



## The contents and use of the additional output file tauvc.dat

Additional output files are obtained by selecting the **Output Settings** option from the **Other Settings** menu to bring up the **Output Settings** dialog box. On this dialog box, the **Print variance-covariance matrices** option should be checked. To illustrate, we first consider a three-level model. The additional output files for the fixed effects in the model (**gamvc.dat** and **gamvcr.dat**) are discussed elsewhere.

### A three-level example:

The file **tauvc.dat** contains the tau matrices and the inverse of the associated information matrix. For a model with  $r$  random effects at level-1, and  $r2$  random effects at level-2, estimated with full maximum likelihood rather than the default restricted maximum likelihood, **tauvc.dat** contains the following:

- $\tau_{\pi}$ , in  $r$  columns and  $r$  rows
- $\tau_{\beta}$ , in  $r$  columns and  $r2$  rows
- the inverse of the information matrix. The dimensions of the matrix are

$$\left( \frac{r(r+1)}{2} + \frac{r2(r2+1)}{2} \right) \times \left( \frac{r(r+1)}{2} + \frac{r2(r2+1)}{2} \right)$$

The screenshot shows the HLM software interface with the following model specifications:

Level	Model Specification
Level-1	LEVEL 1 MODEL (bold: group-mean centering; bold italic: grand-mean centering) MATH = $\pi_0 + \pi_1(\text{YEAR}) + e$
Level-2	LEVEL 2 MODEL (bold: group-mean centering; bold italic: grand-mean centering) $\pi_0 = \beta_{00} + r_0$ $\pi_1 = \beta_{10} + r_1$
Level-3	LEVEL 3 MODEL (bold italic: grand-mean centering) $\beta_{00} = \gamma_{000} + u_{00}$ $\beta_{10} = \gamma_{100} + u_{10}$

The interface also shows a list of variables on the left: INTRCPT1, YEAR, GRADE, MATH, and RETAINED. The bottom right corner of the window is labeled "Mixed".

For the model shown above, tauvc.dat contains the following:

```

0.6404879    0.0467574
0.0467574    0.0112249
0.1653112    0.0170460
0.0170460    0.0110198
0.0006327    0.0000285    0.0000014    -0.0000281    -0.0000014    -0.0000001
0.0000285    0.0000249    0.0000020    -0.0000014    -0.0000011    -0.0000001
0.0000014    0.0000020    0.0000039    -0.0000001    -0.0000001    -0.0000002
-0.0000281   -0.0000014   -0.0000001    0.0013258    0.0001252    0.0000117
-0.0000014   -0.0000011   -0.0000001    0.0001252    0.0000518    0.0000087
-0.0000001   -0.0000001   -0.0000002    0.0000117    0.0000087    0.0000063
0.3014750

```

- For this model,  $r = r_2 = 2$ , so that  $\tau_\pi : 2 \times 2$ ,  $\tau_\beta : 2 \times 2$ , and the information matrix is of dimension  $6 \times 6$ .
- The first two lines contain the estimates of the level-2 variance-covariance matrix,  $\tau_\pi$ , and corresponds with the Tau(pi) matrix as listed in the output file. The elements of this matrix are  $\tau_{\pi_{00}}$ ,  $\tau_{\pi_{10}}$ ,  $\tau_{\pi_{01}}$  and  $\tau_{\pi_{11}}$ , with  $\tau_{\pi_{00}} = 0.6404879$ ,  $\tau_{\pi_{10}} = 0.0467574$ , and  $\tau_{\pi_{11}} = 0.0112249$  respectively.
- The third and fourth lines are the level-3 variance-covariance matrix,  $\tau_\beta$ . We see that  $\tau_{\beta_{00}} = 0.1653112$ ,  $\tau_{\beta_{10}} = 0.0170460$ , and  $\tau_{\beta_{11}} = 0.0110298$ . This matrix is also called the asymptotic covariance matrix of the estimated  $\tau$ -elements.
- The matrices are followed by the inverse of the information matrix,  $\hat{V}$ . The information matrix can be divided into four blocks. The two blocks on the diagonal (see table below) contain the variances of the elements of the matrices at levels 2 and 3 respectively. The off-diagonal blocks, which are the same due to the symmetry of the information matrix, contains the covariances of the elements of the  $\tau$  matrices.

	$\tau_{\pi_{00}}$	$\tau_{\pi_{10}}$	$\tau_{\pi_{11}}$	$\tau_{\beta_{00}}$	$\tau_{\beta_{10}}$	$\tau_{\beta_{11}}$
$\tau_{\pi_{00}}$	0.0006327	0.0000285	0.0000020	-0.0000281	-0.0000014	-0.0000001
$\tau_{\pi_{10}}$	0.0000285	0.0000249	0.0000020	-0.0000014	-0.0000011	-0.0000001
$\tau_{\pi_{11}}$	0.0000014	0.0000020	0.0000039	-0.0000001	-0.0000001	-0.0000002
$\tau_{\beta_{00}}$	-0.0000281	-0.0000014	-0.0000001	0.0013258	0.0001252	0.0000117
$\tau_{\beta_{10}}$	-0.0000014	-0.0000011	-0.0000001	0.0001252	0.0000518	0.0000087
$\tau_{\beta_{11}}$	-0.0000001	-0.0000011	-0.0000001	0.0001252	0.0000518	0.0000087

The square roots of the diagonal elements of this matrix are the standard errors of the  $\tau$  s. For example, if we denote the  $q$ -th diagonal element of the matrix by  $\hat{V}_{qq}$ , it follows that  $\hat{V}_{11}$  is the square of the standard error of the estimated level-2 variance component  $\tau_{\pi_{00}}$ . Similarly,  $\hat{V}_{22}$  is that of  $\tau_{\pi_{10}}$ , and  $\hat{V}_{33}$  is that of  $\tau_{\pi_{11}}$ . The remaining three diagonal elements are the square roots of the standard errors of the level-3 variances and covariance, the  $\tau_{\beta}$  s. The standard errors of  $\tau_{\pi_{00}}$ ,  $\tau_{\pi_{10}}$ ,  $\tau_{\pi_{11}}$ ,  $\tau_{\beta_{00}}$ ,  $\tau_{\beta_{10}}$ , and  $\tau_{\beta_{11}}$  are thus 0.025153529, 0.00498999, 0.001974842, 0.036411537, 0.007197222, and 0.00250998 respectively.

From the almost zero values of all the covariances shown in the off-diagonal blocks, it can be concluded that the elements of the estimated  $\tau_{\pi}$  are uncorrelated with the elements of the estimated  $\tau_{\beta}$ .

Based on the information matrix, hypothesis tests and confidence intervals for the individual  $\tau$  s may be constructed. For example, we can compute the 95% confidence interval for the variance component  $\tau_{\pi_{11}}$  as shown below:

$$\begin{aligned} 95\% \text{ C.I. of } \tau_{\pi_{11}} &= \hat{\tau}_{\pi_{11}} \pm 1.96 \times \left( \hat{V}_{33} \right)^{\frac{1}{2}} \\ &= 0.0112249 \pm 1.96 \times \sqrt{0.0000039} \\ &= 0.0112 \pm 0.0039 \\ &= (0.0081, 0.0151). \end{aligned}$$

In addition, one can test whether significant random variation exists for random coefficient  $\tau_{qq}$  with a null hypothesis of the form

$$H_0 : \tau_{qq} = 0$$

against the alternative hypothesis

$$H_0 : \tau_{qq} \neq 0$$

Rejection of the null hypothesis will lead to the conclusion that there is random variation in  $\tau_{qq}$ . A test for this purpose is based on the standard error of  $\hat{\tau}_{qq}$ , as obtained from the information matrix shown above. The statistic

$$z = \hat{\tau}_{qq} / \left[ \text{Var} \left( \hat{\tau}_{qq} \right) \right]^{\frac{1}{2}}$$

is approximately normally distributed under the large sample theory of maximum likelihood estimates.

The assumption underlying both the confidence interval shown above and the single parameter hypothesis test depends on the quality of the normal approximation. For large samples and non-normal data, the use of the ratio shown above is preferred. If this test is used to test the significance

of a variance component, it is appropriate to use a one-sided test (Hox, 2002).

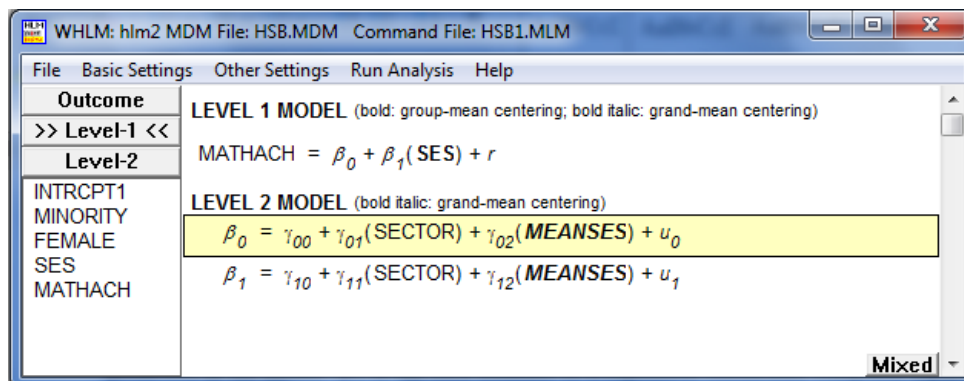
However, when the normality approximation is rather poor, as will be the case when  $\tau_{qq}$  is near zero and a small sample of groups, the distribution of this statistic is non-normal as the sampling distribution of the variances is skewed. As pointed out by Raudenbush & Bryk (2002): “In general, the smaller the variance component, the more data are needed to justify the large-sample normality approximation on which this test is based”. A possible rule of thumb for evaluating the size of  $\tau_{qq}$  is to avoid using this statistic if the contribution of  $\tau_{qq}$  to the total variation in the model is less than 1%. When use of the statistic discussed above is inadvisable, the chi-square test based on the residuals as used by Raudenbush & Bryk (2002), is the appropriate choice. For a two-level model, this  $\chi^2$  is defined as

$$\chi^2 = \sum (\hat{\beta}_j - \beta)^2 / \hat{V}_j$$

with degrees of freedom equal to  $J - p - 1$ , with  $J$  the number of level-2 units, and  $p$  the number of predictors in the model.  $\hat{\beta}_j$  is the OLS estimate of a regression coefficient, computed separately in group  $j$ , with its estimated sampling variance in group  $j$  denoted by  $V_j$ . This is the approach followed in the calculation of the  $\chi^2$  - statistics reported for individual variance components as reported in the HLM output.

### A two-level example:

In the case of a two-level model, the contents of tauvc.dat is dependent on the method of estimation selected. Consider the model



By default, HLM2 uses restricted maximum likelihood. Under restricted maximum likelihood, the **tauvc.dat** file contains the following:

```
2.3799608    0.1905837
0.1905837    0.1489167
36.7031309
```

In the case of full maximum likelihood estimation, however, the inverse of the information matrix is also given:

2.3164620	0.1908846	
0.1908846	0.0621249	
0.1260556	0.0077185	0.0008819
0.0077185	0.0382319	0.0046935
0.0008819	0.0046935	0.0429481
36.7226579		

For more on this topic, please see:

- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods*. Second Edition. Sage. pp. 61-64; 283-284.
- Hox, J. (2002). *Multilevel Analysis. Techniques and Applications*. IEA. pp.42-3.