LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

INSTITUT FÜR STATISTIK

Andrea Wiencierz, Sonja Greven, Helmut Küchenhoff

Restricted Likelihood Ratio Testing in Linear Mixed Models with General Error Covariance Structure

Technical Report Number 109, 2011
Department of Statistics
University of Munich

http://www.stat.uni-muenchen.de

Restricted Likelihood Ratio Testing in Linear Mixed Models with General Error Covariance Structure

Andrea Wiencierz*, Sonja Greven, Helmut Küchenhoff
Department of Statistics, LMU Munich

Abstract

We consider the problem of testing for zero variance components in linear mixed models with correlated or heteroscedastic errors. In the case of independent and identically distributed errors, a valid test exists, which is based on the exact finite sample distribution of the restricted likelihood ratio test statistic under the null hypothesis. We propose to make use of a transformation to derive the (approximate) test distribution for the restricted likelihood ratio test statistic in the case of a general error covariance structure. The proposed test proves its value in simulations and is finally applied to an interesting question in the field of well-being economics.

Keywords: linear mixed model, penalized splines, likelihood ratio test, correlated errors, generalized least squares, SOEP data, subjective well-being

1 Introduction

We are interested in testing for zero variance components in the context of Linear Mixed Models (LMMs). LMMs constitute a flexible class of statistical models that can be applied to a wide range of problems. These models are suitable to analyze, for example, longitudinal or clustered data (e.g. Verbeke and Molenberghs, 2000) as well as spatial data (e.g. Kammann and Wand, 2003). Moreover, LMMs are established as a useful framework for nonparametric models (e.g. Ruppert, Wand and Carroll, 2003).

Within the LMM framework it is possible to test whether individuals or clusters differ from other observation units in a way that cannot be predicted by the considered covariates, or whether an estimated smooth function is different from a polynomial of a given degree. Such test problems are equivalent to testing whether the variance of the corresponding random effect is different from zero. This is a nonstandard test situation because under

*Correspondence should be addressed to: andrea.wiencierz@stat.uni-muenchen.de

the null hypothesis the tested parameter is on the boundary of the parameter space. A valid Restricted Likelihood Ratio Test (RLRT) for this situation has been established by Crainiceanu and Ruppert (2004), who derived the exact finite sample distribution of the RLRT statistic in the case of an LMM with only one random effect's variance component. For LMMs with more than one additional variance components, there are good approximations of the test distribution based on the exact RLRT distribution (Greven et al., 2008; Scheipl, Greven and Küchenhoff, 2008).

The established test procedure has been derived based on the assumption that the errors are independent and identically distributed (i.i.d.) conditional on the random effects. In many practical settings the i.i.d. assumption is not fulfilled. For example, in a panel data set from a social survey, the observations of each respondent are likely to be autocorrelated conditional on the individual random effect, because observations closer in time are likely to be more similar.

The data situation we have in mind is motivated by the micro-econometric analysis in Wunder et al. (2011). Here, unbalanced panel data on subjective well-being were analyzed with a semiparametric mixed model containing a smooth function in age and an individual random intercept. The data used are from the German Socio-Economic Panel Study (SOEP), which are available as scientific use files at http://www.diw.de/en/soep (Wagner, Frick and Schupp, 2007). The question of interest was whether the smooth function is different from a quadratic or cubic polynomial as is typically used in well-being research. In the LMM framework, this question can be answered in testing whether the corresponding random effect's variance is different from zero. Further analyses of the data set indicated that for each individual the errors are correlated over time.

Therefore, we look for a generalization of the established test allowing for more general covariance structures of the errors. Since the exact RLRT null distribution cannot simply be extended to correlated errors and a bootstrap-based test is computationally too demanding (given the large SOEP data set of more than 250 000 observations), we follow another approach: We use the RLRT statistic and propose to use a transformation to approximate the test distribution employing results on the exact distribution in the case of i.i.d. errors.

The paper is organized as follows: In Section 2 we present the methodological framework and the notation used in this paper. First, we introduce the considered class of models and discuss the problem of testing for zero variance components in the case of i.i.d. errors. After this, we present the idea of the transformation we adapt to take the correlation structure of the errors into account, before we propose the test procedure for the case of errors with general covariance structure. The accuracy of the proposed test is then investigated in a simulation study described in Section 3. Moreover, we apply the suggested test to the motivating question regarding the profile of subjective well-being over age in Section 4. Finally, Section 5 gives a short summary and discussion of our findings.

2 Theoretical Background and Methods

2.1 The Linear Mixed Model

We consider LMMs of the following form:

$$\begin{cases} \boldsymbol{y} & = & \boldsymbol{X\beta} + \boldsymbol{Z}_1\boldsymbol{b}_1 + \ldots + \boldsymbol{Z}_S\boldsymbol{b}_S + \boldsymbol{\epsilon}, \\ \boldsymbol{b}_s & \sim & \mathrm{N}(\boldsymbol{0},\, \sigma_s^2\boldsymbol{\Sigma}_s), \quad s = 1, \ldots, S, \\ \boldsymbol{\epsilon} & \sim & \mathrm{N}(\boldsymbol{0},\, \sigma_\epsilon^2\boldsymbol{R}), \\ \boldsymbol{b}_1, & \ldots, & \boldsymbol{b}_S,\ \boldsymbol{\epsilon} \quad \text{independent.} \end{cases} \tag{1}$$

Here, $\boldsymbol{X}$ and $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_S$ are design matrices with full ranks p and $K_1, \ldots, K_S$, respectively. The vector $\boldsymbol{\beta}$ contains fixed covariate effects, whereas $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_S$ are independent random effect vectors. The correlation matrices associated with the random effects vectors, $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_S$, are assumed known. The matrix $\boldsymbol{R}$ describes an error covariance structure. It is often assumed that the errors are uncorrelated, i.e. $\boldsymbol{R} = \boldsymbol{I}_n$, where $\boldsymbol{I}_n$ is the n-dimensional identity matrix and n is the total number of observations. Since there are many situations in practice in which this assumption is not appropriate, a more general class of models is considered here, allowing any covariance structure for the errors to be specified in $\boldsymbol{R}$.

There are two important special cases of LMMs that are often used in applications: the simple random intercept model and mixed model smoothing. In both special cases, the model contains only one random effect vector $\boldsymbol{b}$ of length K with covariance matrix $\mathrm{Cov}(\boldsymbol{b}) = \sigma_b^2\boldsymbol{\Sigma} = \sigma_b^2\boldsymbol{I}_K$.

The random intercept model can be used to model longitudinal or clustered data where it is assumed that the individual or cluster means deviate randomly from the overall mean (e.g. Verbeke and Molenberghs, 2000),

$$y_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta} + b_i + \epsilon_{ij}.$$

In this case, $\boldsymbol{b} = (b_1, \ldots, b_N)'$ is an $N \times 1$-vector, where N is the number of individuals or clusters.

To express a penalized splines model as an LMM, the basis coefficients of the nonparametric model are split into two groups: those which are shrunk by the penalization imposed during the estimation and the unpenalized coefficients. The penalization of the one group of coefficients can equivalently be expressed by assuming a normal distribution with mean zero for those coefficients. Therefore, the penalized splines model can be formulated as an LMM where the random effects vector $\boldsymbol{b}$ is formed by the penalized basis coefficients. These coefficients represent deviations of the modeled function from a global polynomial of degree d.

For example, consider a penalized splines model with truncated power functions as spline basis functions

$$y_i = f(x_i) + \epsilon_i = \sum_{j=0}^{d} \beta_j x_i^j + \sum_{j=1}^{K} b_j (x_i - \kappa_j)_+^d + \epsilon_i, \quad i = 1, \ldots, n,$$

where $\kappa_1 < \ldots < \kappa_K$ are K knots and $(u)_+^d = (\max\{0,\, u\})^d$. To impose smoothness of the estimated curve, high values of the coefficients $b_1, \ldots, b_K$ are penalized. Then, the estimator of $\boldsymbol{\theta} = (\beta_0, \ldots, \beta_d, b_1, \ldots, b_K)$ is given by:

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{B}'\boldsymbol{B} + \frac{1}{\lambda}\boldsymbol{D})^{-1}\boldsymbol{B}'\boldsymbol{y},$$

where $\boldsymbol{B}$ is the design matrix of the spline basis, λ is the inverse of the so-called smoothing parameter, which controls the amount of shrinkage, and

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_K \end{pmatrix}.$$

This penalized splines model can alternatively be formulated as an LMM, where

$$\sum_{j=0}^{d} \beta_j x_i^j = \boldsymbol{x}_i'\boldsymbol{\beta}, \qquad \text{and} \qquad \sum_{j=1}^{K} b_j (x_i - \kappa_j)_+^d = \boldsymbol{z}_i'\boldsymbol{b},$$

and the smoothing parameter is given by $\frac{1}{\lambda} = \frac{\sigma_\epsilon^2}{\sigma_b^2}$. Estimating a smooth function within the LMM framework has the advantages of higher computational efficiency and of an integrated, automated smoothing parameter selection. Detailed information about penalized splines and mixed model smoothing can be found, for example, in Eilers and Marx (1996) and Ruppert, Wand and Carroll (2003).

2.2 Testing for Zero Variance Components in LMMs

We are interested in testing whether a random effect is different from zero. In the case of a random intercept model for longitudinal data, this means determining whether there are significant individual-specific deviations from the population mean. In the penalized splines context, the tested null hypothesis (H_0) is that the smooth function of interest can adequately be described by a polynomial of a given degree d. Since the random effects have expectation zero, the considered testing situation is equivalent to testing if the variance of that random effect is different from zero. Hence, the test has the structure:

$$H_0 : \sigma_s^2 = 0 \qquad \text{vs.} \qquad H_1 : \sigma_s^2 > 0. \tag{2}$$

The considered testing problem is not regular because first, under the null hypothesis the tested parameter lies on the boundary of the parameter space and second, the data are not i.i.d. because of the particular covariance structure of grouped data or due to correlations induced by penalized splines. Both characteristics violate the typical regularity conditions for likelihood ratio testing, therefore, the asymptotic distribution of the Likelihood Ratio Test statistic (LRT) is not the standard χ^2 distribution.

In the case of longitudinal data, however, the vector of observations $\boldsymbol{y}$ can be subdivided into independent subvectors. Considering an increasing number of subvectors, Stram and Lee (1994) derived an asymptotic distribution of the LRT, applying results of Self and Liang (1987). This asymptotic distribution is an equal mixture of two χ^2 distributions. As shown by Morrell (1998), their results also apply to the RLRT statistic, which is the test statistic based on Restricted Maximum Likelihood (REML) estimation of the variance components.

If we consider LMMs of a more general form like, for example, a penalized splines model, such a subdivision of the data is not possible. Therefore, the asymptotic χ^2 mixture distribution does not apply to this case. Crainiceanu and Ruppert (2004) found the exact finite sample distribution of the LRT and of the RLRT in the case of an LMM with only one random effect variance component and i.i.d. errors. They also derived asymptotic distributions of both statistics in this situation. The distribution of the LRT under the null hypothesis can have a very high mass on zero; thus, testing on the basis of the LRT may be impracticable. That is why we will focus on the RLRT in the remainder of the paper.

Consider a test of the given structure for an LMM with only one random effect (vector) and i.i.d. errors. In this case, the exact RLRT distribution under H_0 can be expressed (with $\lambda = \frac{\sigma_b^2}{\sigma_\epsilon^2}$) as (Crainiceanu and Ruppert, 2004):

$$RLRT_n \stackrel{d}{=} \sup_{\lambda \geq 0} \left\{ (n-p) \log \left[1 + \frac{N_n(\lambda)}{D_n(\lambda)} \right] - \sum_{l=1}^{K} \log(1 + \lambda \mu_{l,n}) \right\}, \tag{3}$$

where

$$N_n(\lambda) = \sum_{l=1}^{K} \frac{\lambda \mu_{l,n}}{1 + \lambda \mu_{l,n}} \omega_l^2, \qquad D_n(\lambda) = \sum_{l=1}^{K} \frac{\omega_l^2}{1 + \lambda \mu_{l,n}} + \sum_{l=K+1}^{n-p} \omega_l^2,$$

and $\stackrel{d}{=}$ means equality in distribution. Here, $\mu_{l,n}$, $l = 1, \ldots, K_s$, are the eigenvalues of the matrix $\boldsymbol{\Sigma}_s^{\frac{1}{2}} \boldsymbol{Z}_s' (\boldsymbol{I_n} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}') \boldsymbol{Z}_s \boldsymbol{\Sigma}_s^{\frac{1}{2}}$ and $\omega_l \stackrel{iid}{\sim} \mathrm{N}(0, 1)$, $l = 1, \ldots, n-p$. This distribution only depends on the design matrices of the fixed and random effects, $\boldsymbol{X}$ and $\boldsymbol{Z}_s$, and on the correlation structure within the random effects vector, $\boldsymbol{\Sigma}_s$.

Critical values or p-values of the distribution of $RLRT_n$ can be determined efficiently by simulation, which is implemented in the RLRsim package (Scheipl, 2008) for the statistical software environment R (R Development Core Team, 2011). In case of nuisance variance

components, the above RLRT distribution is a well-performing approximation of the exact test distribution (Greven et al., 2008; Scheipl, Greven and Küchenhoff, 2008).

The previously mentioned results are based on the assumption $\boldsymbol{R} = \boldsymbol{I}_n$. We are interested in a more general setting where the covariance matrix of the errors can be of any possible structure, including heteroscedastic and correlated errors. Since the existing test cannot directly be extended to a general $\boldsymbol{R}$, we suggest to follow the idea of the Generalized Least Squares (GLS) transformation to derive the null distribution of the RLRT in this situation.

2.3 Regression with Correlated Errors

The errors in a simple linear regression model are usually assumed to be i.i.d. with expectation zero and fixed variance σ_ϵ^2. In this situation, the Best Linear Unbiased Estimator (BLUE) of $\boldsymbol{\beta}$ is the Ordinary Least Squares (OLS) estimator, $\hat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$, which is equivalent to the ML estimator in the case of normality.

If the i.i.d. assumption is not fulfilled, the OLS estimator is still unbiased but no longer efficient. In case the covariance matrix of the errors $\mathrm{Cov}(\boldsymbol{\epsilon}) = \boldsymbol{V}$ is known, one can obtain the BLUE as the GLS estimator:

$$\hat{\boldsymbol{\beta}}_{GLS} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y},$$

which is again the ML estimator, under normality. The GLS estimate of $\boldsymbol{\beta}$ can also be obtained as the OLS estimate of a transformed model, where the initial model equation is premultiplied by $\boldsymbol{V}^{-\frac{1}{2}}$, i.e. $\hat{\boldsymbol{\beta}}_{GLS}$ is equivalent to $\hat{\tilde{\boldsymbol{\beta}}}_{OLS}$ of the model

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\epsilon}},$$

where $\tilde{\boldsymbol{y}} = \boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{y}$, $\tilde{\boldsymbol{X}} = \boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{X}$ and $\tilde{\boldsymbol{\epsilon}} = \boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{\epsilon}$. In the transformed model, the error covariance matrix is given by $\mathrm{Cov}(\tilde{\boldsymbol{\epsilon}}) = \boldsymbol{I}_n$. Furthermore, if the covariance matrix of the errors is specified by $\boldsymbol{V} = \sigma_\epsilon^2\boldsymbol{R}$, assuming the covariance structure is known up to a constant, it is possible to base the so-called GLS transformation only on the matrix $\boldsymbol{R}$. Then, the variance of the transformed i.i.d. errors is σ_ϵ^2 instead of one.

If the covariance structure is unknown, so-called Feasible GLS (FGLS) estimators can be used. Many of those are obtained from two-step estimation methods where the parameters of the covariance matrix are estimated in a first step. Then, the data are transformed by the inverse root of the estimated covariance matrix, before OLS estimation is finally applied to the modified data set. FGLS estimators are consistent and asymptotically normal, provided certain regularity conditions are fulfilled (Newey and McFadden, 1994; Wooldridge, 2001, ch. 12). Furthermore, comparative simulation studies indicate that consistent FGLS estimators are more efficient than the OLS estimator, in the situation of linear regressions

with correlated errors as well as in the case of simple variance component models (Rao and Griliches, 1969; Baltagi and Chang, 1994).

In the context of LMMs, parameter estimates are efficiently obtained via (RE)ML estimation. Thus, we are not interested in (F)GLS estimation here, but we adopt the idea of (F)GLS estimation to construct a test for zero variance components of an LMM in the case of a general error covariance structure.

2.4 RLRT in LMMs with General Error Covariance Structure

2.4.1 Known Covariance Structure of the Errors

We use the idea of the GLS transformation to derive the RLRT distribution in the case of a general covariance structure of the errors, i.e. $\boldsymbol{V} = \sigma_\epsilon^2 \boldsymbol{R}$ for some $\boldsymbol{R}$.

Consider first the LMM (1) with one random effect, $S = 1$, $\sigma_S^2 = \sigma_b^2$, and assume that $\boldsymbol{R}$ and $\boldsymbol{\Sigma}$ are known. Model (1) then reduces to

$$\boldsymbol{y} \;=\; \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{\epsilon}; \quad \begin{pmatrix} \boldsymbol{b} \\ \boldsymbol{\epsilon} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \sigma_b^2 \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_\epsilon^2 \boldsymbol{R} \end{pmatrix} \right). \tag{4}$$

We can show that the RLRT for testing (2) in (4) is equivalent to the RLRT for hypothesis (2) in the GLS-transformed model

$$\widetilde{\boldsymbol{y}} \;=\; \widetilde{\boldsymbol{X}}\boldsymbol{\beta} + \widetilde{\boldsymbol{Z}}\boldsymbol{b} + \widetilde{\boldsymbol{\epsilon}}; \quad \begin{pmatrix} \boldsymbol{b} \\ \widetilde{\boldsymbol{\epsilon}} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \sigma_b^2 \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_\epsilon^2 \boldsymbol{I}_n \end{pmatrix} \right), \tag{5}$$

where $\widetilde{\boldsymbol{y}} = \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{y}$, $\widetilde{\boldsymbol{X}} = \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{X}$, $\widetilde{\boldsymbol{Z}} = \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{Z}$ and $\widetilde{\boldsymbol{\epsilon}} = \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{\epsilon}$.

Lemma 1 *The RLRT for hypothesis* (2) *in* (4) *is equivalent to the RLRT for hypothesis* (2) *in the GLS-transformed model* (5).

A proof is given in the appendix. As model (5) has independent and identically distributed errors, the exact RLRT distribution for testing (2) in (4) is thus given by (3) (Crainiceanu and Ruppert, 2004), with the design matrixes $\boldsymbol{X}$ and $\boldsymbol{Z}$ replaced by $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{Z}}$.

2.4.2 Unknown Covariance Structure of the Errors

Now consider the more general case where $\boldsymbol{R} = \boldsymbol{R}(\boldsymbol{\phi})$ may depend on an unknown parameter vector $\boldsymbol{\phi}$, which is estimated simultaneously with the other parameters using REML estimation.

We propose to use the RLRT distribution derived in 2.4.1 as an approximation in this case, substituting the estimate $\widehat{\boldsymbol{R}} = \boldsymbol{R}(\widehat{\boldsymbol{\phi}})$ from the null model to obtain the transformed design matrices $\widetilde{\boldsymbol{X}} = \widehat{\boldsymbol{R}}^{-\frac{1}{2}}\boldsymbol{X}$ and $\widetilde{\boldsymbol{Z}} = \widehat{\boldsymbol{R}}^{-\frac{1}{2}}\boldsymbol{Z}$. Under the null hypothesis, $\widehat{\boldsymbol{\phi}}$ is a consistent

estimator of $\boldsymbol{\phi}$, and this substitution should thus be negligible asymptotically. We investigate the goodness of the approximation for finite sample sizes in the next section, focusing in particular on the observance of the type 1 error of the resulting test.

For models with more than one random effect, $S > 1$, we employ the approximation proposed in Greven et al. (2008), which uses the RLRT in a reduced model with $S = 1$, including only the relevant random effect. The goodness of our approximation for this case is likewise analyzed in the next section.

3 Simulation Study

We set up a simulation study to investigate the performance of the proposed approximation of the RLRT distribution in the case of correlated errors. Motivated by the economic analysis of subjective well-being mentioned in section 1, we consider a longitudinal data set where the errors of each individual follow a first order autoregressive process (AR(1)). The simulation setup covers three different settings. All simulations are conducted in the statistical software environment R (R Development Core Team, 2011). The proposed test is then evaluated by means of simulation-based type-one error rates.

3.1 Simulation Setup

For our simulation study we consider special cases of the following general model:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_1\boldsymbol{b}_1 + \boldsymbol{Z}_2\boldsymbol{b}_2 + \boldsymbol{\epsilon}.$$

The design matrix of the fixed effects, $\boldsymbol{X}$, contains terms of a cubic polynomial in x, $\boldsymbol{Z}_1$ is the design matrix of a random intercept, and $\boldsymbol{Z}_2$ contains cubic truncated polynomial terms completing a penalized splines model for $f(x)$. Finally, $\boldsymbol{\epsilon}$ contains the residuals that are autocorrelated for each individual. We investigate three different test situations:

1. $H_0 : \sigma_1^2 = 0 \quad$ vs. $\quad H_1 : \sigma_1^2 > 0, \quad$ where $\sigma_2^2 = 0$,
   i.e. testing the random intercept in a simple LMM.

2. $H_0 : \sigma_2^2 = 0 \quad$ vs. $\quad H_1 : \sigma_2^2 > 0, \quad$ where $\sigma_1^2 = 0$,
   i.e. testing for a cubic polynomial vs. a smooth function in an LMM designed for smoothing.

3. $H_0 : \sigma_2^2 = 0 \quad$ vs. $\quad H_1 : \sigma_2^2 > 0, \quad$ where $\sigma_1^2 > 0$,
   i.e. testing the variance of the penalized smoothing coefficients in the presence of a nuisance random intercept.

We vary the number of individuals $N \in \{20, 100, 500\}$, the number of observations per individual $n_i \in \{4, 10, 20\}$ as well as the parameter of autocorrelation $\rho \in \{0, 0.4, 0.8\}$. We run 10 000 simulations for each of the 27 scenarios defined by combinations of the varied parameters in each of the three test situations. The data are generated according to the corresponding null model of the test situation. Motivated by our application we set $\boldsymbol{\beta} = (11, -0.7, 0.03, -0.0003)'$ and $\sigma_1^2 = \sigma_\epsilon^2 = 1$. For the penalized splines model, we choose a cubic truncated power function basis with 39 inner knots, which represent the $2.5\,j\%$ quantiles, $j = 1, \ldots, 39$, of the empirical distribution of the simulated x-values.

In each simulation step, both null and alternative model are estimated using the functions $lme()$ and $gls()$ of the R package nlme (Pinheiro and Bates, 2008) with option REML. After this, the RLRT statistic is computed. Then, the GLS transformation matrix is constructed using the estimate $\hat{\rho}_0$ from the null model as correlation parameter ρ. In the considered AR(1) case, the transformation matrix $\boldsymbol{R}^{-\frac{1}{2}}$ is block diagonal with N blocks $\boldsymbol{R}_i^{-\frac{1}{2}}$ of size $n_i \times n_i$:

$$\boldsymbol{R}_i^{-\frac{1}{2}} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ \frac{-\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} & 0 & \ldots & 0 \\ 0 & \frac{-\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \frac{-\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix}.$$

Finally, the design matrices of the null model are transformed and used to obtain the 95% quantile of the RLRT distribution through the RLRTSim() function of the RLRsim package (Scheipl, 2008), and the test decision is taken.

3.2 Results

Table 1 lists the percentages of rejections of the null hypothesis in each scenario. We observe values ranging from 4.57% to 5.91%, which are all reasonably close to 5%. Thus, the estimation uncertainty of the correlation parameter does not considerably affect the accuracy of the test. Furthermore, there is no systematic variation of the rejection rate across the simulated scenarios besides a general slight improvement of accuracy with larger sample sizes n. We conclude that the proposed approximation of the exact RLRT distribution performs sufficiently well.

Scenarios			Test situation		
Parameters			1	2	3
ρ	N	n_i	% rejected	% rejected	% rejected
0.0	20	4	4.86	4.69	5.82
0.0	20	10	5.19	4.96	5.49
0.0	20	20	4.95	5.50	5.07
0.0	100	4	4.85	4.72	5.17
0.0	100	10	4.73	5.03	5.07
0.0	100	20	5.11	4.79	5.30
0.0	500	4	4.82	5.11	4.83
0.0	500	10	4.75	5.30	4.92
0.0	500	20	5.08	5.55	4.96
0.4	20	4	5.26	5.17	5.51
0.4	20	10	5.00	5.21	5.46
0.4	20	20	5.06	5.14	5.10
0.4	100	4	5.31	5.21	5.09
0.4	100	10	4.70	5.31	4.77
0.4	100	20	5.04	5.27	5.69
0.4	500	4	4.84	4.87	5.07
0.4	500	10	5.17	5.05	4.71
0.4	500	20	4.63	5.10	4.97
0.8	20	4	5.58	5.10	5.77
0.8	20	10	5.55	5.91	5.57
0.8	20	20	5.18	5.63	5.34
0.8	100	4	4.57	4.96	5.11
0.8	100	10	4.92	5.61	5.00
0.8	100	20	4.81	5.19	5.36
0.8	500	4	5.15	4.81	4.91
0.8	500	10	5.05	4.84	4.99
0.8	500	20	5.14	5.04	5.36

Table 1: Percentages of rejected tests out of 10 000 simulated tests per scenario.

4 Application: Subjective well-being in Germany

4.1 Background

The work presented here was motivated by an econometric study about subjective well-being (Wunder et al., 2011). One crucial question in this field is how the level of overall life satisfaction evolves with age. Is there a midlife crisis? Does the level of life satisfaction decline when growing old?

In the literature on well-being, it is often assumed that the relationship between life satisfaction and age can be described by a quadratic function. Yet empirical analyses using

models that are quadratic in age find contradictory results: both a U-shaped profile (e.g. Blanchflower and Oswald, 2004) and a reverse U-shaped profile (e.g. Easterlin, 2006) have been reported. Since either shape lacks a convincing foundation in the economic theory, Wunder et al. (2011) argue that the profiles found in these studies are an artefact of the quadratic modeling and suggest to follow a semiparametric approach instead.

Wunder et al. (2011) base their analyses on data from the German SOEP and the British Household Panel Survey (BHPS), two suitable data sources for studies on well-being (Wagner, Frick and Schupp, 2007). For each data set, they estimate an additive mixed effects model including a smooth term in age, several socio-economic control variables and a random intercept for each individual. The estimation of the smooth functions is based on the LMM representation of penalized splines. The resulting profiles appear to be clearly different from a quadratic curve. Moreover, likelihood-based tests and Bonferroni-adjusted confidence intervals indicate that, even though an LMM with a cubic function in age fits the data better than a quadratic one, this parametrization is not flexible enough to adequately describe the entire variability within the data sets.

In their models, Wunder et al. (2011) assume that the errors are i.i.d. conditional on the random intercept. However, individual values of self-reported life satisfaction are likely to be correlated over time. Further investigations indeed confirmed a superior fit for a model with AR(1) errors for each respondent. Thus, the test based on the i.i.d. assumption for the errors is only a rough approximation. Here, we analyze the SOEP data set again, explicitely accounting for the covariance structure of the individual error terms. We fit models including errors that follow an AR(1) process for each respondent and then apply the appropriate test developed in Section 2.

4.2 Analysis of the SOEP data

For the SOEP Study, overall life satisfaction is measured on a discrete scale from 0 (not satisfied) to 10 (completely satisfied). The analyzed data set contains observations of 20 waves between 1986 and 2007. The years 1990 and 1993 are excluded, because one relevant control variable is not available for these years. Altogether, there are 253 044 data points obtained from 33 451 persons. Since the data set is sufficiently large, we can assume an approximate normal distribution for the life satisfaction and apply the LMM framework.

The aim of our analysis is to reveal the relationship between age and self-reported well-being, controlling for relevant socio-economic characteristics and at the same time accounting for the correlation structure of the observations. For this purpose, we test the null hypothesis that the profile of well-being over the life span is adequately described by a cubic polynomial against a smooth alternative with an RLRT. This test has the same structure as the test in the third situation of the simulation study in section 3.

We consider the following models under the null and under the alternative hypotheses:

$$H_0: \quad \boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_1\boldsymbol{b}_1 + \boldsymbol{\epsilon}, \qquad H_1: \quad \boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_1\boldsymbol{b}_1 + \boldsymbol{Z}_2\boldsymbol{b}_2 + \boldsymbol{\epsilon}.$$

Here, $\boldsymbol{X}$ contains terms of a cubic polynomial in age and additional covariates, which have been chosen in accordance with Wunder et al. (2011). The matrix $\boldsymbol{Z}_1$ is the design matrix of a random intercept and $\boldsymbol{Z}_2$ is composed of cubic truncated polynomial terms of the penalized splines model under H_1. Furthermore, the errors in $\boldsymbol{\epsilon}$ are assumed to be autocorrelated for each individual. A detailed list of the covariates included in $\boldsymbol{X}$ can be found in appendix B, together with further estimation results.

To obtain the test decision, we transform the design matrices $\boldsymbol{X}$ and $\boldsymbol{Z}_2$ with the inverse root of the estimated correlation matrix under the null hypothesis and determine the 95% and 99.9% quantiles of the corresponding RLRT statistic. Since the data set is an unbalanced panel, the transformation matrix is block diagonal with $N = 33\,451$ blocks of different sizes $n_i \times n_i$, $n_i \in \{1, \ldots, 20\}$. Furthermore, for some respondents there are missing observations between two observations. We found that, in this case, the individual transformation matrices $\boldsymbol{R}_i^{-\frac{1}{2}}$ are given by:

$$\boldsymbol{R}_i^{-\frac{1}{2}} = \frac{1}{\sqrt{1-\rho^2}} \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \ldots & 0 \\ \frac{-\rho^{\Delta_2}}{f(\Delta_2)} & \frac{1}{f(\Delta_2)} & 0 & \ldots & 0 \\ 0 & \frac{-\rho^{\Delta_3}}{f(\Delta_3)} & \frac{1}{f(\Delta_3)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \frac{-\rho^{\Delta_{n_i}}}{f(\Delta_{n_i})} & \frac{1}{f(\Delta_{n_i})} \end{pmatrix},$$

where $\Delta_h = t_{i,h} - t_{i,h-1}$, $h \in \{2, \ldots, n_i\}$, is the time difference between two subsequent observations and $f(\Delta_h) = \sqrt{1 + \sum_{r=1}^{\Delta_h - 1} \rho^{2r}}$. Using the explicit form above to obtain the transformation matrix requires much less computational effort than other methods to compute $\boldsymbol{R}^{-\frac{1}{2}}$, which is particularly useful when analyzing a large data set as we do here.

Figure 1 shows both estimated curves. There are considerable differences between the profiles estimated from the null and from the alternative model, in particular between 45 and 70 years of age. In contrast to the curve from the cubic model, the semiparametric curve does reveal a midlife crisis, which is located around the age of 50. Moreover, the following temporary recovery before the final decline of life satisfaction is not reflected by the curve estimated from the null model, which is due to the limited flexibility of a cubic polynomial. In the corresponding RLRT, the null hypothesis is rejected at both levels, 95% and 99.9% (i.e. the p-value is less than 0.001). Hence, the test confirms that the profile of well-being over the life span has a complex shape that cannot be captured by a polynomial of a small degree. Neither a U-shaped or reverse U-shaped profile nor a cubic function are suitable to describe the course of life satisfaction along the aging process.

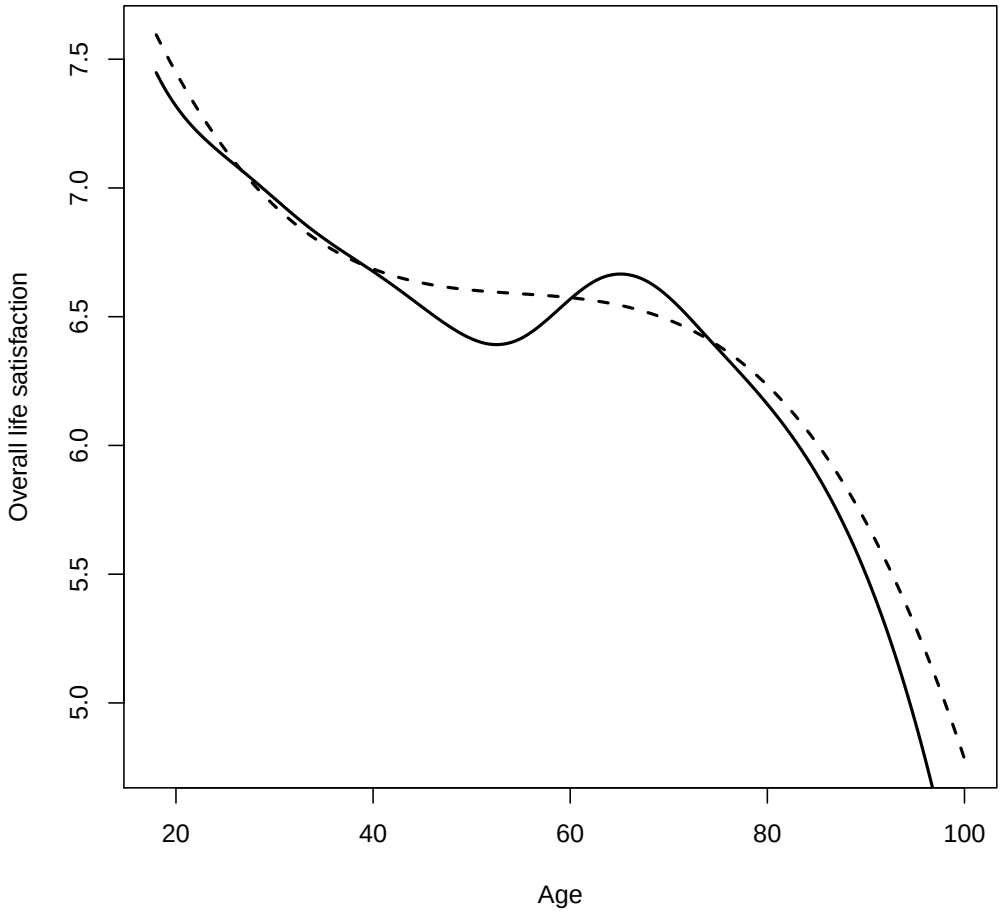

Figure 1: Profile of the overall life satisfaction along the aging process: Curves estimated from the null model (dashed line) and from the alternative model (solid line).

5 Discussion

In the present paper, we established a test for zero variance components in LMMs where the errors are not i.i.d. We used the idea of the GLS transformation to derive the RLRT distribution in the case of a general covariance structure of the errors. If the covariance structure of the errors is unknown and has to be estimated, we proposed to approximate the RLRT distribution by transforming the model according to the covariance parameters estimated from the considered null model.

We set up a simulation study to investigate the performance of the proposed approximation of the RLRT distribution. Since the type one error rates of the simulated tests were reasonably close to 5%, we concluded that the approximation performs sufficiently well. Finally, we applied our test methodology to a micro-econometric analysis on subjective well-being. The test provides an answer to one crucial question of well-being economics:

The profile of overall life satisfaction over the life span has a complex shape, that cannot be captured by a polynomial of a small degree. First, there is a steady decline up to a midlife crisis around the age of 50, then, well-being temporarily rises again, before finally, it declines again from the early sixties on.

The proposed method can be applied to a wide range of practical settings where correlated or heteroscedastic errors occur. An interesting future extension could be to generalize the test methodology to Generalized Linear Mixed Models (GLMMs).

References

Baltagi, B. H. and Chang, Y.-J. (1994). Incomplete panels. A comparative study of alternative estimators for the unbalanced one-way error component regression model. *Journal of Econometrics* **62**, 67–89.

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being over time in britain and the usa. *Journal of Public Economics 88*(7-8), 1359 – 1386.

Crainiceanu, C. and Ruppert, D. (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 66*(1), 165–185.

Easterlin, R. A. (2006). Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology 27*(4), 463 – 482.

Eilers, P. H. C. and Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties. *Statistical Science 11*(2), 89–121.

Greven, S., Crainiceanu, C. M., Küchenhoff, H., and Peters, A. (2008). Restricted likelihood ratio testing for zero variance components in linear mixed models. *Journal of Computational and Graphical Statistics 17*(4), 1–22.

Harville, D. A. (1974). Bayesian inference for variance components using only error contrasts. *Biometrika 61*(2), 383–385.

Kammann, E. E. and Wand, M. P. (2003). Geoadditive models. *Applied Statistics* **52**, 1–18.

Morrell, C. H. (1998). Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood. *Biometrics 54*(4), 1560–1568.

Newey, W. K. and McFadden, D. (1994). Large Sample Estimation and Hypothesis Testing. Volume 4 of *Handbook of Econometrics*, Chapter 36, pp. 2111–2245. Elsevier.

Patterson, H. and Thompson, R. (1971). Recovery of interblock information when the block sizes are unequal. *Biometrika 58*(3), 545–554.

Pinheiro, J. and Bates, D. (2008). *nlme: Linear and Nonlinear Mixed Effects Models*. R package nlme 3.1-89.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna,. Used R versions 2.8.1, 2.10.0 and 2.12.1.

Rao, P. and Griliches, Z. (1969). Small-Sample Properties of Several Two-Stage Regression Methods in the Context of Auto- Correlated Errors. *Journal of the American Statistical Association 64*(325), 253–272.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge: Cambridge University Press.

Scheipl, F. (2008). *RLRsim: Exact (Restricted) Likelihood Ratio tests for mixed and additive models*. R package RLRsim 2.0-2.

Scheipl, F., Greven, S., and Küchenhoff, H. (2008). Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. *Computational Statistics and Data Analysis* **52**, 3283–3299.

Self, S. G. and Liang, K.-Y. (1987). Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. *Journal of the American Statistical Association 82*(398), 605–610.

Stram, D. O. and Lee, J. W. (1994). Variance Components Testing in the Longitudinal Mixed Effects Model. *Biometrics 50*(4), 1171–1177.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.

Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel Study (SOEP) — Scope, Evolution and Enhancements. Volume 127 of *Schmollers Jahrbuch*, pp. 139–169. Duncker and Humblot.

Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Wunder, C., Wiencierz, A., Schwarze, J., and Küchenhoff, H. (2011). Well-Being over the Life Span: Semiparametric Evidence from British and German Longitudinal Data. *Review of Economics and Statistics*. Accepted for publication.

A Proof of Lemma 1

The restricted likelihood for model (4) is the probability density function (pdf) for $\boldsymbol{A}_1\boldsymbol{y}$, where $\boldsymbol{A}_1\boldsymbol{y}$ define error contrasts (Patterson and Thompson, 1971). The restricted likelihood is - up to a constant - independent of the precise error contrast chosen (Harville, 1974). A suitable matrix $\boldsymbol{A}_1$ is, for example, $\boldsymbol{A}_1 = \boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}^{-1}$.

The restricted likelihood for model (5) is the pdf for $\boldsymbol{A}_2\widetilde{\boldsymbol{y}}$, where $\boldsymbol{A}_2\widetilde{\boldsymbol{y}}$ again define suitable error contrasts, for example, $\boldsymbol{A}_2 = \boldsymbol{I}_n - \widetilde{\boldsymbol{X}}(\widetilde{\boldsymbol{X}}'\widetilde{\boldsymbol{X}})^{-1}\widetilde{\boldsymbol{X}}'$. It is

$$\begin{aligned}\boldsymbol{A}_2\widetilde{\boldsymbol{y}} &= (\boldsymbol{I}_n - \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}^{-\frac{1}{2}})\boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{y} \\ &= \boldsymbol{R}^{-\frac{1}{2}}(\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{R}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}^{-1})\boldsymbol{y} = \boldsymbol{R}^{-\frac{1}{2}}\boldsymbol{A}_1\boldsymbol{y}.\end{aligned}$$

Thus, $\boldsymbol{A}_2\widetilde{\boldsymbol{y}}$ is a simple variable transformation of $\boldsymbol{A}_1\boldsymbol{y}$, and the pdf of $\boldsymbol{A}_2\widetilde{\boldsymbol{y}}$ is equal to the pdf of $\boldsymbol{A}_1\boldsymbol{y}$ up to a multiplicative constant, the Jacobian $|\boldsymbol{R}^{-\frac{1}{2}}|$.

For the two restricted likelihood ratio test statistics $RLRT_1$, based on the profile restricted log-likelihood $\ell_1(\lambda)$ for model (4), and $RLRT_2$, based on the profile restricted log-likelihood $\ell_2(\lambda)$ for model (5), where $\lambda = \frac{\sigma_b^2}{\sigma_\epsilon^2}$, we have

$$\begin{aligned}RLRT_1 &= \sup_{\lambda\geq 0}\{2\ell_1(\lambda) - 2\ell_1(0)\} = \sup_{\lambda\geq 0}\{(2\ell_2(\lambda) - \log|\boldsymbol{R}|) - (2\ell_2(0) - \log|\boldsymbol{R}|)\} \\ &= \sup_{\lambda\geq 0}\{2\ell_2(\lambda) - 2\ell_2(0)\} = RLRT_2.\end{aligned}$$

B List of covariates and further estimation results

Covariate	Variable Description
Intercept	global constant
Age	current age of the respondent
Sex: Female	indicator for sex, reference category: male
Disability status: Disabled	indicator for disability, reference category: not disabled
Nights stayed in Hospital	number of nights spent in hospital during the last year
Years of Education	number of years during which education was acquired
log of net houshold income	allows to control for log of net houshold income per person
log of household size	
Nationality: German	indicator of nationality, reference category: not German
Full time employed	employment status, reference category: vocational training or self-employed
Part time employed	
Unemployed	
Single	marital status, reference category: married
Divorced	
Widowed	
Region: West-Germany	indicator for the region in which the respondent lives, reference category: East-Germany
Life expectation	average life expectancy at birth
Attrition in $t+1$	indicator for attrition in $t+1$, reference category: not retired from the panel in $t+1$
Attrition in $t+2$	indicator for attrition in $t+2$, reference category: not retired from the panel in $t+2$
Attrition in $t+3$	indicator for attrition in $t+3$, reference category: not retired from the panel in $t+3$

Table 2: List of covariates considered in the model. The design matrix of the fixed effects, $\boldsymbol{X}$, furthermore includes dummy variables for each year (reference year 1986) in order to control for period effects on the level of subjective well-being.

Control variables	null model			alternative model		
	Estimate	St. Dev.	p-value	Estimate	St. Dev.	p-value
Intercept	7.912	0.315	< 0.001	11.248	3.282	< 0.001
Sex: Female	0.219	0.023	< 0.001	0.234	0.023	< 0.001
Disability status: Disabled	-0.427	0.015	< 0.001	-0.432	0.015	< 0.001
Nights stayed in Hospital	-0.011	0.000	< 0.001	-0.010	0.000	< 0.001
Years of Education	0.042	0.003	< 0.001	0.037	0.003	< 0.001
log of net houshold income	0.462	0.010	< 0.001	0.477	0.010	< 0.001
log of household size	-0.195	0.013	< 0.001	-0.173	0.013	< 0.001
Nationality: German	0.037	0.020	0.070	0.038	0.020	0.063
Full time employed	0.060	0.011	< 0.001	0.102	0.012	< 0.001
Part time employed	0.011	0.013	0.376	0.039	0.013	0.003
Unemployed	-0.599	0.014	< 0.001	-0.564	0.014	< 0.001
Single	-0.221	0.018	< 0.001	-0.183	0.018	< 0.001
Divorced	-0.174	0.020	< 0.001	-0.155	0.020	< 0.001
Widowed	-0.246	0.025	< 0.001	-0.228	0.025	< 0.001
Region: West-Germany	0.537	0.017	< 0.001	0.534	0.017	< 0.001
Life expectation	-0.033	0.004	< 0.001	-0.034	0.004	< 0.001
Attrition in $t+1$	-0.118	0.012	< 0.001	-0.117	0.013	< 0.001
Attrition in $t+2$	-0.137	0.012	< 0.001	-0.135	0.012	< 0.001
Attrition in $t+3$	-0.078	0.012	< 0.001	-0.076	0.012	< 0.001
Covariance estimates						
$\hat{\rho}$			0.249			0.246
$\hat{\sigma}_\epsilon$			1.343			1.341
$\hat{\sigma}_1$			1.068			1.067
Data: $n = 253\,044$ observations, $N = 33\,451$ persons, SOEP 1986–2007 (excl. 1990, 1993)						

Table 3: Estimates of the covariate effects in the null model and in the alternative model, rounded to three digits. The estimates of the year effects are not displayed, since they are not of major interest here.