Journal of Educational and Behavioral Statistics
Summer 2004, Vol. 29, No. 2, pp. 201–218

Asymptotic Effect of Misspecification in the Random Part of the Multilevel Model

Johannes Berkhof
VU University Medical Center

Jarl Kennard Kampen
Katholieke Universiteit Leuven

The authors examine the asymptotic effect of omitting a random coefficient in the multilevel model and derive expressions for the change in (a) the variance components estimator and (b) the estimated variance of the fixed effects estimator. They apply the method of moments, which yields a closed form expression for the omission effect. In practice, the model parameters are estimated by maximum likelihood; however, since the moment estimator and the maximum likelihood estimator are both consistent, the presented expression for the change in the variance components estimator asymptotically holds for the maximum likelihood estimator as well. The results are illustrated with an analysis of mathematics performance data.

Keywords: *between-unit variance proportion, misspecification, moment estimation, multilevel model, random coefficient*

An important aspect of the specification of the multilevel model (Bryk & Raudenbush, 1992; Goldstein, 1995; Longford, 1993) concerns selection of the random coefficients. Since it is not always obvious which random coefficients to select, the final model may contain misspecification in the random part. This affects the inferences that can be made about fixed effects, since the estimated standard errors of the fixed effects depend on the covariance structure of the model. It is important to have insight into the sensitivity of the fixed effects with respect to changes in the specification of the random part. This insight may act as a guide when one is modeling data, and it is also helpful when one is examining the results of a multilevel analysis performed by others. In the latter case, one usually does not have (quick) access to the original data and, for that reason, cannot instantly check the sensitivity of the results by refitting the model under different covariance structures. A good insight, however, still enables one to make a statement about the sensitivity of the reported results.

This article is based on research reported in the first author's PhD dissertation supervised by Tom Snijders and financially supported by Grant ESR 510-78-501 from the Dutch Organization for Scientific Research. Part of the research was done while the first author was employed at the Katholieke Universiteit Leuven and the second author was employed at the Katholieke Universiteit Brussel. We are grateful to Marie-Christine Opdenakker and Jan Van Damme for the data and to Anne Boomsma for helpful comments. This article was accepted under the editorship of Larry Hedges.

In this article, we study the asymptotic effect of omitting a random coefficient on the variance component estimator. We apply a moment estimation procedure because the moment estimator can be written in closed form. In addition, the moment estimator is consistent, so the change in the variance component estimator asymptotically holds for the maximum likelihood (ML) estimator as well. We also derive an expression for the asymptotic change in the estimated variance of the fixed effects estimator obtained when a random coefficient is omitted. This is important because this estimated variance is used in the *t* test for a fixed coefficient.

The literature on misspecification of the random part of the multilevel model usually focuses on the change in the ML estimator. Closed form expressions are then difficult to obtain but can still be derived when the data are balanced (i.e., equal Level 1 design matrices for higher level units) or when the model contains only a random intercept at Level 2 (Berkhof & Snijders, 2001). In the case of the balanced two-level model, Lange and Laird (1989) presented expressions for the effect of omitting a random coefficient on the unrestricted and restricted ML estimator of the variance components. Further results for the two-level random intercept model have been presented by Bryk and Raudenbush (1992, p. 92) and Longford (1993, p. 53), and results for the three-level random intercept model have been presented by Hutchison and Healy (2001) and Moerbeek (2004). The purpose of this article is to present a general method that covers a wide range of multilevel models. Although the general applicability of the method is appealing, the usefulness of the approach largely depends on whether the derived formulas are easy to interpret. For this purpose, we introduce the between-unit variance proportion, which is a generalization of the intraclass correlation coefficient (Bryk & Raudenbush, 1992, p. 18; Goldstein, 1995, p. 19; Snijders & Bosker, 1999, p. 16); the two measures coincide under the two-level random intercept model. After the between-unit variance proportion has been defined, the omission effects can be written in an attractive, simple form.

The article is organized as follows. In the next three sections, we describe the model, derive moment estimators for the model parameters, and present the between-unit variance proportion. In the subsequent sections, we study the effect of omitting a random coefficient in a two-level model and in a model with more than two levels. The final section contains some concluding remarks.

The Multilevel Model

Multilevel models for normally distributed response data can be formulated as special cases of the general linear mixed model (Harville, 1977):

$$\boldsymbol{y} = \boldsymbol{Z}\gamma + \boldsymbol{X}\boldsymbol{u} + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{y}$ is an $N \times 1$ vector of responses, $\boldsymbol{Z}$ is an $N \times c_Z$ design matrix for the fixed effects γ, $\boldsymbol{X}$ is an $N \times c_X$ design matrix for the random effects $\boldsymbol{u}$, and $\boldsymbol{\epsilon}$ is a vector of Level 1 disturbances. The distributional assumptions are $\boldsymbol{u} \sim N_{cX}[\boldsymbol{0}, \boldsymbol{T}(\boldsymbol{\theta})]$ and

$\epsilon \sim N_N(\mathbf{0}, \sigma^2 \boldsymbol{I}_N)$. The matrix $\boldsymbol{T}(\boldsymbol{\theta})$ is linear in $\boldsymbol{\theta}$, the elements of which are variance and covariance parameters. We may decompose $\boldsymbol{X}$ as $\boldsymbol{X} = (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_Q)$ where each $N \times c_{Xq}$ matrix $\boldsymbol{X}_q$ corresponds to a different random coefficient. If we substitute the $\boldsymbol{X}_q$s in Model 1, we obtain

$$y = Z\gamma + \sum_{q=1}^{Q} X_q u_q + \epsilon. \tag{2}$$

The variance of the $c_{Xq} \times 1$ random coefficient $\boldsymbol{u}_q$ is $\text{var}(\boldsymbol{u}_q) = \theta_{qq} \boldsymbol{I}_{cX_q}$ $(q = 1, \ldots, Q)$; the covariance between $\boldsymbol{u}_q$ and $\boldsymbol{u}_s$ is $\text{cov}(\boldsymbol{u}_q, \boldsymbol{u}_s) = \theta_{qs} \boldsymbol{I}_{cX_q}$ $(q \neq s)$ for some $\boldsymbol{X}_q$ and $\boldsymbol{X}_s$ that have the same number of columns and a matrix of zeros otherwise.

Estimation

We examine the effect of omitting a random coefficient on the estimators of σ^2 and θ. This effect is, asymptotically, a model misspecification effect that does not depend on the estimation method but is the same for all consistent estimators. We choose consistent moment estimators since they can be written in closed form. The moment estimators are constructed as follows. We define $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{Z}\gamma$ and estimate σ^2 and θ by setting $\text{tr}(\boldsymbol{rr}')$ and $\text{tr}(\boldsymbol{X}_q'\boldsymbol{rr}'\boldsymbol{X}_s)$ equal to their expected values, yielding $\hat{\sigma}^2$ and $\hat{\theta}$. In practice, we do not observe $\boldsymbol{r}$, since γ is unknown. However, if we replace γ with a consistent estimator, then the consistency of $\hat{\sigma}^2$ and $\hat{\theta}$ is preserved.

We also examine the effect of omitting a random coefficient on the estimated variance of the estimator of γ. Here the choice of the fixed effects estimator matters, also asymptotically. To illustrate this, let us denote the fixed effects estimator in the full (correct) model as $\hat{\gamma}$ and the estimator in the constrained (misspecified) model as $\hat{\gamma}_c$. The estimators $\hat{\gamma}_c$ and $\hat{\gamma}$ may be different if they depend on the estimated variance components, in which case the change in the variance of the fixed effects estimator consists of two components: a misspecification and an efficiency component. The two components sum up to the total change:

$$\begin{aligned} \widehat{\text{var}}_c(\hat{\gamma}_c) - \widehat{\text{var}}(\hat{\gamma}) &= [\widehat{\text{var}}_c(\hat{\gamma}_c) - \widehat{\text{var}}(\hat{\gamma}_c)] + [\widehat{\text{var}}(\hat{\gamma}_c) - \widehat{\text{var}}(\hat{\gamma})] \\ &= \text{misspecification effect} + \text{efficiency loss}. \end{aligned} \tag{3}$$

The efficiency loss arises when, under the correct model, the variance of the estimator $\hat{\gamma}_c$ is larger than the variance of $\hat{\gamma}$ (Longford, 1993, p. 56). This loss does not have to vanish if the sample size goes to infinity.

Efficiency loss is of own interest, but, for now, we focus only on the omission effect on $\widehat{\text{var}}(\hat{\gamma})$ that is not the result of a change in the definition of $\hat{\gamma}$. Regarding the t test for γ, efficiency loss is reflected by a change in the sampling distribution of the p value under prespecified alternatives. Misspecification affects the sampling distribution of the p value under the null hypothesis, which means that the test becomes either liberal or conservative. When studying the misspecification effect on $\hat{\text{var}}(\hat{\gamma})$, we need an estimator of γ for which the efficiency loss equals zero. An estimator for which this holds is the ordinary least squares (OLS) estimator

$$\hat{\gamma}_{\text{ols}} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'y, \tag{4}$$

which is used throughout this article.

Between-Unit Variance Proportion

To obtain an expression for the misspecification effect that is easy to interpret, we introduce a measure that we will label between-unit variance proportion. We derive this measure from the intraclass correlation coefficient (Bryk & Raudenbush, 1992, p. 18; Goldstein, 1995, p. 19; Snijders & Bosker, 1999, p. 16), which is defined as the proportion of the total variance in a two-level random intercept model that is accounted for by Level 2. The intraclass correlation coefficient is a correlation in that the Level 2 variance is equal to the covariance between two responses from the same Level 2 unit.

If we consider the intraclass correlation coefficient as a variance proportion rather than as a correlation, it is straightforward to formulate a variant of the intraclass correlation coefficient for multilevel models beyond the two-level random intercept model. Consider a simple version of Model 1 where $\text{var}(\boldsymbol{u}) = \theta \boldsymbol{I}_J$ and $\text{var}(\epsilon) = \sigma^2 \boldsymbol{I}_N$. The total variance of the data can be expressed as

$$\text{tr}\,[\text{var}(\boldsymbol{y})] = \text{tr}\,(\boldsymbol{X}'\boldsymbol{X})\theta + N\sigma^2. \tag{5}$$

The term $\text{tr}\,(\boldsymbol{X}'\boldsymbol{X})\theta$ is the part of the total variance modeled by predictor $\boldsymbol{X}$. Expressed as a proportion, we have the between-unit variance proportion

$$\phi = \frac{\text{tr}\,(\boldsymbol{X}'\boldsymbol{X})\theta}{\text{tr}\,(\boldsymbol{X}'\boldsymbol{X})\theta + N\sigma^2}. \tag{6}$$

Unlike the intraclass correlation coefficient, the between-unit variance proportion ϕ is not always a correlation.

The parameters θ and σ^2 in Equation 6 are unknown and replaced by their moment estimators, defined in the previous section. We obtain the following expression for the estimated between-unit variance proportion:

$$\hat{f}_{rx} = \left[\frac{\text{tr}\,(\boldsymbol{X}'\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}'\boldsymbol{X})\,N}{\text{tr}\,(\boldsymbol{X}'\boldsymbol{X})\,\text{tr}\,(\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}')} - 1\right] \Big/ \left[\frac{\text{tr}\,(\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}'\boldsymbol{X})N}{\text{tr}\,(\boldsymbol{X}'\boldsymbol{X})^2} - 1\right], \tag{7}$$

where $\hat{r} = y - Z\hat{\gamma}_{\text{ols}}$. The definition of $\hat{f}_{rx}$ can be easily extended to other variables. Suppose we have an $N \times 1$ vector w and an $N \times J$ matrix X. Then we define the between-unit variance proportion as

$$f_{wx} = \left[\frac{\text{tr}(X'ww'X)N}{\text{tr}(X'X)\,\text{tr}(ww')} - 1\right] \Big/ \left[\frac{\text{tr}(X'XX'X)N}{\text{tr}(X'X)^2} - 1\right], \tag{8}$$

If we assume that w is stochastic with mean zero, the between-unit variance proportion f_{wx} may be interpreted as the estimated proportion of the variance of w that is accounted for by predictor X.

Misspecification in Two-Level Models

Omitting a Random Intercept at Level 2

The two-level random intercept model with one fixed effect can be written as

$$y = z\gamma + Xu + \epsilon, \tag{9}$$

where X is a block-diagonal matrix. The jth block of X, corresponding to the jth Level 2 unit ($j = 1, \ldots, J$), consists of an $n \times 1$ vector of ones. We study the asymptotic misspecification effect obtained when n and J tend to infinity. We assume that z is bounded and that $\lim_{N\to\infty} z'z/N$ is equal to positive σ_{zz}. After discarding the random intercept from the model, we retain the single-level model:

$$y = z\gamma + \epsilon_c, \tag{10}$$

The Level 1 variance of the single-level model is denoted by σ_c^2. In what follows, we are interested in the difference between the estimators for σ^2 and σ_c^2. To estimate the parameters in the random part of the random intercept model, we assume that γ is known and equate $\text{tr}(X'rr'X)$ and $\text{tr}(rr')$ to their expected values (where $r = y - z\gamma$). In a second step, we replace r by $\hat{r} = y - Z\hat{\gamma}_{\text{ols}}$ (the bias as a consequence of this substitution tends to zero if z is bounded and $z'z \to \infty$), which gives the following equations:

$$\text{tr}(X'\hat{r}\hat{r}'X) = nN\hat{\theta} + N\hat{\sigma}^2 \tag{11}$$

and

$$\text{tr}(\hat{r}\hat{r}') = N\hat{\theta} + N\hat{\sigma}^2. \tag{12}$$

The parameter in the random part of the single-level model (Equation 10) can be estimated by equating $\text{tr}(rr')$ to its expected value and replacing r by $\hat{r}$, yielding

$$\text{tr}(\hat{r}\hat{r}') = N\hat{\sigma}_c^2. \tag{13}$$

Because we are interested in the differences between the moment estimators $\hat{\sigma}_c^2$ and $\hat{\sigma}^2$, we subtract Equation 12 from Equation 13 and obtain

$$\begin{aligned} \hat{\sigma}_c^2 - \sigma^2 &= \hat{\theta} \\ &= \theta + op(1), \end{aligned} \tag{14}$$

where $op(1)$ is a random variable that converges to zero in probability. Equation 14 holds for any consistent estimator. The expression for the estimated variance of the OLS estimator $\hat{\gamma}_{\text{ols}}$ is in the random intercept model

$$\begin{aligned} \text{var}\left(\hat{\gamma}_{\text{ols}}\right) &= \frac{\hat{\sigma}^2}{z'z} + \frac{\text{tr}(X'zz'X)}{(z'z)^2}\hat{\theta} \\ &= \frac{1}{z'z}\left(\hat{\sigma}^2 + \hat{\theta}\right) + \frac{(n-1)f_{zx}}{z'z}\hat{\theta}, \end{aligned} \tag{15}$$

where f_{zx} is defined according to Equation 8. In the single-level model, the estimated variance of $\hat{\gamma}_{\text{ols}}$ equals

$$\hat{\text{var}}_c(\gamma_{\text{ols}}) = \frac{\hat{\sigma}_c^2}{z'z}. \tag{16}$$

Subtracting $\hat{\text{var}}_c(\hat{\gamma}_{\text{ols}})$ from $\hat{\text{var}}(\hat{\gamma}_{\text{ols}})$ and dividing the difference by $\hat{\text{var}}(\hat{\gamma}_{\text{ols}})$ yields the relative change, the limiting value of which is

$$\begin{aligned} ARC_{\text{ols}} &= \operatorname*{plim}_{n,J\to\infty} \frac{\hat{\text{var}}_c(\hat{\gamma}_{\text{ols}}) - \hat{\text{var}}(\hat{\gamma}_{\text{ols}})}{\hat{\text{var}}(\hat{\gamma}_{\text{ols}})} \\ &= \operatorname*{plim}_{n,J\to\infty} \frac{-f_{zx}\hat{\theta}}{\frac{1}{n-1}\hat{\sigma}^2 + \left(\frac{1}{n-1} + f_{zx}\right)\hat{\theta}}. \end{aligned} \tag{17}$$

The sign of ARC_{ols} is opposite to the sign of f_{zx}. Because it is assumed that the random intercept model is the correct model and the single-level model is the incorrect model, a negative change corresponds to liberal t testing and a positive change corresponds to conservative t testing.

To gain insight into the relation between the asymptotic relative change and z, we consider three cases.

1. z is a group variable. Then all variation in z is between-group variation so that f_{zx} equals 1 and $ARC_{\text{ols}} = -1$, in which case the t test for the fixed effect becomes liberal.
2. z is a group-centered variable. Then $f_{zx} = -1/(n-1)$ and $ARC_{\text{ols}} = \theta/\sigma^2$, in which case the t test for the fixed effect becomes conservative.

3. z is a stochastic variable for which $\text{plim}_{n,J\to\infty}\,\text{tr}(X'zz'X)N/[\text{tr}(X'X)(z'z)] = 1$. Then plim $n\,f_{zx} = 0$ and $ARC_{\text{ols}} = 0$. An example of this type of predictor is a stochastic variable z with zero variance at Level 2 (i.e., no correlation within groups).

These cases show that if the random intercept is omitted, the test for the fixed effect is asymptotically liberal if the predictor with a fixed effect is a group variable and asymptotically conservative if the predictor is a group-centered variable. The omission of a random intercept does not asymptotically affect the test results (when γ is estimated via OLS) if the predictor is a Level 1 variable with zero between-group variance.

Example

We compared the change in the ML estimate of the Level 1 variance with an approximation of the asymptotic change obtained when substituting ML estimates in Equation 14. We used a data set containing mathematics achievement scores of 3,632 students from Belgium (Opdenakker & Van Damme, 2000). The students were nested within 279 classes in 57 secondary schools. The data were not balanced; observations per class varied from 1 to 26, and observations per school varied from 1 to 266. The response variable was the mathematics achievement score (MATH) at the end of the first grade. Administrated variables with a fixed effect were student gender (SEX; boy = −1; girl = 1); standardized motivation score (PM), standardized educational level of father (ELFA); standardized intake mathematics achievement score (MATH0); a class-centered IQ score (IQ1) obtained by centering the standardized IQ score (SIQ) around the class mean; a school-centered IQ score (IQ2) obtained by centering the class mean of SIQ around the school mean of SIQ; and an aggregate IQ score (IQ3) obtained by centering the school mean of IQ around the overall mean.

We specified two different random intercept models by including a random intercept at the class level (Model 2) and at the school level (Model 3). These models, as well as the single-level model (Model 1), were estimated with MLwiN (Goldstein et al., 1998). We first assumed that the correct model is a two-level model with a random intercept at the class level and that the random intercept is omitted. The change in the ML estimate of the Level 1 variance is then 14.91 − 12.68 = 2.23 (Table 1). If the ML estimates are substituted in Equation 14, we obtain a difference of 2.30. We see that Equation 14 and the ML method yielded similar values for the change in the estimates of the Level 1 variance. We then assumed that the correct model is a two-level model with a random intercept at the school level, in which case the change in the ML estimate of the Level 1 variance is 14.91 − 13.65 = 1.26. If we use Equation 14, we obtain a change of 1.83. In the latter comparison (Model 3 vs. Model 1), the computed change is larger than in the former one (Model 2 vs. Model 1). A possible explanation for this is that the number of clusters J is equal to 279 in Model 2 and equal to 57 in Model 3. The precision of the Level 2 variance estimator increases with the number of clusters J (Miller, 1977), and therefore the estimator of the intercept variance tends to be closer to the asymptotic value in Model 2 than in Model 3.

TABLE 1
Mathematics Performance Data: Maximum Likelihood Estimates

Fixed effects	Model 1	Model 2	Model 3	Model 4	Model 5
CONS	18.05	17.98	17.99	17.92	17.97
	(0.064)	(0.114)	(0.227)	(0.225)	(0.202)
SEX	0.21	0.07	0.05	0.04	0.02
	(0.066)	(0.076)	(0.082)	(0.082)	(0.081)
PM	0.21	0.16	0.17	0.16	0.16
	(0.065)	(0.063)	(0.064)	(0.063)	(0.063)
ELFA	0.01	0.04	0.06	0.06	0.05
	(0.067)	(0.065)	(0.066)	(0.066)	(0.065)
MATH0	2.62	2.56	2.58	2.55	2.55
	(0.091)	(0.090)	(0.090)	(0.122)	(0.090)
IQ1	1.05	1.06	1.04	1.03	1.06
	(0.095)	(0.089)	(0.092)	(0.092)	(0.089)
IQ2	2.62	2.45	2.64	2.60	2.53
	(0.167)	(0.256)	(0.161)	(0.162)	(0.218)
IQ3	2.11	2.17	2.00	2.10	2.09
	(0.190)	(0.273)	(0.395)	(0.402)	(0.376)
Variance components					
School level					
Intercept variance	0	0	1.83	1.77	1.20
Slope variance	0	0	0	0.24	0
Class level					
Intercept variance	0	2.30	0	0	1.24
Student level					
Residual variance	14.91	12.68	13.65	13.47	12.69

Note. Standard errors are shown in parentheses.

In addition to the change in the estimated Level 1 variance, we compared the change in the estimated variance of $\hat{\gamma}_{ml}$ with the asymptotic misspecification effect. We computed the relative changes from the estimates shown in Table 1 and multiplied by 100% to obtain the percentage changes in $\hat{\text{var}}(\hat{\gamma}_{ml})$, denoted by PC_{ml}. The results are presented in Table 2. It can be seen that, for the between-cluster variables, the percentage changes are negative. The values, however, substantially deviate from the asymptotic percentage change APC_{ols} (= $ARC_{ols} \times 100\%$), which is −100% for between-cluster variables. The value of PC_{ml} for the group-centered variable IQ1 in Model 2 (14%) is comparable to the approximation of APC_{ols} obtained when substituting the ML estimate in Equation 17, which is $\theta/\sigma^2 = 2.3/12.68 = 18\%$. The value of PC_{ml} for the group-centered variable IQ2 in Model 3 (8%) is also comparable to the approximation of APC_{ols}, which is $1.83/13.65 = 13\%$. The tabulated percentage changes in the other covariates were close to zero and slightly higher than corresponding approximations of APC_{ols} (data not shown).

TABLE 2
Two-Level Random Intercept Model Versus Single-Level Model: Between-Unit Variance Proportions and Percentage Changes in the Estimated Variances of the Fixed Effects Estimator

	Model 2 vs. Model 1		Model 3 vs. Model 1	
	f_{zx}	PC_{ml}	f_{zx}	PC_{ml}
CONS	1.00	−68	1.00	−92
SEX	.50	−25	.53	−35
PM	.06	6	.03	3
ELFA	.12	6	.07	3
MATH0	.32	2	.18	2
IQ1	−.07	14	−.01	7
IQ2	1.00	−57	−.01	8
IQ3	1.00	−52	1.00	−77

These differences are possibly related to the efficiency loss involved in estimating under a misspecified model.

Omitting a Random Slope at Level 2

We assume that the population model is a two-level model with a random intercept and one random slope and that the misspecified model is a random intercept model. The model can be written as

$$y = z\gamma + X_1 u_1 + X_2 u_2 + \epsilon, \tag{18}$$

where $X_1 = \oplus_j\{x_{1j}\}$ and $X_2 = \oplus_j\{x_{2j}\}$ are block-diagonal design matrices for the random intercept u_1 and the random slope u_2. The Level 1 variance, the Level 2 intercept variance, and the Level 2 slope variance are denoted by σ^2, θ_{11}, and θ_{22}. The Level 2 intercept/slope covariance is denoted by θ_{12}. As before, we assume that $n_1 = \cdots = n_J = n$, that z is bounded, and that $\lim_{N\to\infty} z'z/N$ is positive and equal to σ_{zz}. We further assume that $x_2 = (x'_{21}, \ldots, x'_{2J})'$ is bounded and that $\lim_{n\to\infty} x'_{2j}x_{2j}/n = \sigma_{x_2x_2}$ and $\lim_{n\to\infty} x'_{2j}x_{1j}/n = \sigma_{x_1x_2}$, where $\sigma_{x_2x_2}$ is positive. When u_2 is omitted from Model 18, we obtain the constrained model

$$y = z\gamma + X_1 u_{1c} + \epsilon_c, \tag{19}$$

with variance parameters σ_c^2 and θ_{11c}.

We estimate the parameters in the random part of Model 18 by equating $\mathrm{tr}(rr')$, $\mathrm{tr}(X'_1 rr' X_1)$, $\mathrm{tr}(X'_2 rr' X_2)$, and $\mathrm{tr}(X'_1 rr' X_2)$ to their expected values, where $r = y - Z\gamma$. Then we replace $\hat{r}$ with $\hat{r} = y - Z\hat{\gamma}_{ols}$. Expressions for $(\hat{\theta}_{11c} - \hat{\theta}_{11})$ and $(\hat{\sigma}_c^2 - \hat{\sigma}^2)$ as functions of the estimators of θ_{22} and θ_{12} are derived by subtracting the equalities for the two models in which $\mathrm{tr}(\hat{r}\hat{r}')$ and $\mathrm{tr}(X'_1\hat{r}\hat{r}'X_1)$ appear. We obtain

$$\hat{\theta}_{11c} - \hat{\theta}_{11} = f_{x_2x_1}\frac{\mathrm{tr}(X_2'X_2)}{N}\hat{\theta}_{22} + 2\frac{\mathrm{tr}(X_1'X_2)}{N}\hat{\theta}_{12}$$
$$= f_{x_2x_1}\sigma_{x_2x_2}\theta_{22} + 2\sigma_{x_1x_2}\theta_{12} + o_P(1) \tag{20}$$

and

$$\hat{\sigma}_c^2 - \hat{\sigma}^2 = (1 - f_{x_2x_1})\frac{\mathrm{tr}(X_2'X_2)}{N}\hat{\theta}_{22} + 2\left[1 - \frac{\mathrm{tr}(X_1'X_2)}{N}\right]\hat{\theta}_{12}$$
$$= (1 - f_{x_2x_1})\sigma_{x_2x_2}\theta_{22} + 2(1 - \sigma_{x_1x_2})\theta_{12} + o_P(1). \tag{21}$$

The between-unit variance proportion $f_{x_2x_1}$ is of the type illustrated in Equation 8, where the vector $x_2 = (x'_{21}, \ldots, x'_{2J})'$ is substituted for $\boldsymbol{w}$.

From here onward, we will assume that θ_{12} is equal to zero. Under this assumption, the change in the estimated intercept variance is proportional to $f_{x_2x_1}$. This means, for instance, that if x_2 is a group-centered variable (in which case $f_{x_2x_1} < 0$), the change in the estimated intercept variance is negative, and if x_2 is a group variable (in which case $f_{x_2x_1} = 1$), the change in the estimated intercept variance is positive. The change in the estimated Level 1 variance is proportional to $(1 - f_{x_2x_1})$, and thus it is equal to zero if x_2 is a Level 2 variable and positive otherwise.

The variance of $\hat{\gamma}_{\mathrm{ols}}$ is in the random intercept model estimated by

$$\hat{\mathrm{var}}_c(\hat{\gamma}_{\mathrm{ols}}) = \frac{\hat{\sigma}_c^2}{z'z} + \frac{\mathrm{tr}(X_1'zz'X_1)}{(z'z)^2}\hat{\theta}_{11c} \tag{22}$$

and in the random slope model estimated by

$$\hat{\mathrm{var}}(\hat{\gamma}_{\mathrm{ols}}) = \frac{\hat{\sigma}^2}{z'z} + \frac{\hat{\theta}_{11}\,\mathrm{tr}(X_1'zz'X_1)}{(z'z)^2} + \frac{\hat{\theta}_{22}\,\mathrm{tr}(X_2'zz'X_2)}{(z'z)^2}. \tag{23}$$

If we combine Equations 20, 21, 22, and 23, we can derive the asymptotic relative change in the estimated variance of the OLS estimator of γ:

$$ARC_{\mathrm{ols}} = \operatorname*{plim}_{n,J\to\infty}\frac{\hat{\mathrm{var}}_c(\hat{\gamma}_{\mathrm{ols}}) - \hat{\mathrm{var}}(\hat{\gamma}_{\mathrm{ols}})}{\hat{\mathrm{var}}(\hat{\gamma}_{\mathrm{ols}})}$$
$$= \operatorname*{plim}_{n,J\to\infty}\frac{(f_{zx_1}f_{x_2x_1} - f_{zx_2})\sigma_{x_2x_2}\hat{\theta}_{22}}{\frac{1}{n-1}\hat{\sigma}^2 + \left(\frac{1}{n-1} + f_{zx_1}\right)\hat{\theta}_{11} + \left(\frac{1}{n-1} + f_{zx_2}\right)\sigma_{x_2x_2}\hat{\theta}_{22}}. \tag{24}$$

We evaluate Equation 24 for two cases.

1. x_2 is a stochastic predictor with zero variance at Level 2. Then plim $n f_{x_2 x_l} = 0$, and the sign of ARC_{ols} is opposite to the sign of f_{zx_2}. If z is a Level 2 variable, plim $n f_{zx_2} = 0$, and $ARC_{ols} = 0$. If z is a cross-level interaction variable (i.e., of the type $z = x_2 \times w$, where w is a Level 2 variable), $f_{zx_2} = 1$, and $ARC_{ols} = -1$.
2. x_2 is a Level 2 variable. Then $f_{x_2 x_1} = 1$, and the sign of ARC_{ols} is equal to the sign of $f_{zx_1} - f_{zx_2}$. If z is a stochastic predictor with zero variance at Level 2, then plim $n f_{zx_1} = 0$, plim $n f_{zx_2} = 0$, and $ARC_{ols} = 0$.

Example: Continued

We compared the changes in the ML estimates of the variance components with the changes obtained when substituting the ML estimates in Equations 20 and 21. We assumed that the population model is a two-level model with a random intercept and a random slope of MATH0, both at the school level (Model 4). The Level 2 intercept/slope correlation was set equal to zero. The ML estimates of the Level 2 slope variance, Level 2 intercept variance, and Level 1 variance are 0.24, 1.77, and 13.47, respectively (Table 1). When the random slope is omitted, the changes in the ML estimates of the Level 2 intercept variance and the Level 1 variance are .06 and .18. If the changes are computed by substituting ML estimates in Equations 20 and 21 and a value of .18 for the between-unit variance proportion, values of .04 and .20 are obtained. We see that the changes in the ML estimates are comparable to the changes obtained when applying Equations 20 and 21. We also computed the changes in the estimated variances of the ML estimator $\hat{\gamma}_{ml}$. The results are presented in Table 3. The largest negative change of 46% was found for the predictor MATH0, which had an APC_{ols} value of −100%. The other changes in the estimated variances were all close to zero, and the corresponding APC_{ols} values varied between −18% and 0% where n was replaced by the average number of students per school in Equation 24.

TABLE 3
Two-Level Random Slope Model Versus Two-Level Random Intercept Model: Between-Unit Variance Proportions and Percentage Changes in the Estimated Variances of the Fixed Effects Estimator

	Model 4 vs. Model 3		
	f_{zx_1}	f_{zx_2}	PC_{ml}
CONS	1.00	.18	2
SEX	.53	.14	0
PM	.03	.01	3
ELFA	.07	.06	0
MATH0	.18	1.00	−46
IQ1	−.01	.21	0
IQ2	−.01	.17	−1
IQ3	1.00	.45	−3

Misspecification in Random Intercept Models With Q Levels

The Q-level model ($Q > 2$) can be written as

$$y = z\gamma + \sum_{q=1}^{Q} X_q u_q, \tag{25}$$

where X_q is the design matrix for the random intercept at level q ($q = 1, \ldots, Q$). The Level 1 vector of disturbance terms is represented by u_1, and therefore X_1 is an identity matrix. In the models discussed so far, the vector of Level 1 disturbance terms was denoted by ϵ, but this notation is not useful for a Q-level model ($Q > 2$). We denote the variance at level q by θ_q and the number of level $q - 1$ units within each level q unit by n_{q-1}, and we set n_0 equal to 1. Note that n_{q-1} does not carry an index for the unit number, which means that the group sizes of the level q units are assumed to be equal.

The constrained model is obtained by omitting the random intercept at level *m:*

$$y = z\gamma + \sum_{\substack{q=1 \\ q \neq m}}^{Q} X_{qc} u_{qc}. \tag{26}$$

The level q variance in the constrained model is denoted by θ_{qc}.

Omitting the Random Intercept at the Highest Level

When we omit the random intercept at level Q, the change in the moment estimator for θ_{Q-1} is

$$\begin{aligned} \hat{\theta}_{Q-1,c} - \hat{\theta}_{Q-1} &= \hat{\theta}_Q \\ &= \theta_Q + o_P(1). \end{aligned} \tag{27}$$

The moment estimators for the variances at Level 1 to $Q - 2$ do not change. Thus, the level Q variance is fully absorbed by level $Q - 1$. This may give rise to the following interpretation for the intercept variance defined at the highest level of the data hierarchy: It describes not only the variation at this level but also the variation at higher levels that are not included in the model.

We introduce $N_q = \Pi_{i=1}^{q} n_i$ and estimate the variance of $\hat{\gamma}_{\text{ols}}$ in the constrained model by

$$\begin{aligned} \hat{\text{var}}_c(\hat{\gamma}_{\text{ols}}) &= \sum_{q=1}^{Q-1} \frac{\text{tr}(X_q' z z' X_q)}{(z'z)^2} \hat{\theta}_{qc} \\ &= \frac{1}{z'z} \sum_{q=1}^{Q} \hat{\theta}_q + \frac{1}{z'z} \sum_{q=2}^{Q-1} f_{zx_q}(N_{q-1} - 1)\hat{\theta}_q + \frac{f_{zx_{Q-1}}(N_{Q-2} - 1)}{z'z} \hat{\theta}_Q \end{aligned} \tag{28}$$

and in the full model by

$$\hat{\text{var}}(\hat{\gamma}_{\text{ols}}) = \sum_{q=1}^{Q} \frac{\text{tr}(X_q'zz'X_q)}{(z'z)^2}\hat{\theta}_q$$
$$= \frac{1}{z'z}\sum_{q=1}^{Q}\hat{\theta}_q + \frac{1}{z'z}\sum_{q=2}^{Q} f_{zx_q}(N_{q-1}-1)\hat{\theta}_q. \tag{29}$$

The asymptotic relative change in the estimated variance of $\hat{\gamma}_{\text{ols}}$ is

$$ARC_{\text{ols}} = \operatorname*{plim}_{n_1,\ldots,n_Q\to\infty} \frac{\hat{\text{var}}_c(\hat{\gamma}_{\text{ols}}) - \hat{\text{var}}(\hat{\gamma}_{\text{ols}})}{\hat{\text{var}}(\hat{\gamma}_{\text{ols}})}$$
$$= \operatorname*{plim}_{n_1,\ldots,n_Q\to\infty} \frac{[f_{zx_{Q-1}}(N_{Q-2}-1) - f_{zx_Q}(N_{Q-1}-1)]\hat{\theta}_Q}{\sum_{q=1}^{Q}\hat{\theta}_q + \sum_{q=2}^{Q} f_{zx_q}(N_{q-1}-1)\hat{\theta}_q}. \tag{30}$$

We evaluate ARC_{ols} for five cases. For the sake of simplicity, we confine ourselves to a three-level model (i.e., $Q = 3$).

1. z is a Level 1 variable, and the entries are centered around the Level 2 means. Then $f_{zx_2} = -(n_1 - 1)^{-1}, f_{zx_3} = -(N_2 - 1)^{-1}$, and $ARC_{\text{ols}} = 0$.
2. z is a stochastic Level 1 variable with zero variance at Level 2. Then plim $n_1 f_{zx_2} = 0$, plim $N_2 f_{zx_3} = 0$, and $ARC_{\text{ols}} = 0$.
3. z is a Level 2 variable the entries of which are centered around the Level 3 means. Then $f_{zx_2} = 1, f_{zx_3} = -(N_2 - 1)^{-1}$, and $ARC_{\text{ols}} = \theta_3/\theta_2$, in which case the t test becomes conservative.
4. z is a stochastic Level 2 variable with zero variance at Level 3. Then $f_{zx_2} = 1$, plim $n_2 f_{zx_3} = 1$, and $ARC_{\text{ols}} = 0$.
5. z is a Level 3 variable. Then $f_{zx_2} = 1, f_{zx_3} = 1$, and $ARC_{\text{ols}} = -1$, in which case the t test becomes liberal.

These cases show that the estimated variance of $\hat{\gamma}_{\text{ols}}$, asymptotically, is not affected by the omission of Level 3 if z is a Level 1 variable centered around the Level 2 group means or a Level 1 or Level 2 variable with zero variance at Level 3. Furthermore, the omission of Level 3 leads to a liberal test result (when testing H_0: $\gamma = 0$) if z is a Level 3 variable and to a conservative test result if z is a Level 2 variable with entries centered around the Level 3 group means.

Omitting the Random Intercept at an Intermediate Level

When we omit the random intercept at level m ($1 < m < Q$), the changes in the moment estimators for θ_{m-1} and θ_{m+1} are

$$\hat{\theta}_{m+1,c} - \hat{\theta}_{m+1} = \frac{n_{m-1}-1}{n_{m-1}n_m - 1}\hat{\theta}_m$$
$$= o_P(1) \tag{31}$$

and

$$\hat{\theta}_{m-1,c} - \hat{\theta}_{m-1} = \frac{n_{m-1}n_m - n_{m-1}}{n_{m-1}n_m - 1}\hat{\theta}_m$$
$$= \theta_m + o_P(1). \tag{32}$$

The moment estimators of the variances at Level $1, \ldots, m-2, m+2, \ldots, Q$ do not change. Hence, levels $m-1$ and $m+1$ may be regarded as buffers for the other levels. It is also interesting to note that the changes in the level $m-1$ variance and the level $m+1$ variance sum to $\hat{\theta}_m$ and that $\hat{\theta}_m$ is absorbed by level $m-1$ if sample sizes n_m and n_{m-1} tend to infinity.

The estimated variance of $\hat{\gamma}_{\text{ols}}$ in the constrained model is

$$\hat{\text{var}}_c(\hat{\gamma}_{\text{ols}}) = \sum_{\substack{q=1\\q\neq m}}^{Q} \frac{\text{tr}(X_q'zz'X_q)}{(z'z)^2}\hat{\theta}_{qc}$$
$$= \frac{1}{z'z}\sum_{q=1}^{Q}\hat{\theta}_q + \frac{1}{z'z}\sum_{\substack{q=2\\q\neq m}}^{Q} f_{zx_q}(N_{q-1} - 1)\hat{\theta}_q$$
$$+ \frac{1}{z'z}\left[\frac{n_{m-1}n_m - n_{m-1}}{n_{m-1}n_m - 1} f_{zx_{m-1}}(N_{m-2} - 1)\right.$$
$$\left. + \frac{n_{m-1} - 1}{n_{m-1}n_m - 1} f_{zx_{m+1}}(N_m - 1)\right]\hat{\theta}_m. \tag{33}$$

Using Equation 33 and Equation 29, we obtain the following expression for the asymptotic relative change in the estimated variance of $\hat{\gamma}_{\text{ols}}$:

$$ARC_{\text{ols}} = \operatorname*{plim}_{n_1,\ldots,n_Q\to\infty} \frac{\hat{\text{var}}_c(\hat{\gamma}_{\text{ols}}) - \hat{\text{var}}(\hat{\gamma}_{\text{ols}})}{\hat{\text{var}}(\hat{\gamma}_{\text{ols}})}$$
$$= \operatorname*{plim}_{n_1,\ldots,n_Q\to\infty} \frac{1}{\sum_{q=1}^{Q}\hat{\theta}_q + \sum_{q=2}^{Q} f_{zx_q}(N_{q-1} - 1)\hat{\theta}_q}$$
$$\times\left[\frac{n_{m-1}n_m - n_{m-1}}{n_{m-1}n_m - 1} f_{zx_{m-1}}(N_{m-2} - 1)\right.$$
$$\left. + \frac{n_{m-1} - 1}{n_{m-1}n_m - 1}(N_m - 1)f_{zx_{m+1}} - (Nm - 1)f_{zx_m}\right]\hat{\theta}_m. \tag{34}$$

In the following, we examine the three-level model from which we discard the random intercept at Level 2. We evaluate ARC_{ols} for the five cases considered in the previous subsection.

1. z is a Level 1 variable the entries of which are centered around the Level 2 means. Then $f_{zx_2} = -(n_1 - 1)^{-1}$, $f_{zx_3} = -(N_2 - 1)^{-1}$, and $ARC_{ols} = \theta_2/\theta_1$, in which case the t test becomes conservative.
2. z is a stochastic Level 1 variable with zero variance at Level 2. Then plim $n_1 f_{zx_2} = 0$, plim $N_2 f_{zx_3} = 0$, and $ARC_{ols} = 0$.
3. z is a Level 2 variable the entries of which are centered around the Level 3 means. Then $f_{zx_2} = 1$, $f_{zx_3} = -(N_2 - 1)^{-1}$, and $ARC_{ols} = -1$, in which case the t test becomes liberal.
4. z is a stochastic Level 2 variable with zero variance at Level 3. Then $f_{zx_2} = 1$, plim $n_2 f_{zx_3} = 1$, and $ARC_{ols} = -1$, in which case the t test becomes liberal.
5. z is a Level 3 variable. Then $f_{zx_2} = 1$, $f_{zx_3} = 1$, and $ARC_{ols} = 0$.

These results can be summarized as follows. The omission of a random intercept at Level 2, asymptotically, does not affect the test result (when testing H_0: $\gamma = 0$) if z is a Level 3 variable or a Level 1 variable with zero variance at a higher level. If z is a Level 2 variable, the omission of Level 2 leads to a liberal test result, and if z is a Level 1 variable with entries centered around the Level 2 group means, the omission leads to a conservative test result.

Example: Continued

We assumed that the population model is a three-level model with random intercepts at the class and school level 5 (Model 5) and considered the effect of omitting the random intercept at the school level and the effect of omitting the random intercept at the class level. The ML estimates of the Level 1 variance, Level 2 variance, and Level 3 variance in the three-level random intercept model are 12.69, 1.24, and 1.20, respectively (see Table 1). When the random intercept at the school level is omitted, the changes in the ML estimates of the Level 1 and Level 2 variance are 0.0 and 1.1. If the ML estimate of the Level 3 variance is substituted in Equation 27, we obtain values of 0 and 1.2. It can be seen that the values obtained with Equation 27 are similar to the changes in the ML estimates of the Level 1 and Level 2 variance. When the random intercept at the class level is omitted from the three-level model, the changes in the ML estimates of the Level 1 and Level 3 variance are 0.9 and 0.6. If the ML estimate of the two-level variance is substituted in Equations 31 and 32, we obtain values of 1.0 and 0.3. Here we replaced n_1 and n_2 with the average number of students per class (13.0) and the average number of classes per school (4.9). Equations 31 and 32 yielded approximations that were in the same direction as the changes in the ML estimates.

As before, we also compared the changes in the estimated variance of the ML estimator $\hat{\gamma}_{ml}$ with APC_{ols} ($= ARC_{ols} \times 100\%$). For the group-centered variables IQ2 (Model 5 vs. Model 2) and IQ1 (Model 5 vs. Model 3), the changes in the estimated

TABLE 4
Three-Level Random Intercept Model Versus Two-Level Random Intercept Model: Between-Unit Variance Proportions and Percentage Changes in the Estimated Variances of the Fixed Effects Estimator

			Model 5 vs. Model 2	Model 5 vs. Model 3
	f_{zx_2}	f_{zx_3}	PC_{ml}	PC_{ml}
CONS	1.00	1.00	−68	26
SEX	.50	.53	−12	2
PM	.06	.03	0	3
ELFA	.12	.07	0	3
MATH0	.32	.18	0	0
IQ1	−.07	−.01	0	7
IQ2	1.00	−.01	38	−45
IQ3	1.00	1.00	−47	10

variance of $\hat{\gamma}_{\text{ml}}$ were smaller than the corresponding approximations of APC_{ols} (see Table 4): The values for IQ2 were 38% and 1.20/1.24 = 97%, respectively, and the values for IQ1 were 7% and 1.24/12.69 = 10%. For the variables CONS and IQ3 (Model 5 vs. Model 2) and IQ2 (Model 5 vs. Model 3), the value of APC_{ols} is −100%, and the changes in the estimated variance of $\hat{\gamma}_{\text{ml}}$ lie between −70% and −40%.

Concluding Remarks

The purpose of this article was to examine the asymptotic effect of omitting a random coefficient using analytical tools. We derived expressions for the change in the moment estimator of the variance components and for the change in the estimated variance of the fixed effects estimator. Both expressions turned out to be rather simple functions of between-unit variance proportions.

The expression for the asymptotic change in the moment estimator of the variance components may be useful when interpreting the variance terms in the model. We showed, for instance, that ignoring a level in a Q-level random intercept model may help us understand how the variance components at the different levels are related to each other. The expression for the change in the estimated variance of the fixed effects estimator sharpens one's insight into the sensitivity of the inferences about the fixed effects. The formulas can be applied when evaluating the results of a multilevel analysis performed by other researchers and when there is no access to the raw data. On the basis of the between-unit variance proportions, we made statements about whether the t test tends to be liberal, conservative, or exact when sample sizes are large.

Some limitations of the study can be mentioned. First, we confined ourselves to models with only one predictor and studied the asymptotic misspecification effect. This approach yields transparent formulas, and asymptotic control of Type I error in the *t* test for fixed effects is a common criterion for evaluating the performance of a test. However, it is also important to know how accurate the formulas are in a more realistic setting with moderate sample sizes and many correlated predictors. A comparison with a real data set indicated that the change in the ML estimates of the variance components is often comparable to the change computed from the formulas presented here. However, the presented asymptotic change in the variance of the estimated fixed effects can substantially deviate from the change in the variance of the ML estimates. The presented asymptotic change merely indicates whether the *t* test will become conservative or liberal, but the absolute change is not very informative. Second, we considered only the omission of a random coefficient, but it is also interesting to study the effect of including an extra random coefficient. Expressions for inclusion effects can be derived in a manner analogous to the expressions presented in this article. Third, we confined ourselves to studying misspecification in some basic multilevel models. With the technique presented here, however, it is also possible to derive expressions for the omission effect in multivariate multilevel models and multilevel models with crossed random effects.

References

Berkhof, J., & Snijders, T. A. B. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics, 26,* 133–153.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Application and data analysis methods.* Newbury Park, CA: Sage.

De Leeuw, J., & Kreft, I. G. G. (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics, 11,* 57–85.

Dempster, A. P., Rubin, D. B., & Tsutakawa, R. K. (1981). Estimation in covariance components models. *Journal of the American Statistical Association, 76,* 341–353.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. *Biometrika, 73,* 43–56.

Goldstein, H. (1995). *Multilevel statistical models.* London: Arnold.

Goldstein, H., Rasbash, J., Plewis, I., Draper, D., Browne, W., Yang, M., Woodhouse, G., & Healy, M. (1998). *A user's guide to MLwiN.* London: Multilevel Models Project, Institute of Education, University of London.

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association, 72,* 320–340.

Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics, 9,* 226–252.

Hutchison, D., & Healy, M. (2001). The effect on variance component estimates of ignoring a level in a multilevel model. *Multilevel Modelling Newsletter, 13,* 4–5.

Lange, N., & Laird, N. M. (1989). The effect of covariance structure on variance estimation in balanced growth-curve models with random parameters. *Journal of the American Statistical Association, 84,* 241–247.

Longford, N. T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika, 74,* 817–827.

Longford, N. T. (1993). *Random coefficient models.* Oxford, England: Oxford Science.

Miller, J. J. (1977). Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. *Annals of Statistics, 5,* 746–762.

Moerbeek, M. (2004). The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research, 39*(1), 129–149.

Opdenakker, M. C., & Van Damme, J. (2000). Effects of schools, teaching staff, and classes on achievement and well-being in secondary education: Similarities and differences between school outcomes. *School Effectiveness and School Improvement, 11,* 165–196.

Searle, S. R. (1968). Another look at Henderson's methods of estimating variance components. *Biometrics, 24,* 749–788.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage.

Authors

JOHANNES BERKHOF is a Researcher, Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, P.O. Box 7057, 10007MB Amsterdam, the Netherlands; h.berkhof@vumc.nl. His areas of specialization are random effects modeling, Bayesian model checking and model-based cost-effectiveness analyses in medical research.

JARL KENNARD KAMPEN is a Research Fellow, Public Management Institute, Faculty of Social Sciences, Katholieke Universiteit, Leuven, Van Evenstraat 2A, 3000 Leuven, Belgium; jarl.kampen@soc.kuleuven.ac.be. His areas of specialization are measurement of public sector performance, e-government, and the evaluation of data quality of social scientific surveys.