Level-1 variance: heteroscedasticity/testing homogeneity

Implicit in hierarchical linear models are assumptions concerning the distributions of the error or residual terms at the various levels of the hierarchy. In most cases, it is assumed that the errors in the level-1 model are normally distributed with expected mean zero and equal variance σ^2. Violation of this assumption is not without consequences: if the level-1 variances are assumed to be equal but are really unequal, the point estimation of the level-2 coefficients will not be biased. However, these estimates will be inefficient and the associated standard errors will be biased. As such, it is advisable to check the validity of the assumptions after model specification by performing a test for the homogeneity of the level-1 variances.

Causes of heterogeneous level-1 variance

Heterogeneity of the level-1 variances may have several causes:

- The omission of one or more important level-1 variables. Heterogeneity of variance would result if the excluded variable were distributed with unequal variance across groups.
- Fixing or omitting the effects of a level-1 predictor that is random or non-randomly varying.
- Bad data. Extreme data values due to, for example, bad coding, may inflate the variance for a group, and overall significant heterogeneity of variance may consequently be observed.
- The use of non-normal data with heavy tails. Parametric tests are sensitive to the presence of more extreme observations than expected, and heavy tails (kurtosis) can cause a significant test for heterogeneity of variance.

A good example of predictor-dependent residual variance is given by Snijders & Bosker, (1999), page 111-112, which is given below:

The data set used in this illustration contains information on grade 8 students in elementary schools in The Netherlands. Information on 2287 students clustered within 131 schools is available, with class sizes ranging from 4 to 35. The outcome variable of interest is the score on a language test. Predictors already added to the model include the socio-economic status and an indicator (COMB) representing whether a class is multi-grade or not. The predictor variable IQ is on a scale, with mean 0 and standard deviation 2.7 while the predictor GS indicates the group size. The model specification is as follows:

$$SCORE_{ij} = \beta_{0j} + \beta_{1j}(IQ_{ij}) + \beta_{2j}(SES_{ij}) + \beta_{3j}(GENDER_{ij}) + r_{ij}$$

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}(\overline{IQ}_{.j}) + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} \\ \beta_{3j} &= \gamma_{30} \end{aligned}$$

Note that, through the inclusion of IQ, GS and COMB in both level-2 equations main effects and cross-level interactions are included in the model. Also note that the slope of IQ is allowed to vary randomly over the groups, while a non-randomly varying slope is assumed for socio-economic status (SES) and the gender of the student (GENDER). In the table below, results are given for two models: one where the standard assumption of homogeneous error variance is made, and a second model where it is assumed that the level-1 variance is gender dependent and can be expressed as

$$\text{level-1 variance} = \sigma_0^2 + 2\sigma_{01}\left(GENDER_{ij}\right)$$

The first model shown is a homoscedastic model while the second is a gender-dependent heteroscedastic model.

Fixed effect	Model 1 (homoscedastic)		Model 2 (heteroscedastic)	
	Coefficient	S.E.	Coefficient	S.E.
γ_{00}	39.53	0.31	39.53	0.31
γ_{10}	2.268	0.081	2.264	0.081
γ_{20}	0.152	0.014	0.151	0.014
γ_{30}	2.64	0.26	2.64	0.26
γ_{01}	1.02	0.32	1.01	0.32

Random effect	Model 1 (homoscedastic)		Model 2 (heteroscedastic)	
	Coefficient	S.E.	Coefficient	S.E.
$\text{var}(u_{0j})$	8.27	1.33	8.24	1.33
$\text{var}(u_{1j})$	0.169	0.088	0.171	0.088
$\text{cov}(u_{1j}, u_{0j})$	-0.76	0.25	-0.78	0.26
σ_0^2	37.56	1.17	38.72	1.67
σ_{01}			-1.21	1.17
Deviance	15005.5		15004.4	

From the table of fixed effects, it follows that the fixed effect of gender is significant. After controlling for IQ and SES, girls are expected so score 2.64 higher than boys in the test.

According to the model for the level-1 variance specified for the heteroscedastic model

$$\text{level-1 variance} = \sigma_0^2 + 2\sigma_{01}\left(GENDER_{ij}\right)$$

the residual variance for boys is 38.72 and (38.72 - 2(1.21)) = 36.30 for girls. The average of these two residual variances is 37.51, which is very close to the residual variance estimated under the homoscedastic model. In this sample, where the gender groups are approximately equal in size, this is to expected. When the difference between the residual variances is tested for statistical significance, the chi-square obtained from the deviances is 1.1, with 1 degree of freedom and thus not significant.

However, when it is investigated whether the residual variance is dependent on IQ instead of gender, significant heterogeneity is encountered. The conclusion is reached that the scores of the less intelligent students are more variable than the scores of the more intelligent.

Failure to take heterogeneity present in the residual variance into account may lead to a misspecified model and consequently to incorrect parameter estimates and standard errors.

Testing homogeneity of level-1 variance

A full discussion of the process of testing for heterogeneity using HLM and the residual files produced by the program is given on pages 207-211 of the Bryk & Raudenbush text. In short, the test is based on the standardized measure of dispersion for each group j:

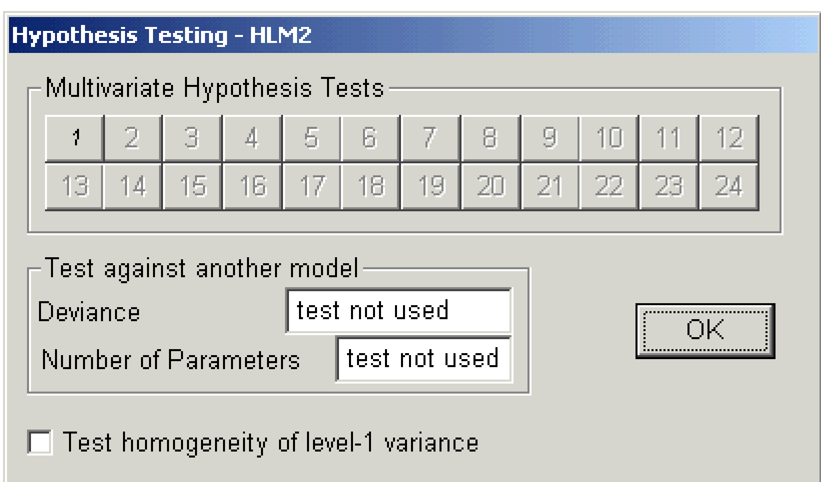
$$d_j = \frac{\ln\left(S_j^2\right) - \left[\sum f_j \ln\left(S_j^2\right) / \sum f_j\right]}{\left(2/f_j\right)^{1/2}}$$

Use is made of the simple and commonly used test statistic for homogeneity

$$H = \sum d_j^2$$

which has a large sample chi-square distribution with $J-1$ degrees of freedom under the homogeneity hypothesis. This test is appropriate when the data are normal and sample size per unit are 10 or more.
The homogeneity test is a function of the OLS residuals (OLSRSVAR in the residual file). There is a $n(j)-r$ factor (r being the number of random effects) in the denominator used to compute this. A such, the degrees of freedom reported by HLM will be based on the number of level-2 units where $n(j)$ exceeds 2.

The option is available when the **Hypothesis Testing** option is selected from the **Optional Settings** menu. To test this hypothesis, check the box next to **Test homogeneity of level-1 variance** on the **Hypothesis Testing** dialog box as shown below.



After analysis, the following output is written to the output file:

```
Test of homogeneity of level-1 variance
 ----------------------------------------
 Chi-square statistic = 245.76576
 Number of degrees of freedom = 159
 P-value = 0.000
```

If the chi-square statistic is significant, as is the case in this example, it indicates that the null hypothesis of homogeneity of the level-1 variance is rejected. Note that this option is only available for two level models analyzed with HLM2. HLM3, HMLM, HMLM2, and HCM2 do not currently have this feature.

Modeling level-1 variance as a function of predictors

In addition to this, HLM also offers the opportunity to model the level-1 variance as a function of predictors in the MDM file. Again, this option is only available for linear and non-linear two-level (HLM2) models. The **Heterogeneous sigma^2** dialog box is accessed by selecting **Estimation Settings** option from the **Other Settings** menu. When the **heterogeneous sigma^2** button is clicked on the **estimation Settings – HLM2** dialog box, the **Heterogeneous sigma^2: Predictors of level-1 variance** dialog box appears. In the example below, the residual variance for the model described in **hsb1.mlm** is modeled as a function of the level-1 predictor FEMALE.

In the HLM output file, this is represented as part of the model specification at the top of the output file:

```
Var(R) = Sigma_squared and
log(Sigma_squared) = alpha0 + alpha1(FEMALE)
```

At the end of the output file, additional results for the heterogeneous model are printed, in addition to the standard results for the homogeneous model. A test for statistical significance based on the deviances for the two models is also automatically performed by HLM, with results as shown below.

```
RESULTS FOR HETEROGENEOUS SIGMA-SQUARED
(macro iteration 4)
Var(R) = Sigma_squared and
log(Sigma_squared) = alpha0 + alpha1(FEMALE)
Model for level-1 variance
-----------------------------------------------------------------
Standard
Parameter Coefficient Error Z-ratio P-value
-----------------------------------------------------------------
INTRCPT1 ,alpha0 3.66581 0.024717 148.308 0.000
FEMALE ,alpha1 -0.12121 0.033936 -3.572 0.001
-----------------------------------------------------------------
Summary of Model Fit
----------------------------------------------------------------
Model Number of Deviance
Parameters
----------------------------------------------------------------
1. Homogeneous sigma_squared 10 46494.59260
2. Heterogeneous sigma_squared 11 46482.02598
----------------------------------------------------------------
Model Comparison Chi-square df P-value
```

```
-------------------------------------------------------------------
Model 1 vs Model 2 12.56662    1   0.001
```