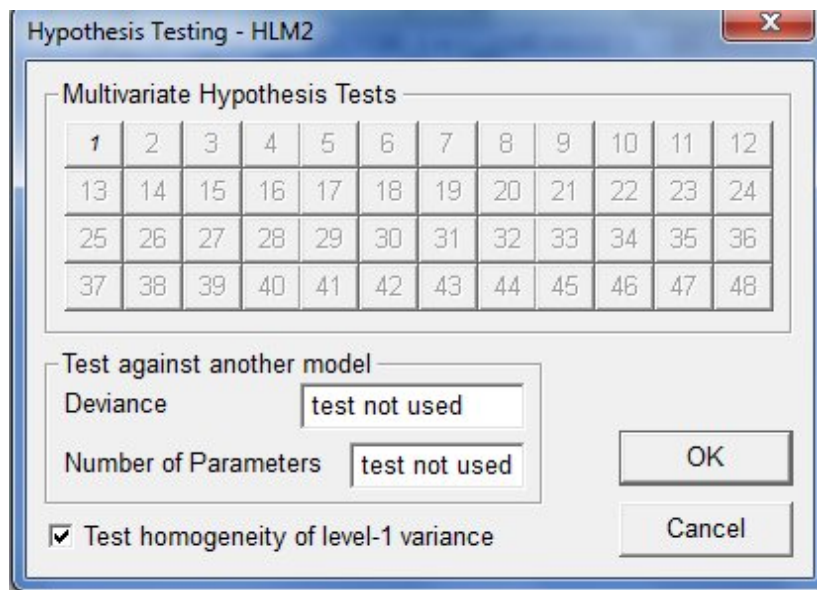
Figure 1.2 First ten cases in HSB2.SAV

2. Creating the command file

In all the models fitted thus far, it was assumed that the errors in the level-1 model have equal variance σ^2 . Should that not be the case and in reality the level-1 variance varies randomly over level-2 units, the impact on inference about the level-2 effects will be mild. However, if the variances are a function of level-1 or level-2 predictors, the consequences may be more serious. In a previous example we have already modeled the level-1 variance as a function of a level-1 predictor. In this example, we simply test the validity of the homogeneity assumption. The command file used as point of departure is shown below (HSB3.HLM).



Select the **Hypothesis Testing** option from the **Other Settings** menu on the main menu bar to access the **Hypothesis Testing – HLM2** dialog box.



Check the box for **Test homogeneity of level-1 variance** and click **OK** to return to the main window. Save the additional instructions to the command file using the **File, Save** command before running the analysis.

3. Interpreting the output

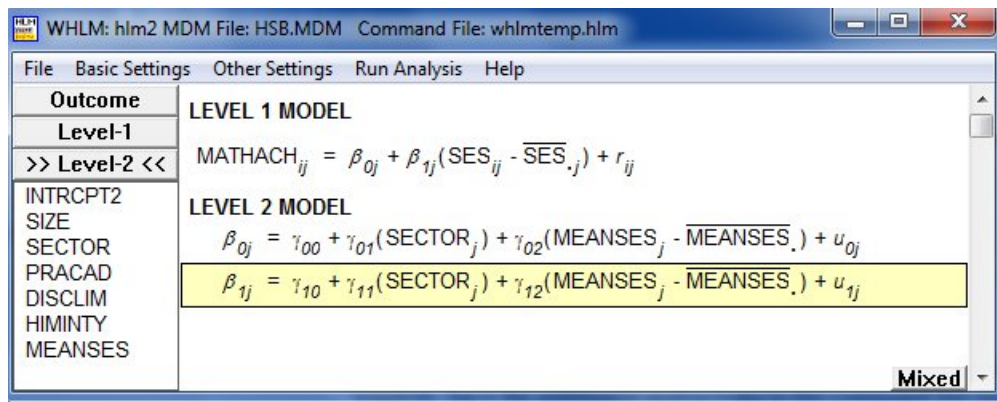
Results of the homogeneity test are given at the end of the output file.

Test of homogeneity of level-1 variance
 χ^2 statistic = 244.08638
 degrees of freedom = 159
 p-value = 0.000

The results indicate that the statistic is very significant and that heterogeneity of level-1 variance exists between the 160 schools. Closer scrutiny of the level-1 model is advisable before continuing additional analyses. It may also be that a few unusual schools account for most of the observed heterogeneity.

In the case of these data, a heterogeneous model such as fitted in a previous model is one option.

In some cases, heterogeneous level-1 variances may indicate the need to specify certain level-1 coefficients as randomly varying at level-2. To see whether allowing the SES slope to vary randomly over the level-2 units play a role, we rerun this model with both random intercept and slope. This model is shown below.



Test of homogeneity of level-1 variance
 χ^2 statistic = 245.76576
 degrees of freedom = 159
 p-value = 0.000

Results for the random intercept and slope model does not differ much from that of the random slope model. Allowing the SES slope to vary randomly over the level-2 units does not seem to play a role in terms of the heterogeneity of the level-1 variance between the schools.