The Annals of Statistics
1989, Vol. 17, No. 1, 268–281

ALL ADMISSIBLE LINEAR ESTIMATORS OF THE VECTOR OF GAMMA SCALE PARAMETERS WITH APPLICATION TO RANDOM EFFECTS MODELS

BY ROGER H. FARRELL, WITOLD KLONECKI AND STEFAN ZONTEK

Cornell University, Polish Academy of Sciences and Polish Academy of Sciences

The paper is devoted to the problem of simultaneous estimation of scale and natural parameters of the multiparameter gamma distribution under a quadratic loss. The vector of the scale parameters is assumed to range over a certain subset of the Cartesian product $\mathscr{R}_+^n$ of n positive half lines. We identify the class of all linear admissible estimators for the scale parameters and show that all linear estimators of the natural parameters are inadmissible.

Since the problem of invariant quadratic estimation of variance components in balanced random effects normal models leads to a problem of linear estimation of parametric functions of gamma scale parameters restricted to subsets of $\mathscr{R}_+^n$ being considered in this paper, some results on admissibility of invariant quadratic estimators of variance components are also established.

1. Introduction. Let $Y = (Y_1, \dots, Y_n)'$ be a random vector, where the Y_i are independent gamma (α_i, θ_i) random variables, having density [on $(0, \infty)$]

$$y_i^{\alpha_i - 1} \exp\{-y_i/\theta_i\}/\theta_i^{\alpha_i}\Gamma(\alpha_i).$$

For notational convenience set $A = \operatorname{diag}(\alpha_1, \dots, \alpha_n)$ and $\theta = (\theta_1, \dots, \theta_n)'$. Here A is a known p.d. matrix and θ is the vector of scale parameters taking values in

$$\Theta = \{H'\sigma\colon \sigma \in \mathscr{R}^k, \sigma > 0\}.$$

The symbol $a > 0$ ($a \geq 0$), where $a \in \mathscr{R}^k$, means that all components of the vector a are positive (nonnegative) and H stands for a $k \times n$, $k \leq n$, matrix of rank k with nonnegative entries and having the property that $H'\sigma > 0$ when $\sigma > 0$. The closure $\overline{\Theta}$ of Θ is a finite cone generated by the k linearly independent column vectors of matrix H', which lay in the Cartesian product $\mathscr{R}_+^n$ of n positive half lines.

Under the above setup

$$\begin{aligned} EY &= A\theta, \\ \operatorname{cov} Y &= A \operatorname{diag}(\theta_1^2, \dots, \theta_n^2). \end{aligned} \tag{1.1}$$

When $\overline{\Theta}$ is the entire convex cone of vectors $\theta \geq 0$, the model is called unrestricted, otherwise it is called a restricted model.

Received February 1987; revised May 1988.

AMS 1980 *subject classifications*. Primary 62C15; secondary 62F10.

Key words and phrases. Gamma scale and natural parameters, simultaneous estimation, squared error loss, linear estimators, admissibility, inadmissibility, variance components.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics.*
www.jstor.org

We assume that the loss of estimating $F'\theta$ by a random vector δ, where F is any $n \times m$ matrix, is

$$(\delta - F'\theta)'\Lambda(\delta - F'\theta), \tag{1.2}$$

where Λ is a n.n.d. $m \times m$ matrix. If $\Lambda = I_m$, I_m being the $m \times m$ identity matrix, then (1.2) becomes the ordinary squared error loss. Throughout the paper we assume that Λ is p.d. unless indicated otherwise.

We denote by $\mathscr{N}(F)$ the class of all estimators of $F'\theta$ having finite risk and by $\mathscr{L}(F)$ the subclass of all linear estimators in $\mathscr{N}(F)$. Moreover, $\mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F)$ and $\mathscr{L}_{\mathrm{ad}|\mathscr{N}}(F)$ stand for the class of linear estimators admissible within $\mathscr{L}(F)$ and $\mathscr{N}(F)$, respectively. We shall delete F when it is the identity matrix, i.e., when the parameter θ itself is being estimated.

As usual $\mathbf{1}_k$ stands for the k-vector of 1's and $\mathbf{0}_{k\times l}$ for the $k \times l$ 0 matrix. The indices will be deleted when they can be guessed from the context. For any $n \times n$ matrix M, the symbol $M > 0$ ($M \geq 0$) means that M is p.d. (n.n.d.), $|M|$ denotes the determinant of M, M^+ the Moore–Penrose g-inverse of M, $\operatorname{tr} M$ the trace of M and M_d a diagonal matrix with the ith diagonal element equal to the ith diagonal element of M. Throughout the paper P stands for an $n \times n$ n.n.d. diagonal matrix such that $\operatorname{tr} P = 1$ and $\alpha = \operatorname{tr} A$.

We make extensive use of a result due to Klonecki and Zontek (1987) [see also Zontek (1988)] which provides a complete characterization of $\mathscr{L}_{\mathrm{ad}|\mathscr{L}}$, i.e., of the class of all linear admissible estimators of θ within the class of all linear estimators under loss (1.2).

THEOREM 1.1. *The class $\mathscr{L}_{\mathrm{ad}|\mathscr{L}}$ is given by*

$$\{L_G'Y\colon L_G = (I + GA)^{-1}G,\ G \in \bar{\mathscr{G}}\},$$

where $\bar{\mathscr{G}}$ is the closure of a set $\mathscr{G}$ defined by

$$\mathscr{G} = \{\Delta_d^{-1}\Delta\colon \Delta \in \Omega\},$$

and where Ω is the convex hull

$$\Omega = \operatorname{conv}\{H'\sigma\sigma'H\colon \sigma \in \mathscr{R}^k, \sigma > 0\}.$$

This theorem can be established by showing that for every matrix G in $\mathscr{G}$ the corresponding estimator for $L'Y$ is a unique Bayes estimator within $\mathscr{L}$ and that, if $L_{G_r} \to L$ as $r \to \infty$, where $\{G_r\} \subset \mathscr{G}$, then $L'Y$ is admissible within $\mathscr{L}$.

The main result in this paper gives a complete characterization of $\mathscr{L}_{\mathrm{ad}|\mathscr{N}}$, i.e., of the class of all linear estimators of θ admissible within the class of all estimators of θ. Using the notation appearing in Theorem 1.1 it can be stated as follows.

THEOREM 1.2. *For all $n \geq 1$, $\mathscr{L}_{\mathrm{ad}|\mathscr{N}} = \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$, where*

$$\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}} = \{L_G'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}\colon \operatorname{rank} G = 1,\ G \in \bar{\mathscr{G}}\}.$$

One can easily verify that $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ consists of all estimators of the form

$$\frac{\mathbf{1}'KY}{1+\alpha}K^{-1}\mathbf{1},$$

where K may be any p.d. $n \times n$ diagonal matrix such that $K^{-1}\mathbf{1}$ belongs to $\overline{\Theta}$.

Clearly, for $n = 1$ the set $\mathscr{L}_{\mathrm{ad}|\mathscr{N}}$ consists of one element $(1/(1+\alpha))Y$ so that the assertion of this theorem reduces to the celebrated result of Karlin (1958).

For $n \geq 2$ the theorem is a direct consequence of Theorem 3.1(i) and Theorem 3.2 given in Section 3. The first theorem asserts that all estimators in $\mathscr{L} \setminus \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ are inadmissible. The second theorem states that every estimator in $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ is admissible among the class of all estimators, which turns out to be a consequence of Karlin's theorem, mentioned above. The first result is established by showing that for every estimator $L'Y$ in $\mathscr{L} \setminus \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ there exists an improved estimator of the form

$$L'Y + cY_1^{p_1}Y_2^{p_2} \cdots Y_n^{p_n}, \tag{1.3}$$

where $P = \mathrm{diag}(p_1, \dots, p_n) \geq 0$, $\mathrm{tr}\, P = 1$, while $c \in \mathscr{R}^n$. This class of estimators is broader than the class of estimators,

$$(I + A)^{-1}\left[Y + c(Y_1Y_2 \cdots Y_n)^{1/n}\mathbf{1}\right], \qquad c \in \mathscr{R},$$

suggested by Das Gupta (1986) for improving the standard estimator $(I + A)^{-1}Y$ of θ within the unrestricted model under loss (1.2). The estimators suggested by Das Gupta as well as the broader class (1.3) permit exact analytical representation of the risk, as is shown in the sequel.

The inadmissibility of the standard estimator of θ for $n \geq 2$ was established by Berger (1980) by finding improved estimators of a different structure than (1.3).

Using a well known lemma due to Shinozaki (1975) [see also Rao (1976)] we establish in Section 4 some results concerning characterizations of admissible linear estimators for parametric functions $F'\theta$ and extend a result of Das Gupta and Sinha (1986). Since the models with restrictions cover models encountered for invariant quadratic estimation of variance components in random and mixed models, in Section 5 some results on admissible estimation of variance components will be also formulated. In the notation of model (1.1) the vector of the variance components is represented by the vector parameter σ or, equivalently, by $F'\theta$ with F being equal to $H'(HH')^{-1}$. Finally, in Section 6 we show that if $n \geq 2$, then all linear estimators based on $(1/Y_1, \dots, 1/Y_n)'$ of the vector of natural parameters $\eta = (1/\theta_1, \dots, 1/\theta_n)'$ are inadmissible. This latter result extends theorems due to Berger (1980) and Das Gupta (1986).

2. Basic lemma. To show that a linear estimator $L'Y$ of $F'\theta$ may be improved by an estimator of form (1.3) it suffices to show that there exist a matrix P and a vector $c \in \mathscr{R}^m$ such that the risk difference between (1.3) and $L'Y$ is nonnegative for all $\sigma > 0$ and positive for at least one such σ.

To write the formula for the risk function in a compact form we introduce the following notation. For any vector $a = (a_1, \dots, a_n)'$ and any n.n.d. diagonal matrices $B = \operatorname{diag}(b_1, \dots, b_n)$ and $P = \operatorname{diag}(p_1, \dots, p_n)$ let

$$a_P = \prod_{i=1}^{n} a_i^{p_i}$$

and let

$$\gamma_{B,P} = \prod_{i=1}^{n} \frac{\Gamma(b_i + p_i)}{\Gamma(p_i)}.$$

Then the risk difference between $L'Y$ and

$$L'Y + \kappa c Y_P, \tag{2.1}$$

where $\kappa \in \mathscr{R}$, can be expressed as

$$\Delta R(\sigma) = \kappa \gamma_{A,P} \theta_P [2\sigma' q_H - \kappa \gamma_{A+P,P} c' \Lambda c \theta_P],$$

where

$$q_H = H[F - (A + P)L]\Lambda c. \tag{2.2}$$

Now let $\tau = \sup \theta_P$ with the supremum taken over

$$\Sigma_H = \{\sigma \in \mathscr{R}^k: \sigma' q_H = 1, \sigma > 0\},$$

say, and notice that there exists a $\kappa > 0$ such that $\Delta R(\sigma) > 0$ for every σ in Σ_H if and only if τ is finite. In fact, if τ is finite, then the desired inequality holds when

$$0 < \kappa < \frac{2}{\tau \gamma_{A+P,P} c' \Lambda c},$$

because for such κ and $\sigma \in \Sigma_H$,

$$\Delta R(\sigma) > \kappa \gamma_{A,P} \theta_P [2 - \kappa \gamma_{A+P,P} c' \Lambda c \tau] > 0.$$

Since under the adopted assumptions $\Delta R(t\sigma) = t^2 \Delta R(\sigma)$ for all vectors $\sigma > 0$ and all positive numbers t, one can easily see that $\Delta R(\sigma) > 0$ for every $\sigma > 0$ if and only if $\Delta R(\sigma) > 0$ for every $\sigma \in \Sigma_H$.

Putting $q_H = (q_1, \dots, q_k)'$ and $H = (h_{ij})$ we can hence formulate the following result which is basic for the discussion in the sequel.

LEMMA 2.1. *There exists a constant $\kappa > 0$ such that estimator* (2.1) *is better than $L'Y$ for $F'\theta$ if and only if $q_H \geq' 0$ and if $q_i = 0$ and $h_{ij} = 0$ for some $1 \leq i \leq k$ and $1 \leq j \leq n$, then $p_j = 0$.*

REMARK 2.1. For an unrestricted model with $H = I_n$, $\Lambda = I_m$ and $P = (1/\alpha)A$, formula (2.2) becomes

$$q = \left[F - \left(1 + \frac{1}{\alpha}\right)AL\right]c, \qquad c \in \mathscr{R}^m.$$

If $q > 0$, then $L'Y$ is inadmissible. For $m = 1$, i.e., for $L, F \in \mathscr{R}^n$, this sufficient condition for inadmissibility was obtained by Das Gupta and Sinha

(1986). Notice that if $L = (I + GA)^{-1}GF$, where $F \in \mathscr{R}^n$, then $q > 0$ implies that $Fc > 0$ and that $Lc > 0$, where $c \in \mathscr{R}$ (see Theorem 2.1).

These developments show that improvements of linear estimators of $F'\theta$ are possible only in some directions and with limited shifts. The following result which holds for any n.n.d. matrix Λ indicates directions in which improvements are possible for linear admissible estimators $L'Y$ belonging to

$$\{(LF)'Y\colon L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}\} \setminus \{(LF)'Y\colon L'Y \in \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}\}.$$

THEOREM 2.1. *If $L'Y + cY_P$ dominates $L'Y$ and if P is p.d., then $HF\Lambda c > 0$.*

PROOF. As usual let G be a matrix in $\overline{\mathscr{G}}$ and let rank $G > 1$. First we show that if $h(G)c > 0$, where

$$h(G) = H\left[I - (A + P)(I + GA)^{-1}G\right]H'(HH')^{-1},$$

then $c > 0$.

From the definition of $\overline{\mathscr{G}}$ it follows that there exists a sequence $\{w_r\}$ of p.d. matrices with nonnegative entries in Ω such that

$$G_r = (H'w_rH)_d^{-1}H'w_rH \to G$$

as $r \to \infty$, so that $h(G_r)c > 0$ for sufficiently large r. Thus without loss of generality we may assume that

$$G = (H'wH)_d^{-1}H'wH,$$

where $w \in \Omega$.

Now we define a matrix S by

$$S = P^{-1}(H'wH)_d - H'wH$$

and notice that

$$P^{1/2}(H'wH)_d^{-1/2}S(H'wH)_d^{-1/2}P^{1/2} = I_n - T,$$

where

$$T = P^{1/2}(H'wH)_d^{-1/2}H'wH(H'wH)_d^{-1/2}P^{1/2}.$$

The matrix T is n.n.d. and $\operatorname{tr} T = 1$. Since rank $G > 1$, this ensures that its eigenvalues are in $[0, 1)$, so that S is p.d. From Theorem 12.2.9 in Graybill (1983) it then follows that all entries of S^{-1} are nonnegative.

Next notice that $h(G)$ can be written as $h(G) = UW$, where

$$U = I_k - HP(H'wH)_d^{-1}H'w,$$

while

$$W = H(I_n + AG)^{-1}H'(HH')^{-1}.$$

Since the inverses of these matrices

$$U^{-1} = I_k + HS^{-1}H'w$$

and

$$W^{-1} = I_k + HA(H'wH)_d^{-1}H'w$$

have only nonnegative entries, we see that all coordinates of c must be positive if $h(G)c > 0$. To end the proof suppose that

$$\left[(I + GA)^{-1}GF\right]'Y + cY_P$$

dominates

$$\left[(I + GA)^{-1}GF\right]'Y$$

and that P is p.d. Then by Lemma 2.1

$$q_H = H\left[F - (A + P)(I + GA)^{-1}GF\right]\Lambda c > 0.$$

Since $q_H = h(G)HF\Lambda c$, the first part of the proof leads to the desired conclusion. □

REMARK 2.2. For $n = 2$ and $m = 1$ the assertion of this lemma may be strengthened. If $L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F)$, where $F \in \mathscr{R}^2$, and if $L'Y + cY_P$ dominates $L'Y$, then $Fc > 0$.

3. Estimation of the scale parameter vector θ. In this section we establish the main results of the paper. The first theorem asserts that linear estimators that are not in $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ are inadmissible among the class of all estimators. The second one states that the remaining linear estimators, i.e., the estimators in $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ are admissible among the class of all estimators under the quadratic loss function.

Throughout this section we assume that Λ is the identity matrix.

THEOREM 3.1. (i) *There exists an estimator of form* (2.1) *dominating a linear estimator $L'Y$ of θ if and only if*

$$L'Y \in \mathscr{L} \setminus \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}.$$

(ii) *If $L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}} \setminus L^0_{\mathrm{ad}|\mathscr{L}}$, then for every p.d. matrix P there exists an estimator of the form* (2.1) *dominating $L'Y$.*

PROOF. (i) First suppose that $L'Y \in \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$. Then for any matrix P and any vector $c \in \mathscr{R}^n$ it follows from (2.2) that $q_H = HQc$, where

$$Q = I - \frac{1}{1 + \alpha}(A + P)K\mathbf{1}\mathbf{1}'K^{-1}, \tag{3.1}$$

while $K^{-1}\mathbf{1} \in \overline{\Theta}$, K being a p.d. diagonal matrix. From the fact that $K^{-1}\mathbf{1} > 0$, it follows that there exists a nonzero point $\sigma = (\sigma_1, \dots, \sigma_k)'$ such that $\sigma \geq 0$ and

$K^{-1}\mathbf{1} = H'\sigma$. For such a point $\sigma' q_H = 0$ holds, because

$$\mathbf{1}'K^{-1}(A + P)K\mathbf{1} = 1 + \alpha.$$

Let σ_{s_i}, $i = 1, \dots, t$, $1 \le t \le k$, be the nonzero elements of σ. The assumption that $q_H = (q_1, \dots, q_k)' \ge 0$ would then imply that $q_{s_i} = 0$ for $i = 1, \dots, t$. But this and the fact thats $(h_{s_1 j}, \dots, h_{s_t j})' \ne 0$ for $j = 1, \dots, n$ would in turn imply that $P = 0$ by the second condition of Lemma 2.1. Since this contradicts the assumption that $\operatorname{tr} P = 1$, no estimator in $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ can be dominated by an estimator of form (2.1).

Now let $L'Y \in \mathscr{L}$ and suppose that there are no matrices P and no vectors c fulfilling the conditions of Lemma 2.1. In this case $|I - (A + P)L| = 0$ for every matrix P, since all entries of matrix H are nonnegative. This implies that L must be of the form

$$L = \frac{1}{1 + \alpha} K_0 \mathbf{1}\mathbf{1}' K_0^{-1}, \tag{3.2}$$

where K_0 is a nonsingular diagonal matrix. In fact, let $W(X) = |I - XL|$, where $X = A + P$. Asserting that $|I - (A + P)L| = 0$ for all matrices P is the same as asserting that $W(X) = 0$ when $\operatorname{tr} X = 1 + \alpha$ and $X \ge A$. Now this latter statement implies that $W(X) = a(\operatorname{tr} X - 1 - \alpha)$, where $a \in \mathscr{R}$. And, since $W(0) = 1$, we obtain that $W(X) = 1 - \operatorname{tr} X/(1 + \alpha)$ for all diagonal matrices X. From this we may infer that all principal minors of L of order larger than 1 must be equal to 0 and that all diagonal entries of L are equal to $1/(1 + \alpha)$. Hence L must be of form (3.2) as asserted.

To end the proof of part (i) it suffices to show that under the adopted assumption either $K_0^{-1}\mathbf{1} \in \overline{\Theta}$ or $-K_0^{-1}\mathbf{1} \in \overline{\Theta}$.

Defining Q_0 analogously to (3.1) with K replaced by K_0 one can easily verify that $\ker(Q_0' H') = \{0\}$ and, consequently, that $\mathscr{R}(HQ_0) = \mathscr{R}^k$, when $K^{-1}\mathbf{1} \notin \mathscr{R}(H')$. Here, for any matrix M, $\ker(M)$ denotes the kernel and $\mathscr{R}(M)$ the column space of M.

In such a case there exists for every matrix P a vector c such that $q_H = HQ_0 c > 0$, which in view of Lemma 2.1 contradicts the assumption.

When $K_0^{-1}\mathbf{1}$ or $-K_0^{-1}\mathbf{1}$ belongs to $\mathscr{R}(H') \setminus \overline{\Theta}$, there must exist a vector $g > 0$ such that $g' K_0^{-1}\mathbf{1} = 0$. Because $K_0^{-1}\mathbf{1} \in \ker(Q_0')$, this implies that $g \in \mathscr{R}(Q_0)$ which also leads to a contradiction.

(ii) Again making use of the fact that all entries of H are nonnegative, it is enough by virtue of Lemma 2.1 to show that if $L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}} \setminus \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ and if P is p.d., then $|I - (A + P)L| \ne 0$. Suppose that this determinant is equal to 0. Expressing L in terms of a matrix G in $\overline{\mathscr{G}}$ we get

$$|I - (A + P)(I + GA)^{-1}G| = 0.$$

Since

$$I - (A + P)(I + GA)^{-1}G = (I - PG)(I + AG)^{-1},$$

this condition reduces to $|I - PG| = 0$. From the definition of $\mathscr{G}$ it now follows that the characteristic roots of PG are nonnegative and that their sum must be equal to 1. This in turn implies that PG must be of rank 1. Therefore, G must be also of rank 1, since P is p.d. But in such a case $L'Y \in \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ which contradicts the assumption that $L'Y \notin \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$. □

REMARK 3.1. It should be noticed that the assertions of Theorem 3.1 may not be valid for parametric functions $F'\theta$. The following illustrates this.

Consider model (1.1) with $n = 2$, $A = H = I_2$. Take $L = (1, 1)'$, and $F = (4, 0)'$. Observe that $L'Y = Y_1 + Y_2$ is not admissible for $F'\theta$ within the class of all linear estimators $\mathscr{L}(F)$, but since

$$F - (A + P)L = (2 + p_2, -(1 + p_2))',$$

the estimator $Y_1 + Y_2$ cannot be improved by an estimator of form (2.1). To see that Theorem 3.1(ii) does not extend to parametric functions consider the same model as above, but now treat $Y_1 + Y_2$ as an estimator of $F'\theta = 2\theta_1 + (4/3)\theta_2$, i.e., take $F = (2, \frac{4}{3})'$. In this case this estimator belongs to $\mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F) \setminus \mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}(F)$. In fact, in the representation of Theorem 1.1 it corresponds to the matrix

$$G = \begin{bmatrix} 1 & 0 \\ 2/3 & 1 \end{bmatrix}.$$

Lemma 2.1 immediately gives that this estimator can be improved by an estimator of form (2.1) if and only if $(p_2, \frac{1}{3} - p_2)' > 0$. Now this condition is fulfilled if and only if $0 < p_2 < \frac{1}{3}$.

This remark and Theorem 3.1 show that the class of estimators suggested by Das Gupta (1986) is broad enough to improve on inadmissible estimators of θ, but not for improving on inadmissible linear estimators of parametric functions.

THEOREM 3.2. *All estimators in* $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ *are admissible for* θ *within the class* $\mathscr{N}$ *of all estimators.*

PROOF. As noted in Section 1 every estimator in $\mathscr{L}^0_{\mathrm{ad}|\mathscr{L}}$ can be presented as

$$\delta^0 = \frac{\mathbf{1}'KY}{1 + \alpha} K^{-1}\mathbf{1}.$$

where $K = \mathrm{diag}(k_1, \dots, k_n)$ is a p.d. matrix such that $K^{-1}\mathbf{1} \in \overline{\Theta}$.

The parameter set may not be all of $\mathscr{R}^n_+$. If δ is as good as δ^0, then the risk function $R(\delta, \theta)$ is a lower semicontinuous function of θ which, since $R(\delta^0, \theta)$ is continuous, satisfies $R(\delta, \theta) \le R(\delta^0, \theta)$ for all $\theta \in \overline{\Theta}$.

By construction $\theta = H'\sigma$. The inequality $R(\delta, \theta) \le R(\delta^0, \theta)$ implies that $R(\delta, H'\sigma) \le R(\delta^0, H'\sigma)$. The inequality holds for all parameters θ if and only if it holds for all $\sigma > 0$. The following proof thus establishes admissibility of δ^0 as an estimator of $H'\sigma$, $\sigma > 0$.

Now let

$$\delta^* = \frac{1}{1 + \alpha} \mathbf{1}'KY.$$

At point $\theta = \theta_\lambda = \lambda K^{-1}\mathbf{1}$, where $\lambda > 0$, the random variable $\mathbf{1}'KY$ has a gamma (α, λ) distribution so that by Karlin's theorem δ^* is an admissible estimator of λ.

On the other hand, the risk of any estimator $\delta = (\delta_1, \dots, \delta_n)'$ at θ_λ can be written as

$$R(\delta, \theta_\lambda) = aE\left[\sum_{i=1}^{n} \frac{1}{a}\frac{1}{k_i^2}(k_i\delta_i - \lambda)^2\right],$$

where $a = \sum_{i=1}^{n} k_i^{-2}$.

Applying Jensen's inequality to the expression in the big parentheses we obtain

$$R(\delta, \theta_\lambda) \geq aE\left(\frac{1}{a}\sum_{i=1}^{n}\frac{1}{k_i}\delta_i - \lambda\right)^2$$

and the inequality is strict unless $k_i\delta_i = k_j\delta_j$ for all $i, j = 1, \dots, n$.

Since, as we have already noted, δ^* is admissible for λ and since the quadratic risks of δ^0 and δ^* are related at $\theta = \theta_\lambda$ by $R(\delta^0, \theta_\lambda) = aR(\delta^*, \lambda)$, it follows that if, say, δ dominates δ^0, then necessarily

$$R(\delta^*, \lambda) = E\left(\frac{1}{a}\sum_{i=1}^{n}\frac{1}{k_i}\delta_i - \lambda\right)^2.$$

Consequently, $\delta_i = (1/k_i)\delta^*$ for all i with probability 1, so that $\delta = \delta^0$ with probability 1. But this contradicts the assumption that δ dominates δ^0 and concludes the proof of the theorem. □

4. Some remarks on estimating parametric functions. Now we show that a number of results on estimating parametric functions $F'\theta$ considered in Section 2 follow from Section 3 as consequences of a lemma due to Shinozaki and some results of Klonecki and Zontek.

The relevant Shinozaki's lemma (1975) states that if a random vector δ is admissible under loss function (1.2) with respect to a p.d. matrix $\Lambda = \Lambda_1$, then it is also admissible under (1.2) with respect to any n.n.d. matrix $\Lambda = \Lambda_2$. In view of this lemma it is obvious that for any matrix $n \times m$ matrix F

$$\{(LF)'Y\colon L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{N}}\} \subset \mathscr{L}_{\mathrm{ad}|\mathscr{N}}(F) \tag{4.1}$$

and

$$\{(LF)'Y\colon L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}\} \subset \mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F). \tag{4.2}$$

The reversed inclusions in (4.1) and (4.2) do not hold in general. To see that the reversed inclusion in (4.1) may not hold, consider an unrestricted model with $H = I_n$ and let $F = (1, 0, \dots, 0)' \in \mathscr{R}^n$. Then $(1/(1 + \alpha_1))F'\bar{Y}$ is admissible for $F'\theta$. Now suppose that for some matrix G in the relevant set $\bar{\mathscr{G}}$,

$$\frac{1}{1 + \alpha_1}F = (I + GA)^{-1}GF.$$

From this one can deduce that rank $G > 1$, and Theorem 3.1 implies then that $[(I + GA)^{-1}G]'Y$ cannot be admissible for θ.

A sufficient condition for both the inverse inclusions to hold, i.e., for

$$\mathscr{L}_{\mathrm{ad}|\mathscr{N}}(F) = \{(LF)'Y\colon L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{N}}\} \tag{4.3}$$

and

$$\mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F) = \{(LF)'Y\colon L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}\}, \tag{4.4}$$

is that rank $HF =$ rank H. In fact, let $M'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{N}}(F)$. Since rank $HF =$ rank H implies that $HF(HF)^{+} = I_k$, it follows from Shinozaki's lemma mentioned above that

$$H'(F'H')^{+}M'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{N}}.$$

Since $\mathscr{R}(M') \subset \mathscr{R}(F'H')$ by Theorem 3.1(i) in Klonecki and Zontek (1988), it is clear that $M = M(HF)^{+}HF$. The assertion follows now by using once more Shinozaki's lemma. Relation (4.4) may be established by using similar arguments.

REMARK 4.1. An estimator $L'Y$ is in the set

$$\{(MF)'Y\colon M'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{N}}\},$$

where $L = (l_1, \dots, l_n)' \in \mathscr{R}^n$ and $F = (f_1, \dots, f_n)' \in \mathscr{R}^n$, if and only if

$$\sum_{i=1}^{n} \frac{f_i}{l_i} = 1 + \alpha$$

and either $L \in \overline{\Theta}$ and $L > 0$ or $-L \in \overline{\Theta}$ and $-L > 0$.

For $n = 2$ there exists an estimator of form (2.1) that dominates $L'Y \in \mathscr{L}_{\mathrm{ad}|\mathscr{L}}(F)$, where $L > 0$, if and only if

$$\sum_{i=1}^{n} \frac{f_i}{l_i} > 1 + \alpha.$$

This latter result does not extend to $n > 2$.

REMARK 4.2. In a recent paper Das Gupta and Sinha (1986) remarked correctly that $(\alpha/(1 + \alpha))\mathbf{1}'Y$ may be an admissible estimator of $(A\mathbf{1})'\theta$ within the unrestricted model with $H = I_n$. This follows straightforwardly from Remark 4.1 for $F = A\mathbf{1}$ and $L = (\alpha/(1 + \alpha))\mathbf{1}$.

REMARK 4.3. It may be of some interest to mention that in the multivariate normal case, i.e., when Y is distributed as $N(\mu, I)$, where $\mu \in \mathscr{R}^n$, an expression similar to (4.3) holds for any $n \times m$ matrix F and for all $n \geq 1$ [see Zontek (1986)].

5. Variance components. From the work of Farrell (1969), Olsen, Seely and Birkes (1976), LaMotte (1976) and Anderson, Henderson, Pukelsheim and

Searle (1984) it is known that for balanced as well as for some unbalanced random and mixed effects models there exists for estimating variance components a minimal sufficient statistic Y having a multivariate gamma distribution with mean vector and covariance matrix of the form (1.1) and with $\sigma_1, \dots, \sigma_k$ representing the variance components involved. The matrices H and A are determined by the model considered.

The linear admissible estimators of σ, i.e., of $F'\theta$ with F being equal to $H'(HH')^{-1}$, within the relevant reduced model Y have the following form

$$[(I + GA)^{-1}GH'(HH')^{-1}]'Y, \tag{5.1}$$

where $G \in \overline{\mathscr{G}}$, while $\overline{\mathscr{G}}$ is defined as in Theorem 1.1.

This is an admissible estimator if and only if rank $G = 1$. Moreover, since rank HF = rank H, the set of estimators (5.1) with G ranging over the subset of all matrices of $\overline{\mathscr{G}}$ with rank 1 represents the class of all admissible estimators of σ among the class of all linear estimators within the reduced model Y.

Example 5.1. For the one-way balanced random effects normal model, i.e., in case when the observed random vector X is distributed as $N_{uv}(\mathbf{1}\mu, J\sigma_1 + I\sigma_2)$, where $\mu \in \mathscr{R}$, $\sigma_1, \sigma_2 > 0$, while J is the Kronecker product of the $u \times u$ identity matrix I and the $v \times v$ matrix $\mathbf{11}'$, the relevant minimal sufficient invariant statistic for σ becomes

$$Y = \frac{1}{2}\begin{pmatrix} X'\left(\frac{1}{v}J - \frac{1}{uv}\mathbf{11}'\right)X \\ X'\left(I - \frac{1}{v}J\right)X \end{pmatrix}.$$

The corresponding matrices A and H are, respectively,

$$A = \tfrac{1}{2}\operatorname{diag}(u-1, u(v-1))$$

and

$$H = \begin{bmatrix} v & 0 \\ 1 & 1 \end{bmatrix}.$$

From the above it follows that a linear estimator of σ is admissible within the class of all translation invariant estimators if and only if it has the structure

$$\frac{4}{2+uv}\left(\frac{1}{g}Y_1 + Y_2\right)\begin{bmatrix} (g-1)/v \\ 1 \end{bmatrix},$$

where $g \geq 1$.

In case when rank $G > 1$ every estimator (5.1) can be improved by an estimator of the form (2.1), but no individual coordinate of the improved estimator may be better than the corresponding individual coordinate of (5.1). In fact, in view of Theorem 2.1 there may exist an improved estimator of the form (2.1) for $f'\sigma$ or, equivalently, for $[H'(HH')^{-1}f]'\theta$, where f is any nonzero vector in $\mathscr{R}^k$, only in cases when all coordinates of f are different from 0 and have the same sign. Now $f = (0, \dots, 0, 1, 0, \dots, 0)'$ does not meet this condition.

It is yet an unresolved problem whether within the reduced model Y individual components of a nonnegative estimator of σ corresponding to a matrix G with rank greater than 1 are admissible. If there would exist such an estimator with admissible components, one could claim to have established the presence of the Stein effect for estimating variance components.

6. Estimation of the vector η of natural parameters. In this section we confine our attention to unrestricted gamma models. Without loss of generality we may then assume that H is the identity matrix.

Using the same methods as in Sections 1 and 2 we show that for $n \geq 2$ no linear estimator $L'Z$, where $Z = (1/Y_1, \dots, 1/Y_n)'$, is admissible for the vector $\eta = (1/\theta_1, \dots, 1/\theta_n)'$ of natural parameters under squared error loss,

$$(L'Z - \eta)'\Lambda(L'Z - \eta),$$

where Λ is a p.d. $n \times n$ matrix. To ensure the existence of the second moments we need now to assume that $A - 2I > 0$.

To begin with we formulate a theorem analogous to Theorem 1.1 which provides a complete characterization of the class $\mathscr{K}_{\mathrm{ad}|\mathscr{K}}(F)$ of all linear admissible estimators $L'Z$ of $F'\eta$, where as above F may be any $n \times m$ matrix, within the class $\mathscr{K}(F)$ of all linear estimators of $F'\eta$. The proof is similar to that of Theorem 1.1 and will be omitted.

THEOREM 6.1. *The class* $\mathscr{K}_{\mathrm{ad}|\mathscr{K}}(F)$ *is given by*

$$\left\{(L_G F)'Z\colon L_G = (A - I)(A - 2I)[I + G(A - 2I)]^{-1}G,\, G \in \bar{\mathscr{G}}\right\},$$

where $\bar{\mathscr{G}}$ *is the closure of a set* $\mathscr{G}$ *defined by*

$$\mathscr{G} = \left\{\Delta_d^{-1}\Delta\colon \Delta \in \Omega\right\},$$

while

$$\Omega = \operatorname{conv}\{\eta\eta'\colon \eta \in \mathscr{R}^n,\, \eta > 0\}.$$

In exactly the same way as Lemma 2.1, one can establish the following.

LEMMA 6.1. *There exists a number* $\kappa > 0$ *such that*

$$L'Z + \kappa c Z_P, \tag{6.1}$$

where $c \in \mathscr{R}^m$, *while* P *is an n.n.d. diagonal matrix and* $\operatorname{tr} P = 1$, *dominates the estimator* $L'Z$ *of* $F'\eta$ *if and only if*

$$q = [(A - I - P)F - L]c \geq 0$$

and, if the ith coordinate of q *is* 0, *then also the ith diagonal element of* P *must be* 0.

Now put $X = \operatorname{diag}(x_1, \ldots, x_n) = A - I - P$. Since the coefficient appearing at $x_1 x_2 \cdots x_n$ in the polynomial $W(x) = |L - X|$ is equal to 1, we see that $W(x)$ cannot be identically 0 when $\operatorname{tr} X = \alpha - n - 1$ and $X \geq A - 2I$. This shows that there exists a matrix P having the required properties such that

$$|A - I - P - L| = 0.$$

This fact yields the following main result of this section.

THEOREM 6.2. *If $n \geq 2$, then for every linear estimator $L'Z$ of η there exists an estimator of the form* (5.1) *which dominates $L'Z$.*

This theorem extends some work of Berger (1980). He has shown that the standard estimator $(A - 2I)Z$ of η which is admissible within the class of linear estimators, is inadmissible within the class all estimators of η. Das Gupta (1986) has found a class of estimators of a simple form dominating this standard estimator. The class of estimators defined by (6.1) is slightly broader than the class considered by Das Gupta.

REMARK 6.1. Since every linear estimator of η may be improved uniformly by (6.1), the assertion of Theorem 6.2 is valid when η ranges over any nonempty subset of $\{\nu \in \mathscr{R}^n: \nu > 0\}$.

To end this section we shall formulate some results analogous to those presented in Theorem 3.1(ii) and Theorem 3.2.

THEOREM 6.3. (i) *If $L'Z \in \mathscr{K}_{\mathrm{ad}|\mathscr{X}}$, then for every p.d. matrix P there exists a vector c in $\mathscr{R}^n$ such that for sufficiently small positive values of κ the estimator* (6.1) *dominates $L'Z$.*

(ii) *If $L'Z + cZ_P$ improves upon $L'Z$ in $\mathscr{K}_{\mathrm{ad}|\mathscr{X}}(F)$, where P is p.d., then $F\Lambda c > 0$.*

PROOF. Let G be a matrix in $\overline{\mathscr{G}}$ corresponding to the matrix L. Noting that

$$(A - I - P)F - L = (A - I)(I - PG^*)[I + (A - 2I)G]^{-1}F,$$

where

$$G^* = (A - I)^{-1}[I + (A - 2I)G],$$

and that G^* is a matrix of rank n belonging to $\mathscr{G}$, the proof may be concluded by using the methods developed in Section 2.

Acknowledgments. We wish to thank a referee for useful suggestions that led to improvement of the presentation of the original version of the paper. We thank L. D. Brown for his interest and helpful comments.

REFERENCES

ANDERSON, R. D., HENDERSON, H. V., PUKELSHEIM, F. and SEARLE, S. R. (1984). Best estimation of variance components from balanced data with arbitrary kurtosis. *Math. Operationsforsch. Statist. Ser. Statist.* **15** 163–176.

BERGER, J. (1980). Improving on inadmissible estimators in continuous exponential families with applications to simultaneous estimation of gamma scale parameters. *Ann. Statist.* **8** 545–571.

DAS GUPTA, A. (1986). Simultaneous estimation in the multiparameter gamma distribution under weighted quadratic losses. *Ann. Statist.* **14** 206–219.

DAS GUPTA, A. and SINHA, B. K. (1986). Estimation in the multiparameter exponential family: Admissibility and inadmissibility results. *Statist. Decisions* **4** 101–130.

FARRELL, R. H. (1969). On the Bayes character of the standard model II analysis of variance test. *Ann. Math. Statist.* **40** 1094–1097.

GRAYBILL, F. A. (1983). *Matrices with Applications in Statistics*, 2nd ed. Wadsworth, Belmont, Calif.

KARLIN, S. (1958). Admissibility for estimation with quadratic loss. *Ann. Math. Statist.* **29** 404–436.

KLONECKI, W. and ZONTEK, S. (1987). On admissible invariant estimation of variance components which dominate unbiased invariant estimators. *Statistics* **18** 483–498.

KLONECKI, W. and ZONTEK, S. (1988). On the structure of admissible linear estimators. *J. Multivariate Anal.* **24** 11–30.

LAMOTTE, L. R. (1976). Invariant quadratic estimators in random one-way ANOVA model. *Biometrics* **32** 793–804.

OLSEN, A., SEELY, J. and BIRKES, D. (1976). Invariant quadratic unbiased estimation for two variance components. *Ann. Statist.* **4** 878–890.

RAO, C. R. (1976). Estimation of parameters in a linear model. *Ann. Statist.* **4** 1023–1037.

SHINOZAKI, N. (1975). A study of generalized inverse of matrix and estimation with quadratic loss. Ph.D. dissertation, Keio Univ., Japan.

ZONTEK, S. (1986). Characterizations of linear admissible estimators in Gauss–Markov model under normality. *Linear Statistical Inference. Lecture Notes in Statist.* **35** 311–317. Springer, Berlin.

ZONTEK, S. (1988). On admissibility of limits of unique locally best linear estimators with applications to variance components models. *Probab. Math. Statist.* To appear.

ROGER H. FARRELL
DEPARTMENT OF MATHEMATICS
WHITE HALL
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-7901

WITOLD KLONECKI
STEFAN ZONTEK
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
KOPERNIKA 18
51 617 WROCŁAW
POLAND