

RECURSIVE LINEAR MIXED MODELS: SOME RESULTS ON SHARED PARAMETER ESTIMATION

ROLAND FRIED*

*Department of Statistics, University of Dortmund,
Vogelthsweg 87, 44227 Dortmund, Germany*

(Received 11 February 2000; In final form 1 February 2001)

Linear mixed models have become an important tool for the analysis of longitudinal data as they allow modeling of the covariance structure of the observations and handle missing data. Recursive linear models allow interactions between the endogenous variables in multivariate regression analysis. Parameter restrictions cause difficulties in any linear model. In a recursive linear model, the analytical properties of combined estimators of regression coefficients which are shared by some variables are hard to derive. We give some insights in the simpler case of separate estimation and derive the analytical distributions of two-stage generalized-least-squares estimators for shared coefficients. These estimators are used as a benchmark in a simulation study for the maximum likelihood estimators. The setting of this study is characterized by few (uniform) random effects, many regressors and a large number of unbalanced observations. The maximum likelihood estimators compare favorably to the two-stage estimators and we suppose them to be satisfactory if the number of degrees of freedom is large.

Keywords: Longitudinal data; Maximum likelihood estimation; Random-effects; Missing data; Simulation study

1. INTRODUCTION

In longitudinal studies the observations are unlikely to be independent. Linear mixed models proposed by Laird and Ware (1982), which are also called variance components models (Searle *et al.*, 1992), model dependencies by means of normally distributed random effects.

*Tel.: ++49 (0)231 755 3129, Fax: ++49 (0)231 755 5305, e-mail: fried@statistik.uni-dortmund.de

Missing data can also be successfully handled in this context (e.g., Jennrich and Schluchter, 1986). However, efficient estimation of the model parameters may become difficult. Maximum likelihood (ML) estimators of the variance components, *i.e.*, of the variances of the random effects, are usually biased downward as they do not account for the estimation of the fixed effects. Therefore, the computationally slightly more expensive technique of restricted maximum likelihood (REML) has been proposed for bias correction (e.g., Searle *et al.*, 1992; or Diggle *et al.*, 1996). Generally, for the computation gradient-type (Newton-Raphson) or EM-type algorithms are needed (e.g., Lindstrom and Bates, 1988).

Data from multivariate longitudinal studies can be modeled by simultaneous-equations models (Hsiao, 1986). These models are characterized by additive interaction terms for including dependencies between the endogenous variables in a system of structural equations. Often two-stage generalized-least-squares (2S GLS) estimators are applied for the unknown parameters. Triangular systems of interactions are called recursive (Wermuth, 1980; Andersson and Perlman, 1998). If there are common explanatory variables then it is interesting to determine whether some of the regression coefficients are identical in distinct equations. While the finite-sample behavior of the ML estimators is tractable for separate estimations, things become complicated in the shared parameter case. The combination of information gained in distinct equations results in non-standard bilinear-forms and prevents an easy deduction of the analytic behavior.

We analyse the ML estimators in a special linear recursive mixed model *via* a simulation study. By means of this study, we compare the ML estimators with the analytically simpler two-stage GLS estimators. Although the latter have desirable properties like unbiasedness in a model with known covariance structure, the former compare favorably to them. The setting of the study was designed for longitudinal data observed to inspect forest damages (Fried, Eichhorn and Paar, 2000). Its characteristics are many explanatory variables, a large number of unbalanced observations and a parsimonious structure of variance components (uniform effects for "individuals"). As the number of degrees of freedom (DF) is huge, the superiority of REML in comparison to ML in case of few DF (e.g., Yau and McGilchrist, 1996) should not be relevant here.

In Section 2 recursive linear mixed models are introduced in the context of the forest data mentioned. In Section 3 we give formula for separate as well as for shared ML estimation, discuss theoretical properties of the ML estimators in case of separate estimation and transfer these results to the two-stage GLS estimators. In Section 4, the results of our simulation study are presented. For exemplification, we analyse the forest data in Section 5. We conclude with a discussion of our findings.

2. RECURSIVE LINEAR MIXED MODELS

Suppose there are U variables indexed by $u=1, \dots, U$, which have been measured repeatedly ($r=1, \dots, n_i$) for individuals $i=1, \dots, I$. Then we have random variables $Y_{u,i,r}$ and corresponding values $x_{u,i,r,j}$ of m explanatory variables $j=1, \dots, m$. In our model every variable u acts as a further explanatory variable for each of the variables u' with $u' > u$ (but not *vice versa*). Often the measurements of a variable are correlated for the same individual since some of the individuals are 'high responders' and others are 'low responders'. This can be modeled using independent random effects $w_{u,i} \sim N(0, \tau_u^2)$, which may be weighted by measurements $d_{u,i,r}$ of additional explanatory variables (which may be distinct for $u=1, \dots, U$). Furthermore, $\varepsilon_{u,i,r} \sim N(0, \sigma_u^2)$ are independent measurement errors. Then a recursive linear mixed model can be defined by

$$\begin{aligned}
 Y_{1,i,r} &= \sum_{j=1}^m x_{1,i,r,j} \beta_{1j} + d_{1,i,r} w_{1,i} + \varepsilon_{1,i,r} \\
 Y_{2,i,r} &= \alpha_{21} Y_{1,i,r} + \sum_{j=1}^m x_{2,i,r,j} \beta_{2j} + d_{2,i,r} w_{2,i} + \varepsilon_{2,i,r} \\
 &\vdots \quad \quad \quad \vdots \\
 Y_{U,i,r} &= \sum_{u=1}^{U-1} \alpha_{Uu} Y_{u,i,r} + \sum_{j=1}^m x_{U,i,r,j} \beta_{Uj} + d_{U,i,r} w_{U,i} + \varepsilon_{U,i,r}
 \end{aligned}
 \tag{1}$$

Let $Y_{u,i} = (Y_{u,i,1}, \dots, Y_{u,i,n_i})'$ be the vector of all observations of variable u for individual i . In case of a balanced design, *i.e.*,

$n_1 = n_2 = \dots = n_U = n$, we could more generally add $\mathbf{H}_{u,i} \mathbf{w}_{u,i}$ to $\mathbf{Y}_{u,i}$, where the $\mathbf{w}_{u,i}$ are k -variate random effects and the $\mathbf{H}_{u,i}$ are known $n \times k$ -matrices. Assuming $\mathbf{w}_{u,i} \sim N(\mathbf{0}, \tau_u^2 \mathbf{I})$ we would get $\text{Cov}(\mathbf{Y}_{u,i}) = \sigma_u^2 \boldsymbol{\Sigma}_u = \tilde{\sigma}_u^2 \mathbf{I} + \tau_u^2 \mathbf{H}_{u,i} \mathbf{H}'_{u,i}$. In Fried (1999) a model of this type was used to inspect the hypothesis of an increase of humus disintegration in central European forests (Ulrich, 1981). This new kind of forest damage is indicated by high amounts of dissolved organic carbon 'DOC' in the soil solution, which are complexed with aluminium. For an investigation of this hypothesis in field measurements, interesting substances are monitored in the seepage water at $n = 20$ sites in $U = 2$ soil depths in a small part of the forest near Zierenberg, a village in the region of Kassel, Germany (Eichhorn, 1995). We use the same setting for our simulation study.

In Fried, Eichhorn and Paar (2000) an extensive exploratory analysis of these monitoring data measured between October 1989 and December 1994 was done for modeling DOC. Explanatory variables as well as systematic spatial and temporal effects were inspected. The resulting model is a recursive linear mixed model with the logarithms of DOC (used for stabilizing variances) in the two depths $u = 1, 2$ being treated as the distinct dependent variables $Y_{u,t,s}$, where $t \in \{1, \dots, T\}$ is the time of observation and $s \in \{1, \dots, n\}$ is the site. The time points are regarded as $T = 56$ individuals and the sites correspond to between 1 and 20 repeated measurements. The reason for this is that strong weather effects $\mathbf{w}_{u,t} \sim N(\mathbf{0}, \tau_u^2)$ consisting of temperature and rainfall cause strong uniform ($d_{u,t,s} = 1$) correlations among all simultaneous observations (corresponding to the same individual t), while dependencies between the observation times were found to be negligible. Temporal interactions are apparently very short term and there are only a few observations in short time delays. The aim of the data analysis is to assess the relationships of DOC with the following explanatory variables: sulphate SO_4 , chloride Cl , iron Fe , aluminium Al , magnesium Mg , calcium Ca , potassium K , oxonium H . Fixed seasonal effects are included by a general sine function of time $g(u, t) = \sin(2\pi t/26)\beta_{u,10} + \cos(2\pi t/26)\beta_{u,11}$ with periodicity one year (26 observation times per annum according to a minimal time lag of two weeks). As the spatial variability in mean level does not follow a systematic pattern, dummy variables $\beta_{u,10+s}$, $s = 2, \dots, 20$, indicating the site are included. Dependencies between the layers are considered

in a directed manner by adding an autoregressive term $\alpha Y_{1,t,s}$ to $Y_{2,t,s}$. Hence, the model is

$$\begin{aligned}
 Y_{1,t,s} &= \sum_{j=1}^9 x_{1,t,s,j} \beta_{1,j} + g(1,t) + \beta_{1,10+s} + w_{1,t} + \varepsilon_{1,t,s} \\
 Y_{2,t,s} &= \alpha Y_{1,t,s} + \sum_{j=1}^9 x_{2,t,s,j} \beta_{2,j} + g(2,t) + \beta_{2,10+s} + w_{2,t} + \varepsilon_{2,t,s}
 \end{aligned}
 \tag{2}$$

Under the assumption of independent errors $\varepsilon_{u,t,s} \sim N(0, \bar{\sigma}_u^2)$, the covariance matrix of $Y_{u,t}$ is $\sigma_u^2 \Sigma_u = \bar{\sigma}_u^2 \mathbf{I} + \tau_u^2 \mathbf{1}$ with $\mathbf{1}$ being an $n \times n$ -matrix with all entries equal to one, and $\bar{\sigma}_u^2$ is the unknown error variance.

Since even rainy days did not cause seepage water at all sites, there is a large amount of missing values causing an unbalanced design. As most biochemical processes predominantly happen in the upper layer, measurements in the lower depth were only taken when the upper one was observable. Thus, there are more observations (923) in the upper layer than in the lower one (734). If there had been no missing values, 2720 observations in lags of two weeks would have been available for each depth, *i.e.*, about 65.9% and 72.8% of the observations are missing respectively. The vector $Z_{u,t} = M_{u,t} Y_{u,t}$ denotes the actually observed values, where $M_{u,t}$ is a $(n_{u,t} \times n)$ -matrix having 0/1 entries. We assume that data is missing at random, so the location does not contain information. The values of the covariates corresponding to $Z_{u,t}$ are given by the $n_{u,t} \times q$ -matrix $X_{u,t}$, which also contains the systematic temporal and spatial influences. We have $E(Z_{u,t}) = X_{u,t} \beta$ in case of $\alpha = 0$.

It is possible that the effects of the covariates in the lower depth differ from the effects in the upper depth because of distinct soil consistencies. This is an interesting aspect of the analysis. The design matrices $X_{u,t}$ and the regression coefficients β can be organized in such a way that the equations for $u = 1, 2$ share some of the unknown parameters. Especially we consider $\beta_{1,j} = \beta_{2,j}$ for $j = 1, \dots, 8$, *i.e.*, equality of the effects of the covariates. This can be achieved by setting $\beta = (\beta_{1,1}, \dots, \beta_{1,8}, \beta_{1,9}, \dots, \beta_{1,p}, \beta_{2,9}, \dots, \beta_{2,p})'$ (with $p = 30$, *i.e.*, β is a vector with $q = 52$ elements) and putting the covariates $x_{u,t,s,j}$, $j = 1, \dots, 8$, in the first columns of $X_{u,t}$.

If all parameters are estimated separately for both parts $u=1,2$, some simplifications are possible since the design matrix \mathbf{X}_t partitions for each observation time t into the block matrices

$$\begin{pmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{X}}_{1,t} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{X}}_{2,t} \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} \quad (3)$$

Here, the number q of columns of \mathbf{X}_t is twice the number p of explanatory variables, which is also the number of columns of $\tilde{\mathbf{X}}_{1,t}$ and $\tilde{\mathbf{X}}_{2,t}$.

3. MAXIMUM LIKELIHOOD ESTIMATION

Now we study the computation of the ML estimators and their analytical properties. We set $\mathbf{V}_{u,t} = \mathbf{M}_{u,t} \boldsymbol{\Sigma}_u \mathbf{M}'_{u,t}$ and $v = \sigma_2^2 / \sigma_1^2$. Furthermore, let $\mathbf{Z}_{1,2,t} = \mathbf{M}_{2,t} \mathbf{Y}_{1,t}$ be the vector of all components of $\mathbf{Z}_{1,t}$ for which $u=2$ was also observed. We have $\mathbf{Z}_{2,t}(\alpha) = \mathbf{Z}_{2,t} - \alpha \mathbf{Z}_{1,2,t} \sim N(\mathbf{X}_{2,t} \boldsymbol{\beta}, \sigma_2^2 \mathbf{V}_{2,t})$. The number of observations of variable u is $N_u = \sum_{t=1}^T n_{u,t}$ and $N = N_1 + N_2$ is the total number of observations. We assume that the overall design matrix has full rank $q < N$.

3.1. Calculation of the ML Estimators

In model (2), the likelihood function is the product of the densities

$$\begin{aligned} \prod_{t=1}^T f_t(\mathbf{z}_{1,t}, \mathbf{z}_{2,t}) &= \prod_{t=1}^T f_{1,t}(\mathbf{z}_{1,t}) f_{2,t}(\mathbf{z}_{2,t} | \mathbf{Z}_{1,t} = \mathbf{z}_{1,t}) \\ &= \prod_{t=1}^T \phi(\mathbf{z}_{1,t}; \mathbf{X}_{1,t} \boldsymbol{\beta}, \sigma_1^2 \mathbf{V}_{1,t}) \phi(\mathbf{z}_{2,t}; \mathbf{z}_{1,2,t} \alpha + \mathbf{X}_{2,t} \boldsymbol{\beta}, \sigma_2^2 \mathbf{V}_{2,t}) \end{aligned} \quad (4)$$

where $\phi(\cdot; \boldsymbol{\mu}, \boldsymbol{\Gamma})$ denotes the multivariate Gaussian density with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Gamma}$. This is the same as the likelihood of a linear mixed model with the observations of variable $u=1$ considered to be

fixed in the equation $u = 2$, *i.e.*,

$$\begin{pmatrix} \mathbf{Z}_{1,t} \\ \mathbf{Z}_{2,t} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{1,t} & \mathbf{0} \\ \mathbf{X}_{2,t} & \mathbf{z}_{1,2,t} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \alpha \end{pmatrix} + \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon}_{1,t} \\ \boldsymbol{\varepsilon}_{2,t} \end{pmatrix} \quad (5)$$

In consequence, the ML estimators can be calculated by standard statistical software for linear mixed models (compare Wolfinger *et al.*, 1994; or Keselman *et al.*, 1999).

3.2. Separate Estimations of the Regression Coefficients

For separate estimation of the regression coefficients using the design matrix (3), we get the following formulae for the ML estimators by some matrix manipulations (see Fried, 1999):

$$\hat{\boldsymbol{\beta}}_1 = \sum_{t=1}^T \tilde{\mathbf{A}}_{1,t} \mathbf{Z}_{1,t}, \quad (6)$$

$$\hat{\alpha} = \frac{\sum_{t=1}^T \mathbf{Z}'_{1,2,t} \mathbf{V}_{2,t}^{-1} \left(\mathbf{Z}_{2,t} - \tilde{\mathbf{X}}_{2,t} \sum_{i=1}^T \tilde{\mathbf{A}}_{2,i} \mathbf{Z}_{2,i} \right)}{\sum_{t=1}^T \mathbf{Z}'_{1,2,t} \mathbf{V}_{2,t}^{-1} \left(\mathbf{Z}_{1,2,t} - \tilde{\mathbf{X}}_{t,2} \sum_{i=1}^T \tilde{\mathbf{A}}_{2,i} \mathbf{Z}_{1,2,i} \right)}, \quad (7)$$

$$\hat{\boldsymbol{\beta}}_2(\hat{\alpha}) = \sum_{t=1}^T \tilde{\mathbf{A}}_{2,t} \mathbf{Z}_{2,t}(\hat{\alpha}), \quad (8)$$

$$\hat{\sigma}_1^2 = \frac{1}{N_1} \sum_{t=1}^T \mathbf{Z}'_{1,t} \mathbf{V}_{1,t}^{-1} \left(\mathbf{Z}_{1,t} - \tilde{\mathbf{X}}_{1,t} \sum_{i=1}^T \tilde{\mathbf{A}}_{1,i} \mathbf{Z}_{1,i} \right), \quad (9)$$

$$\hat{\sigma}_2^2 = \frac{1}{N_2} \left[\sum_{t=1}^T \mathbf{Z}'_{2,t} \mathbf{V}_{2,t}^{-1} \left(\mathbf{Z}_{2,t} - \tilde{\mathbf{X}}_{2,t} \sum_{i=1}^T \tilde{\mathbf{A}}_{2,i} \mathbf{Z}_{2,i} \right) - \right. \quad (10)$$

$$\left. \hat{\alpha} \sum_{t=1}^T \mathbf{Z}'_{1,2,t} \mathbf{V}_{2,t}^{-1} \left(\mathbf{Z}_{2,t} - \tilde{\mathbf{X}}_{2,t} \sum_{i=1}^T \tilde{\mathbf{A}}_{2,i} \mathbf{Z}_{2,i} \right) \right], \quad (11)$$

where

$$\tilde{A}_{u,t} = \left(\sum_{\bar{i}=1}^T \tilde{X}'_{u,\bar{i}} V_{u,\bar{i}}^{-1} \tilde{X}_{u,\bar{i}} \right) \tilde{X}'_{u,t} V_{u,t}^{-1}, \quad u = 1, 2, \quad t = 1, \dots, T. \quad (12)$$

The formulae for $\hat{\alpha}$, $\hat{\beta}_2(\hat{\alpha})$ and $\hat{\sigma}_2^2$ hold almost surely since we have to consider the possibility of collinearity, which occurs with probability 0. The estimates depend on the remaining variance components, which can be estimated by inserting these formulae in the likelihood and maximizing the profile likelihood numerically. We used a truncated Newton algorithm for this.

If the remaining variance components are known, *i.e.*, Σ_u is known, $u = 1, 2$, these estimators have the typical Gaussian and χ^2 -distributions. In general, the distributions of the estimators in a linear model with the variance components being replaced by their estimates can be treated as good approximations of the distributions in the mixed case if the number of degrees of freedom, which is $N_u - p$ in our case, is large in comparison to the number of variance components (see Diggle *et al.*, 1996, p. 93ff). For stating these distributional results let $\tilde{X}_u = (\tilde{X}'_{u,1}, \dots, \tilde{X}'_{u,T})'$ be the total design matrix for variable u , $V_u = \text{diag}(V_{u,1}, \dots, V_{u,T})$ and

$$\tilde{P}_u = \tilde{X}_u \tilde{A}_u = \tilde{X}_u (\tilde{X}'_u V_u^{-1} \tilde{X}_u)^{-1} \tilde{X}'_u V_u^{-1}, \quad u = 1, 2,$$

the matrix of the perpendicular projection onto the column space of \tilde{X}_u with respect to the scalar product induced by V_u^{-1} .

PROPOSITION 1 *In case of separate estimation of the regression coefficients, normally distributed errors and known correlation matrix Σ_u we have*

(a)

$$\hat{\beta}_1 \sim N(\beta_1, \sigma_1^2 (\tilde{X}'_1 V_1^{-1} \tilde{X}_1)^{-1}),$$

(b)

$$\frac{N_1 \hat{\sigma}_1^2}{\sigma_1^2} \sim \chi_{N_1 - p}^2, \quad \text{and } \hat{\beta}_1 \text{ and } \hat{\sigma}_1^2 \text{ are independent,}$$

(c)

$$\hat{\alpha} \Big|_{\mathbf{z}_1 = \mathbf{z}_1} \sim N \left(\alpha, \frac{\sigma_2^2}{\mathbf{z}'_{1,2} (\mathbf{I} - \tilde{\mathbf{P}}_2)' \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{z}_{1,2}} \right) a.s.,$$

(d)

$$\hat{\boldsymbol{\beta}}_2(\hat{\alpha}) \Big|_{\mathbf{z}_1 = \mathbf{z}_1} \sim N \left(\boldsymbol{\beta}_2, \sigma_2^2 \left[(\tilde{\mathbf{X}}_2' \mathbf{V}_2^{-1} \tilde{\mathbf{X}}_2)^{-1} + \frac{\tilde{\mathbf{A}}_2 \mathbf{z}_{1,2} \mathbf{z}'_{1,2} \tilde{\mathbf{A}}_2'}{\mathbf{z}'_{1,2} (\mathbf{I} - \tilde{\mathbf{P}}_2)' \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{z}_{1,2}} \right] \right) a.s.,$$

(e)

$$\frac{N_2 \hat{\sigma}_2^2}{\sigma_2^2} \sim \chi_{N_2 - p - 1}^2 \text{ independently from } \mathbf{Z}_1,$$

(f)

$(\hat{\boldsymbol{\beta}}_2(\hat{\alpha})', \hat{\alpha})'$ and $\hat{\sigma}_2^2$ are conditionally independent given \mathbf{Z}_1 .

A short proof is given in the Appendix. As usually, the moments stated in the previous proposition are valid without assuming normality.

It is well-known that ML estimators of variance components in linear mixed models are biased downward since they do not take into account the loss in degrees of freedom resulting from the estimation of the fixed effects (see Harville, 1977; or Yau and McGilchrist, 1996, compare also parts (b) and (e) of the preceding proposition).

The idea of restricted maximum likelihood (REML) is to overcome this bias by maximizing a modified likelihood based on linearly independent error contrasts, so that the influence of the fixed effects is eliminated (Harville, 1974; Diggle *et al.*, 1994, p. 64ff). In a recursive linear mixed model with separate parameters this can be accomplished using linearly transformed data $\mathbf{C}_1 \mathbf{Z}_1$ and $\mathbf{C}_2 \mathbf{Z}_2$ for estimation of the variance components, where \mathbf{C}_1 and \mathbf{C}_2 are $(N_1 - p) \times N_1$ - and $(N_2 - p - 1) \times N_2$ -matrices respectively with $\mathbf{C}_1 \tilde{\mathbf{X}}_1 = \mathbf{0}$ and $\mathbf{C}_2 (\tilde{\mathbf{X}}_2, \mathbf{z}_{1,2}) = \mathbf{0}$. Such matrices always exist, *e.g.*, \mathbf{C}_1 can be derived

by eliminating any p rows of the matrix $\mathbf{I} - \tilde{\mathbf{X}}_1(\tilde{\mathbf{X}}_1'\tilde{\mathbf{X}}_1)\tilde{\mathbf{X}}_1'$. The REML estimators of the variance components do not depend on the particular choice of \mathbf{C}_1 and \mathbf{C}_2 since for any choice the modified negative profile log-likelihood for $u=1$ and data $\mathbf{C}_1\mathbf{z}_1$ is (up to a constant)

$$\begin{aligned} \ell_1^* &= (N_1 - p)\ln(\sigma_1^2) + \ln|\mathbf{V}_1| + \ln|\tilde{\mathbf{X}}_1'\mathbf{V}_1^{-1}\tilde{\mathbf{X}}_1| \\ &\quad + \sigma_1^{-2}(\mathbf{z}_1 - \tilde{\mathbf{X}}_1\hat{\boldsymbol{\beta}}_1)'\mathbf{V}_1^{-1}(\mathbf{z}_1 - \tilde{\mathbf{X}}_1\hat{\boldsymbol{\beta}}_1). \end{aligned}$$

In comparison, the ML estimators of the variance components in part $u=1$ are calculated maximizing the negative profile log-likelihood

$$\ell_1 = N_1\ln(\sigma_1^2) + \ln|\mathbf{V}_1| + \sigma_1^{-2}(\mathbf{z}_1 - \tilde{\mathbf{X}}_1\hat{\boldsymbol{\beta}}_1)'\mathbf{V}_1^{-1}(\mathbf{z}_1 - \tilde{\mathbf{X}}_1\hat{\boldsymbol{\beta}}_1).$$

In the formulae for the (modified) profile likelihood ℓ_2 (ℓ_2^*) we have to replace $\tilde{\mathbf{X}}_1$ by $(\tilde{\mathbf{X}}_2, \mathbf{z}_{1,2})$ and the number of covariates p by $p+1$. The REML estimators for $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and α are calculated using the same formulae as the ML estimators, but instead of the ML estimators of the variance components we insert the REML estimators. These statements follow immediately from Diggle *et al.* (1994, p. 64ff) using the same basic arguments as for the proof of Proposition 1. In conclusion, the REML estimators of the variances are the unbiased estimators $N_1\hat{\sigma}_1^2/(N_1 - p)$ and $N_2\hat{\sigma}_2^2/(N_2 - p - 1)$.

Although REML estimators have advantages w.r.t. bias, neither of both approaches is generally superior in the mean squared error sense. The reason for this is that ML estimators usually have smaller variances. In our case, the variance of the REML estimator of σ_1^2 and σ_2^2 is $2\sigma_1^4/(N_1 - p)$ and $2\sigma_2^4/(N_2 - p - 1)$ respectively, while the variance of the ML estimator is $2(N_1 - p)\sigma_1^4/N_1^2$ and $2(N_2 - p - 1)\sigma_2^4/N_2^2$ respectively. The root of the mean squared error RMSE is $\sqrt{2/(N_1 - p)}\sigma_1^2$ and $\sqrt{2/(N_2 - p - 1)}\sigma_2^2$ for the REML estimators, and $\sqrt{2(N_1 - p + p^2)/N_1^2}\sigma_1^2$ and $\sqrt{2(N_2 - p - 1 + (p + 1)^2)/N_2^2}\sigma_2^2$ for the ML estimators. Hence, the comparison depends on the number of fixed effects and on the number of observations. Harville (1977) noticed that in linear mixed models with several variance components the comparison may even depend on the true parameter values. Asymptotically, both approaches have the same efficiency for the estimation of the residual variance.

3.3. Shared Regression Coefficients

In case of shared regression coefficients we get the following formulae for the ML estimators setting $\mathbf{X}_u = (\mathbf{X}'_{u,1}, \dots, \mathbf{X}'_{u,T})'$, $\mathbf{V}_u = \text{diag}(\mathbf{V}_{u,1}, \dots, \mathbf{V}_{u,T})$, $\mathbf{X} = (\mathbf{X}'_1, \mathbf{X}'_2)'$, $\mathbf{V} = \text{diag}(\mathbf{V}_1, \nu \mathbf{V}_2)$, $\mathbf{A}_u = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}'_u \mathbf{V}_u^{-1}$, and $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$.

$$\hat{\alpha} = \frac{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{Z}_2 - \mathbf{X}_2 \mathbf{A} \mathbf{Z})}{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{Z}_{1,2} - \mathbf{X}_2 \mathbf{A}_2 \mathbf{Z}_{1,2})}, \tag{13}$$

$$\hat{\beta}(\hat{\alpha}) = \mathbf{A} \mathbf{Z}(\hat{\alpha}), \tag{14}$$

$$\hat{\sigma}_1^2 = \frac{1}{N} \left[\mathbf{Z}' \mathbf{V}^{-1} (\mathbf{Z} - \mathbf{X} \mathbf{A} \mathbf{Z}) - \frac{\hat{\alpha}}{\hat{\nu}} \mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{Z}_2 - \mathbf{X}_2 \mathbf{A} \mathbf{Z}) \right] \tag{15}$$

The distributions of the ML estimators are difficult to handle analytically here, because the formulae cannot be partitioned for $u = 1, 2$. They involve bilinear forms of non-standard type.

For this reason we also consider the following two-stage GLS estimators, which are more tractable. In the first stage, α , σ_1^2 and σ_2^2 (as well as all other variance components) are estimated by GLS using only the corresponding partition (compare to the ML estimators in Section 3.2). Thereafter, the estimator $\check{\alpha}$ of α is inserted in the estimation equation (14) for β .

$$\check{\alpha} = \frac{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_2}{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_{1,2}} \tag{16}$$

$$\check{\sigma}_1^2 = \frac{1}{N_1} \mathbf{Z}'_1 \mathbf{V}_1^{-1} (\mathbf{Z}_1 - \tilde{\mathbf{P}}_1 \mathbf{Z}_1) \tag{17}$$

$$\check{\sigma}_2^2 = \frac{1}{N_2} \left[\mathbf{Z}'_2 \mathbf{V}_2^{-1} (\mathbf{Z}_2 - \tilde{\mathbf{P}}_2 \mathbf{Z}_2) - \check{\alpha} \mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{Z}_2 - \tilde{\mathbf{P}}_2 \mathbf{Z}_2) \right] \tag{18}$$

$$\check{\beta} = \mathbf{A} \mathbf{Z}(\check{\alpha}) \tag{19}$$

$\tilde{\mathbf{P}}_u$ is the same projection matrix as is used for separate estimations. The main difference between these estimators and the ML estimators is that for the two-stage GLS estimators of α and the variance components distinct preliminary estimates of β are implicitly used

for $u=1$ and $u=2$, which are based on the corresponding data only. Only β is estimated from all observations. Thus, the ML estimators use the sample more efficiently.

The two-stage GLS estimators have similar properties to the ML estimators in the separate parameter case. For this result, which is formulated for known correlation matrices, *i.e.*, known variance components, let $\mathbf{X}_{1,2}$ and $\mathbf{V}_{1,2}$ be the design and the covariance matrix of $\mathbf{Z}_{1,2}$ respectively. In case there are variance components which have to be estimated from the data it should provide a useful approximation if the number of degrees of freedom is large.

PROPOSITION 2 *In the recursive linear model with shared parameters we have in case of known Σ_u*

(a)

$$\begin{aligned} E(\tilde{\beta}) &= \beta \\ \text{Var}(\tilde{\beta}) &= \sigma_1^2 (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \\ &\quad + \sigma_2^2 \frac{\mathbf{A}_2 (\mathbf{V}_{1,2} + \mathbf{X}_{1,2} \beta \beta' \mathbf{X}'_{1,2}) \mathbf{A}'_2}{\text{trace}[\mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) (\mathbf{V}_{1,2} + \mathbf{X}_{1,2} \beta \beta' \mathbf{X}'_{1,2})]} \end{aligned}$$

(b)

$$\frac{N_1 \tilde{\sigma}_1^2}{\sigma_1^2} \sim \chi_{N_1-p}^2$$

(c)

$$\tilde{\alpha} \Big|_{\mathbf{z}_1=\mathbf{z}_1} \sim N \left(\alpha, \frac{\sigma_2^2}{\mathbf{z}'_{1,2} (\mathbf{I} - \tilde{\mathbf{P}}_2)' \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{z}_{1,2}} \right) \text{ a.s.}$$

(d)

$$\frac{N_2 \tilde{\sigma}_2^2}{\sigma_2^2} \sim \chi_{N_2-p-1}^2 \text{ independently from } \mathbf{Z}_1.$$

(e)

$(\tilde{\beta}', \tilde{\alpha})'$ and $\tilde{\sigma}_2^2$ are conditionally independent given \mathbf{Z}_1 .

Again, the first and second moments are valid without assuming normality.

4. SIMULATION STUDY

In view of the analytical intractability of the distributions of the ML estimators we judge their reliability in a simulation study. We use the setting described in Section 2 since our primary interest is the *DOC* data. The unknown parameters are varied systematically and 200 realizations are generated for every parameter set. For every set the mean, the standard deviation and the 2.5%- as well as the 97.5%-percentile of the ML estimates corresponding to 95% confidence limits are calculated. We consider the following parameter values.

1. Three vectors β of regression coefficients:

- $\beta^{(1)} = \mathbf{0}$ (no effects of the covariates),
- $\beta^{(2)}$ with components $\beta_j^{(2)} = j, j = 1, \dots, p$ (effects of different strength),
- $\beta^{(3)}$ with components in $\{0, \pm 1, \pm 2, \pm 10\}$ (positive as well as negative effects);

2. Three vectors of variance components:

$$\theta = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \bar{\sigma}_1 \\ \bar{\sigma}_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0.0 \\ 0.0 \\ 0.3 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.2 \\ 0.4 \\ 0.3 \\ 0.2 \end{pmatrix} \right\},$$

i.e., $\bar{\sigma}_1$ and $\bar{\sigma}_2$ fixed, corresponding to

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \sigma_1^2 \\ \sigma_2^2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0.0 \\ 0.0 \\ 0.09 \\ 0.04 \end{pmatrix}, \begin{pmatrix} 0.1 \\ 0.5 \\ 0.1 \\ 0.08 \end{pmatrix}, \begin{pmatrix} 4/13 \\ 0.8 \\ 0.13 \\ 0.2 \end{pmatrix} \right\}$$

3. Two values $\alpha \in \{0.0, 0.1\}$

All possible combinations are studied. In the following the vectors θ are denoted in the order mentioned above as $\theta^{(i)}, i = 1, 2, 3$.

For a simulation study of this magnitude (923 + 734 observations and additional random effects have to be generated for each realization, *i.e.*, more than 6,000,000 pseudo random numbers altogether) a generator with high periodicity is needed. We apply a

composite inversion generator with moduli $2^{31}-1$ and $2^{31}-61$. Eichenauer-Herrmann and Emmerich (1995) showed that such generators composed of two or more generators retain most of the favorable properties of the individuals. Inversion generators have a couple of advantages in comparison with classical linear generators (Eichenauer-Herrmann *et al.*, 1998), and these advantages increase even more after transformation to non-uniform distributions (Leydold *et al.*, 1998). For the transformation to normal deviates we use the common method of Box and Muller (1958). The simulation results for the bias B , the root of the mean squared error $RMSE$, and the lower and upper confidence limits LCL and UCL relative to the corresponding parameter $\vartheta \in \{\alpha, \rho_1, \rho_2, \sigma_1^2, \sigma_2^2\}$ are reported in Table I. Table II provides the same statistics for some exemplary regression coefficients. In the upper part the results for $\alpha = 0.0$ and in the lower part those for $\alpha = 0.1$ are given.

In our simulations a negative bias of the ML estimators of the variance components, which is well-known for linear mixed models, becomes obvious, too. For the correlations ρ_1 and ρ_2 both the negative bias and the mean squared error seem to be determined mainly by the correlation itself. They are smaller for correlations near zero or one. The confidence limits show an increasing asymmetry with higher correlations. For σ_1^2 the bias is between 2% and 3% of the true parameter value, and for σ_2^2 it is between 2% and 6%.

We compare these results to the analytical results for two-stage GLS estimators in a recursive model with known correlation matrix, which have been stated in Proposition 2. These formulae should be approximately valid in our setting since we have 893 and 704 degrees of freedom in the upper and lower part respectively, and there is only one additional variance component beside the error variance in each part.

In both cases the bias of the ML estimator found in the simulations is similar to the analytical bias of the two-stage GLS $\check{\sigma}_1^2$ and $\check{\sigma}_2^2$. For the numbers of observations given here, the bias of the latter is approximately 3.4% and 4.4% respectively, even in case of known correlations. The root of the mean squared error of $\hat{\sigma}_1^2$ is smaller than the RMSE of $\check{\sigma}_1^2$ for all parameter values considered here. It is between $0.025\sigma_1^2$ and $0.045\sigma_1^2$, whereas the RMSE of $\check{\sigma}_1^2$ is $0.046\sigma_1^2$ in case of known correlations. Bias correction of the two-stage GLS by using a REML-type estimate would increase the RMSE slightly to $0.047\sigma_1^2$.

TABLE I ML estimators of the variance components

<i>i</i>	<i>j</i>	α			ρ_1			ρ_2			σ_1^2			σ_2^2		
		<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>UCL</i>	<i>B</i>	<i>SE</i>	<i>UCL</i>	<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>UCL</i>
1	1	-.00076	.02375	-.00236	.01271	-.00543	.01515	-.00305	.00483	-.00138	.00232					
1	2	-.00007	.00056	-.00094	.01143	-.00351	.01372	-.00313	.00531	-.00147	.00246					
1	3	.00002	.00040	-.00344	.01191	-.00352	.01518	-.00261	.00514	-.00146	.00253					
2	1	.00189	.02588	-.00365	.03133	-.01355	.05556	-.00413	.06678	-.00372	.00873					
2	2	.00014	.00071	-.00290	.02880	-.01137	.05473	-.00274	.00541	-.00438	.00929					
2	3	.00002	.00042	-.00897	.02731	-.00928	.05441	-.00332	.00600	-.00383	.00846					
3	1	.00262	.02447	-.00571	.04757	-.00821	.03410	-.00357	.01022	-.01116	.03095					
3	2	.00000	.00077	-.01315	.05087	-.00717	.03099	-.00483	.01014	-.01056	.02842					
3	3	-.00004	.00041	-.01142	.05116	-.01037	.03414	-.00558	.01086	-.01188	.03047					
1	1	-.00536	.02721	-.00204	.01173	-.00292	.01408	-.00291	.00501	-.00134	.00245					
1	2	.00001	.00059	-.00470	.01260	-.00357	.01461	-.00228	.00444	-.00158	.00256					
1	3	.00000	.00033	-.00446	.01280	-.00358	.01500	-.00232	.00433	-.00157	.00250					
2	1	-.00071	.02714	-.00647	.02739	-.00629	.05226	-.00389	.00630	-.00348	.00902					
2	2	-.00003	.00075	-.00859	.02878	-.01217	.05094	-.00406	.00627	-.00397	.00831					
2	3	.00001	.00043	-.00666	.03174	-.00310	.05258	-.00371	.00586	-.00275	.00769					
3	1	.00300	.02564	-.00908	.04929	-.00786	.03485	-.00535	.01059	-.00872	.03107					
3	2	.00008	.00076	-.00877	.04936	-.00272	.02675	-.00421	.01023	-.00770	.02260					
3	3	-.00002	.00041	-.01205	.04625	-.00575	.02366	-.00534	.01058	-.00960	.02080					
<i>i</i>	<i>j</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>					
1	1	-.04436	.04647	-.02477	.02010	-.03101	.02835	-.01034	.00455	-.00464	.00196					
1	2	-.00110	.00111	-.02145	.02371	-.02754	.02595	-.01159	.00524	-.00531	.00290					
1	3	-.00076	.00077	-.02404	.02021	-.02969	.03291	-.01180	.00691	-.00581	.00258					
2	1	-.04372	.05035	-.05863	.05809	-.13907	.08891	-.01461	.00553	-.01886	.01186					
2	2	-.00131	.00146	-.05762	.05643	-.12812	.09022	-.01105	.00712	-.01909	.01198					
2	3	-.00072	.00092	-.05902	.05246	-.10933	.09315	-.01361	.00556	-.01714	.01174					

TABLE I (Continued)

<i>i</i>	<i>j</i>	ρ_1			ρ_2			σ_1^2			σ_2^2		
		LCL	UCL	LCL	LCL	UCL	LCL	LCL	UCL	LCL	UCL	LCL	UCL
3	1	-.04611	.05636	-.11105	.08746	-.09137	.04845	-.02126	.01619	-.06125	.04877		
3	2	-.00143	.00157	-.10303	.09411	-.09184	.04504	-.02147	.01777	-.06106	.04548		
3	3	-.00091	.00069	-.10408	.09289	-.08973	.04726	-.02357	.01371	-.06612	.04670		
1	1	-.06406	.04739	-.02352	.02221	-.02764	.02593	-.01073	.00490	-.00496	.00301		
1	2	-.00139	.00103	-.02506	.01948	-.03159	.02619	-.00876	.00606	-.00578	.00275		
1	3	-.00070	.00061	-.02576	.01939	-.03197	.02844	-.00895	.00563	-.00578	.00218		
2	1	-.05539	.05128	-.05905	.04275	-.10912	.09573	-.01430	.00475	-.01661	.01434		
2	2	-.00143	.00145	-.05849	.05654	-.11352	.07915	-.01305	.00576	-.01719	.01055		
2	3	-.00078	.00088	-.07062	.06001	-.12373	.09676	-.01355	.00516	-.01598	.01234		
3	1	-.04714	.05415	-.11429	.09017	-.08819	.05501	-.02151	.01211	-.06579	.05297		
3	2	-.00149	.00172	-.09983	.08547	-.06228	.04554	-.02253	.01709	-.04724	.03028		
3	3	-.00090	.00076	-.10054	.07945	-.05724	.03510	-.02310	.01178	-.04703	.02683		

Bias *B* and root of the mean squared error (here briefly SE) when $(\theta, \beta) = (\theta^{(i)}, \beta^{(i)})$.
 Relative 95% confidence limits LCL and UCL when $(\theta, \beta) = (\theta^{(i)}, \beta^{(i)})$.

TABLE II ML estimators of the regression coefficients

<i>i</i>	<i>j</i>	<i>Cl</i>			<i>H</i>			<i>Mg</i>			<i>Na</i>			<i>Fe</i>			<i>Al</i>		
		<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>UCL</i>	<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>LCL</i>	<i>B</i>	<i>SE</i>	<i>LCL</i>
1	1	.0001	.0024	-.08	10.78	-.0002	.0046	-.0003	.0118	.0095	.1459	.0015	.0559						
1	2	-.0001	.0023	.97	11.47	.0003	.0052	.0008	.0115	.0190	.1489	-.0015	.0540						
1	3	-.0001	.0025	-.54	11.10	.0002	.0049	-.0001	.0126	-.0067	.1619	-.0003	.0577						
2	1	-.0002	.0028	.35	10.68	.0000	.0053	.0011	.0126	-.0167	.1580	.0026	.0622						
2	2	-.0007	.0028	-.70	11.45	.0002	.0052	-.0012	.0123	-.0049	.1638	-.0007	.0592						
2	3	.0005	.0030	.64	10.98	-.0004	.0053	.0025	.0122	.0172	.1621	.0032	.0580						
3	1	.0001	.0029	-.19	12.65	-.0002	.0049	.0001	.0126	-.0043	.1607	.0026	.0584						
3	2	-.0003	.0033	1.04	11.66	.0001	.0050	-.0004	.0123	-.0166	.1804	.0061	.0603						
3	3	.0001	.0030	-.50	11.63	.0000	.0048	-.0014	.0137	.0161	.1623	-.0029	.0605						
1	1	.0003	.0024	.71	10.34	-.0002	.0050	-.0009	.0108	.0007	.1511	-.0066	.0493						
1	2	.0000	.0025	.40	10.98	.0000	.0048	-.0001	.0115	.0040	.1383	-.0010	.0583						
1	3	.0001	.0025	.46	10.91	.0000	.0048	-.0001	.0115	.0087	.1379	-.0007	.0579						
2	1	.0000	.0029	-.14	12.01	.0000	.0050	.0012	.0135	.0009	.1661	.0036	.0649						
2	2	-.0002	.0031	-.43	11.55	.0001	.0050	-.0006	.0140	-.0105	.1498	.0023	.0585						
2	3	.0002	.0029	.22	12.39	.0003	.0053	-.0001	.0125	.0035	.1623	-.0012	.0613						
3	1	.0001	.0030	.70	12.40	.0001	.0048	-.0015	.0139	.0025	.1715	.0085	.0651						
3	2	-.0003	.0031	.27	10.98	.0000	.0051	-.0004	.0132	-.0089	.1513	-.0027	.0674						
3	3	.0001	.0030	.56	12.44	-.0003	.0055	.0018	.0131	-.0103	.1424	.0023	.0635						
<i>i</i>	<i>j</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>	<i>LCL</i>	<i>UCL</i>						
1	1	-.0045	.0045	-23.05	20.15	-.0097	.0082	-.0221	.0234	-.2949	.2849	-.1005	.1222						
1	2	-.0042	.0047	-21.92	26.68	-.0099	.0108	-.0233	.0244	-.3100	.2868	-.1118	.1178						
1	3	-.0048	.0052	-24.24	22.72	-.0100	.0101	-.0242	.0240	-.3404	.3180	-.1000	.1183						
2	1	-.0061	.0058	-23.12	21.49	-.0095	.0115	-.0270	.0237	-.3665	.2740	-.1083	.1358						
2	2	-.0065	.0048	-23.53	21.76	-.0101	.0110	-.0245	.0234	-.3253	.3069	-.1170	.1277						
2	3	-.0061	.0066	-20.38	22.23	-.0105	.0123	-.0213	.0257	-.2720	.3402	-.1191	.1317						

TABLE II (Continued)

i	j	Cl		H		Mg		Na		Fe		Al	
		LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL	LCL	UCL
3	1	-.0057	.0073	-20.99	27.28	-.0089	.0090	-.0228	.0255	-.3355	.3224	-.1092	.1175
3	2	-.0061	.0063	-20.99	24.59	-.0094	.0094	-.0241	.0257	-.3780	.3485	-.1058	.1353
3	3	-.0068	.0053	-22.46	22.76	-.0104	.0095	-.0277	.0241	-.3285	.3563	-.1053	.1152
1	1	-.0046	.0060	-18.74	23.03	-.0100	.0107	-.0215	.0235	-.2786	.3029	-.0963	.1130
1	2	-.0057	.0047	-21.11	22.74	-.0095	.0112	-.0226	.0223	-.2735	.3013	-.1196	.1216
1	3	-.0056	.0048	-21.02	22.59	-.0097	.0111	-.0228	.0226	-.2735	.2939	-.1205	.1217
2	1	-.0057	.0061	-26.01	20.60	-.0110	.0095	-.0212	.0278	-.3247	.3460	-.1244	.1324
2	2	-.0060	.0060	-21.66	22.84	-.0092	.0093	-.0257	.0278	-.3352	.2563	-.1053	.1151
2	3	-.0057	.0060	-26.81	25.10	-.0100	.0110	-.0219	.0258	-.3311	.3466	-.1262	.1228
3	1	-.0059	.0061	-24.99	22.77	-.0106	.0089	-.0268	.0248	-.4041	.3434	-.1046	.1287
3	2	-.0063	.0063	-21.16	22.98	-.0083	.0105	-.0252	.0277	-.3329	.2845	-.1462	.1339
3	3	-.0054	.0066	-25.03	24.04	-.0105	.0115	-.0210	.0265	-.2622	.2612	-.1314	.1106

Bias B and root of the mean squared error (here briefly SE) when $(\theta, \beta) = (\theta^{(i)}, \beta^{(i)})$.
 Relative confidence limits LCL and UCL when $(\theta, \beta) = (\theta^{(i)}, \beta^{(i)})$.

The ML estimator of σ_2^2 does better than the two-stage GLS for most of the parameter combinations considered here. While the ML approach appears to be superior for small and intermediate correlations ρ_2 , the two-stage GLS might do significantly better for high positive correlations. For $\rho_2 = 0.8$ the *RMSE* of $\hat{\sigma}_2^2$ amounts up to $0.15\sigma_2^2$. This is more than the *RMSE* of $\check{\sigma}_2^2$, which is about $0.066\sigma_2^2$ for known correlations and the number of observations given here. Using the REML-type version of two-stage GLS would diminish the *RMSE* to 0.053.

The deterioration of $\hat{\sigma}_2^2$ for very high correlation ρ_2 might be due to numerical instability because of a very fiat likelihood function. Having an analytical formula at hand as we have for $\check{\sigma}_2^2$ might help to overcome this problem.

The variances of the two-stage GLS of both the autoregressive parameter α and the regression coefficients $\beta_{u,j}$ depend on the true regression coefficients, compare Proposition 2. Hence, in a simulation study a huge number of parameter combinations had to be used to get general insights. Instead we only comment on the results of our simulations for the ML estimators. However, we feel that ML estimation typically should be better because of the more efficient use of the sample.

For $\hat{\alpha}$ no systematic bias is evident. There are no obvious effects of the variance components, but β seems to be influential. Both the mean squared error and the absolute bias are largest when the regression coefficients are zero. This behavior can also be recognized from the confidence limits. In general the percentiles are rather symmetrical to the true α .

The ML estimator $\hat{\beta}$ provides symmetric confidence limits, and there are both small positive and negative deviations of the sample mean from the true parameter without systematic patterns. Thus, we do not find an obvious, systematic bias. The only discernible behavior is the usual slight increase of *RMSE* with the variances. Other influences of the parameters are not visible. The results must not be compared between distinct parameters since they are shaped by the design matrices (Cook and Weisberg, 1982). In particular, the large *RMSE* for H results from a bad-conditioned design for this covariate.

The findings mentioned above were checked by further simulations for a couple of other values of θ . These results are not reported here since we did not detect remarkable differences.

5. EXAMPLE: THE *DOC*-DATA

Now we estimate the model parameters for the *DOC* data presented in Section 2 for both the separate and the shared parameter case and compare the results. In view of our simulation study, we consider ML estimation to be reliable in our example and prefer it to two-stage GLS estimation. Nevertheless, the results of both approaches are reported in Tables III and IV for the reason of comparison. Most of the differences between the ML and the two-stage GLS estimates are quite small, the large difference for oxonium *H* is due to the bad design for this variable and the resulting high variance of the estimation. The estimates $\hat{\rho}_1$ and $\hat{\rho}_2$ signal positive correlations, particularly in the lower part.

Comparing the shared and the separate estimations of the regression coefficients, we see some rather large differences and some similar values. The value of the likelihood ratio test statistic is 32.41, which can be compared to the χ^2_3 -distribution (e.g., Whittle, 1953). We find the 0.9995-percentile of this distribution in Hald (1967) to be 29.7. Thus, the result is highly significant, and we conclude that the coefficients are distinct and estimate them separately henceforth. Detailed results of our data analysis can be found in Fried, Eichhorn and Paar (2000) and in Fried (1999). We do not address the question here whether some of the covariates have the same coefficients in both layers. The likelihood ratio test which has been applied discriminates between all of the covariates effects being equal or none being equal. We suppose that either there are differences between both layers

TABLE III Variance components for the shared parameter model

<i>Parameter</i>	α	ρ_1	ρ_2	σ_1^2	σ_2^2
MLE	.075837	.16593	.42173	.035582	.042745
2S GLS	.058230	.18890	.42177	.035686	.041903

ML (first line) and two-stage GLS estimates (second line) in the shared parameter model.

TABLE IV Estimated regression coefficients

	SO_4	Cl	H	Mg	Ca	K	Na	Fe	Al	\sin	\cos
ML	.0133	-.0049	.140	-.0114	-.0176	.0222	.0414	.8462	.2839	-.1141	-.0210
2S GLS	.0136	-.0070	1.297	-.0125	-.0161	.0226	.0392	.7632	.3217	-.1177	-.0192
20 cm	.0104	-.0055	-1.810	.0078	-.0242	.0253	-.0046	.8698	.2684	-.1846	-.1082
60 cm	.0193	-.0031	-85.406	-.0214	-.0166	.0084	.0668	.6783	.4595	-.0986	-.0237

In most of the cases the ML and the two-stage GLS estimates for the shared parameter model (first and second line) are rather similar. The bad design for H causes a high variance. In the separate parameter model (third and fourth line), some of the differences between the layers are rather large.

influencing the biochemical mechanism as a whole, or such differences do not exist.

6. CONCLUSION

We have inspected the behavior of the ML estimators in a special multi-response linear mixed model, which has a recursive structure of interactions between the dependent variables. The distributions of separate ML estimators for each equation can be approximated analytically by assuming the variance components to be fixed. However, in case of shared coefficients the ML estimators are difficult to handle.

In the former, unrestricted case the ML estimators of the regression coefficients and the variances are approximately Gaussian and χ^2 -distributed respectively if the number of variance components is small relative to the number of observations. It is well-known that the ML estimators of the variance components in linear mixed models typically have a negative bias, which increases with the number of explanatory variables. Nevertheless, the ML estimators are supposed to lead to satisfactory results if the number of observations is large in comparison to the total number of model parameters, *i.e.*, regression coefficients and variance components.

For the latter case of shared parameters we have done a simulation study. As we do not know the analytical finite-sample properties of the ML estimators, we have also proposed analytically easier two-stage GLS estimators, which are (approximately) unbiased. For the design of our study, we have found the ML estimators to be reliable. We prefer them for our data analysis since their performance is superior to the performance of the alternatives in the mean squared error sense in most of the cases. This superiority is in accordance to the more efficient sample use by ML estimation.

Acknowledgements

The author is grateful to Prof. Dr. P. Spellucci and Dr. R. Felkel from the Department of Mathematics, Darmstadt University of Technology, Germany, for their numerical support.

References

- Andersson, S. A. and Perlman, M. D. (1998) Normal Linear Regression Models with Recursive Graphical Markov Structure. *J. Multivariate Analysis*, **66**, 133–187.
- Box, G. E. P. and Muller, M. E. (1958) A note on the generation of random normal deviates. *Annals of Mathematical Statistics*, **29**, 610–611.
- Cook, R. D. and Weisberg, S. (1982) *Residuals and Influence in Regression*. Chapman & Hall, New York.
- Diggle, P. J., Liang, K.-Y. and Zeger, S. L. (1996) *Analysis of Longitudinal Data*. Clarendon Press, Oxford.
- Eichenauer-Herrmann, J. and Emmerich, F. (1995) A review of compound methods for pseudorandom number generation. In: *Proceedings of the 1st Salzburg Minisymposium on Pseudorandom Number Generation and Quasi Monte Carlo Methods*, Hellekalek, P., Larcher, G. and Zinterhof, P. (Eds.), Technical Report Series, Austrian Center for Parallel Computation.
- Eichenauer-Herrmann, J., Herrmann, E. and Wegenkittl, S. (1998) A survey of quadratic and inverse congruential pseudorandom numbers. In: *Monte Carlo and Quasi-Monte Carlo Methods 1996*, Niederreiter, H., Hellekalek, P., Larcher, G. and Zinterhof, P. (Eds.), Lecture Notes in Statistics, Vol. 127, Springer, New York.
- Eichhorn, J. (1995) Stickstoffsättigung und ihre Auswirkungen auf das Buchenwaldökosystem der Fallstudie Zierenberg. Habilitationsschrift, *Berichte des Forschungszentrums Wald-ökosysteme Reihe A*, **124**, Göttingen, Germany.
- Fried, R. (1999) *Räumlich-zeitliche Modellierung der Kohlenstoffkonzentration im Waldbodensickerwasser zur Untersuchung der Hypothese der Humusdisintegration*. Ph. D.-thesis, Department of Mathematics, Darmstadt University of Technology, D-64289 Darmstadt, Germany; also Logos-Verlag, Berlin.
- Fried, R., Eichhorn, J. and Paar, U. (2000) Exploratory Analysis and a Stochastic Model for Humusdisintegration. *Environmental Monitoring and Assessment*, in press.
- Hald, A. (1967) *Statistical Tables and Formulas*. 7th edition. John Wiley & Sons, New York.
- Harville, D. A. (1974) Bayesian Inference for Variance Components Using Only Error Contrasts. *Biometrika*, **61**, 383–385.
- Harville, D. A. (1977) Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems (with discussion). *J. Americ. Statist. Assoc.*, **72**, 320–340.
- Hsiao, C. (1986) *Analysis of panel data*. Cambridge University Press, Cambridge.
- Jennrich, R. I. and Schluchter, M. D. (1986) Unbalanced Repeated-Measures Models with Structured Covariance Matrices. *Biometrics*, **42**, 805–820.
- Keselman, H. J., Algina, J., Kowalchuk, R. K. and Wolfinger, R. D. (1999) The Analysis of Repeated Measurements: A Comparison of Mixed-Model Satterwaite F Tests and a Nonpooled Adjusted Degrees of Freedom Multivariate Test. *Commun. Statist.-Theory Meth.*, **28**, 2967–2999.
- Laird, N. M. and Ware, J. H. (1982) Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963–974.
- Leydold, J., Leeb, H. and Hörmann, W. (1998) *Higher-Dimensional Properties of Non-Uniform Pseudo-Random Variates*. Technical Report, TU Wien.
- Lindstrom, M. J. and Bates, D. M. (1988) Newton-Raphson and EM Algorithms for Linear Mixed-Effects Models for Repeated-Measures Data. *J. Americ. Statist. Assoc.*, **83**, 1015–1022.
- Searle, S. R., Casella, G. and McCulloch, C. E. (1992) *Variance Components*. John Wiley & Sons, New York.
- Ulrich, B. (1981) Theoretische Betrachtung des Ionenkreislaufs in Waldökosystemen. *Z. Pflanzenernährung Bodenkd.*, **144**, 647–659.
- Wermuth, N. (1980) Linear Recursive Equations, Covariance Selection, and Path Analysis. *J. Americ. Statist. Assoc.*, **75**, 963–972.

- Whittle, P. (1953) The Analysis of Multiple Stationary Time Series. *Journal of the Royal Statistical Society B*, **15**, 125–139.
- Wolfinger, R., Tobias, R. and Sall, J. (1994) Computing Gaussian Likelihoods and Their Derivatives for General Linear Mixed Models. *SIAM J. Sci. Comput.*, **15**, 1294–1310.
- Yau, K. K. W. and McGilchrist, C. A. (1996) Simulation Study of the GLMM Method Applied to the Analysis of Clustered Survival Data. *J. Statist. Comput. Simul.*, **55**, 189–200.

A. RESULTS ON DISTRIBUTIONS

Proof of Proposition 1 If we only look at equation u and condition for $u=2$ on the observations of part 1, then we have an usual Gauss-Markov model. ML estimators are the same as in this Gauss-Markov model. Hence, (a) and (b) are standard results from the theory of linear models.

For $u=2$ we get analogous statements for $(\hat{\beta}'_2, \hat{\alpha})'$ and $\hat{\sigma}_2^2$ conditional on the observations of $u=1$ with a number of degrees of freedom which is diminished by one (which already proves f). Some matrix manipulations show that the conditional distributions are

$$\hat{\alpha}|_{z_1=z_1} \sim N\left(\alpha, \frac{\sigma_2^2}{z'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)' \mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2)z_{1,2}}\right) a.s.$$

$$\hat{\beta}_2|_{z_1=z_1} \sim N\left(\beta_2, \sigma_2^2 \left[(\tilde{\mathbf{X}}_2' \mathbf{V}_2^{-1} \tilde{\mathbf{X}}_2)^{-1} + \frac{\tilde{\mathbf{A}}_2 z_{1,2} z'_{1,2} \tilde{\mathbf{A}}_2'}{z'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) z_{1,2}} \right] \right) a.s.,$$

so we have (c) and (d). Since the conditional distribution of $\hat{\sigma}_2^2$ does not depend on the observations of $u=1$, (e) is proved, and (f) is unconditionally valid, too. ■

Proof of Proposition 2 The estimators in the first stage are the same as the ML estimators in the separate parameter case, thus (b), (c) and (d) follow in the same way as in the Proof of Proposition 1.

For proving (a) and (e) we first state $(\mathbf{I} - \tilde{\mathbf{P}}_u)\mathbf{X}_u = \mathbf{0}$.

For $u=1$, this follows from $(\mathbf{I} - \tilde{\mathbf{P}}_1)\mathbf{X}_1 = (\mathbf{I} - \tilde{\mathbf{X}}_1(\tilde{\mathbf{X}}_1' \mathbf{V}_1^{-1} \tilde{\mathbf{X}}_1)^{-1} \tilde{\mathbf{X}}_1' \mathbf{V}_1)(\tilde{\mathbf{X}}_1, \mathbf{0}) = (\tilde{\mathbf{X}}_1, \mathbf{0}) - \tilde{\mathbf{X}}_1(\mathbf{I}, \mathbf{0}) = \mathbf{0}$. For $u=2$, this is also true since \mathbf{X}_2 has the same columns as $(\tilde{\mathbf{X}}_2, \mathbf{0})$, but in another order. Hence we get $(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{X}_2$ from $(\mathbf{I} - \tilde{\mathbf{P}}_2)(\tilde{\mathbf{X}}_2, \mathbf{0})$ by ordering the columns. The latter, however, is equal to $\mathbf{0}$ using the same argument as for $u=1$.

Hence we get

$$\begin{aligned}
 & \text{Cov} \left[\begin{pmatrix} \mathbf{I} - \tilde{\mathbf{P}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \tilde{\mathbf{P}}_2 \end{pmatrix} \mathbf{Z}(\alpha), \mathbf{AZ}(\alpha) \right] & (20) \\
 &= \begin{pmatrix} \mathbf{I} - \tilde{\mathbf{P}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \tilde{\mathbf{P}}_2 \end{pmatrix} \sigma_1^2 \mathbf{V} \mathbf{A}' \\
 &= \sigma_1^2 \begin{pmatrix} \mathbf{I} - \tilde{\mathbf{P}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - \tilde{\mathbf{P}}_2 \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \\
 &= \mathbf{0}. & (21)
 \end{aligned}$$

Moreover, doing some straightforward matrix manipulations we get

$$\tilde{\alpha} = \alpha + \frac{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_2(\alpha)}{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_{1,2}} a.s. \tag{22}$$

$$\check{\beta} = \mathbf{AZ}(\alpha) - (\tilde{\alpha} - \alpha) \mathbf{A}_2 \mathbf{Z}_{1,2} \tag{23}$$

$$N_2 \check{\sigma}_2^2 = \mathbf{Z}_2(\alpha)' \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_2(\alpha) - \frac{[\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_2(\alpha)]^2}{\mathbf{Z}'_{1,2} \mathbf{V}_2^{-1} (\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_{1,2}} a.s. \tag{24}$$

Now, for Gaussian errors we see from (21) that $\check{\sigma}_1^2$ and

$$\mathbf{AZ}(\alpha) \sim N(\beta, \sigma_1^2 (\mathbf{X} \mathbf{V}^{-1} \mathbf{X})^{-1}) \tag{25}$$

are independent as $\check{\sigma}_1^2$ is a function of $(\mathbf{I} - \tilde{\mathbf{P}}_1) \mathbf{Z}_1$. We see from (22) and (24) that, given $\mathbf{Z}_1 = \mathbf{z}_1$, both $\tilde{\alpha}$ and $\check{\sigma}_2^2$ are functions of $(\mathbf{I} - \tilde{\mathbf{P}}_2) \mathbf{Z}_2(\alpha)$ only, which is conditionally independent from $\mathbf{A}_2 \mathbf{Z}_2(\alpha)$ given \mathbf{Z}_1 because of (21). Thus $\mathbf{AZ}(\alpha)$ and $(\tilde{\alpha}, \check{\sigma}_2^2)$ are conditionally independent given \mathbf{Z}_1 . Since $\tilde{\alpha}$ and $\check{\sigma}_2^2$ are also conditionally independent given \mathbf{Z}_1 and $\mathbf{AZ}(\alpha) = \mathbf{A}_1 \mathbf{Z}_1 + \mathbf{A}_2 \mathbf{Z}_2(\alpha)$ we have (e).

From (23) and part (c) we get

$$\begin{aligned}
 E(\check{\beta} | \mathbf{Z}_1) &= E(\mathbf{AZ}(\alpha) | \mathbf{Z}_1) - E(\tilde{\alpha} - \alpha | \mathbf{Z}_1) \mathbf{A}_2 \mathbf{Z}_{1,2} \\
 &= \beta \quad a.s.,
 \end{aligned}$$

implying the marginal unbiasedness of $\check{\beta}$ stated in part (a).

From (23) the covariance matrix of $\check{\beta}$ is

$$E(\check{\beta}\check{\beta}') = E(\mathbf{AZ}(\alpha)\mathbf{Z}(\alpha)'\mathbf{A}') + E((\check{\alpha} - \alpha)^2\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2) \\ - E(\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}(\alpha)'\mathbf{A}'(\check{\alpha} - \alpha)) - E(\mathbf{AZ}(\alpha)\mathbf{Z}'_{1,2}\mathbf{A}'_2(\check{\alpha} - \alpha)) \quad (26)$$

The last two terms equal $\mathbf{0}$ as

$$[\mathbf{Z}'_{1,2}\mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_{1,2}] \cdot E(\mathbf{AZ}(\alpha)\mathbf{Z}'_{1,2}\mathbf{A}'_2(\check{\alpha} - \alpha)|\mathbf{Z}_1) \\ = E([\mathbf{A}_1\mathbf{Z}_1 + \mathbf{A}_2\mathbf{Z}_2(\alpha)] \\ \cdot [\mathbf{Z}'_{1,2}\mathbf{A}'_2\mathbf{Z}'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_2(\alpha)|\mathbf{Z}_1]) \\ = \mathbf{A}_1\mathbf{Z}_1\mathbf{Z}'_{1,2}\mathbf{A}'_2\mathbf{Z}'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1}E[(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_2(\alpha)|\mathbf{Z}_1] \\ + \text{Cov}[\mathbf{A}_2\mathbf{Z}_2(\alpha), (\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1})(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_2(\alpha)|\mathbf{Z}_1] \\ = \mathbf{0}$$

because of (21). For the first two terms we have

$$E(\mathbf{AZ}(\alpha)\mathbf{Z}(\alpha)'\mathbf{A}') = \sigma_1^2\mathbf{A}\mathbf{V}\mathbf{A}' + \beta\beta' \\ = \sigma_1^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} + \beta\beta', \\ E((\check{\alpha} - \alpha)^2\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2) = E(E([\check{\alpha} - \alpha]^2|\mathbf{Z}_1) \cdot \mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2) \\ = E(\text{Var}(\check{\alpha}|\mathbf{Z}_1) \cdot \mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2) \\ = \sigma_2^2 E\left(\frac{\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2}{\mathbf{Z}'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_{1,2}}\right)$$

The numerator is a function of $\mathbf{A}_2\mathbf{Z}_{1,2}$ and the denominator is a function of $(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_{1,2}$. Because of $\mathbf{A}_2(\mathbf{I} - \tilde{\mathbf{P}}_2) = \mathbf{0}$ they are independent and we can calculate their expectations separately:

$$E(\mathbf{A}_2\mathbf{Z}_{1,2}\mathbf{Z}'_{1,2}\mathbf{A}'_2) = \mathbf{A}_2(\mathbf{V}_{1,2} + \mathbf{X}_{1,2}\beta\beta'\mathbf{X}'_{1,2})\mathbf{A}_2 \\ E(\mathbf{Z}'_{1,2}(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2)\mathbf{Z}_{1,2}) = \text{trace}[(\mathbf{I} - \tilde{\mathbf{P}}_2)'\mathbf{V}_2^{-1}(\mathbf{I} - \tilde{\mathbf{P}}_2) \\ (\mathbf{V}_{1,2} + \mathbf{X}_{1,2}\beta\beta'\mathbf{X}'_{1,2})]$$

Using $\text{Var}(\check{\beta}) = E(\check{\beta}\check{\beta}')$ and (26) shows the formula for the variance matrix of $\check{\beta}$ because of the unbiasedness of $\check{\beta}$. ■

Copyright of Journal of Statistical Computation & Simulation is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.