CIRJE-F-19

Bayes, Minimax and Nonnegative Estimators of Variance Components under Kullback-Leibler Loss

Tatsuya Kubokawa, The University of Tokyo
A. K. Md. Ehsanes Saleh, Carleton University
Yoshihiko Konno, Chiba University

September 1998

Bayes, Minimax and Nonnegative Estimators of Variance Components under Kullback-Leibler Loss

Tatsuya Kubokawa, A.K.Md. Ehsanes Saleh and Yoshihiko Konno
University of Tokyo, Carleton University and Chiba University

In a balanced one-way model with random effects, the simultaneous estimation of the variance components are considered under the intrinsic Kullback-Leibler loss function. The uniformly minimum variance unbiased (UMVU) or ANOVA estimators are known to have a drawback of taking negative values. The paper shows the minimaxity of the ANOVA estimators of the variance components and obtains classes of minimax estimators. Out of these classes, two types of minimax and nonnegative estimators are singled out, and they are characterized as empirical Bayes and generalized Bayes estimators. Also a residual maximum likelihood (REML) estimator is interpreted as an empirical Bayes rule. The risk performances of the derived estimators are investigated based on simulation experiments. An extension to the general mixed linear model with two components of variances is studied, and nonnegative estimators improving on the ANOVA estimators are given.

AMS 1991 subject classifications. Primary 62J10; secondary 62F10, 62F11.
Key words and phrases: Variance components model, unbalanced mixed linear model, point estimation, ANOVA estimator, empirical Bayes rule, generalized Bayes rule, residual maximum likelihood estimator, Kullback-Leibler information loss.

1 Introduction

Mixed linear models or variance components models have been effectively and extensively employed in practical data-analysis. When the statistical inference for regression coefficients or treatment effects is implemented, estimators of the variance components are used to get two-stage procedures such that two-stage generalized least squares (2GLS) estimators and 2GLS tests (for instance, see Fuller and Battese(1973), Peixoto and Harville(1986), Battese *et al.*(1988) and Rao *et al.*(1993)).

In the estimation of the variance components, it is well known that every unbiased estimator of the 'between' component of variance possesses the crucial drawback of taking negative values with a positive probability. Much effort has been devoted to this issue and reasonable procedures eliminating this undesirable property have been proposed: see Hartung (1981), Mathew (1984), Mathew *et al.*(1992) and their references. Of these, LaMotte (1973) showed that with the exception of the error variance, unbiased nonnegative quadratic estimators of variance components do not exist. On the other hand, Kleffe and Rao (1986) demonstrated that nonnegative biased quadratic estimators of the 'between' component of variance fail the minimum condition of consistency as each number of replication remains fixed, but the number of blocks or groups tend to infinity. These let us consider nonnegative estimators other than the quadratic estimators.

A simple modification of an undesirable unbiased estimator is a truncation of it at zero, yielding the uniform improvement as noted by Herbach (1959), Thompson (1962) and others. The truncated estimator, however, seems still unpleasant because the 'between' component of variance needs to be estimated by zero with a probability. Some other reasonable procedures have been proposed by Portnoy (1971), Chow and Shao (1988), Mathew *et al.*(1991), Kubokawa (1995) and Kubokawa *et al.*(1993b). From the aspects of Bayesness, minimaxity and admissibility, however, decision-theoretical researches have not been developed for the estimation of the 'between' component of variance.

In this paper, our interest is focused on the estimation of the variance components in a decision-theoretic framework. Especially our scope is to construct classes of minimax estimators improving on the ANOVA estimators, and to find out empirical Bayes and generalized Bayes estimators within the classes.

The mixed linear model we first treat in the present paper is the simple one-way random effect model with equal replications:

$$y_{ij} = \mu + v_i + e_{ij}, \quad i = 1, \ldots, k, \quad j = 1, \ldots, n, \tag{1.1}$$

where v_i's and e_{ij}'s are independent random variables, v_i being normally distributed with mean 0 and variance σ_v^2, $\mathcal{N}(0, \sigma_v^2)$ and e_{ij} having $\mathcal{N}(0, \sigma_e^2)$. Let $\overline{y}_{i\cdot} = \sum_{j=1}^n y_{ij}/n$, $\overline{y}_{\cdot\cdot} = \sum_{i=1}^k \sum_{j=1}^n y_{ij}/(nk)$, $S_1 = \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \overline{y}_{i\cdot})^2$ and $S_2 = n\sum_{i=1}^k (\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})^2$. The statistics $\overline{Y}_{\cdot\cdot}$, S_1 and S_2 are the minimal sufficient and are mutually independently distributed as $\overline{y}_{\cdot\cdot} \sim \mathcal{N}(\mu, (nk)^{-1}(\sigma_e^2 + n\sigma_v^2))$,

$$S_1 \sim \sigma_e^2 \chi_{\nu_1}^2 \quad \text{and} \quad S_2 \sim (\sigma_e^2 + n\sigma_v^2)\chi_{\nu_2}^2, \tag{1.2}$$

for

$$\nu_1 = k(n-1) \quad \text{and} \quad \nu_2 = k - 1,$$

where χ_ν^2 designates a chi squared distribution with ν degrees of freedom. Then we want to estimate the variance components σ_v^2 and σ_e^2 based on S_1 and S_2.

For comparing estimators of σ_e^2 and σ_v^2, the squared errors loss or quadratic loss have been mainly utilized in the literature. For instance, the quadratic loss is given by $L_q(\hat{\sigma}_e^2/\sigma_e^2) = (\hat{\sigma}_e^2/\sigma_e^2 - 1)^2$ in estimation of σ_e^2. It may be, however, inappropriate to employ $L_q(\hat{\sigma}_e^2/\sigma_e^2)$ because $L_q(\hat{\sigma}_e^2/\sigma_e^2)$ penalizes the under-estimate less than the over-estimate as seen from the fact that $\lim_{t\to 0} L_q(t) = 1$ and $\lim_{t\to\infty} L_q(t) = \infty$. This is also a reason why the best multiple of S_1 in terms of the mean squared error (MSE) is given by $(\nu_1 + 2)^{-1}$, that is, the unbiased estimator $\hat{\sigma}_e^{2UB}$ is not the best relative to the MSE.

In this paper, instead of the quadratic loss, we employ the Kullback-Leibler loss function in the simultaneous point estimation of the variance components σ_e^2 and σ_v^2:

$$
\begin{aligned}
L(\omega; \hat{\sigma}_e^2, \hat{\sigma}_v^2; \nu_1, \nu_2, n) &= \nu_1 \left\{ \frac{\hat{\sigma}_e^2}{\sigma_e^2} - \log \frac{\hat{\sigma}_e^2}{\sigma_e^2} - 1 \right\} \\
&\quad + \nu_2 \left\{ \frac{\hat{\sigma}_e^2 + n\hat{\sigma}_v^2}{\sigma_e^2 + n\sigma_v^2} - \log \frac{\hat{\sigma}_e^2 + n\hat{\sigma}_v^2}{\sigma_e^2 + n\sigma_v^2} - 1 \right\}
\end{aligned}
\tag{1.3}
$$

for a couple of unknown parameters $\omega = (\sigma_e^2, \sigma_v^2)$. This loss is defined in the case of $\hat{\sigma}_e^2 + n\hat{\sigma}_v^2 > 0$. This can be verified to be a convex function of $\hat{\sigma}_e^2/\sigma_e^2$ and $\hat{\sigma}_v^2/\sigma_v^2$. In fact (1.3) can be derived from the Kullback-Leibler information loss

$$
\int\int \log \left\{ \frac{f(s_1, s_2 | \hat{\sigma}_e^2, \hat{\sigma}_v^2)}{f(s_1, s_2 | \sigma_e^2, \sigma_v^2)} \right\} f(s_1, s_2 | \hat{\sigma}_e^2, \hat{\sigma}_v^2) ds_1 ds_2,
$$

where $f(s_1, s_2 | \sigma_e^2, \sigma_v^2)$ is a joint density of S_1 and S_2. With respect to the Kullback-Leibler loss, the ANOVA estimators $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is the best in the sense of minimizing the risk among estimators $(\hat{\sigma}_e^2, \hat{\sigma}_v^2) = (aS_1, n^{-1}(bS_2 - aS_1))$ for constants a and b, where

$$
\begin{cases}
\hat{\sigma}_e^{2UB} = \dfrac{S_1}{\nu_1}, \\[2ex]
\hat{\sigma}_v^{2UB} = \dfrac{1}{n}\left(\dfrac{S_2}{\nu_2} - \dfrac{S_1}{\nu_1} \right),
\end{cases}
\tag{1.4}
$$

which are the uniformly minimum variance unbiased (UMVU) estimators of σ_e^2 and σ_v^2.

In Section 2, it is first shown that the ANOVA estimator $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is minimax relative to the Kullback-Leibler loss (1.3). We next construct classes of minimax estimators improving on $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$. Out of the classes, we develop two types of nonnegative and minimax estimators $(\hat{\sigma}_e^{2EB}(a), \hat{\sigma}_v^{2EB}(a))$ and $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$, where

$$
\begin{cases}
\hat{\sigma}_e^{2EB}(a) = \min \left\{ \dfrac{S_1}{\nu_1}, \dfrac{S_1 + aS_2}{\nu_1 + \nu_2} \right\}, \\[2ex]
\hat{\sigma}_v^{2EB}(a) = \dfrac{1}{n}\left[\max \left\{ \dfrac{S_2}{\nu_2}, \dfrac{S_1/a + S_2}{\nu_1 + \nu_2 - 2} \right\} - \min \left\{ \dfrac{S_1}{\nu_1}, \dfrac{S_1 + aS_2}{\nu_1 + \nu_2} \right\} \right],
\end{cases}
\tag{1.5}
$$

$$
\begin{cases}
\hat{\sigma}_e^{2GB}(b) = \dfrac{S_1}{\nu_1 + \nu_2} \dfrac{\int_0^{bS_2/S_1} x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2} dx}{\int_0^{bS_2/S_1} x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2+1} dx}, \\[2ex]
\hat{\sigma}_v^{2GB}(b) = \dfrac{1}{n}\left[\dfrac{S_2}{\nu_1 + \nu_2} \dfrac{\int_0^{bS_2/S_1} x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2} dx}{\int_0^{bS_2/S_1} x^{\nu_2/2}/(1+x)^{(\nu_1+\nu_2)/2+1} dx} - \hat{\sigma}_e^{2GB}(b) \right],
\end{cases}
\tag{1.6}
$$

where a and b are suitable constants satisfying

$$\frac{\nu_2}{\nu_2+2} \le a \le 1 \quad \text{and} \quad 1 \le b \le \frac{\nu_2+2}{\nu_2}.$$

It is interesting to note that $(\hat{\sigma}_e^{2EB}(a), \hat{\sigma}_v^{2EB}(a))$ and $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ for $0 < a \le 1$ and $0 < b \le 1$ can be derived as empirical Bayes and generalized Bayes rules as shown in Section 2.2. This implies that the case of $b = 1$ in (1.6) satisfies both the minimaxity and the generalized Bayesness. Putting $a = 1$ in (1.5) yields

$$\begin{cases} \hat{\sigma}_e^{2REML} = \hat{\sigma}_e^{2EB}(1) = \min\left\{\frac{S_1}{\nu_1}, \frac{S_1+S_2}{\nu_1+\nu_2}\right\}, \\ \hat{\sigma}_v^{2REML} = \hat{\sigma}_v^{2EB}(1) = \frac{1}{n}\max\left\{\frac{S_2}{\nu_2} - \frac{S_1}{\nu_1}, 0\right\}, \end{cases} \tag{1.7}$$

which are known as the residual (or restricted) maximum likelihood (REML) estimators.

The risk performances of some nonnegative estimators including the above ones are investigated based on simulation experiments. Section 3 deals with an extension to general mixed linear models with two variance components and constructs classes of estimators improving on ANOVA estimators.

2 Bayes and Minimax Estimation

2.1 Construction of a Class of Minimax Estimators

We now construct a class of minimax estimators improving on the ANOVA estimator $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ given by (1.4) relative to the Kullback-Leibler loss (1.3).

From the form of the ANOVA estimator, we may consider the estimators of the general type

$$\begin{cases} \hat{\sigma}_e^2(\psi) = S_1\psi\left(\frac{S_2}{S_1}\right) \\ \hat{\sigma}_v^2(\phi,\psi;n) = \frac{1}{n}\left\{S_2\phi\left(\frac{S_1}{S_2}\right) - S_1\psi\left(\frac{S_2}{S_1}\right)\right\}. \end{cases} \tag{2.1}$$

Then the risk function of $(\hat{\sigma}_e^2(\psi), \hat{\sigma}_v^2(\phi,\psi;n))$ relative to the Kullback-Leibler loss (1.3) is written by

$$\begin{aligned} & R(\omega; \hat{\sigma}_e^2(\psi), \hat{\sigma}_v^2(\phi,\psi;n); \nu_1, \nu_2, n) \\ = \;& \nu_1 R_1(\omega; S_1\psi\left(\frac{S_2}{S_1}\right)) + \nu_2 R_2(\omega; S_2\phi\left(\frac{S_1}{S_2}\right); n), \end{aligned} \tag{2.2}$$

where

$$R_1(\omega; S_1\psi\left(\frac{S_2}{S_1}\right)) = E\left[\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}\right) - \log\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}\right) - 1\right],$$

$$R_2(\omega; S_2\phi\left(\frac{S_1}{S_2}\right); n) = E\left[\frac{S_2}{\sigma_e^2+n\sigma_v^2}\phi\left(\frac{S_1}{S_2}\right) - \log\frac{S_2}{\sigma_e^2+n\sigma_v^2}\phi\left(\frac{S_1}{S_2}\right) - 1\right].$$

Hence the original problem under the loss (1.3) is decomposed into two problems of estimating σ_e^2 and $\sigma_e^2+n\sigma_v^2$ in terms of the risks $R_1(\omega; S_1\psi)$ and $R_2(\omega; S_2\phi; n)$, respectively.

We begin with noting the minimaxity of the ANOVA estimator (1.4) under the Kullback-Leibler loss (1.3). In the above transformed problem, this issue is equivalent to the minimaxity of the estimator $(\nu_1^{-1}S_1, \nu_2^{-1}S_2)$. Since the parameter space is restricted by the inequality $\sigma_e^2 < \sigma_e^2 + n\sigma_v^2$, the Bayes estimator of each parameter is not simple, and so the minimaxity of (1.4) is not trivial. However we can verify that the Bayes risk of the Bayes estimator converges the constant risk of the estimator $(\nu_1^{-1}S_1, \nu_2^{-1}S_2)$ when the prior distribution approaches a noninformative (improper) prior. The details of the proof are given in Section 4.

Proposition 1. *The ANOVA estimator* $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ *given by (1.4) is minimax relative to the Kullback-Leibler loss (1.3).*

We next construct a class of minimax estimators, namely, improving on the ANOVA estimator. The *Integral-Expression-of-Risk-Difference* (IERD) method given by Kubokawa (1994a,b) is useful for deriving sufficient conditions for the domination, and we can get the following main results.

Theorem 1. *Assume that*
(a) $\psi(w)$ *is nondecreasing and* $\lim_{w\to\infty}\psi(w) = \nu_1^{-1}$,
(b) $\psi(w) \geq \psi_0(w)$ *where*

$$\psi_0(w) = \frac{1}{\nu_1+\nu_2}\frac{\int_0^w x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2}dx}{\int_0^w x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2+2)/2}dx}. \tag{2.3}$$

Then $R_1(\omega; S_1\psi(S_2/S_1)) \leq R_1(\omega; \nu_1^{-1}S_1)$ *uniformly for every* ω.

Theorem 2. *Assume that*
(a) $\phi(w)$ *is nondecreasing and* $\phi(0) = \nu_2^{-1}$,
(b) $\phi(w) \leq \phi_0(w)$ *where*

$$\phi_0(w) = \frac{1}{\nu_1+\nu_2}\frac{\int_0^{1/w} x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2}dx}{\int_0^{1/w} x^{\nu_2/2}/(1+x)^{(\nu_1+\nu_2+2)/2}dx}. \tag{2.4}$$

Then $R_2(\omega; S_2\phi(S_1/S_2); n) \leq R_2(\omega; \nu_2^{-1}S_2; n)$ *uniformly for every* ω.

The proofs are technical and deferred to Section 4. Combining Theorems 1 and 2 gives

Corollary 1. *The estimator* $(\hat{\sigma}_e^2(\psi), \hat{\sigma}_v^2(\phi, \psi; n))$ *given by (2.1) is minimax relative to the Kullback-Leibler loss (1.3) if the shrinkage functions* ϕ *and* ψ *satisfy the conditions (a) and (b) of both Theorems 1 and 2.*

When we find minimax estimators by using the above results, the following lemma is useful. For positive constants α and β, let

$$\begin{aligned} h_1(z;\alpha,\beta) &= \frac{\int_0^z x^\alpha/(1+x)^\beta dx}{\int_0^z x^\alpha/(1+x)^{\beta+1} dx}, \\ h_2(z;\alpha,\beta) &= \frac{\int_0^z x^\alpha/(1+x)^\beta dx}{\int_0^z x^{\alpha+1}/(1+x)^{\beta+1} dx}. \end{aligned}$$

Lemma 1.
(i) $h_1(z;\alpha,\beta)$ *is increasing in* z *with* $\lim_{z\to\infty} h_1(z;\alpha,\beta) = \beta/(\beta-\alpha-1)$, *and*

$$h_1(z;\alpha,\beta) \le 1 + \frac{\alpha+1}{\alpha+2} z.$$

(ii) $h_2(z;\alpha,\beta)$ *is decreasing in* z *with* $\lim_{z\to\infty} h_2(z;\alpha,\beta) = \beta/(\alpha+1)$, *and*

$$h_2(z;\alpha,\beta) \ge 1 + \frac{\alpha+2}{\alpha+1}\frac{1}{z}.$$

Proof. For the monotonicity of $h_1(z;\alpha,\beta)$, the derivative $(d/dz)h_1(z;\alpha,\beta)$ is proportional to

$$\begin{aligned} &\frac{z^\alpha}{(1+z)^\beta} \int_0^z \frac{x^\alpha}{(1+x)^{\beta+1}} dx - \frac{z^\alpha}{(1+z)^{\beta+1}} \int_0^z \frac{x^\alpha}{(1+x)^\beta} dx \\ &= \int_0^z \frac{x^\alpha z^\alpha (z-x)}{(1+x)^{\beta+1}(1+z)^{\beta+1}} dx, \end{aligned}$$

which is positive, and thus $h_1(z;\alpha,\beta)$ is increasing. It is easy to show that $h_1(z;\alpha,\beta)$ converges $\beta/(\beta-\alpha-1)$ as z tends to infinity. Finally the inequality given in (i) is derived as follows:

$$\begin{aligned} h_1(z;\alpha,\beta) - 1 &= \frac{\int_0^z x^{\alpha+1}/(1+x)^{\beta+1} dx}{\int_0^z x^\alpha/(1+x)^{\beta+1} dx} \\ &\le \frac{\int_0^z x^{\alpha+1} dx}{\int_0^z x^\alpha dx} = \frac{\alpha+1}{\alpha+2} z. \end{aligned} \tag{2.5}$$

Here the inequality (2.5) is expressed by

$$E^*\left[X(1+X)^{-\beta-1}\right] \leq E^*[X]\,E^*\left[(1+X)^{-\beta-1}\right], \tag{2.6}$$

where $E^*[\cdot]$ stands for the expectation with respect to the probability $P^*(A) = \int_A x^\alpha/(1+x)^{\beta+1}dx/\int_0^z x^\alpha/(1+x)^{\beta+1}dx$. Since X and $(1+X)^{-\beta-1}$ are monotone to the opposite directions, the inequality (2.6) holds and we get the inequality (2.5). The results for $h_2(z;\alpha,\beta)$ can be verified by the same arguments. Especially, the inequality in (ii) follows from the fact that

$$\begin{aligned} h_2(z;\alpha,\beta) - 1 &= \frac{\int_0^z x^\alpha/(1+x)^{\beta+1}dx}{\int_0^z x^{\alpha+1}/(1+x)^{\beta+1}dx} \\ &\geq \frac{\int_0^z x^\alpha dx}{\int_0^z x^{\alpha+1}dx} = \frac{\alpha+2}{\alpha+1}\frac{1}{z}. \end{aligned}$$

Therefore we get the results of Lemma 1. □□

2.2 Empirical Bayes and Generalized Bayes Estimators with Minimaxity

We shall develop empirical Bayes and generalized Bayes estimators out of the class of minimax estimators given by Section 2.1.

Let $\eta = 1/\sigma_e^2$ and $\xi = \sigma_e^2/(\sigma_e^2 + n\sigma_v^2)$ and note that ξ is constrained by $0 < \xi < 1$. When prior distribution $\pi(\eta,\xi)$ of (η,ξ) is supposed, the Bayes estimators under the loss (1.3) are generally given by

$$\begin{cases} \hat{\sigma}_e^{2B} = \dfrac{1}{E^\pi[\eta|S_1,S_2]}, \\ \hat{\sigma}_v^{2B} = \dfrac{1}{n}\Big[\dfrac{1}{E^\pi[\eta\xi|S_1,S_2]} - \hat{\sigma}_e^{2B}\Big], \end{cases} \tag{2.7}$$

where $E^\pi[\cdot|S_1,S_2]$ stands for an expectation with respect to the posterior distribution $\pi(\eta,\xi|S_1,S_2)$ of (η,ξ) given S_1 and S_2.

For the purpose of derivation of empirical Bayes procedures, let a be a given positive constant less than or equal to one, and suppose the prior distribution $\pi(\eta,\xi) = \eta^{-1}I(\xi = \xi_0)$ where $I(\cdot)$ denotes the indicator function, $P(\xi = \xi_0) = 1$ and ξ_0 is an unknown constant such that $0 < \xi_0 \leq a$. Namely the supposed prior information about ξ is that ξ is unknown and in the interval $(0,a]$. Then the posterior density of η given S_1 and S_2, and the marginal density of S_1 and S_2 are given by

$$\begin{aligned} &\text{(posterior density)} \propto \eta^{(\nu_1+\nu_2)/2-1}e^{-\frac{1}{2}(S_1+\xi_0 S_2)\eta} \\ &\text{(marginal density)} \propto \xi_0^{\nu_2/2}[S_1+\xi_0 S_2]^{-(\nu_1+\nu_2)/2}S_1^{\nu_1/2-1}S_2^{\nu_2/2-1}. \end{aligned}$$

Hence the Bayes estimators of σ_e^2 and σ_v^2 are

$$\begin{cases} \hat{\sigma}_e^{2B}(\xi_0) = \dfrac{S_1 + \xi_0 S_2}{\nu_1 + \nu_2}, \\ \hat{\sigma}_v^{2B}(\xi_0) = \dfrac{1}{n}\Big[\dfrac{S_1 + \xi_0 S_2}{(\nu_1 + \nu_2)\xi_0} - \hat{\sigma}_e^{2B}(\xi_0)\Big]. \end{cases}$$

Since ξ_0 is unknown, it should be estimated from the marginal density. Noting that $0 < \xi_0 < a$, it is seen that the maximum likelihood estimator of ξ_0 is written by $\hat{\xi}_0^{ML} = \min\{\nu_2 S_1/(\nu_1 S_2),\ a\}$, which is substituted in the above Bayes estimators so as to obtain the empirical Bayes rules:

$$\begin{aligned} \hat{\sigma}_e^{2EB}(a) &= \hat{\sigma}_e^{2B}(\hat{\xi}_0^{ML}) \\ &= \min\left\{\frac{S_1}{\nu_1}, \frac{S_1 + aS_2}{\nu_1 + \nu_2}\right\} \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} \hat{\sigma}_v^{2EB}(a) &= \hat{\sigma}_v^{2B}(\hat{\xi}_0^{ML}) \\ &= \frac{1}{n}\Big[\max\left\{\frac{S_2}{\nu_2}, \frac{S_1/a + S_2}{\nu_1 + \nu_2}\right\} - \hat{\sigma}_e^{2EB}(a)\Big]. \end{aligned} \tag{2.9}$$

This estimator is also rewritten by

$$\begin{aligned} \hat{\sigma}_v^{2EB}(a) &= \frac{1 - \hat{\xi}_0^{ML}}{n\hat{\xi}_0^{ML}} \frac{S_1 + \hat{\xi}_0^{ML} S_2}{\nu_1 + \nu_2} \\ &= \frac{1}{n}\max\left\{\frac{\nu_1}{\nu_2}\frac{S_2}{S_1} - 1, a^{-1} - 1\right\} \cdot \min\left\{\frac{S_1}{\nu_1}, \frac{S_1 + aS_2}{\nu_1 + \nu_2}\right\}, \end{aligned}$$

so that $\hat{\sigma}_v^{2EB}(a)$ is positive almost everywhere as long as a is strictly less than one. Especially putting $a = 1$ yields

$$\hat{\sigma}_e^{2REML} = \min\left\{\frac{S_1}{\nu_1}, \frac{S_1 + S_2}{\nu_1 + \nu_2}\right\}, \tag{2.10}$$

$$\begin{aligned} \hat{\sigma}_v^{2REML} &= \frac{1}{n}\max\left\{\frac{\nu_1}{\nu_2}\frac{S_2}{S_1} - 1, 0\right\} \cdot \min\left\{\frac{S_1}{\nu_1}, \frac{S_1 + S_2}{\nu_1 + \nu_2}\right\} \\ &= \frac{1}{n}\max\left\{\frac{S_2}{\nu_2} - \frac{S_1}{\nu_1}, 0\right\}, \end{aligned} \tag{2.11}$$

which are known as the Residual (or Restricted) Maximum Likelihood (REML) estimators. So it is interesting to point out that the REML estimators are derived as the empirical Bayes rules.

The minimaxity of the empirical Bayes estimators is argued based on Theorems 1 and 2 and Lemma 1. From the inequalities in Lemma 1, it is seen that

$$\begin{aligned}\psi_0(w) &\le \frac{1}{\nu_1+\nu_2}\left(1+\frac{\nu_2}{\nu_2+2}w\right),\\ \phi_0(w) &\ge \frac{1}{\nu_1+\nu_2}\left(1+\frac{\nu_2+2}{\nu_2}w\right).\end{aligned} \tag{2.12}$$

Putting

$$\begin{aligned}\psi^T(w;a) &= \min\left\{\frac{1}{\nu_1},\frac{1+aw}{\nu_1+\nu_2}\right\},\\ \phi^T(w;a) &= \max\left\{\frac{1}{\nu_2},\frac{1+w/a}{\nu_1+\nu_2}\right\},\end{aligned}$$

we can check that $\psi^T(w;a)$ and $\phi^T(w;a)$, respectively, satisfy the conditions (a) and (b) of Theorems 1 and 2 if

$$\nu_2/(\nu_2+2)\le a\le 1. \tag{2.13}$$

As one of drawbacks of the REML estimator $\hat{\sigma}_v^{2REML}$ given by (2.11), the parameter σ_v^2 is estimated by zero when $\nu_2^{-1}S_2<\nu_1^{-1}S_1$. Alternative estimators may be provided by putting $a=\nu_2/(\nu_2+2)$ and we denote $\hat{\sigma}_e^{2EB}=\hat{\sigma}_e^{2EB}(\nu_2/(\nu_2+2))$ and $\hat{\sigma}_v^{2EB}=\hat{\sigma}_v^{2EB}(\nu_2/(\nu_2+2))$, which is also expressed by

$$\hat{\sigma}_v^{2EB}=\frac{1}{n}\max\left\{\frac{\nu_1}{\nu_2}\frac{S_2}{S_1}-1,\frac{2}{\nu_2}\right\}\cdot\min\left\{\frac{S_1}{\nu_1},\frac{S_1+\nu_2(\nu_2+2)^{-1}S_2}{\nu_1+\nu_2}\right\}.$$

This demonstrates that $(\hat{\sigma}_e^{2EB},\hat{\sigma}_v^{2EB})$ is a positive, minimax and empirical Bayes procedure for (σ_e^2,σ_v^2). The risk performances of the empirical Bayes estimators are investigated in Section 2.3 for various values of a. The investigation reveals that the empirical Bayes estimator violates the minimaxity for small a.

For deriving the generalized Bayes estimators, suppose the improper prior distribution

$$\pi(\eta,\xi)d\eta d\xi=\eta^{-1}\xi^{-1}d\eta d\xi I(0<\xi<b)$$

for suitable constant b. Then the posterior density of (η,ξ) given S_1 and S_2 is proportional to

$$\xi^{\nu_2/2-1}\eta^{(\nu_1+\nu_2)/2-1}e^{-\frac{1}{2}(S_1+\xi S_2)\eta}I(0<\xi<b),$$

so that from (2.7), the generalized Bayes estimator of (σ_e^2,σ_v^2) is given by $(\hat{\sigma}_e^{2GB}(b),\hat{\sigma}_v^{2GB}(b))$ where

$$\begin{aligned}\hat{\sigma}_e^{2GB}(b) &= S_1\psi_0(bS_2/S_1) \qquad (2.14)\\ &= \frac{S_1}{\nu_1+\nu_2}\frac{\int_0^{bS_2/S_1}x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2}dx}{\int_0^{bS_2/S_1}x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2+1}dx},\\ \hat{\sigma}_v^{2GB}(b) &= n^{-1}[S_2\phi_0(S_1/(bS_2))-S_1\psi_0(bS_2/S_1)] \qquad (2.15)\\ &= \frac{1}{n}\Big[\frac{S_2}{\nu_1+\nu_2}\frac{\int_0^{bS_2/S_1}x^{\nu_2/2-1}/(1+x)^{(\nu_1+\nu_2)/2}dx}{\int_0^{bS_2/S_1}x^{\nu_2/2}/(1+x)^{(\nu_1+\nu_2)/2+1}dx}-\hat{\sigma}_e^{2GB}(b)\Big].\end{aligned}$$

Although the generalized Bayes estimator has a complicated form including the ratio of integrals, it can be expressed by the incomplete beta functions ratio $I_x(\cdot,\cdot)$, for the integrals are written as

$$\begin{aligned}\int_0^w z^\alpha/(1+z)^{\alpha+\beta}dz &= \int_0^{w/(1+w)} x^\alpha(1-x)^{\beta-2}dx \\ &= B(\alpha+1,\beta-1)I_{w/(1+w)}(\alpha+1,\beta-1).\end{aligned}$$

When a table of values of the incomplete beta functions ratio is available, one can compute them in a practical use.

For the minimaxity of the generalized Bayes estimator, from Theorems 1 and 2 and Lemma 1, it is seen that $\psi_0(bS_2/S_1)$ and $\phi_0(S_1/(bS_2))$, respectively, satisfy the conditions (a) and (b) of Theorems 1 and 2 as long as b is greater than or equal to one. Hence $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ is minimax for $b \geq 1$. On the other hand, it is the generalized Bayes for $b \leq 1$. Hence for $b = 1$, $(\hat{\sigma}_e^{2GB}(1), \hat{\sigma}_v^{2GB}(1))$ is the generalized Bayes and minimax rule. When $b > 1$, the support of the prior distribution of ξ is beyond the parameter space, that is, it is not an appropriate generalized Bayes rule. Nevertheless, it is minimax and smooth, and so from a frequentist's stand-point, we can take it into account as one of possible candidates. Since $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ approaches the ANOVA estimator $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ as b tends to infinity, $\hat{\sigma}_v^{2GB}(b)$ may happen to take negative values unless b is small. Here from (2.12),

$$\begin{aligned}\hat{\sigma}_v^{2GB}(b) &= \frac{1}{n}\Big[S_2\phi_0(S_1/(bS_2)) - S_1\psi_0(bS_2/S_1)\Big] \\ &\geq \frac{1}{n(\nu_1+\nu_2)}\left[S_2\left(1+\frac{\nu_2+2}{\nu_2 b}\frac{S_1}{S_2}\right) - S_1\left(1+\frac{\nu_2 b}{\nu_2+2}\frac{S_2}{S_1}\right)\right] \\ &= \frac{1}{n(\nu_1+\nu_2)}\left\{\frac{\nu_2+2}{\nu_2 b}-1\right\}\left\{S_1+\frac{\nu_2 b}{\nu_2+2}S_2\right\},\end{aligned}$$

which implies that $\sigma_v^{2GB}(b)$ is positive almost everywhere when

$$1 \leq b \leq (\nu_2+2)/\nu_2 = 1 + 2/\nu_2. \tag{2.16}$$

Namely, $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ is positive and minimax for $1 \leq b \leq 1 + 2/\nu_2$ and is generalized Bayes for $b = 1$. The risk performances of the estimators $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ are numerically investigated in the next section for various values of b. This study will show that $(\hat{\sigma}_e^{2GB}(b), \hat{\sigma}_v^{2GB}(b))$ violates the minimaxity when b is less than 1.

2.3 Simulation Studies

It is of interest to investigate the risk behaviors of several estimators given in Section 2.2. For the sake of simplicity, we denote $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$, $(\hat{\sigma}_e^{2EB}(\nu_2/(\nu_2+c)), \hat{\sigma}_v^{2EB}(\nu_2/(\nu_2+c)))$ and $(\hat{\sigma}_e^{2GB}(1+c/\nu_2), \hat{\sigma}_v^{2GB}(1+c/\nu_2))$ by UB, EB(c) and GB(c), respectively. We provide the results of Monte Carlo simulation for the risks relative to the Kullback-Leibler loss where the values of the risks are given by average values of the risk functions based on 50,000 replications.

We first treat the model (1.1) with $(n,k)=(2,6)$, that is, $(\nu_1,\nu_2)=(6,5)$, and $\sigma_e^2=1$. The range of σ_v^2 is taken between 0 and 14. Figure 1 reports the risk behaviors of the ANOVA estimator UB and the empirical Bayes estimators EB(c) for $c=0,2,4$ and 6. From (2.13), the minimaxity of EB(c) is guaranteed for $0\le c\le 2$, and EB(0) and EB(2) correspond to $(\hat{\sigma}_e^{2REML},\hat{\sigma}_v^{2REML})$ and $(\hat{\sigma}_e^{2EB},\hat{\sigma}_v^{2EB})$, respectively, which are also denoted by REML and EB. Figure 1 reveals that

(1) EB(4) is minimax while EB(6) violates the minimaxity,

(2) for larger c, EB(c) has a larger risk at $\sigma_v^2=0$, and the risk of EB(c) is minimized at σ_v^2 being farther from zero,

(3) the minimum values of the risks of EB(c) are almost the same for $0\le c\le 6$.

Figure 2 reports the risk behaviors of the ANOVA estimator UB, the generalized Bayes estimators GB(-1) and GB(0) and the minimax estimators GB(2), GB(5) and GB(10). From the figure, we see that

(1) the generalized Bayes estimator GB(-1) is not minimax and has the largest risk at $\sigma_v^2=0$,

(2) the generalized Bayes estimator GB(0) has no risk reduction at $\sigma_v^2=0$ while it is the best among minimax estimators in a large range of σ_v^2,

(3) for larger c, the point of σ_v^2 minimizing the risk of GB(c) is closer to zero.

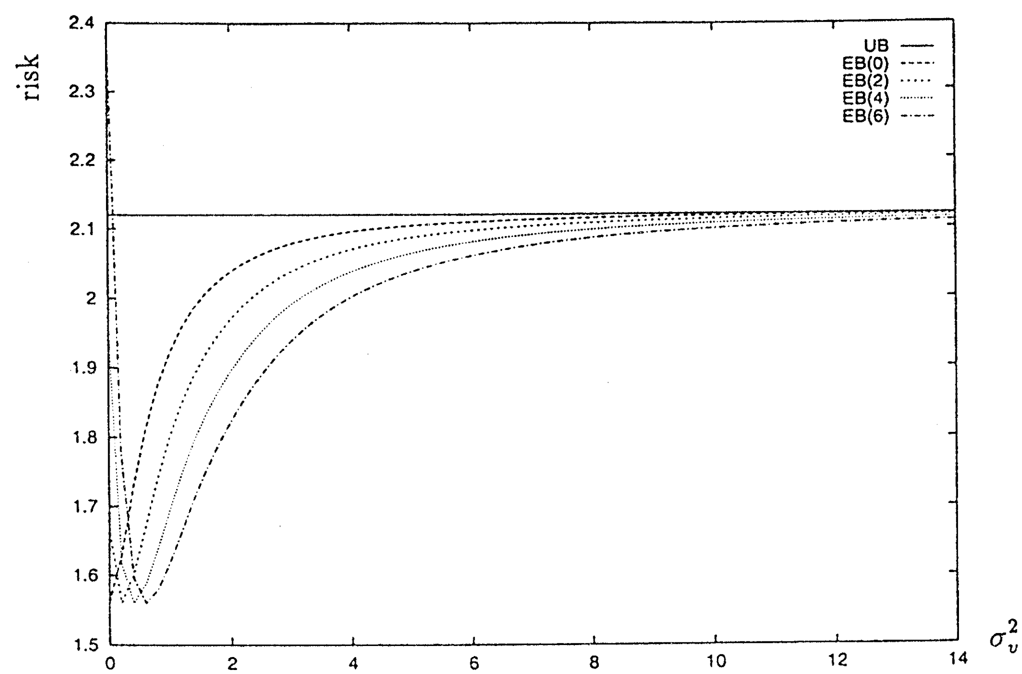

Figure 1. Risks of the ANOVA Estimator UB and the Empirical Bayes Estimators EB(0), EB(2), EB(4) and EB(6)

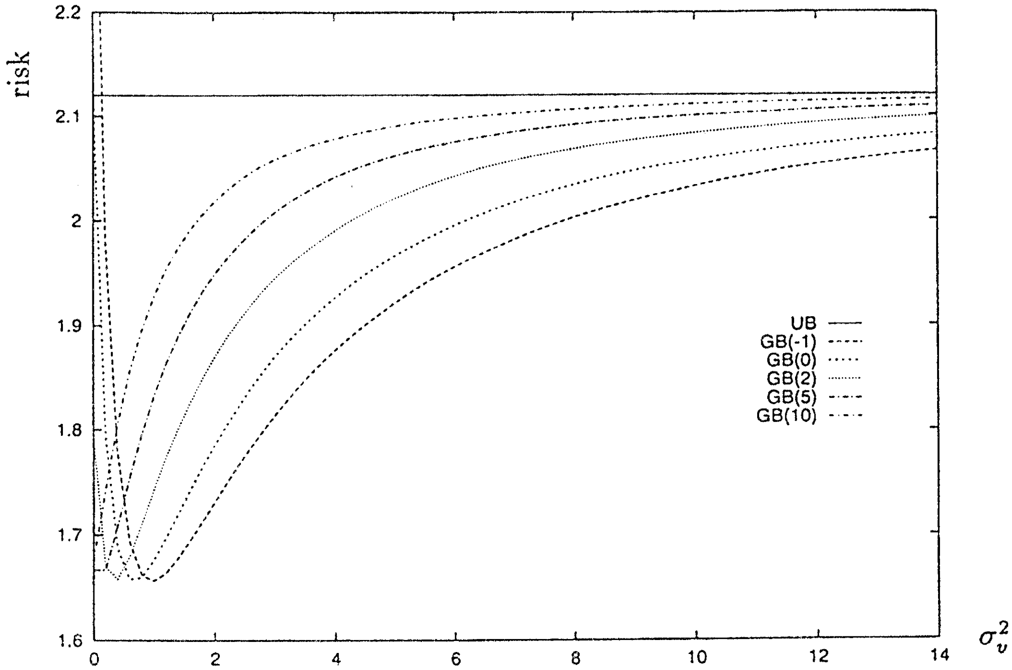

Figure 2. Risks of the ANOVA Estimator UB, the Generalized Bayes Estimators GB(-1) and GB(0) and the Smooth Estimators GB(2), GB(10) and GB(15)

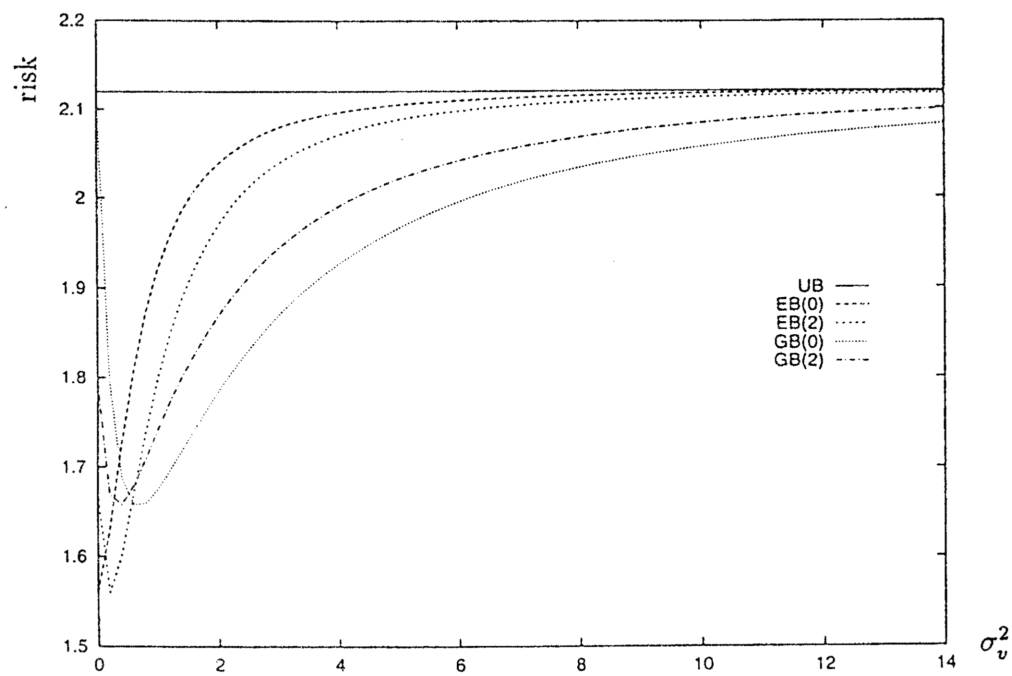

Figure 3. Risks of the Estimators UB, REML, EB, GB(0), GB(2)

Table 1. Risks of Estimators UB, REML, EB, GB(0) and GB(2) under the Kullback-Leibler Loss in the Balanced Cases

σ_v^2		0.0	0.01	0.05	0.1	0.5	1.0	4.0	9.0	16.0
	UB	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09
$n = 3$	REML	1.47	1.47	1.49	1.53	1.82	1.96	2.06	2.07	2.08
$k = 6$	EB	1.62	1.60	1.52	1.48	1.67	1.87	2.04	2.07	2.08
	GB(0)	2.09	2.05	1.88	1.75	1.57	1.69	1.95	2.05	2.07
	GB(2)	1.69	1.67	1.61	1.57	1.65	1.80	2.00	2.06	2.08
	UB	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17
$n = 6$	REML	1.36	1.37	1.40	1.45	1.77	1.92	2.07	2.13	2.15
$k = 3$	EB	1.79	1.71	1.50	1.40	1.53	1.73	1.98	2.10	2.13
	GB(0)	2.17	2.07	1.78	1.59	1.45	1.61	1.89	2.06	2.11
	GB(2)	1.47	1.45	1.42	1.42	1.66	1.82	2.02	2.11	2.14

It is still open to obtain a minimax and generalized Bayes estimator possessing a significant improvement at $\sigma_v^2 = 0$.

Figure 3 gives the risk performances of the five estimators UB, REML, EB, GB(0) and GB(2). The empirical Bayes estimators REML and EB present the large improvements for σ_v^2 close to zero, while the significant risk reductions of GB(0) and GB(2) yield for σ_v^2 far from zero. The maximum improvements of EB(c)'s are larger than those of GB(c)'s. Table 1 provides the values of the risks of these five estimators for $(n, k) = (3, 6)$ and $(6, 3)$, and shows that the estimators have the similar risk performances as in Figure 3.

The empirical Bayes estimator EB(2) is always positive, minimax and has a significant improvement for small σ_v^2, which demonstrates that EB(2) is a superior procedure, and we want to propose the use of it.

3 An Extension to Mixed Linear Models

We extend the results of Section 2.1 to the general mixed linear models with two variance components:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}\boldsymbol{v} + \boldsymbol{e} \tag{3.1}$$

where $\boldsymbol{y}$ is an N-vector of observations, $\boldsymbol{X}$ is an $N \times p_1$ known matrix with $\text{rank}(\boldsymbol{X}) = r$, $\boldsymbol{\beta}$ is a p_1-vector of parameters, $\boldsymbol{U}$ is an $N \times p_2$ known matrix, $\boldsymbol{v}$ and $\boldsymbol{e}$ are independent random p_2- and N-vectors, respectively, with $\boldsymbol{v} \sim \mathcal{N}_{p_2}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{I}_{p_2})$ and $\boldsymbol{e} \sim \mathcal{N}_N(\boldsymbol{0}, \sigma_e^2 \boldsymbol{I}_N)$.

It is reasonable that the variance components σ_e^2 and σ_v^2 are estimated based on the statistics that are invariant under the group of transformation $\boldsymbol{y} \to \boldsymbol{y} + \boldsymbol{X}\boldsymbol{a}$ where $\boldsymbol{a}$ is any p_1-vector. For the purpose, consider an $(N - r) \times N$ matrix $\boldsymbol{P}$ such that $\boldsymbol{P}\boldsymbol{X} = \boldsymbol{0}$

and $\boldsymbol{PP}' = \boldsymbol{I}_{N-r}$. Letting $\boldsymbol{x} = \boldsymbol{Py}$, we see that $\boldsymbol{x} \sim \mathcal{N}_{N-r}(\boldsymbol{0}, \sigma_v^2 \boldsymbol{PUU}'\boldsymbol{P}' + \sigma_e^2 \boldsymbol{I}_{N-r})$. Consider the spectral decompositions $\boldsymbol{PUU}'\boldsymbol{P}' = \sum_{i=1}^{\ell} \lambda_i \boldsymbol{E}_i$, where $\mathrm{rank}(\boldsymbol{E}_i) = m_i$ and $\sum_{i=1}^{\ell} m_i = \mathrm{rank}(\boldsymbol{PUU}'\boldsymbol{P}')$. Assume that $N - r - \sum_{i=1}^{\ell} m_i > 0$, and let

$$\nu_1 = N - r - \sum_{i=1}^{\ell} m_i \quad \text{and} \quad \nu_2 = \sum_{i=1}^{\ell} m_i. \tag{3.2}$$

Let $\boldsymbol{E}_{\ell+1} = \boldsymbol{I}_{N-r} - \sum_{i=1}^{\ell} \boldsymbol{E}_i$. The rank of $\boldsymbol{E}_{\ell+1}$ is ν_1. We thus get the quadratic statistics $\boldsymbol{x}'\boldsymbol{E}_i\boldsymbol{x}$ for $i = 1, \ldots, \ell + 1$ with

$$\begin{aligned} \boldsymbol{x}'\boldsymbol{E}_{\ell+1}\boldsymbol{x} &\sim \sigma_e^2 \chi_{\nu_1}^2, \\ \boldsymbol{x}'\boldsymbol{E}_i\boldsymbol{x} &\sim (\sigma_e^2 + \lambda_i \sigma_v^2)\chi_{m_i}^2, \quad i = 1, \ldots, \ell. \end{aligned}$$

For the sake of convenience, let

$$\begin{aligned} S_1 &= \boldsymbol{x}'\boldsymbol{E}_{\ell+1}\boldsymbol{x}, \quad S_2 = \sum_{i=1}^{\ell} \boldsymbol{x}'\boldsymbol{E}_i\boldsymbol{x} \quad \text{and} \\ \lambda &= \sum_{i=1}^{\ell} \lambda_i m_i / \sum_{i=1}^{\ell} m_i = \sum_{i=1}^{\ell} \lambda_i m_i / \nu_2. \end{aligned} \tag{3.3}$$

The quadratic statistics S_1 and S_2 can be given through a usual method when we deal with the following special model of (3.1):

$$y_{ij} = \boldsymbol{x}_{ij}'\boldsymbol{\beta} + v_i + e_{ij}, \quad j = 1, \ldots, n_i, \ i = 1, \ldots, k \tag{3.4}$$

where $\boldsymbol{x}_{ij} = (x_{ij0}, x_{ij1}, \ldots, x_{ij,p-1})'$ with $x_{ij0} \equiv 1$ is a vector of known covariates, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})'$ is a vector of unknown regression coefficients, $v_i \sim \mathcal{N}(0, \sigma_v^2)$, $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$, and v_i's are independent of e_{ij}'s. This model is dealt with in Battese *et al.*(1988) for prediction of county crop areas (small areas) using survey and satellite data. It is also known as an error component model in econometrics. In this model, as shown in Fuller and Battese (1973), Battese *et al.*(1988) and Rao *et al.*(1993), S_1 and S_2 can be represented as

$$S_1 = \sum_i \sum_j \hat{e}_{ij}^2 \quad \text{and} \quad S_2 = \sum_i \sum_j \hat{u}_{ij}^2 - S_1,$$

where $\{\hat{e}_{ij}\}$ are the residuals from the OLS regression of $y_{ij} - \overline{y}_{i\cdot}$ on $\{x_{ij1} - \overline{x}_{i\cdot 1}, \ldots, x_{ij,p-1} - \overline{x}_{i\cdot p-1}\}$ without the intercept term for $\overline{y}_{i\cdot} = \sum_j y_{ij}/n_i$ and $\overline{x}_{i\cdot\ell} = \sum_j x_{ij\ell}/n_i$, and $\{\hat{u}_{ij}\}$ are the residuals from the OLS regression of y_{ij} on $\{x_{ij1}, \ldots, x_{ij,p-1}\}$ with the intercept term. Also $\nu_2 = k - 1$ and λ defined in (3.3) is written by

$$\lambda = \left\{ N - \mathrm{tr}\left[(\boldsymbol{X}'\boldsymbol{X})^{-} \Sigma_{i=1}^{k} n_i^2 \overline{\boldsymbol{x}}_{i\cdot} \overline{\boldsymbol{x}}_{i\cdot}' \right] \right\} / (k-1).$$

In the unbalanced case with $\boldsymbol{\beta} = (\beta_0, 0, \ldots, 0)'$, S_1 and S_2 are further simplified as $S_1 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{i\cdot})^2$ and $S_2 = \sum_{i=1}^{k} n_i (\overline{y}_{i\cdot} - \overline{y}_{\cdot\cdot})^2$ for total mean $\overline{y}_{\cdot\cdot}$.

The lack of completeness of the quadratic statistics in the general mixed linear model causes various unbiased estimators such as, for instance, the MINQUE introduced by Rao (1971a,b) and the ANOVA estimators given by Henderson (1953). The unbiased estimators have the fundamental defect of taking on negative values as stated in Section 1.

The present section gives a note that the results of Section 2.1 can be directly applicable for getting nonnegative estimators improving on the ANOVA estimators which can be derived by the well-known Henderson (1953) method (III). The ANOVA estimators are of the forms

$$\begin{aligned}\hat{\sigma}_e^{2UB} &= \frac{S_1}{\nu_1}, \\ \hat{\sigma}_v^{2UB} &= \frac{1}{\lambda}\left\{\frac{S_2}{\nu_2} - \frac{S_1}{\nu_1}\right\},\end{aligned}$$

where ν_1, ν_2 and λ are defined in (3.2) and (3.3). Similar to Section 2.1, we consider the loss function $L(\omega; \hat{\sigma}_e^2, \hat{\sigma}_v^2; \nu_1, \nu_2, \lambda)$ given by (1.3) and the estimators $(\hat{\sigma}_e^2(\psi), \hat{\sigma}_v^2(\phi, \psi; \lambda))$ given by (2.1), where $L(\omega; \hat{\sigma}_e^2, \hat{\sigma}_v^2; \nu_1, \nu_2, \lambda)$ is different from an original Kullback-Leibler loss in the general setup. Then all the results in Section 2.1 still hold in the general mixed linear models by replacing ν_1, ν_2 and n in Section 2.1 with ν_1, ν_2 and λ, respectively, given by (3.2) and (3.3).

Theorem 3. *Under the same conditions as in Theorem 1, the dominance result*

$$R_1(\omega; S_1\psi(S_2/S_1)) \leq R_1(\omega; \nu_1^{-1}S_1)$$

holds uniformly for every ω.

Theorem 4. *Under the same conditions as in Theorem 2, the dominance result*

$$R_2(\omega; S_2\phi(S_1/S_2); \lambda) \leq R_2(\omega; \nu_2^{-1}S_2; \lambda)$$

uniformly for every ω.

The proofs are given in Section 4. By replacing n with λ in the estimators given in Section 2, we can consider the four estimators $(\hat{\sigma}_e^{2REML}, \hat{\sigma}_v^{2REML})$, $(\hat{\sigma}_e^{2EB}, \hat{\sigma}_v^{2EB})$, $(\hat{\sigma}_e^{2GB}(1), \hat{\sigma}_v^{2GB}(1))$ and $(\hat{\sigma}_e^{2GB}(1 + 2/\nu_2), \hat{\sigma}_v^{2GB}(1 + 2/\nu_2))$, which are simply denoted by REML, EB, GB(0) and GB(2). Also the ANOVA estimator $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is indicated by UB. Combining Theorems 3, 4 and the arguments as in Section 2 shows that REML, EB, GB(0) and GB(2) are nonnegative estimators improving upon UB, although these do not necessarily possess Bayesian properties like Section 2.2.

We investigate the risk behaviors of these nonnegative and improved estimators through Monte Carlo simulation. We consider the following two models: One is the model (3.4) with $\boldsymbol{\beta} = (\beta_0, 0, \ldots, 0)$ and $\sigma_e^2 = 1$, that is, $y_{ij} = \beta_0 + v_i + e_{ij}$, $j = 1, \ldots, n_i$,

Table 2. Risks of Estimators UB, REML, EB, GB(0) and GB(2) under the Kullback-Leibler Loss in the Unbalanced Cases

σ_v^2		0.0	0.01	0.05	0.1	0.5	1.0	4.0	9.0	16.0
	UB	2.08	2.08	2.09	2.09	2.12	2.13	2.14	2.15	2.15
replications	REML	1.43	1.43	1.47	1.56	1.95	2.07	2.14	2.15	2.15
(3,3,5,	EB	1.61	1.56	1.44	1.43	1.82	2.01	2.13	2.15	2.15
5,7,7)	GB(0)	2.08	1.99	1.74	1.58	1.66	1.87	2.09	2.15	2.15
	GB(2)	1.62	1.59	1.52	1.50	1.79	1.97	2.12	2.15	2.15
	UB	2.07	2.07	2.09	2.14	2.45	2.59	2.75	2.82	2.83
replications	REML	1.40	1.40	1.48	1.61	2.26	2.51	2.74	2.81	2.83
(1,1,1,	EB	1.56	1.53	1.42	1.45	2.12	2.44	2.72	2.81	2.83
13,13,13)	GB(0)	2.07	1.96	1.67	1.54	1.92	2.29	2.67	2.81	2.83
	GB(2)	1.59	1.55	1.49	1.52	2.09	2.41	2.71	2.81	2.83
	UB	2.09	2.09	2.09	2.10	2.18	2.25	2.36	2.42	2.44
replications	REML	1.47	1.47	1.47	1.49	1.74	1.97	2.28	2.41	2.44
(1,1,1,	EB	1.62	1.60	1.55	1.51	1.58	1.80	2.22	2.40	2.43
1,1,13)	GB(0)	2.09	2.07	1.97	1.87	1.61	1.68	2.04	2.34	2.40
	GB(2)	1.68	1.67	1.63	1.60	1.62	1.78	2.14	2.37	2.42

$i = 1, \ldots, k$. Table 2 reports the average values of the risks of the above five estimators based on 50,000 replications for $(n_1, n_2, n_3, n_4, n_5, n_6) = (3, 3, 5, 5, 7, 7)$, $(1, 1, 1, 13, 13, 13)$ and $(1, 1, 1, 1, 1, 13)$ and $\sigma_v^2 = 0.0$, 0.01, 0.05, 0.1, 0.5, 1.0, 4.0, 9.0 and 16.0. The other model treated here is (3.4) with $p = 2$ and $n_1 = \cdots = n_k = n$, that is,

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + v_i + e_{ij},\ i = 1, \ldots, k,\ j = 1, \ldots, n$$

where $\sigma_e^2 = 1$ and $\{x_{ij}\}$ are generated from $\mathcal{N}(10, \sigma_x^2)$ for $\sigma_x = 5$. Table 3 presents the risk behaviors of the five estimators for $(n, k) = (2, 4)$, $(2, 10)$ and $(10, 3)$. Tables 2 and 3 reveal that the estimators REML, EB, GB(0) and GB(2) have risk performances similar to those indicated in Section 2.3. It is also seen that the risks of UB is increasing in σ_v^2. Since the estimator EB is always positive and has a good risk performance for small σ_v^2, it can be employed for a practical use.

4 Proofs

We first show the minimaxity of the ANOVA estimator $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$. From the arguments as around (2.2), the minimaxity of $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is equivalent to the minimaxity of $(S_1/\nu_1, S_2/\nu_2)$ in terms of the risk $\nu_1 R_1(\eta, \theta; \widehat{\eta^{-1}}) + \nu_2 R_2(\eta, \theta; \widehat{\theta^{-1}})$ where $R_1(\eta, \theta; \widehat{\eta^{-1}}) = E\left[L(\eta\widehat{\eta^{-1}})\right]$ and $R_2(\eta, \theta; \widehat{\theta^{-1}}) = E\left[L(\theta\widehat{\theta^{-1}})\right]$ for $L(z) = z - \log z - 1$, $\eta = \sigma_e^{-2}$ and

Table 3. Risks of Estimators UB, REML, EB, GB(0) and GB(2) under the Kullback-Leibler Loss in the Simple Regression Models for $\sigma_x = 5$

σ_v^2		0.0	0.01	0.05	0.1	0.5	1.0	4.0	9.0	16.0
	UB	2.21	2.21	2.21	2.22	2.22	2.23	2.25	2.27	2.27
$n=2$	REML	1.63	1.63	1.64	1.64	1.75	1.87	2.10	2.22	2.25
$k=4$	EB	1.80	1.79	1.74	1.70	1.64	1.72	1.98	2.18	2.23
	GB(0)	2.21	2.18	2.11	2.04	1.78	1.73	1.83	2.03	2.12
	GB(2)	1.83	1.82	1.79	1.77	1.73	1.77	1.93	2.11	2.18
	UB	2.07	2.07	2.07	2.07	2.07	2.07	2.08	2.09	2.10
$n=2$	REML	1.55	1.55	1.55	1.58	1.81	1.97	2.08	2.09	2.10
$k=10$	EB	1.61	1.60	1.56	1.55	1.71	1.90	2.07	2.09	2.10
	GB(0)	2.07	2.05	1.95	1.85	1.64	1.71	1.95	2.08	2.10
	GB(2)	1.83	1.81	1.76	1.71	1.66	1.77	1.99	2.09	2.10
	UB	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17	2.17
$n=10$	REML	1.34	1.34	1.41	1.50	1.87	2.00	2.10	2.14	2.15
$k=3$	EB	1.79	1.66	1.40	1.34	1.65	1.85	2.04	2.12	2.14
	GB(0)	2.17	2.00	1.62	1.44	1.55	1.75	1.99	2.10	2.13
	GB(2)	1.43	1.41	1.39	1.43	1.77	1.93	2.07	2.13	2.15

$\theta = (\sigma_e^2 + n\sigma_v^2)^{-1}$. If the parameter (η, θ) was spanned on whole the space $\boldsymbol{R}^+ \times \boldsymbol{R}^+$ for $\boldsymbol{R}^+ = \{x \in \boldsymbol{R};\ x > 0\}$, the proof of the minimaxity would be trivial. Since the parameter space of (η, θ) is restricted into $\{(\eta, \theta) \in \boldsymbol{R}^+ \times \boldsymbol{R}^+;\ \eta \geq \theta\}$ in this problem, we need to give a proof for the minimaxity.

Proof of Proposition 1. Suppose that $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is not minimax relative to the Kullback-Leibler loss (1.3). Then there exists an estimator (δ_e, δ_v) such that

$$\begin{aligned}\sup_{\eta\geq\theta} R(\eta, \theta; \delta_e, \delta_v) &< \sup_{\eta\geq\theta} R(\eta, \theta; \hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB}) \\ &= \nu_1 R_1(1, 1; \nu_1^{-1} S_1) + \nu_2 R_2(1, 1; \nu_2^{-1} S_2). \end{aligned} \tag{4.1}$$

Let $\xi = \sigma_e^2/(\sigma_e^2 + n\sigma_v^2) = \theta/\eta$ with $0 < \xi < 1$. Suppose that η and ξ are independently distributed as $\eta \sim \mathcal{G}amma(b/2, a/2)$ and $\xi \sim \mathcal{B}eta(c/2, 1)$, that is,

$$\pi(\eta, \xi) = \frac{1}{\Gamma(a/2)} \left(\frac{b}{2}\right)^{a/2} \eta^{a/2-1} e^{-b\eta/2} \times \frac{c}{2} \xi^{c/2-1}, \tag{4.2}$$

for positive constants a, b and c. Then the posterior distribution of (η, ξ) given $(S_1, S_2) = (s_1, s_2)$ is proportional to

$$\pi(\eta, \xi | s_1, s_2) \propto \eta^{(\nu_1+\nu_2+a)/2-1} \xi^{(\nu_2+c)/2-1} e^{-(s_1+b+\xi s_2)\eta/2}.$$

Denoting the Bayes estimators of η^{-1} and θ^{-1} by δ_1^B and δ_2^B, we see that

$$\begin{aligned}\int_0^1\int_0^\infty &\left\{\nu_1 R_1(\eta,\xi\eta;\delta_1^B)+\nu_2 R_2(\eta,\xi\eta;\delta_2^B)\right\}\pi(\eta,\xi)d\eta d\xi\\ &\le \sup_{\eta\ge\theta} R(\eta,\theta;\delta_e,\delta_v). \qquad (4.3)\end{aligned}$$

From the inequalities (4.1) and (4.3), the contradiction yields if we can establish that

$$\begin{aligned}\lim_{a\to 0,b\to 0,c\to 0}\int_0^1\int_0^\infty &\left\{\nu_1 R_1(\eta,\xi\eta;\delta_1^B)+\nu_2 R_2(\eta,\xi\eta;\delta_2^B)\right\}\pi(\eta,\xi)d\eta d\xi\\ &= \nu_1 R_1(1,1;\nu_1^{-1}S_1)+\nu_2 R_2(1,1;\nu_2^{-1}S_2). \qquad (4.4)\end{aligned}$$

We first show that

$$\lim_{a\to 0,b\to 0,c\to 0} r_1(\pi,\delta_1^B) = R_1(1,1,\nu_1^{-1}S_1), \qquad (4.5)$$

where

$$r_1(\pi,\delta_1^B)=\int\int R_1(\eta,\eta\xi;\delta_1^B)\pi(\eta,\xi)d\eta d\xi.$$

The Bayes estimator δ_1^B of η^{-1} is given by

$$\delta_1^B=\delta_1^B(a,b,c)=(S_1+b)\psi_{c,a}\left(\frac{S_2}{S_1+b}\right),$$

where

$$\psi_{c,a}(w)=\frac{1}{\nu_1+\nu_2+a}\frac{\int_0^w t^{(\nu_2+c)/2-1}/(1+t)^{(\nu_1+\nu_2+a)/2}dt}{\int_0^w t^{(\nu_2+c)/2-1}/(1+t)^{(\nu_1+\nu_2+a)/2+1}dt},$$

and $r_1(\pi,\delta_1^B)$ is written by

$$\begin{aligned}r_1(\pi,\delta_1^B) &= \int\cdots\int L\left((s_1+b)\psi_{c,a}\left(\frac{s_2}{s_1+b}\right)\eta\right)\eta^2\xi\\ &\quad\times f_1(\eta s_1)f_2(\eta\xi s_2)\pi(\eta,\xi)ds_1 ds_2 d\eta d\xi,\end{aligned}$$

where $f_i(\cdot)$ is a density of $\chi^2_{\nu_i}$ for $i=1,2$. Making the transformations $z=s_2/(s_1+b)$ and $y=(s_1+b)\eta$ with $dz=(s_1+b)^{-1}ds_2$ and $dy=(s_1+b)d\eta$, we have

$$\begin{aligned}r_1(\pi,\delta_1^B) &= \int\cdots\int L\left((s_1+b)\psi_{c,a}(z)\eta\right)\eta^2\xi(s_1+b)\\ &\quad\times f_1(\eta s_1)f_2(\eta\xi(s_1+b)z)\pi(\eta,\xi)ds_1 dz d\eta d\xi\\ &= \int\cdots\int L\left(\psi_{c,a}(z)y\right)\frac{y^2}{(s_1+b)^2}\xi\\ &\quad\times f_1\left(\frac{s_1}{s_1+b}y\right)f_2(\xi yz)\pi\left(\frac{y}{s_1+b},\xi\right)ds_1 dz dy d\xi\\ &= C_1\int\cdots\int L\left(\psi_{c,a}(z)y\right)\frac{s_1^{\nu_1/2-1}}{(s_1+b)^{(\nu_1+a)/2}}\\ &\quad\times y^{(\nu_1+\nu_2+a)/2-1}\xi^{(\nu_2+c)/2-1}z^{\nu_2/2-1}e^{-(1+\xi z)y/2}ds_1 dz dy d\xi,\end{aligned}$$

where

$$C_1 = C_1(\nu_1, \nu_2, a, b, c) = \frac{2^{-(\nu_1+\nu_2+a)/2-1} b^{a/2} c}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)\Gamma(a/2)}.$$

By using the equation

$$\int_0^\infty \frac{s_1^{\nu_1/2-1}}{(s_1+b)^{(\nu_1+a)/2}} ds_1 = \frac{\Gamma(\nu_1/2)\Gamma(a/2)}{b^{a/2}\Gamma\left((\nu_1+a)/2\right)},$$

$r_1(\pi, \delta_1^B)$ is rewritten by

$$\begin{aligned} r_1(\pi, \delta_1^B) &= C_2 \int_0^1 \int_0^\infty \int_0^\infty L\left(\psi_{c,a}(z)y\right) \\ &\quad \times y^{(\nu_1+\nu_2+a)/2-1} \xi^{(\nu_2+c)/2-1} z^{\nu_2/2-1} e^{-(1+\xi z)y/2} dz dy d\xi, \end{aligned}$$

where

$$C_2 = C_2(\nu_1, \nu_2, a, c) = \frac{2^{-(\nu_1+\nu_2+a)/2-1} c}{\Gamma((\nu_1+a)/2)\Gamma(\nu_2/2)}.$$

Making a tend to zero, we see that $r_1(\pi, \delta_1^B)$ approaches

$$\begin{aligned} r_1^*(\pi, \delta_1^B) &= \frac{2^{-(\nu_1+\nu_2)/2-1} c}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^1 \int_0^\infty \int_0^\infty L\left(\psi_{c,0}(z)y\right) \\ &\quad \times y^{(\nu_1+\nu_2)/2-1} \xi^{(\nu_2+c)/2-1} z^{\nu_2/2-1} e^{-(1+\xi z)y/2} dz dy d\xi. \end{aligned}$$

Making the transformations $x = \xi z$ and $\tau = \xi^{a/2}$ in turn with $dx = \xi dz$ and $d\tau = (c/2)\xi^{c/2-1} d\xi$ gives that

$$\begin{aligned} r_1^*(\pi, \delta_1^B) &= \frac{2^{-(\nu_1+\nu_2)/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^1 \int_0^\infty \int_0^\infty L\left(\psi_{c,0}\left(x\tau^{-2/c}\right) y\right) \\ &\quad \times y^{(\nu_1+\nu_2)/2-1} x^{\nu_2/2-1} e^{-(1+x)y/2} dx dy d\tau. \end{aligned}$$

Here note that

$$\lim_{c\to 0} \psi_{c,0}\left(x\tau^{-2/c}\right) = \lim_{w\to\infty} \psi_{0,0}(w) = \nu_1^{-1}.$$

Applying the Lebesgue's dominated convergence theorem, we get

$$\begin{aligned} \lim_{c\to 0} r_1^*(\pi, \delta_1^B) &= \frac{2^{-(\nu_1+\nu_2)/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^1 \int_0^\infty \int_0^\infty L\left(\nu_1^{-1} y\right) \\ &\quad y^{(\nu_1+\nu_2)/2-1} x^{\nu_2/2-1} e^{-(1+x)y/2} dx dy d\tau \\ &= \frac{2^{-\nu_1/2}}{\Gamma(\nu_1/2)} \int_0^\infty L\left(\nu_1^{-1} y\right) y^{\nu_1/2-1} e^{-y/2} dy \\ &= R_1(1, 1; \nu_1^{-1} S_1), \end{aligned}$$

which proves (4.5).

On the other hand, the Bayes estimator δ_2^B of θ^{-1} is given by

$$\delta_2^B = S_2 \phi_{c,a}\left(\frac{S_1+b}{S_2}\right),$$

where

$$\phi_{c,a}(w) = \frac{1}{\nu_1+\nu_2+a} \frac{\int_0^{1/w} t^{(\nu_2+c)/2-1}/(1+t)^{(\nu_1+\nu_2+a)/2}dt}{\int_0^{1/w} t^{(\nu_2+c)/2}/(1+t)^{(\nu_1+\nu_2+a)/2+1}dt}.$$

Let $r_2(\pi, \delta_2^B) = \int\int R_2(\eta, \eta\xi; \delta_2^B)\pi(\eta,\xi)d\eta d\xi$. In order to assert that

$$\lim_{a\to 0, b\to 0, c\to 0} r_2(\pi, \delta_2^B) = R_2(1, 1, \nu_2^{-1}S_2), \tag{4.6}$$

the same arguments as in the proof of (4.5) gives that

$$\begin{aligned} r_2(\pi, \delta_2^B) &= C_2 \int_0^1 \int_0^\infty \int_0^\infty L\left(\phi_{c,a}(z^{-1})\xi z y\right) \\ &\quad \times y^{(\nu_1+\nu_2+a)/2-1}\xi^{(\nu_2+c)/2-1}z^{\nu_2/2-1}e^{-(1+\xi z)y/2}dzdyd\xi, \end{aligned}$$

which converges

$$\begin{aligned} r_2^*(\pi, \delta_2^B) &= \frac{2^{-(\nu_1+\nu_2)/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^1 \int_0^\infty \int_0^\infty L\left(\phi_{c,0}\left(x^{-1}\tau^{2/c}\right)xy\right) \\ &\quad \times y^{(\nu_1+\nu_2)/2-1}x^{\nu_2/2-1}e^{-(1+x)y/2}dxdyd\tau \end{aligned}$$

as a tends to zero. Since $\lim_{c\to 0}\phi_{c,0}(x^{-1}\tau^{2/c}) = \phi_{c,0}(0) = \nu_2^{-1}$, the Lebesgue's dominated convergence theorem can be applied to get

$$\begin{aligned} \lim_{c\to 0} r_2^*(\pi, \delta_2^B) &= \frac{2^{-(\nu_1+\nu_2)/2}}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \int_0^1 \int_0^\infty \int_0^\infty L\left(\nu_2^{-1}xy\right) \\ &\quad \times y^{(\nu_1+\nu_2)/2-1}x^{\nu_2/2-1}e^{-(1+x)y/2}dxdyd\tau \\ &= \frac{2^{-\nu_2/2}}{\Gamma(\nu_2/2)} \int_0^\infty L\left(\nu_2^{-1}v\right)v^{\nu_2/2-1}e^{-v/2}dv \qquad (4.7) \\ &= R_2(1, 1; \nu_2^{-1}S_2), \end{aligned}$$

where the transformation $v = yx$ with $dv = ydx$ is made in the second equality in (4.7). Combining (4.5) and (4.7) proves (4.4) and it contradicts the inequality (4.1). Therefore the minimaxity of $(\hat{\sigma}_e^{2UB}, \hat{\sigma}_v^{2UB})$ is established. □□

We next prove Theorems 3 and 4 in the general mixed linear models (3.1) since Theorems 1 and 2 are included by Theorems 3 and 4.

The key tool for these proofs is the *Integral-Expression-of-Risk-Difference* (IERD) method given by Kubokawa(1994a,b) in estimation of a scale parameter. The IERD method is, through the fundamental theorem of calculus, to give an integral-expression for a difference of risks of two estimators. For other instances in which the IERD method was applied, see Kubokawa *et al.* (1993a, 94).

Proof of Theorems 1 and 3. Since $\lim_{w\to\infty}\psi(w) = \nu_1^{-1}$, from the IERD method of Kubokawa (1994a,b), we have

$$
\begin{aligned}
&R_1(\omega; S_1\nu_1^{-1}) - R_1\left(\omega; S_1\psi\left(\frac{S_2}{S_1}\right)\right) \\
&= E\left[\left\{\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}t\right) - \log\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}t\right) - 1\right\}\Bigg|_{t=1}^{\infty}\right] \\
&= E\left[\int_1^\infty \frac{d}{dt}\left\{\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}t\right) - \log\frac{S_1}{\sigma_e^2}\psi\left(\frac{S_2}{S_1}t\right) - 1\right\}dt\right].
\end{aligned}
\tag{4.8}
$$

Let $v = S_1/\sigma_e^2$, $u_i = \boldsymbol{x}'\boldsymbol{E}_i\boldsymbol{x}/(\sigma_e^2 + \lambda_i\sigma_v^2)$, and $\theta_i = 1 + \lambda_i\sigma_v^2/\sigma_e^2$, and denote the density functions of v and u_i by f and g_i, respectively. Carrying out the differentiation in (4.8) gives

$$
\begin{aligned}
&E\left[\int_1^\infty \left\{\frac{S_1}{\sigma_e^2} - \frac{1}{\psi(S_2t/S_1)}\right\}\frac{S_2}{S_1}\psi'\left(\frac{S_2}{S_1}t\right)dt\right] \\
&= \int\cdots\int\int_1^\infty\left\{v - \frac{1}{\psi(\Sigma\theta_i u_i v/t)}\right\}(\Sigma\theta_i u_i/v)\,\psi'(\Sigma\theta_i u_i t/v)\,dt \\
&\qquad \times f(v)\Pi_i g_i(u_i)dv\Pi_i du_i.
\end{aligned}
$$

Making the transformations $(t/v)u_i = w_i$ and $1/t = x$ in order, we observe that the r.h.s. of (4.8) is equal to

$$
\begin{aligned}
&\int\cdots\int\int_1^\infty\left\{v - \frac{1}{\psi(\Sigma\theta_i w_i)}\right\}(\Sigma\theta_i w_i/t)\,\psi'(\Sigma\theta_i w_i t) \\
&\qquad \times\left(\frac{v}{t}\right)^\ell f(v)\Pi_i g_i(vw_i/t)\,dtdv\Pi_i dw_i \\
&= \int\cdots\int\left\{v - \frac{1}{\psi(\Sigma\theta_i w_i)}\right\}(\Sigma\theta_i w_i)\,\psi'(\Sigma\theta_i w_i) \\
&\qquad \times v^\ell f(v)\int_0^1 x^{\ell-1}\Pi_i g_i(w_i v x)\,dxdv\Pi_i dw_i.
\end{aligned}
\tag{4.9}
$$

Since $\psi'(w) \geq 0$, it is concluded that the r.h.s. of (4.9) is nonnegative if

$$
\psi(\Sigma\theta_i w_i) \geq \frac{\int_0^\infty v^\ell f(v)\int_0^1 x^{\ell-1}\Pi_i g_i(w_i v x)dxdv}{\int_0^\infty v^{\ell+1} f(v)\int_0^1 x^{\ell-1}\Pi_i g_i(w_i v x)dxdv}.
\tag{4.10}
$$

Since $\theta_i \geq 1$ and $\psi'(w) \geq 0$, it follows that $\psi(\Sigma\theta_i w_i) \geq \psi(\Sigma w_i)$, which, from (4.10), gives the sufficient condition that $\psi(\Sigma w_i)$ is greater than or equal to the r.h.s. of (4.10).

Integrating out the r.h.s. of (4.10) with respect to v yields $\psi_0(\Sigma w_i)$ given by (2.3). Hence the inequality (4.10) is guaranteed by the condition (b) of Theorems 1 and 3, which are established. □□

Proof of Theorems 2 and 4. Since $\phi(0) = \nu_2^{-1}$, observe that

$$\begin{aligned} & R_2(\omega; S_2\nu_2^{-1}) - R_2(\omega; S_2\phi\left(\frac{S_1}{S_2}\right)) \qquad (4.11) \\ = & -E\left[\int_0^1 \frac{d}{dt}\left\{\frac{S_2}{\sigma_e^2+\lambda\sigma_v^2}\phi\left(\frac{S_1}{S_2}t\right) - \log\frac{S_2}{\sigma_e^2+\lambda\sigma_v^2}\phi\left(\frac{S_1}{S_2}t\right) - 1\right\}dt\right] \\ = & \int\cdots\int\int_0^1\left\{\frac{1}{\phi(vt/\Sigma\theta_i u_i)} - \frac{\Sigma\theta_i u_i}{1+\lambda\tau}\right\}\frac{v}{\Sigma\theta_i u_i}\phi'\left(\frac{vt}{\Sigma\theta_i u_i}\right)dt \\ & \times f(v)\Pi_i g_i(u_i)dv\Pi_i du_i \end{aligned}$$

for $\tau = \sigma_v^2/\sigma_e^2$. Making the transformations $(t/\Sigma\theta_i u_i)v = w$ and $w(1/t) = y$ in order, we can rewrite (4.11) as

$$\begin{aligned} & \int\cdots\int\int_0^1\left\{\frac{1}{\phi(w)} - \frac{\Sigma\theta_i u_i}{1+\lambda\tau}\right\}\phi'(w)w\Sigma\theta_i u_i/t^2 \\ & \qquad \times f(\Sigma\theta_i u_i w/t)\Pi_i g_i(u_i)dt\Pi_i du_i dw \\ = & \int\cdots\int\int_w^\infty\left\{\frac{1}{\phi(w)} - \frac{\Sigma\theta_i u_i}{1+\lambda\tau}\right\}\phi'(w)\Sigma\theta_i u_i \qquad (4.12) \\ & \qquad \times f(\Sigma\theta_i u_i y)\Pi_i g_i(u_i)dy\Pi_i du_i dw \end{aligned}$$

so that since $\phi'(w) \geq 0$, the l.h.s. of (4.11) is nonnegative if

$$\phi(w) \leq \frac{\int\cdots\int\int_w^\infty(\Sigma\theta_i u_i)f(\Sigma\theta_i u_i y)\Pi_i g_i(u_i)dy\Pi_i du_i}{\int\cdots\int\int_w^\infty(\Sigma\theta_i u_i)^2/(1+\lambda\tau)f(\Sigma\theta_i u_i y)\Pi_i g_i(u_i)dy\Pi_i du_i}. \qquad (4.13)$$

Letting $s = \Sigma_{i=1}^{\ell} u_i$ and $z_i = u_i/s$, we see that

$$\begin{aligned} s & \sim \chi^2_{\nu_2}, \\ z_i & \sim Beta(m_i/2, \Sigma_{j\neq i}m_j/2) \end{aligned}$$

and s and z_i are independent. Let $Q = \Sigma\theta_i z_i$, being independent of s. The r.h.s. of (4.13) can be rewritten as

$$\begin{aligned} & \frac{E^Q[\int\int_w^\infty Qsf(Qsy)g(s)dyds]}{E^Q[\int\int_w^\infty Q^2s^2/(1+\lambda\tau)f(Qsy)g(s)dyds]} \qquad (4.14) \\ = & \frac{E^Q[\int_{Qw}^\infty\int sf(sx)g(s)dsdx]}{E^Q[Q/(1+\lambda\tau)\int_{Qw}^\infty\int s^2f(sx)g(s)dsdx]}, \end{aligned}$$

where $g(s)$ is a density of $\chi^2_{\nu_2}$. Since Q and $\int_{Qw}^{\infty} \int s^2 f(sx)g(s)dsdx$ are monotone in the opposite directions, we can show the following inequality for the denominator of the r.h.s. of (4.14):

$$E^Q\left[\frac{Q}{1+\lambda\tau}\int_{Qw}^{\infty}\int s^2 f(sx)g(s)dsdx\right] \leq E^Q\left[\frac{Q}{1+\lambda\tau}\right]E^Q\left[\int_{Qw}^{\infty}\int s^2 f(sx)g(s)dsdx\right]. \tag{4.15}$$

Here observe that

$$\begin{aligned} E^Q\left[\frac{Q}{1+\lambda\tau}\right] &= \frac{1}{1+\lambda\tau}+\frac{1}{1+\lambda\tau}\sum_{i=1}^{\ell}\lambda_i\tau E[z_i] \\ &= \frac{1}{1+\lambda\tau}+\frac{1}{1+\lambda\tau}\frac{\Sigma\lambda_i\tau m_i}{\nu_2}=1, \end{aligned} \tag{4.16}$$

since $\lambda = \sum \lambda_i m_i/\nu_2$. Combining (4.13), (4.14), (4.15) and (4.16) gives a sufficient condition as

$$\phi(w) \leq \frac{E^Q[\int_{Qw}^{\infty}\int sf(sx)g(s)dsdx]}{E^Q[\int_{Qw}^{\infty}\int s^2 f(sx)g(s)dsdx]}. \tag{4.17}$$

Furthermore the r.h.s. of (4.17) can be shown to be greater than or equal to

$$\inf_Q\left\{\frac{\int_{Qw}^{\infty}\int sf(sx)g(s)dsdx}{\int_{Qw}^{\infty}\int s^2 f(sx)g(s)dsdx}\right\}. \tag{4.18}$$

By integrating out the numerator and denominator of (4.18) with respect to s, it is seen that this quantity is expressed as

$$\begin{aligned} & \frac{1}{\nu_1+\nu_2}\frac{\int_{Qw}^{\infty} x^{\nu_1/2-1}/(1+x)^{(\nu_1+\nu_2)/2}dx}{\int_{Qw}^{\infty} x^{\nu_1/2-1}/(1+x)^{(\nu_1+\nu_2+2)/2}dx} \\ &= \frac{1}{\nu_1+\nu_2}\frac{\int_0^{1/Qw} z^{\nu_2/2-1}/(1+z)^{(\nu_1+\nu_2)/2}dz}{\int_0^{1/Qw} z^{\nu_2/2}/(1+z)^{(\nu_1+\nu_2+2)/2}dx} \\ &= \phi_0(Qw), \end{aligned} \tag{4.19}$$

where $\phi_0(w)$ is defined by (2.4) and the first equality of (4.19) can be obtained by making the transformation $x = z^{-1}$. From Lemma 1, it follows that $\phi_0(w)$ is nondecreasing, so that $\phi_0(Qw) \geq \phi_0(w)$ since $Q \geq 1$. Combining this inequality, (4.17), (4.18) and (4.19), we get the sufficient condition that $\phi(w) \leq \phi_0(w)$, which is just the condition (b), and the proofs of Theorems 2 and 4 are complete. □□

Acknowledgment. This research has been supported by the NSERC grant no. A3088, CANADA, during the visit of the first and the third authors at Carleton University, Ottawa, 1994. The research of the first author was also supported in part by a grant from the Shirakawa Institute of Animal Genetics and the Japan Livestock Technology Association, by the Ministry of Education, Japan, Grant No. 08780216, No. 09780214 and by a grant from the Research Institute for the Japanese Economy, University of Tokyo. I am grateful to Mr. M. Ushijima for his help in the simulation experiments.

REFERENCES

Battese, G.E., Harter, R.M. and Fuller, W.A.(1988). An error-components model for prediction of county crop areas using survey and satellite data. *J. Amer. Statist. Assoc.*, **83**, 28-36.

Chow, S.C. and Shao, J.(1988). A new procedure for the estimation of variance components. *Statist. Probab. Letters*, **6**, 349-355.

Fuller, W.A. and Battese, G.E.(1973). Transformations for estimation of linear models with nested-error structure. *J. Amer. Statist. Assoc.*, **68**, 626-632.

Hartung, J.(1981). Nonnegative minimum biased invariant estimation in variance component models. *Ann. Statist.*, **9**, 278-292.

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

Herbach, L.H.(1959). Properties of model II-type analysis of variance tests, A: Optimum nature of the F-test for model II in the balanced case. *Ann. Math. Statist.*, **30**, 939-959.

Kleffe, J. and Rao, J.N.K.(1986). The existence of asymptotically unbiased nonnegative quadratic estimates of variance components in ANOVA models. *J. Amer. Statist. Assoc.* **81**, 692-698.

Kubokawa, T.(1994a). A unified approach to improving equivariant estimators. *Ann. Statist.* **22**, 290-299.

Kubokawa,T.(1994b). Double shrinkage estimation of ratio of scale parameters. *Ann. Inst. Statist. Math.*, **46**, 95-116.

Kubokawa, T.(1995). Estimation of variance components in mixed linear models. *J. Multivariate Anal.*, **53**, 210-236.

Kubokawa, T. and Robert, C.P.(1994). New perspectives on linear calibration. *J. Multivariate Anal.*, **51**, 178-200.

Kubokawa, T., Robert, C. and Saleh, A.K.Md.E. (1992). Empirical Bayes estimation of the variance parameter of a normal distribution with unknown mean under an entropy loss. *Sankhya* (Ser. A), **54**, 402-410.

Kubokawa, T., Robert, C.P. and Saleh, A.K.Md.E.(1993a). Estimation of noncentrality parameters. *Canad. J. Statist.*, **21**, 45-57.

Kubokawa, T., Saleh, A.K.Md.E. and Makita, S.(1993b). On improved positive estimators of variance components. *Statistics Decisions*, Supplement Issue 3, 1-16.

Kubokawa, T. and Srivastava, M.S. (1996). Double shrinkage estimators of ratio of variances. In *Multidimensional Statistical Analysis and Theory of Random Matrices* (Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, Editors: A.K. Gupta and V.L. Girko), 139-154, VSP, Netherlands.

LaMotte, L.R.(1973). Invariant quadratic estimators in the random, one-way ANOVA model. *Biometrics* **32**, 793-804.

Mathew, T.(1984). On nonnegative quadratic unbiased estimability of variance components. *Ann. Statist.*, **12**, 1566-1569.

Mathew, T., Sinha, B.K. and Sutradhar, B.C.(1991). Nonnegative estimation of random effects variance components in balanced mixed models. *Proceedings of the International Symposium on Nonparametric Statistics and Related Topics* (ed. A.K.Md.E. Saleh).

Mathew, T., Sinha, B.K. and Sutradhar, B.C.(1992). Nonnegative estimation of variance components in unbalanced mixed models with two variance components. *J. Multivariate Anal.*, **42**, 77-101.

Peixoto, J.L. and Harville, D.A.(1986). Comparisons of alternative predictors under the balanced one-way random model. *J. Amer. Statist. Assoc.*, **81**, 431-436.

Portnoy, S.(1971). Formal Bayes estimation with application to a random effect model. *Ann. Math. Statist.* **42**, 1379-1402.

Rao, C.R.(1971a). Estimation of variance and covariance components-MINQUE theory. *J. Multivariate Anal.*, **1**, 257-275.

Rao, C.R.(1971b). Minimum variance quadratic unbiased estimation of variance components. *J. Multivariate Anal.*, **1**, 445-456.

Rao, J.N.K., Sutradhar, B.C. and Yue, K.(1993). Generalized least squares F test in regression analysis with two-stage cluster samples. *J. Amer. Statist. Assoc.*, **88**, 1388-1391.

Robert, C.P. (1994). *The Bayesian Choice : A Decision-Theoretic Motivation.* Springer-Verlag, New York.

Thompson, W.A.Jr.(1962). The problem of negative estimates of variance components. *Ann. Math. Statist.*, **33**, 273-289.