

## Calculating posterior variances

The posterior variances of the Empirical Bayes estimates are given in the residual files produced by HLM2 and HLM3. These posterior variances are calculated according to equation (3.44) in *Hierarchical Linear Models* (Raudenbush & Bryk, Sage, 2002).

$$V_j^* = (V_j^{-1} + \tau_{00}^{-1})^{-1} + (1 - \lambda_j)^2 \text{Var}(\hat{\gamma}_{00})$$

where

- $V_j^{-1}$  is the error variance, defined as  $\sigma^2 / n_j$ ,
- $\tau_{00}$  is the parameter variance,
- $\lambda_j$  is the reliability of the estimator, defined as  $\lambda_j = \frac{\tau_{00}}{\tau_{00} + V_j}$ , and
- $\hat{\gamma}_{00}$  is the common estimate of  $\beta_{0j}$ .

Note that in the case of restricted maximum likelihood estimation, the formula above is used as shown. When full maximum likelihood estimation is used, the second term is equal to 0. In this case, the formula is equivalent to that given in Equation 4.17 in *Multilevel Analysis* (Snijders & Bosker, Sage, 1999).

Using Snijders & Bosker, p 59 we have

$$\lambda_j = \frac{\tau_{00}}{\tau_{00} + \sigma^2 / n_j}$$

We also have equation (4.15), the estimation of variance of the Empirical Bayes if the uncertainty due to the estimation of  $\gamma_{00}$  is neglected:

$$\text{var}\left(\hat{\beta}_{0j}^{EB} - \beta_{0j}\right) = (1 - \lambda_j) \tau_{00}.$$

When  $\lambda_j$  is substituted in (4.15), we get

$$\text{var}\left(\hat{\beta}_{0j}^{EB} - \beta_{0j}\right) = \left(\frac{\sigma^2 + n_j \tau_{00}}{\tau_{00} \sigma^2}\right)^{-1/2} = \left(\frac{\tau_{00} \sigma^2}{\sigma^2 + n_j \tau_{00}}\right)^{1/2}$$

Also (see (4.17))

$$\begin{aligned} S.E.\left(\hat{\beta}_{0j}^{EB}\right) &= \left(\tau_{00}^{-1} + n_j \sigma^{-2}\right)^{-1/2} \\ &= \left(\frac{\sigma^2 \tau_{00}}{\sigma^2 + n_j \tau_{00}}\right)^{1/2} \end{aligned}$$

In Raudenbush & Bryk,  $\lambda_j$  is defined as (see also p. 34, just after (3.6))

$$\begin{aligned} \lambda_j &= \frac{\tau_{00}}{\tau_{00} + V_j} \\ &= \frac{\tau_{00}}{\tau_{00} + \sigma^2 n_j} \end{aligned}$$

From (3.8) we have

$$\begin{aligned} V_j^* &= \left(V_j^{-1} + \tau_{00}^{-1}\right)^{-1} + (1 - \lambda_j)^2 \text{Var}\left(\hat{\gamma}_{00}\right) \\ &= \left(\frac{\sigma^2}{n_j + \tau_{00}^{-1}}\right)^{-1} + (1 - \lambda_j)^2 \text{Var}\left(\hat{\gamma}_{00}\right). \end{aligned}$$

After rewriting the first term of the equation, we have

$$\begin{aligned} V_j &= \left(\frac{\sigma^2}{n_j + \tau_{00}^{-1}}\right)^{-1} + \frac{n_j \tau_{00}^2}{(\sigma^2 + n_j \tau_{00}^{-1})^2} \text{Var}\left(\hat{\gamma}_{00}\right) \\ &= \left[\left(\frac{\sigma^2 \tau_{00}}{\sigma^2 + n_j \tau_{00}}\right)^{1/2}\right]^{-2} + \frac{n_j \tau_{00}^2}{(\sigma^2 + n_j \tau_{00}^{-1})^2} \text{Var}\left(\hat{\gamma}_{00}\right). \end{aligned}$$

The first term is now equal to the definition of  $S.E.\left(\hat{\beta}_{0j}^{EB}\right)$  given by Snijders & Bosker.

The difference between the two approaches (and thus the values obtained using the Snijders & Bosker approximation and the residual files in HLM) is in the second term of the formula, where the variation in  $\hat{\gamma}_{00}$  is taken into account. When REML is used in HLM, the complete equation should be used, while for

full ML estimation only the first term is applicable, making results of calculation under ML in HLM equivalent to those obtained using the Snijders & Bosker formula where they assume the use of ML.