AN ABSTRACT OF THE THESIS OF

Paul James McClellan for the degree of Doctor of Philosophy in Statistics presented on November 16, 1983.

Title: Maximin-Efficient Admissible Linear Unbiased Estimation in Mixed Linear Models

Redacted for privacy

Abstract approved: Justus F. Seely

Linear unbiased estimation of linear parametric functions of fixed parameters in mixed linear models with two variance components is considered. Such models may be written as $E(Y) = X\beta$ , $\mathrm{Cov}(Y) = \theta_1 V_1 + \theta_2 V_2$, where X is an n x p known matrix, $\beta$ is a p-vector of fixed but unknown parameters, $\theta_1$ and $\theta_2$ are fixed but unknown nonnegative variance components, and $V_1$ and $V_2$ are known n x n nonnegative definite matrices.

The class $\bar{\mathbf{A}}_\pi$ of admissible linear unbiased estimators (ALUEs) for a parametric function $\pi'\beta$ is presented. The class $\bar{\mathcal{E}}_\pi$ of ALUEs for $\pi'\beta$ having maximum minimum efficiency within $\bar{\mathbf{A}}_\pi$ is obtained by considering four mutually exclusive and exhaustive cases. Such maximin-efficient ALUEs for $\pi'\beta$ with respect to the mixed linear model may be computed as the best linear unbiased estimators (BLUEs) with respect to certain artificial fixed linear models.

The efficiency curves of the maximin-efficient ALUE and selected other ALUEs for five examples are presented.

It is shown that the maximin-efficient ALUE remains admissible and retains its efficiency within the larger class of all unbiased estimators when the data Y follows a multivariate normal distribution.

# Maximin-Efficient Admissible Linear Unbiased Estimation in Mixed Linear Models

by

Paul James McClellan

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed November 16, 1983

Commencement June 1984

APPROVED:

Redacted for privacy

Professor of Statistics in charge of major

Redacted for privacy

Chairman of department of Statistics

Redacted for privacy

Dean of Graduate School

Date thesis is presented November 16, 1983

Typed by Frances Gallivan for Paul James McClellan

ACKNOWLEDGEMENTS

I wish to express my appreciation to Dr. Justus F. Seely for his guidance and encouragement throughout my graduate studies.

I also express my gratitude toward my parents, Paul and Valerie, for their steady and enthusiastic encouragement of education.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

MAXIMIN-EFFICIENT ADMISSIBLE LINEAR UNBIASED ESTIMATION IN MIXED LINEAR MODELS

# I. INTRODUCTION

This thesis addresses the problem of estimating linear parametric functions of fixed parameters in mixed linear models with two variance components. These models may be written as

$$Y = X\beta + B_1 b_1 + B_2 b_2 \tag{1.1}$$

where $Y$ is an n-vector of observations; $X$, $B_1$, and $B_2$ are known $n \times p$, $n \times q_1$, and $n \times q_2$ matrices, respectively; $\beta$ is a p-vector of fixed but unknown parameters; and $b_1$ and $b_2$ are $q_1$- and $q_2$-vectors of random variables, respectively. It is assumed that $b_1$ and $b_2$ have mean vectors and covariance structure

$$E(b_i) = 0 \quad , \quad \text{Cov}\,(b_i) = \theta_i I_{q_i} \quad , \quad i = 1,2 \quad ,$$

where $\theta_1$ and $\theta_2$ are fixed but unknown nonnegative parameters (variance components) and $I_q$ denotes the $q \times q$ identity matrix. Furthermore, $b_1$ and $b_2$ are assumed to be uncorrelated.

The first and second moment assumptions of model (1.1) are summarized by

$$E(Y) = X\beta \quad , \quad \text{Cov}\,(Y) = \theta_1 V_1 + \theta_2 V_2 \quad , \tag{1.2}$$

where $V_1 = B_1 B_1'$ and $V_2 = B_2 B_2'$ are known nonnegative definite matrices. Throughout all but the last chapter of this thesis we shall make no additional distributional assumptions concerning the terms of (1.1).

Mixed linear models with two variance components arise in several ways, as illustrated by the following examples.

Example 1.1 (heteroscedastic error model)

Suppose $\{Y_{ij}\ ,\ i = 1, 2\ ,\ j = 1,\ldots,n_i\}$ is a collection of uncorrelated random variables such that $Y_{ij}$ has mean $\mu$ and variance $\theta_i$. Letting $Y = (Y_{11},\ldots,Y_{1n_1},\ldots,Y_{2n_2})'$ we may write this model in the form of (1.2) where $n = n_1 + n_2$, $X = \underline{1}_n$ is the n-vector of 1's, $\beta = \mu$,

$$V_1 = \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } V_2 = \begin{bmatrix} 0 & 0 \\ 0 & I_{n_2} \end{bmatrix}.$$

Example 1.2 (random one-way model)

Suppose $\{Y_{ij},\ i = 1,\ldots,a,\ j=1,\ldots,n_i\}$ is a collection of random variables such that $Y_{ij} = \mu + a_i + e_{ij}$ where $\mu$ is the fixed general mean, the $a_i$ are uncorrelated random factor effects, and the $e_{ij}$ are uncorrelated random errors. Also, suppose the $a_i$ and the $e_{ij}$ are uncorrelated with means zero and variances $\theta_1$ and $\theta_2$, respectively. Letting $n = n. = \Sigma n_i$ and $Y = (Y_{11},\ldots,Y_{1n_1},\ldots,Y_{an_a})'$ we may write this model as (1.1) with $X = \underline{1}_n$, $\beta = \mu$, $b_1 = (a_1,\ldots,a_a)'$, $b_2 = (e_{11},\ldots,e_{1n_1},\ldots,e_{an_a})'$, $B_2 = I_n$, and appropriately defined $B_1$. Or we may write it as (1.2) with $V_2 = I_n$ and $V_1 = B_1B_1'$.

Example 1.3 (mixed two-way additive model)

Suppose $\{Y_{ijk}, i = 1,\ldots,a, j = 1,\ldots,d, k = 1,\ldots,n_{ij}\}$ is a collection of random variables such that $Y_{ijk} = \mu + \alpha_i + d_j + e_{ijk}$, where $\mu$ is the fixed general mean, the $\alpha_i$ are fixed effects of one factor, the $d_j$ are uncorrelated random effects of a second factor, and the $e_{ijk}$ are uncorrelated random errors. Suppose also that the $d_j$ and the $e_{ijk}$ are uncorrelated with means zero and variances $\theta_1$ and $\theta_2$, respectively. Letting $n = n_{..} = \Sigma\Sigma n_{ij}$ and Y denote the n-vector of $Y_{ijk}$'s we may write this model as (1.1) with $\beta = (\mu,\alpha_1,\ldots,\alpha_a)'$, $b_1 = (d_1,\ldots,d_d)'$, $b_2 = (e_{111},\ldots,e_{11n_{11}},\ldots, e_{adn_{ad}})'$, $B_2 = I_n$, and appropriately defined X and $B_1$. Or we may write it as (1.2) with $V_2 = I_n$ and $V_1 = B_1B_1'$.

Example 1.4 (mixed two-factor nested model)

Let $\{Y_{ijk}, i = 1,\ldots,a, j = 1,\ldots,d_i, k = 1,\ldots,n_{ij}\}$ be a collection of random variables such that $Y_{ijk} = \mu + \alpha_i + d_{ij} + e_{ijk}$, where $\mu$ is the fixed general mean, the $\alpha_i$ are fixed effects of the first factor, the $d_{ij}$ are uncorrelated random effects of the second factor, and the $e_{ijk}$ are uncorrelated random errors. Here the levels of the second factor, indexed by j, are nested within the levels of the first factor, indexed by i. Let the $d_{ij}$ and the $e_{ijk}$ be uncorrelated with means zero and variances $\theta_1$ and $\theta_2$, respectively. We may write this model as (1.1) or (1.2) with $\beta = (\mu,\alpha_1,\ldots,\alpha_a)'$, $b_1$ and $b_2$ having elements $d_{ij}$ and $e_{ijk}$ ordered lexicographically, and appropriately defined Y, X, $B_1$, and $B_2$.

Example 1.5 (sampling a Wiener process with drift)

Olsen (1973) considers a Wiener process $\{B(t)\}$ with drift

$$f(t) = \sum_{j=1}^{p} \beta_j f_j(t)$$

observed with error $e(t)$ at times $t_0, t_1, \ldots, t_n$ to obtain the model $Y(t_i) = B(t_i) + e(t_i)$, $i = 0,1,\ldots,n$. The $B(t_i)$ and the $e(t_i)$ are uncorrelated random variables with means $f(t_i)$ and zero and variances $\theta_1 t_i$ and $\theta_2$, respectively. Defining $Y$, $X$, and $\beta$ to have elements $Y_i$, $x_{ij}$, and $\beta_j$ where $Y_i = Y(t_i) - Y(t_{i-1})$ and $x_{ij} = f_j(t_i) - f_j(t_{i-1})$ for $i = 1,\ldots,n$ and $j = 1,\ldots,p$ , $V_1 = \mathrm{diag}(t_i - t_{i-1})$, and $V_2$ to be the $n \times n$ tridiagonal matrix with diagonal elements 2 and off-diagonal elements -1, we obtain model (1.2), as pointed out by Olsen.

Example 1.6 (mean vector follows a first-order Markov process)

LaMotte and McWhorter (1978) consider a model which may be written $Y_i = x_i' \beta_i + e_i$ , $\beta_i = \beta_{i-1} + u_i$, for $i = 1,\ldots,n$ where $Y_i$ and $e_i$ are scalars and $x_i$, $\beta_i$, and $u_i$ are p-vectors. The $x_i$ are known and $\beta_0$ is fixed but unknown. It is assumed that the $e_i$ and the $u_i$ are independent and multivariate normal with means zero and covariance structures $\theta_1 I_n$ and $\theta_2 I_n \otimes D$ where $\otimes$ denotes the Kronecker product and $D = \mathrm{diag}(d_{rr})$ is known, with $d_{rr} \geqslant 0$ for $r=1,\ldots,p$. LaMotte and McWhorter point out that by successive substitution the model may be written as (1.2) with $Y = (Y_1,\ldots,Y_n)'$ , $X = (x_1,\ldots,x_n)'$ , $\beta = \beta_0$, $V_1 = I_n$, and $V_2$ having elements

$$v_{ij} = \min(i,j) \sum_{r=1}^{p} d_{rr} x_{ri} x_{rj} \quad .$$

With respect to model (1.1) or (1.2), we wish to estimate a parametric function of the form $\pi'\beta$. Under certain conditions (see, for example, Seely and Zyskind, 1971), a best linear unbiased estimator (BLUE) will exist for a (linearly) estimable $\pi'\beta$. However, such a BLUE will not exist in general. An alternative estimation procedure is therefore needed.

When the ratio of the variance components $\theta_1$ and $\theta_2$ in model (1.2) is known, a BLUE exists for every estimable $\pi'\beta$. One approach to estimating $\pi'\beta$ in model (1.2) is first to estimate $\theta_1$ and $\theta_2$ using one of several procedures. These estimates are then treated as known values of $\theta_1$ and $\theta_2$ to obtain a corresponding "BLUE" for $\pi'\beta$. Kackar and Harville (1981) and Seely and Hogg (1982) discuss this approach for estimating $\pi'\beta$ and show that when the data vector Y is symmetrically distributed about its mean and when any of the standard variance component estimation procedures are used, the resulting two-stage estimator for $\pi'\beta$ is unbiased.

Another approach to estimating $\pi'\beta$ is to assume that Y has a multivariate normal distribution and to estimate $\pi'\beta$ as $\pi'\hat{\beta}$, where $\hat{\beta}$ is the maximum likelihood estimate for $\beta$. Kackar and Harville (1981) and Seely and Hogg (1982) show that $\pi'\hat{\beta}$ is then unbiased for $\pi'\beta$. Harville (1977) includes a discussion of maximum likelihood estimation of fixed effects in his review of maximum likelihood approaches to variance component estimation.

A disadvantage of both above approaches to fixed effect estimation is that the resulting estimators for $\pi'\beta$ are nonlinear and their distributional properties are difficult to obtain. Our goal is to obtain a satisfactory linear estimator for $\pi'\beta$, if possible, in which case the distribution of the estimator would be considerably simplified.

Birkes, Seely, and Azzam (1981) compare the efficiencies of admissible linear unbiased estimators (ALUEs) of the mean in a random one-way model. They show that there exists a unique ALUE with maximum minimum efficiency. The efficiency of this estimator compares favorably with efficiencies of competing linear estimators of the mean in example data sets. Finally, this maximin-efficient ALUE remains admissible and retains its efficiency within the class of all unbiased estimators when the data is assumed to follow a multivariate normal distribution.

The purpose of this thesis is to characterize the class of maximin-efficient ALUEs for an estimable $\pi'\beta$ with respect to model (1.1) or model (1.2). This will provide a basis from which to conduct performance comparisons of these estimators with competing linear and nonlinear estimators.

Chapter II introduces some definitions and notation that will be used throughout the remainder of the thesis. It also presents three lemmas that will be used later.

Chapter III considers the special case where $p = 1$ and characterizes the class of maximin-efficient ALUEs for $\beta$.

Chapter IV characterizes the class of maximin-efficient ALUEs for an estimable $\pi'\beta$ when $p \geqslant 1$.

Chapters V and VI apply the results of Chapter IV to the mixed two-factor nested model and the mixed two-way additive model.

Example data are used to illustrate these results.

Finally, Chapter VII discusses some alternative estimation procedures and shows that a maximin-efficient ALUE for $\pi'\beta$ remains admissible and retains its efficiency within the larger class of unbiased estimators for $\pi'\beta$ when the data is assumed to follow a multivariate normal distribution.

# II. DEFINITIONS, NOTATION, AND PRELIMINARY RESULTS

In this thesis, $Y$ will denote an n-vector of observations with mean vector $X\beta$ and covariance matrix $V$. Here $X$ will be a known $n \times p$ matrix with $0 < \text{rank}(X) \leqslant p \leqslant n$ and $\beta$ will be a p-vector of fixed but unknown parameters. $V$ will be a member of a known set $\mathcal{V}$ of (symmetric) nonnegative definite matrices.

We shall denote the expectation and covariance structure of the random vector $Y$ by $E(Y)$ and $\text{Cov}(Y)$, respectively. Hence we have

$$E(Y) = X\beta \quad , \quad \text{Cov}(Y) = V \in \mathcal{V} \tag{2.1}$$

A linear estimator is any function of $Y$ of the form $t'Y$ where $t$ is a known element of n-dimensional Euclidean space $R^n$. We shall denote the expectation and variance of the estimator $t'Y$ by $E(t'Y)$ and $\text{var}(t'Y|V)$, respectively. If the set $\mathcal{V}$ is parametrized, we may replace $V$ by its parameters in $\text{var}(t'Y|V)$.

A linear parametric function of $\beta$ is any function of the form $\pi'\beta$, where $\pi$ is a known element of $R^p$. If it is obvious by context, the adjective "linear" will be omitted. A parametric function $\pi'\beta$ is said to be (linearly) estimable with respect to (2.1) if there exists a linear estimator $t'Y$ such that $E(t'Y) = \pi'\beta$. In this case, $t'Y$ is said to be a linear unbiased estimator (LUE) for $\pi'\beta$. The class of LUEs for $\pi'\beta$ will be denoted by $\bar{\mathfrak{L}}_\pi$. We shall denote by $\mathfrak{L}_\pi$ the set of coefficient vectors $t \in R^n$ such that $t'Y \in \bar{\mathfrak{L}}_\pi$. If it is obvious by context, the subscript "$\pi$" will be omitted.

An element of $t'Y$ of $\bar{\mathfrak{L}}_\pi$ is said to be as good as $s'Y \in \bar{\mathfrak{L}}_\pi$ if $\mathrm{var}(t'Y|V) \leqslant \mathrm{var}(s'Y|V)$ for all $V \in \mathcal{V}$, and $t'Y$ is said to be better than $s'Y$ or to beat $s'Y$ if $t'Y$ is as good as $s'Y$ and $\mathrm{var}(t'Y|V_0) < \mathrm{var}(s'Y|V_0)$ for at least one $V_0 \in \mathcal{V}$. An admissible linear unbiased estimator (ALUE) for $\pi'\beta$ is any $t'Y \in \bar{\mathfrak{L}}_\pi$ that is not beaten by any other element of $\bar{\mathfrak{L}}_\pi$. The class of ALUEs for $\pi'\beta$ with respect to (2.1) will be denoted by $\bar{\mathfrak{A}}_\pi(\mathcal{V})$ and the corresponding set of coefficient vectors will be denoted by $\bar{A}_\pi(\mathcal{V})$. If $\pi$ or $\mathcal{V}$ is obvious by context, we may omit it from the notation. If there exists a $t'Y \in \bar{\mathfrak{L}}_\pi$ that is as good as every other element of $\bar{\mathfrak{L}}_\pi$, then $t'Y$ is called a best linear unbiased estimator (BLUE) for $\pi'\beta$. If a BLUE for $\pi'\beta$ exists, then $\bar{\mathfrak{A}}_\pi$ consists of all BLUEs for $\pi'\beta$.

It is easy to show that $\bar{\mathfrak{L}}_\pi$ and the variance of any element of $\bar{\mathfrak{L}}_\pi$ are invariant under nonsingular linear transformations of $Y$. Hence $\bar{\mathfrak{A}}_\pi$ is also invariant under nonsingular linear transformations of $Y$. Also, as pointed out by Olsen, Seely, and Birkes (1976), the definitions of "as good as," "better than," and admissibility remain unchanges if $\mathcal{V}$ is replaced by $[\mathcal{V}]$, the smallest closed convex cone containing $\mathcal{V}$. Hence $\bar{\mathfrak{A}}_\pi(\mathcal{V}) = \bar{\mathfrak{A}}_\pi([\mathcal{V}])$.

Let $\pi'\beta$ be estimable and let $V \in \mathcal{V}$. Set

$$v_M(V) = \inf_{t \in \mathfrak{L}_\pi} \mathrm{var}(t'Y|V).$$

Then the efficiency within $\bar{\mathfrak{L}}_\pi$ of any $s'Y \in \bar{\mathfrak{L}}_\pi$ when $\mathrm{Cov}(Y) = V$ is

$$\begin{aligned} \mathrm{Eff}(s'Y|V) &= v_M(V)/\mathrm{var}(s'Y|V) \quad , \quad \text{if } \mathrm{var}(s'Y|V) > 0 \quad , \\ &= 1 \qquad\qquad\qquad\qquad\quad , \quad \text{otherwise.} \end{aligned}$$

For notational convenience, when the random vector $Y$ is understood we will sometimes write $\mathrm{Eff}(s'Y|V)$ as $\mathrm{Eff}(s|V)$. Then we have $0 \leqslant \mathrm{Eff}(s|V) \leqslant 1$, for all $s \in \mathfrak{L}_\pi$ and all $V \in \mathfrak{V}$.

A maximin-efficient ALUE for $\pi'\beta$ with respect to (2.1) is any $t'Y \in \overline{\mathfrak{A}}_\pi$ such that

$$\inf_{V \in \mathfrak{V}} \mathrm{Eff}(t|V) \geqslant \inf_{V \in \mathfrak{V}} \mathrm{Eff}(s|V)$$

for all $s \in \mathfrak{A}_\pi$. We shall denote the class of maximin-efficient ALUEs for $\pi'\beta$ with respect to (2.1) by $\overline{\mathfrak{E}}_\pi(\mathfrak{V})$ and the corresponding set of coefficient vectors by $\mathfrak{E}_\pi(\mathfrak{V})$. If $\pi$ or $\mathfrak{V}$ is obvious by context, we may omit it from the notation.

Since $\overline{\mathfrak{L}}_\pi$, $\overline{\mathfrak{A}}_\pi$, and the variance of any element of $\overline{\mathfrak{L}}_\pi$ is invariant under nonsingular linear transformations of $Y$, so is the efficiency of any element of $\overline{\mathfrak{L}}_\pi$. Hence $\overline{\mathfrak{E}}_\pi$ is also invariant under nonsingular linear transformations of $Y$.

We shall denote by $\mathrm{co}(\mathfrak{V})$, the smallest cone containing $\mathfrak{V}$, defined here as $\mathrm{co}(\mathfrak{V}) = \{cV : V \in \mathfrak{V} \text{ and } c \geqslant 0\}$. We have $\mathfrak{V} \subset \mathrm{co}(\mathfrak{V}) \subset [\mathfrak{V}]$. Let $s \in \mathfrak{L}_\pi$ and $V \in \mathfrak{V}$. Then

$$\begin{aligned} \mathrm{Eff}(s'Y|cV) &= \mathrm{Eff}(s'Y|V) && \text{for } c > 0 \quad , \\ &= 1 && \text{for } c = 0 \quad . \end{aligned}$$

So $\inf_{V \in \mathfrak{V}} \mathrm{Eff}(s'Y|V) = \inf_{V \in \mathrm{co}(\mathfrak{V})} \mathrm{Eff}(s'Y|V)$. Since $\mathfrak{A}_\pi(\mathfrak{V}) = \mathfrak{A}_\pi([\mathfrak{V}]) = \mathfrak{A}_\pi(\mathrm{co}(\mathfrak{V}))$, we obtain

$$\overline{\mathfrak{E}}_\pi(\mathfrak{V}) = \overline{\mathfrak{E}}_\pi(\mathrm{co}(\mathfrak{V})). \tag{2.2}$$

It is the purpose of this thesis to characterize $\overline{\mathcal{E}}_{\pi}(\mathcal{V})$ for a general mixed linear model with two variance components. By equation (2.2) we may replace $\mathcal{V}$ by $\text{co}(\mathcal{V})$ for this purpose.

We shall denote the column space, null space, and rank of a matrix H by $\underline{R}(H)$, $\underline{N}(H)$, and $\underline{r}(H)$. Orthogonality will be defined with respect to the usual inner product. If $\mathcal{S} \subset R^n$, then the set orthogonal to $\mathcal{S}$ within $R^n$ will be denoted $\mathcal{S}^{\perp}$. We shall use the abbreviations p.d. and n.n.d. for positive definite and nonnegative definite, respectively. Finally, we shall denote by $H^+$ the unique Moore-Penrose inverse of a matrix H. If H is nonsingular, then $H^+ = H^{-1}$, the inverse of H. We shall use several well-known properties of the Moore-Penrose inverse of a symmetric matrix.

Before introducing our first specific model, we present three lemmas that will be frequently used later. The first two easily follow from well-known properties of n.n.d. matrices and the Moore-Penrose inverse of a symmetric matrix.

<u>Lemma 2.1</u>: Let $V_1$ and $V_2$ be n.n.d. matrices and let $0 \leq c \leq 1$. Then $W(c) = cV_1 + (1-c)V_2$ is also n.n.d.. If $\underline{R}(X) \subset \underline{R}(V_1) \cap \underline{R}(V_2)$ then $\underline{R}(X) \subset \underline{R}(W(c))$. If $V_2$ is p.d., then so is $W(c)$ for $0 \leq c < 1$.

<u>Lemma 2.2</u>: Let $t \in \underline{R}(V)$ where V is n.n.d.. If $t \neq 0$ then $t'Vt > 0$ and $t'V^+t > 0$.

<u>Lemma 2.3</u>: Let $\underline{R}(X) \subset \underline{R}(V)$ where V is n.n.d.. Then $\underline{R}(X'V^+X) = \underline{R}(X'V^+) = \underline{R}(X')$.

Proof: $\underline{R}(X'V^+X) \subset \underline{R}(X'V^+) \subset \underline{R}(X')$. Suppose $a \in \underline{N}(X'V^+X)$. Then $X'V^+Xa = 0 \Rightarrow a'X'V^+Xa = 0 \Rightarrow Xa = 0$ by Lemma 2.2. But $Xa = 0 \Rightarrow a \in \underline{N}(X)$ so $\underline{N}(X'V^+X) \subset \underline{N}(X) \Rightarrow \underline{R}(X') \subset \underline{R}(X'V^+X)$. ∎

# III. ONE FIXED PARAMETER

In this chapter we consider model (1.1) with p = 1, which can be written as

$$Y = x\beta + B_1 b_1 + B_2 b_2 \tag{3.1}$$

where Y, $B_1$, $b_1$, $B_2$, and $b_2$ are as in (1.1), x is a known nonzero n-vector, and $\beta$ is a fixed but unknown scalar parameter. The first and second moment assumptions are

$$E(Y) = x\beta \quad , \quad Cov(Y) = \theta_1 V_1 + \theta_2 V_2 \tag{3.2}$$

where $V_i = B_i B_i'$ , i = 1,2, are known n.n.d. matrices. We shall take $\mathcal{V} = \{\theta_1 V_1 + \theta_2 V_2 : \theta_i \geqslant 0 \quad , \quad i = 1,2\}$.

We wish to characterize the class $\overline{\mathcal{E}}$ of maximin-efficient ALUEs for $\beta$. We will denote the corresponding set of coefficient vectors by $\mathcal{E}$. By (2.2), our results will hold for any $\mathcal{V}$ such that $co(\mathcal{V}) = \{\theta_1 V_1 + \theta_2 V_2 : \theta_i \geqslant 0 \quad , \quad i = 1,2\}$. We initially assume that $\underline{r}(x,V_1,V_2) = n$.

We begin by stating a necessary and sufficient condition for a BLUE for $\beta$ to exist. We then characterize the class $\overline{\mathcal{A}}$ of ALUEs for $\beta$ using some results from Azzam (1981) and, more recently, by LaMotte (1982). We obtain $\overline{\mathcal{E}}$ from $\overline{\mathcal{A}}$ in four mutually exclusive cases. Finally, we drop the assumption that $\underline{r}(x,V_1,V_2) = n$.

By Corollary 2.1 of Seely and Zyskind (1971), $t'Y$ is a BLUE for its expectation $E(t'Y)$ if and only if $V_i t \in \underline{R}(x)$ , $i = 1,2$. Since $\underline{r}(x,V_1,V_2) = n$, by Corollary 1.3 of that same paper, BLUEs are unique; that is: if $t'Y$ and $s'Y$ are BLUEs and $E(t'Y) = E(s'Y)$, then $t = s$.

Suppose a BLUE $t'Y$ for $\beta$ exists. Then $V_i t \in \underline{R}(x)$ , $i = 1,2$, and $t'x = 1$. Suppose, on the other hand, that $t$ satisfies $V_i t \in \underline{R}(x)$ , $i = 1,2$. If $t'x = 0$ then $t'Y$ is the BLUE for $E(t'Y) = 0$ and hence $t = 0$. If $t'x \neq 0$, then $(t'x)^{-1}t'Y$ is the BLUE for $\beta$. We conclude that there exists a BLUE for $\beta$ if and only if there exists a nonzero $t \in R^n$ satisfying $V_i t \in \underline{R}(x)$ , $i = 1,2$, in which case $(t'x)^{-1}t'Y$ is the unique BLUE for $\beta$.

From this we see that the least-squares estimator $\hat{\beta} = (x'x)^{-1}x'Y$ for $\beta$ is the BLUE for $\beta$ if and only if $x$ is an eigenvector of both $V_1$ and $V_2$. These results were first obtained by Zyskind (1967) when $\mathfrak{V} = \{V\}$ and $V$ is a known n.n.d. matrix.

Let $\bar{\mathfrak{L}}$ denote the class of LUEs for $\beta$ and $\mathfrak{L}$ denote the corresponding set of coefficient vectors. Let $V$ and $W$ be n.n.d. matrices and define

$$\mathfrak{B}(V) = \{t \in \mathfrak{L} : t'Vt = \min_{a\in\mathfrak{L}} a'Va\} ,$$

$$\bar{\mathfrak{B}}(V) = \{t'Y : t \in \mathfrak{B}(V)\} ,$$

$$\mathfrak{B}(V,W) = \{t \in \mathfrak{B}(V) : t'Wt = \min_{a\in\mathfrak{B}(V)} a'Wa\} , \text{ and}$$

$$\bar{\mathfrak{B}}(V,W) = \{t'Y : t \in \mathfrak{B}(V,W)\} .$$

Then by Propositions 3.2.3 and 3.2.5 and Theorem 4.3.1 of Azzam (1981) we have:

$$\mathfrak{B}(V) \quad = \{t \in \mathfrak{L}: Vt \in \underline{R}(x)\} \tag{3.3}$$

$$\mathfrak{B}(V,W) \quad = \{t \in \mathfrak{B}(V): Wt \in \underline{R}(x,V)\} \tag{3.4}$$

$$t_0 \in \mathfrak{B}(V) \Rightarrow \mathfrak{B}(V) \quad = t_0 + \underline{N}(x') \cap \underline{N}(V) \tag{3.5}$$

$$t_0 \in \mathfrak{B}(V,W) \Rightarrow \mathfrak{B}(V,W) = t_0 + \underline{N}(x') \cap \underline{N}(V) \cap \underline{N}(W) \tag{3.6}$$

$$\overline{\mathfrak{A}} \quad = \overline{\mathfrak{B}}(V_1,V_2) \cup \overline{\mathfrak{B}}(V_2,V_1) \cup \bigcup_{0<c<1} \overline{\mathfrak{B}}(cV_1 + (1-c)V_2) \tag{3.7}$$

Since $\beta$ is estimable it can be shown that both $\mathfrak{B}(V)$ and $\mathfrak{B}(V,W)$ are nonempty. Hence, since $\underline{r}(x,V_1,V_2) = n$, (3.5) and (3.6) imply that each set on the right-hand side of (3.7) contains exactly one element.

Lemma 3.1: $\mathfrak{B}(V_1,V_2) = \mathfrak{B}(V_1) \cap \mathfrak{A}$ .

Proof: Clearly $\mathfrak{B}(V_1,V_2) \subset \mathfrak{B}(V_1) \cap \mathfrak{A}$ by (3.7). Suppose $t \in \mathfrak{B}(V_1) \cap \mathfrak{A}$ but $t \notin \mathfrak{B}(V_1,V_2)$. Let $t_0 \in \mathfrak{B}(V_1,V_2)$. Then $t_0'V_1t_0 = t'V_1t$ and $t_0'V_2t_0 < t'V_2t \Rightarrow t \notin \mathfrak{A}$, a contradiction. ∎

By Corollary 2.1 of Seely and Zyskind (1971), $t'Y$ is BLUE for $\beta$ if and only if $t \in \mathfrak{B}(V_1) \cap \mathfrak{B}(V_2)$. But a BLUE is admissible, so by Lemma 3.1 $t'Y$ is BLUE for $\beta$ if and only if $t \in \mathfrak{B}(V_1,V_2) \cap \mathfrak{B}(V_2,V_1)$, in which case $t'Y$ is the unique BLUE for $\beta$ and every set in (3.7) is identical.

We seek the class $\overline{\mathfrak{E}}$ of maximin-efficient ALUEs for $\beta$. For this purpose it is convenient to consider the following mutually exclusive cases:

(3a) $x \notin \underline{R}(V_1,V_2)$

(3b) $x \in \underline{R}(V_1,V_2)$ , $x \notin \underline{R}(V_1)$, $x \notin \underline{R}(V_2)$

(3c) $x \notin \underline{R}(V_1)$ , $x \in \underline{R}(V_2)$ or $x \in \underline{R}(V_1)$ , $x \notin \underline{R}(V_2)$

(3d) $x \in \underline{R}(V_1) \cap \underline{R}(V_2)$

Since we have assumed $\underline{r}(x, V_1,V_2) = n$, we have $\underline{r}(V_1,V_2) = n$ for cases (3b)-(3d). We consider case (3a) first:

<u>Theorem 3.2</u>: Consider model (3.2) where $\underline{r}(x,V_1,V_2) = n$ and $x \notin \underline{R}(V_1,V_2)$. Then there exists a unique BLUE for $\beta$ having zero variance for all covariance matrices $V \in \mathfrak{V}$.

<u>Proof</u>: Since $x \notin \underline{R}(V_1,V_2)$, $x'(I-P_M)x > 0$ where $P_M$ denotes the orthogonal projection matrix onto $\underline{R}(V_1,V_2)$. Let $\tilde{\beta} = (x'(I-P_M)x)^{-1}x'(I-P_M)Y$. Then $\tilde{\beta}$ is unbiased for $\beta$ and $\text{var}(\tilde{\beta}|V) = 0$ for all $V \in \mathfrak{V}$. Hence $\tilde{\beta}$ is a BLUE for $\beta$. BLUEs are unique because $\underline{r}(x,V_1,V_2) = n$. ▯

While considering the remaining cases, we make use of some well-known results for partitioned linear models. These results are summarized below.

Consider the fixed effects model

$$Z = A\alpha + D\delta + e \tag{3.8}$$

where Z is an n-vector of observations, A and D are known n x a and n x d matrices, respectively, $\alpha$ and $\delta$ are a- and d-vectors of fixed but unknown parameters, and e is an n-vector of random errors such that

$$E(e) = 0 \quad , \quad Cov(e) = I_n \quad . \tag{3.9}$$

We then have

$$E(Z) = A\alpha + D\delta \quad , \quad Cov(Z) = I_n \quad . \tag{3.10}$$

The following theorem summarizes some well-known properties of model (3.8) or (3.10). Note that part (a) does not depend upon the covariance structure of Z.

<u>Theorem 3.3</u>: Let $P_D$ denote the orthogonal projection matrix onto $\underline{R}(D)$. Then, with respect to model (3.8) or (3.10), we have:

(a) A parametric function $\lambda'\alpha$ is estimable if and only if $\lambda \in \underline{R}(A'(I-P_D)A)$.

(b) There exists a random vector $\hat{\alpha}$ satisfying $A'(I-P_D)A\hat{\alpha} = A'(I-P_D)Y$, and for any such $\hat{\alpha}$, $\lambda'\hat{\alpha}$ is the unique BLUE for any estimable $\lambda'\alpha$.

Consider the artificial model

$$Y = x\beta + B_1\delta_1 + B_2\delta_2 + e \tag{3.11}$$

where Y, x, $\beta$, $B_1$, and $B_2$ are the same as in model (3.1) but where $\delta_1$ and $\delta_2$ are $q_1$- and $q_2$-vectors of fixed but unknown parameters and e is an n-vector of random errors satisfying (3.9). Combining Theorems 3.2 and 3.3 we obtain

<u>Corollary 3.4</u>: Let $\underline{r}(x,V_1,V_2) = n$. Then $\beta$ is estimable with respect to artificial model (3.11) if and only if $x \notin \underline{R}(V_1,V_2)$. In this case, there exists a unique BLUE for $\beta$ with respect to

this artificial model and it is also the unique BLUE for $\beta$ with respect to model (3.2)

Recall that with respect to model (3.2) the efficiency within $\overline{\mathfrak{L}}$ of any $s'Y \in \overline{\mathfrak{L}}$ when $\mathrm{Cov}(Y) = V$ is

$$\begin{aligned} \mathrm{Eff}(s|V) &= v_M(V)/s'Vs \quad , \text{ if } s'Vs > 0 \\ &= 1 \quad , \text{ if } s'Vs = 0 \end{aligned} \tag{3.12}$$

where $v_M(V) = \inf_{t\in\mathfrak{L}} t'Vt$. Note that $v_M(V) = t_0'Vt_0$ for any $t_0 \in \mathfrak{B}(V)$ and

$$\begin{aligned} v_M(V) &= (x'V^+x)^{-1} \quad , \text{ if } x \in \underline{R}(V) \quad , \\ &= 0 \quad , \text{ otherwise} \quad . \end{aligned}$$

<u>Lemma 3.5</u>: Consider model (3.2) where $x \notin \underline{R}(V)$ for some $V \in \mathfrak{V}$. Then, for any $s \in \mathfrak{L}$ we have

$$\begin{aligned} \mathrm{Eff}(s|V) &= 0 \quad , \text{ if } s'Vs > 0 \quad , \\ &= 1 \quad , \text{ if } s'Vs = 0 \quad . \end{aligned}$$

<u>Proof</u>: Since $x \notin \underline{R}(V)$ we have $\mathfrak{B}(V) = \{t \in \mathfrak{L}: t'Vt = 0\}$ and $v_M(V) = 0$, so the lemma follows from (3.12). ▮

<u>Corollary 3.6</u>: Consider model (3.2) where $x \notin \underline{R}(V_1)$, $x \notin \underline{R}(V_2)$, but $\underline{r}(V_1, V_2) = n$. Then the minimum efficiency of every LUE for $\beta$ is zero and $\overline{\mathfrak{E}} = \overline{\mathfrak{A}}$.

<u>Proof</u>: If $t \in \mathfrak{L}$ then $\mathrm{Eff}(t|V_i) = 0$ for $i = 1$ or $2$. Otherwise, by Lemma 3.5, $t'V_1t = t'V_2t = 0 \Rightarrow t \in \underline{N}(V_1) \cap \underline{N}(V_2) \Rightarrow t = 0$ $\Rightarrow t \notin \mathfrak{L}$, a contradiction. ▮

From Corollary 3.6 we see that the criterion of maximin efficiency does not reduce the class of ALUEs for $\beta$ in case (3b). Thus the criterion of maximin efficiency is not helpful in selecting a LUE for $\beta$ in model (3.2) under case (3b).

We introduce two useful artificial models:

$$E(Y) = x\beta + B_1\delta_1 \quad , \quad Cov(Y) = V_2 \quad , \text{ and} \tag{3.13}$$

$$E(Y) = x\beta + B_2\delta_2 \quad , \quad Cov(Y) = V_1 \quad . \tag{3.14}$$

Model (3.13) is obtained from model (3.1) by replacing $b_1$ by a $q_1$-vector $\delta_1$ of fixed but unknown parameters. Similarly, model (3.14) is obtained from model (3.1) by replacing $b_2$ by a $q_2$-vector $\delta_2$ of fixed but unknown parameters. Both models (3.13) and (3.14) possess a unique BLUE for every estimable parametric function of fixed parameters (e.g., Seely and Zyskind, 1971).

Lemma 3.7: Consider model (3.2) where $\underline{r}(x,V_1,V_2) = n$ and $x \notin \underline{R}(V_1)$. Then $\overline{\mathfrak{B}}(V_1)$ is the class of LUEs for $\beta$ and $\overline{\mathfrak{B}}(V_1,V_2)$ consists of the unique BLUE for $\beta$, both with respect to artificial model (3.13).

Proof: Since $x \notin \underline{R}(V_1)$ and $V_i = B_iB_i'$ , $i = 1,2$, we have

$$\begin{aligned}\mathfrak{B}(V_1) &= \{t \in \mathfrak{L}: V_1t = 0\} \\ &= \{t \in R^n: x't = 1 \quad , \quad B_1't = 0\} \quad \text{and}\end{aligned}$$

$$\begin{aligned}\mathfrak{B}(V_1,V_2) &= \{t \in \mathfrak{B}(V_1): V_2t \in \underline{R}(x,V_1)\} \\ &= \{t \in R^n: x't = 1 \quad , \quad B_1't = 0 \quad , \quad V_2t \in \underline{R}(x,B_1)\} \quad .\end{aligned}$$

It is easy to see that $\overline{\mathfrak{B}}(V_1)$ is the class of LUEs for $\beta$ with respect to artificial model (3.13). By Theorem 3 of Zyskind (1967), $\overline{\mathfrak{B}}(V_1,V_2)$ is then the class of BLUEs for $\beta$ with respect to artificial model (3.13) and it consists of only one element by (3.6). ∎

When $\underline{r}(V_1,V_2) = n$ and $c \in (0,1)$ the matrix $cV_1 + (1-c)V_2$ is p.d. and $\overline{\mathfrak{B}}(cV_1 + (1-c)V_2)$ consists of the unique BLUE for $\beta$ with respect to the artificial model

$$E(Y) = x\beta \quad , \quad Cov(Y) = cV_1 + (1-c)V_2 \quad , \quad c \in (0,1) \quad . \tag{3.15}$$

We can now characterize the set $\overline{\mathcal{E}} = \overline{\mathcal{A}}$ in case 3(b) as a set of BLUEs for $\beta$ with respect to artificial models (3.13)-(3.15).

Theorem 3.8: Consider model (3.2) where $x \notin \underline{R}(V_1)$ , $x \notin \underline{R}(V_2)$, but $\underline{r}(V_1,V_2) = n$. Then $\beta$ is nonestimable with respect to artificial model (3.11) but is estimable with respect to artificial models (3.13)-(3.15). $\overline{\mathfrak{B}}(V_1,V_2)$, $\overline{\mathfrak{B}}(V_2,V_1)$, and $\overline{\mathfrak{B}}(cV_1 + (1-c)V_2)$ for $c \in (0,1)$ consist of the unique BLUE for $\beta$ with respect to artificial models (3.13), (3.14), and (3.15), respectively.

Now Example 1.1 falls under case (3b). Let

$$t_c'Y = [(1-c)Y_{1.} + cY_{2.}] \,/\, [(1-c)n_1 + cn_2] \text{ for } c \in [0,1]$$

where $Y_{i.} = \sum_j Y_{ij}$ . Then

$$\begin{aligned}\overline{\mathfrak{B}}(V_1,V_2) &= \{t'Y : x't = 1 \quad , \quad V_1t = 0 \quad , \quad V_2t \in \underline{R}(x,V_1)\} \\ &= \{t'Y : x't = 1 \quad , \quad t \in \underline{R}(V_2) \cap \underline{R}(x,V_1)\} \\ &= \{t_1'Y\} \qquad , \qquad \text{and}\end{aligned}$$

$$\overline{\mathfrak{B}}(V_2,V_1) = \{t_0'Y\} \quad .$$

Applying Theorem 3.8 we obtain

$$\overline{\mathfrak{B}}(cV_1 + (1-c)V_2) = \{t_c'Y\} \quad , \quad c \in (0,1).$$

Finally, from Corollary 3.5 and (3.7) we have

$$\overline{\mathfrak{E}} = \overline{\mathfrak{A}} = \{t_c'Y \quad , \quad c \in [0,1]\} \quad .$$

The admissible class $\overline{\mathfrak{A}}$ for this example can also be found in LaMotte (1977).

We now turn our attention to cases (3c) and (3d). For $\gamma \in R^1$ let $W(\gamma) = \gamma V_1 + (1-\gamma)V_2$. Then $\mathfrak{V} = \{\theta W(\gamma) : \theta \geqslant 0 \quad , \gamma \in [0,1]\} = co(\mathfrak{W})$, where $\mathfrak{W} = \{W(\gamma) : \gamma \in [0,1]\}$. Then $\overline{\mathfrak{A}}(\mathfrak{V}) = \overline{\mathfrak{A}}(\mathfrak{W})$ and $\overline{\mathfrak{E}}(\mathfrak{V}) = \overline{\mathfrak{E}}(\mathfrak{W})$ and for the purposes of determining $\overline{\mathfrak{A}}$ and $\overline{\mathfrak{E}}$ we may replace $\mathfrak{V}$ by $\mathfrak{W}$:

<u>Corollary 3.9</u>: Consider model (3.2). For $\gamma \in R^1$ let $W(\gamma) = \gamma V_1 + (1-\gamma)V_2$. Then $t \in \mathfrak{E}$ if and only if $t \in \mathfrak{A}$ and

$$\inf_{\gamma\in[0,1]} \mathrm{Eff}(t|W(\gamma)) \geqslant \inf_{\gamma\in[0,1]} \mathrm{Eff}(s|W(\gamma))$$

for all $s \in \mathfrak{A}$.

We have $\underline{r}(V_1,V_2) = n$ in cases (3c) and (3d). Then $W(\gamma)$ is p.d. and for $s \in \mathfrak{A}$ we have

$$\mathrm{Eff}(s|W(\gamma)) = 1 \; / \; [x'W(\gamma)^{-1}x \; s'W(\gamma)s] > 0$$

and is a continuous function of $\gamma$ for $\gamma \in (0,1)$. To determine the behavior of $\mathrm{Eff}(s|W(\gamma))$ at $\gamma = 0$ and $\gamma = 1$, we use a theorem due to Seely (1983a). This theorem may also be obtained from more general results in Seely (1983b).

Theorem 3.10 (Seely): Let A and B be $n \times n$ n.n.d. matrices such that $\underline{r}(A,B) = n$. Let $x \in R^n$, $x \neq 0$, and $t \in \mathfrak{B}(A,B)$. Then

(a) $\lim_{\theta \to 0^+} x'(\theta A + B)^{-1}x = x'B^+x \quad \text{if } x \in \underline{R}(B)$ ,

$= +\infty \quad \text{if } x \notin \underline{R}(B)$ .

(b) $\lim_{\theta \to \infty} x'(\theta A + B)^{-1}x = 0 \quad \text{if } x \in \underline{R}(A)$ ,

$= 1/t'Bt \quad \text{if } x \notin \underline{R}(A)$ .

Corollary 3.11: Consider model (3.2) where $\underline{r}(V_1,V_2) = n$ and let $s \in \mathfrak{A}$. Then $\mathrm{Eff}(s|W(\gamma))$ is a continuous function of $\gamma$ for $\gamma \in [0,1]$.

Proof: By symmetry and the paragraph preceding Theorem 3.10, it suffices to show that $\mathrm{Eff}(s|W(\gamma))$ is continuous at $\gamma = 1$. Suppose $x \in \underline{R}(V_1)$. Then $v_M(V_1) = 1/x'V_1^+x > 0 \Rightarrow s'V_1s > 0$ and

$$\begin{aligned}
\lim_{\gamma \to 1} \mathrm{Eff}(s|W(\gamma)) &= \lim_{\gamma \to 1}[x'(V_1 + \tfrac{1-\gamma}{\gamma}V_2)^{-1}x \; s'(V_1 + \tfrac{1-\gamma}{\gamma}V_2)s]^{-1} \\
&= \lim_{\theta \to 0^+}[x'(V_1 + \theta V_2)^{-1}x \; s'(V_1 + \theta V_2)s]^{-1} \\
&= [x'V_1^+x \; s'V_1s]^{-1} \qquad \text{(Theorem 3.10a)} \\
&= v_M(V_1)/s'V_1s \;=\; \mathrm{Eff}(s|V_1) > 0 \quad .
\end{aligned}$$

Now suppose $x \notin \underline{R}(V_1)$; then $v_M(V_1) = 0$. If $s'V_1s > 0$ then, also by Theorem 3.10a,

$$\lim_{\gamma\to 1}\mathrm{Eff}(s|W(\gamma)) = 0 = v_M(V_1)/s'V_1s = \mathrm{Eff}(s|V_1) \quad .$$

If $s'V_1s = 0$ then $s \in \mathfrak{B}(V_1) \cap \mathbb{A} = \mathfrak{B}(V_1,V_2)$ by Lemma 3.1. Also $s'V_2s > 0$ (otherwise $s = 0$) and

$$\begin{aligned}\lim_{\gamma\to 1}\mathrm{Eff}(s|W(\gamma)) &= \lim_{\gamma\to 1}[x'(\tfrac{\gamma}{1-\gamma}V_1 + V_2)^{-1}x\; s'V_2s]^{-1} \\ &= \lim_{\theta\to\infty}[x'(\theta V_1 + V_2)^{-1}x\; s'V_2s]^{-1} \\ &= s'V_2s/s'V_2s = 1 = \mathrm{Eff}(s|V_1)\end{aligned}$$

by Theorem 3.10b. ∎

Now consider case (3c). Suppose first that $x \notin \underline{R}(V_1)$ and $x \in \underline{R}(V_2)$. The following theorem says that the unique element of $\overline{\mathfrak{B}}(V_1,V_2)$ is the unique maximin-efficient ALUE for $\beta$:

Theorem 3.12: Consider model (3.2) where $x \notin \underline{R}(V_1)$ , $x \in \underline{R}(V_2)$ , and $\underline{r}(V_1,V_2) = n$. Then $\overline{\mathcal{E}} = \overline{\mathfrak{B}}(V_1,V_2)$.

Proof: Let $s \in \mathbb{A}$. If $s \notin \mathfrak{B}(V_1,V_2)$ then by Lemma 3.1 and the proof of Corollary 3.11 $s \notin \mathfrak{B}(V_1)$ and

$$\inf_{\gamma\in[0,1]}\mathrm{Eff}(s|W(\gamma)) = \mathrm{Eff}(s|V_1) = 0 \quad .$$

If $s \in \mathfrak{B}(V_1,V_2)$, then $s'V_2s > 0$ and $\mathrm{Eff}(s|W(\gamma)) > 0$ for all $\gamma \in [0,1]$. By Corollary 3.11, $\mathrm{Eff}(s|W(\gamma))$ is continuous for $\gamma \in [0,1]$ so

$$\inf_{\gamma\in[0,1]}\mathrm{Eff}(s|W(\gamma)) > 0 \quad .$$

$\mathfrak{B}(V_1,V_2)$ has only one element, so $\mathcal{E} = \mathfrak{B}(V_1,V_2)$. ∎

Corollary 3.13: Consider model (3.2) where $x \notin \underline{R}(V_1)$ , $x \in \underline{R}(V_2)$ and $\underline{r}(V_1,V_2) = n$. Then $\beta$ is estimable with respect to artificial model (3.13) but is nonestimable with respect to artificial model (3.14) and $\bar{\mathcal{E}} = \bar{\mathfrak{B}}(V_1,V_2)$ consists of the unique BLUE for $\beta$ with respect to artificial model (3.13).

Proof: Let $P_i$ denote the orthogonal projection matrix onto $\underline{R}(V_i) = \underline{R}(B_i)$ for $i = 1,2$. Then $x'(I-P_1)x \neq 0$ and $x'(I-P_2)x = 0$. The estimability statements then follow from Theorem 3.3a. The rest follows by Lemma 3.7 and Theorem 3.12. ∎

Note that the roles of $V_1$ and $V_2$ may be reversed in Theorem 3.12 and Corollary 3.13 to obtain results for case (3c) when $x \in \underline{R}(V_1)$ and $x \notin \underline{R}(V_2)$.

Finally, we consider case (3d). That is, $x \in \underline{R}(V_1) \cap \underline{R}(V_2)$ and $\underline{r}(V_1,V_2) = n$. Then $H = V_1 + V_2$ is p.d..

Let $Z = H^{-\frac{1}{2}}Y$. Then we have the transformed model

$$E(Z) = w\beta \quad , \quad \mathrm{Cov}(Z) = \theta_1 U + \theta_2 V \quad , \qquad (3.17)$$

where $w = H^{-\frac{1}{2}}x$ , $U = H^{-\frac{1}{2}}V_1H^{-\frac{1}{2}}$ and $V = H^{-\frac{1}{2}}V_2H^{-\frac{1}{2}}$ are n.n.d., and $w \in \underline{R}(U) \cap \underline{R}(V)$. Since H is p.d., $H^{-\frac{1}{2}}$ is nonsingular. As pointed out in Chapter 2, the classes of LUEs, ALUEs, and maximin-efficient ALUEs and the variance and efficiency of any LUE are invariant under nonsingular linear transformations of Y. Hence the classes $\bar{\mathcal{L}}$, $\bar{\mathcal{A}}$, and $\bar{\mathcal{E}}$ with respect to model (3.2) are the same as those classes with respect to model (3.17). Also, the variance and the efficiency within $\bar{\mathbb{A}}$ of any element of $\bar{\mathcal{L}}$ are the same with respect to models (3.2) and (3.17).

We now obtain $\bar{\mathbb{A}}$ and characterize the unique element of $\bar{\mathcal{E}}$ with respect to model (3.17). These results will then be re-expressed in terms of model (3.2).

Now

$$U + V = H^{-\frac{1}{2}}(V_1 + V_2)H^{-\frac{1}{2}} = H^{-\frac{1}{2}}HH^{-\frac{1}{2}} = I \qquad . \tag{3.18}$$

Then $UU + VU = U \Rightarrow VU$ is symmetric $\Rightarrow$ U and V commute. These facts will be repeatedly used below.

Lemma 3.14: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. For $c \in [0,1]$ let $W(c) = cU + (1-c)V$. Then $\bar{\mathbb{A}} = \{\hat{\beta}(c) : c \in [0,1]\}$ where, for $c \in [0,1]$,

$$\hat{\beta}(c) = (w'W(c)^+w)^{-1}\, w'W(c)^+Z \qquad .$$

Furthermore, there exists a BLUE for $\beta$ if and only if w is an eigenvector of both U and V, in which case $\hat{\beta} = (w'w)^{-1}w'Z$ is the unique BLUE for $\beta$.

Proof: From (3.7) we have

$$\bar{\mathbb{A}} = \bar{\mathbb{B}}(U,V) \cup \bar{\mathbb{B}}(V,U) \cup \bigcup_{0<c<1} \bar{\mathbb{B}}(W(c)) \qquad .$$

For $c \in (0,1)$, $W(c)$ is p.d. and $\bar{\mathbb{B}}(W(c)) = \{\hat{\beta}(c)\}$. By Proposition 3.2.11 of Azzam (1981), $\bar{\mathbb{B}}(U,V) = \bar{\mathbb{B}}(U,U + V) = \bar{\mathbb{B}}(U,I)$. Since $\hat{\beta}(1) \in \bar{\mathbb{B}}(U,I)$ and $\bar{\mathbb{B}}(U,I)$ contains exactly one element by (3.6), we have $\bar{\mathbb{B}}(U,V) = \{\hat{\beta}(1)\}$. Similarly, $\bar{\mathbb{B}}(V,U) = \{\hat{\beta}(0)\}$.

From the discussion following Lemma 3.1, $t'Y$ is a BLUE for $\beta$ if and only if $t \in \mathfrak{B}(U,V) \cap \mathfrak{B}(V,U)$, in which case this BLUE is unique. Suppose a BLUE for $\beta$ exists. Then

$$\begin{aligned}
\hat{\beta}(1) = \hat{\beta}(0) \;&\Rightarrow\; (w'U^+w)^{-1}U^+w = (w'V^+w)^{-1}V^+w && \text{(uniqueness)}\\
&\Rightarrow\; (w'V^+w)w = (w'U^+w)UV^+w && (w \in \underline{R}(U))\\
&\qquad\qquad = (w'U^+w)(I-V)V^+w && (U + V = I)\\
&\Rightarrow\; (w'V^+w + w'U^+w)w = (w'U^+w)V^+w && (w \in \underline{R}(V))\\
&\Rightarrow\; Vw = [(w'U^+w) \,/\, (w'V^+w + w'U^+w)]w
\end{aligned}$$

and w is an eigenvector of V. Similarly, w is an eigenvector of U.

Suppose now that w is an eigenvector of U and V. Then $Uw = \lambda w$ and $Vw = \nu w$, where $\lambda, \nu > 0$, so $U^+w = (1/\lambda)w$, $V^+w = (1/\nu)w$, and $\hat{\beta}$ is the unique BLUE for $\beta$. ∎

We wish to characterize the maximin-efficient ALUEs for $\beta$ with respect to model (3.17), if such estimators exist. By Corollary 3.9, we need only compare the efficiencies of ALUEs for covariance matrices lying in the set $\mathfrak{M} = \{W(\gamma) : \gamma \in [0,1]\}$, where $W(\gamma) = \gamma U + (1-\gamma)V$. For $c, \gamma \in [0,1]$ we shall denote the efficiency of the ALUE $\hat{\beta}(c)$ when $\mathrm{Cov}(Z) = W(\gamma)$ by $\mathrm{Eff}(c|\gamma)$. Then we wish to determine the values $c^* \in [0,1]$ such that

$$\inf_{\gamma \in [0,1]} \mathrm{Eff}(c^*|\gamma) \geq \inf_{\gamma \in [0,1]} \mathrm{Eff}(c|\gamma) \tag{3.19}$$

for all $c \in [0,1]$.

Let $c \in [0,1]$. By Lemmas 2.1 and 2.2 and equation (3.12) we have

$$\mathrm{Eff}(c|\gamma) = \frac{(w'W(c)^+w)^2}{(w'W(\gamma)^+w)(w'W(c)^+W(\gamma)W(c)^+w)} \tag{3.20}$$

for $\gamma \in [0,1]$. Since $w \in \underline{R}(U) \cap \underline{R}(V)$, the proof of Corollary 3.11 shows that $\mathrm{Eff}(c|\gamma)$ is a continuous and positive function of $\gamma$ for $\gamma \in [0,1]$.

To find the $c^*$ satisfying (3.19) we first investigate the behavior of $\mathrm{Eff}(c|\gamma)$ as a function of $\gamma$ for fixed c and then as a function of c for fixed $\gamma$. In doing this, we use the simultaneous spectral decomposition of U and V.

Let $0 \leqslant \lambda_1 < \ldots < \lambda_m$ denote the m distinct eigenvalues of U with multiplicities $r_1, \ldots, r_m$, respectively. For $k = 1, \ldots, m$ let $E_k$ denote the orthogonal projection matrix onto the subspace of eigenvectors of U corresponding to $\lambda_k$. Since $U + V = I$, we have

$$U = \sum_{k=1}^{m} \lambda_k E_k \qquad \text{and} \qquad V = \sum_{k=1}^{m} \nu_k E_k$$

where $\nu_k = 1 - \lambda_k$ , $0 \leqslant \lambda_1 < \ldots < \lambda_m \leqslant 1$ , and $1 \geqslant \mu_1 > \ldots > \mu_m \geqslant 0$.

Define $r = \min\,\{k:\lambda_k > 0\}$ and $s = \max\,\{k:\mu_k > 0\}$. Then $k \in [r,s] \Leftrightarrow 0 < \lambda_k, \mu_k < 1$. Also,

$$U = \sum_{k=r}^{m} \lambda_k E_k \qquad \text{and} \qquad V = \sum_{k=1}^{s} \nu_k E_k \qquad .$$

If $r > s$, then $R(U) \cap \underline{R}(V) = \{0\}$. But $w \neq 0$ and $w \in \underline{R}(U) \cap \underline{R}(V)$, so $r \leqslant s$ and

$$\underline{R}(U) \cap \underline{R}(V) = \sum_{k=r}^{s} \underline{R}(E_k).$$

If $r = s$, then $w \in \underline{R}(E_k)$ and w is an eigenvector of U and V. In this case $\hat{\beta} = (w'w)^{-1}w'Y$ is the unique BLUE for $\beta$ by Lemma 3.14.

We suppose, for now, that $r < s$. Let

$$\varphi_k(c) = c\lambda_k + (1 - c)(1 - \lambda_k).$$

Then $\varphi_k(c) > 0$ for all $c \in [0,1]$ and $k \in [r,s]$, and

$$\begin{aligned} W(c) &= \sum_{k=1}^{m} \varphi_k(c)E_k && \text{for } c \in [0,1] \quad , \\ W(c)^+ &= \sum_{k=1}^{s} \varphi_k(0)^{-1}E_k && \text{for } c = 0 \\ &= \sum_{k=1}^{m} \varphi_k(c)^{-1}E_k && \text{for } c \in (0,1) \\ &= \sum_{k=r}^{m} \varphi_k(1)^{-1}E_k && \text{for } c = 1 \quad . \end{aligned}$$

Since $w \in \underline{R}(U) \cap \underline{R}(V) = \sum_{k=r}^{s} \underline{R}(E_k)$ ,

$$\begin{aligned} w'W(\gamma)^+w &= \sum_{k=r}^{s} \varphi_k(\gamma)^{-1}w'E_kw \qquad \text{and} \\ w'W(c)^+W(\gamma)W(c)^+w &= \sum_{k=r}^{s} \varphi_k(\gamma)\varphi_k(c)^{-2}w'E_kw \end{aligned} \tag{3.21}$$

for all $c,\gamma \in [0,1]$.

Lemma 3.15: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. Suppose $w$ is not an eigenvector of $U$ and $V$ and let $c \in [0,1]$ be fixed. Then $\mathrm{Eff}(c|\gamma)$ is a continuously differentiable function of $\gamma$ on $[0,1]$, is strictly increasing on $[0,c)$, and is strictly decreasing on $(c,1]$.

Proof: For $c,\gamma \in [0,1]$ we may write

$$\mathrm{Eff}(c|\gamma) = f(c)^2 \,/\, [f(\gamma)g_c(\gamma)] > 0 \quad ,$$

where

$$f(\gamma) = w'W(\gamma)^{+}w = \sum_{k=r}^{s}\varphi_k(\gamma)^{-1}w'E_kw \qquad \text{and}$$

$$g_c(\gamma) = w'W(c)^{+}W(\gamma)W(c)^{+}w = \sum_{k=r}^{s}\varphi_k(\gamma)\varphi_k(c)^{-2}w'E_kw \qquad .$$

Now $\varphi_k(\gamma) > 0$ and $E_k$ is n.n.d. for $k \in [r,s]$. Since w is not an eigenvector of U and V, $w'E_jw > 0$ and $w'E_kw > 0$ for some j,k such that $r \leqslant j < k \leqslant s$. Then for fixed $c \in [0,1]$ both $f(\gamma)$ and $g_c(\gamma)$ are positive and differentiable functions on some open interval containing [0,1]. For fixed $c \in [0,1]$, $\mathrm{Eff}(c|\gamma)$ is then differentiable with respect to $\gamma$ on [0,1] with

$$\frac{\partial}{\partial\gamma}\mathrm{Eff}(c|\gamma) = -f(c)^2[f(\gamma)g_c(\gamma)]^{-2}\frac{\partial}{\partial\gamma}[f(\gamma)g_c(\gamma)].$$

This has the same sign as $-\frac{\partial}{\partial\gamma}[f(\gamma)g_c(\gamma)]$.

Here

$$f'(\gamma) = \sum_{k=r}^{s}(1 - 2\lambda_k)\varphi_k(\gamma)^{-2}w'E_kw \qquad \text{and}$$

$$g_c'(\gamma) = -\sum_{k=r}^{s}(1 - 2\lambda_k)\varphi_k(c)^{-2}w'E_kw \quad , \quad \text{so}$$

$$-\frac{\partial}{\partial\gamma}[f(\gamma)g_c(\gamma)] = -f'(\gamma)g_c(\gamma) - f(\gamma)g_c'(\gamma)$$

$$= \sum_{k=r}^{s}\sum_{j=r}^{s}\{-(1 - 2\lambda_k)\varphi_j(\gamma) + \varphi_k(\gamma)(1 - 2\lambda_j)\}$$

$$\times\, \varphi_k(\gamma)^{-2}\varphi_j(c)^{-2}w'E_kww'E_jw$$

$$= \sum_{k=r}^{s}\sum_{j=r}^{s}(\lambda_k - \lambda_j)\varphi_k(\gamma)^{-2}\varphi_j(c)^{-2}w'E_kww'E_jw$$

$$= \sum_{k=r+1}^{s}\sum_{j=r}^{k-1}\{\varphi_k(c)^2\varphi_j(\gamma)^2 - \varphi_k(\gamma)^2\varphi_j(c)^2\}(\lambda_k - \lambda_j)$$

$$\times\, \varphi_k(\gamma)^{-2}\varphi_k(c)^{-2}\varphi_j(\gamma)^{-2}\varphi_j(c)^{-2}w'E_kww'E_jw \qquad .$$

Now { } = $[\varphi_k(c)\varphi_j(\gamma) - \varphi_k(\gamma)\varphi_j(c)][\varphi_k(c)\varphi_j(\gamma) + \varphi_k(\gamma)\varphi_j(c)]$ ,

$$\varphi_k(c)\varphi_j(\gamma) - \varphi_k(\gamma)\varphi_j(c) = (c - \gamma)(\lambda_k - \lambda_j) \quad , \text{ and}$$

$$\psi_{kj}(\gamma,c) = \varphi_k(c)\varphi_j(\gamma) + \varphi_k(\gamma)\varphi_j(c) > 0$$

for $\gamma,c \in [0,1]$ and $k,j \in [r,s]$. Then since $w'E_kww'E_jw > 0$ for some $j,k$ such that $r \leqslant j < k \leqslant s$ ,

$$-\frac{\partial}{\partial\gamma}[f(\gamma)g_c(\gamma)] = (c - \gamma)\sum_{k=r+1}^{s}\sum_{j=r}^{k-1} \psi_{kj}(\gamma,c)(\lambda_k - \lambda_j)^2$$

$$\times\ \varphi_k(\gamma)^{-2}\varphi_k(c)^{-2}\varphi_j(\gamma)^{-2}\varphi_j(c)^{-2}w'E_kww'E_jw$$

$$\begin{aligned} &> 0 \qquad \text{for } \gamma \in [0,c) \\ &= 0 \qquad \text{for } \gamma = c \\ &< 0 \qquad \text{for } \gamma \in (c,1] \quad . \end{aligned}$$

Furthermore, $\frac{\partial}{\partial\gamma}[f(\gamma)g_c(\gamma)]$, hence $\frac{\partial}{\partial\gamma}\text{Eff}(c|\gamma)$ is continuous for $\gamma \in [0,1]$. ∎

<u>Lemma 3.16</u>: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. Suppose $w$ is not an eigenvector of $U$ and $V$ and let $\gamma \in [0,1]$ be fixed. Then $\text{Eff}(c|\gamma)$ is a continuously differentiable function of $c$ on $[0,1]$, is strictly increasing on $[0,\gamma)$, and is strictly decreasing on $(\gamma,1]$.

<u>Proof</u>: For $c,\gamma \in [0,1]$ we may write

$$\text{Eff}(c|\gamma) = f(c)^2 \,/\, [f(\gamma)h_\gamma(c)] > 0 \qquad ,$$

where $h_\gamma(c) = g_c(\gamma)$ and where $f(c)$ and $g_c(\gamma)$ are defined in the proof of Lemma 3.15.

For the same reasons as were given in the proof of Lemma 3.15, for fixed $\gamma \in [0,1]$ both $f(c)$ and $h_\gamma(c)$ are positive and differentiable functions on some open interval containing $[0,1]$. For fixed $\gamma \in [0,1]$, $\mathrm{Eff}(c|\gamma)$ is then differentiable with respect to c on $[0,1]$ with

$$\frac{\partial}{\partial c}\mathrm{Eff}(c|\gamma) = f(c)[f(\gamma)h_\gamma(c)^2]^{-1}[2f'(c)h_\gamma(c) - f(c)h_\gamma'(c)] \quad ,$$

which has the same sign as $2f'(c)h_\gamma(c) - f(c)h_\gamma'(c)$.

Here

$$h_\gamma'(c) = 2\sum_{k=r}^{s}(1 - 2\lambda_k)\varphi_k(\gamma)\varphi_k(c)^{-3}w'E_kw$$

so

$$\begin{aligned}
&2f'(c)h_\gamma(c) - f(c)h_\gamma'(c) \\
&\quad = 2 \sum_{k=r}^{s} \sum_{j=r}^{s} \{(1 - 2\lambda_k)\varphi_j(c) - \varphi_k(c)(1 - 2\lambda_j)\} \\
&\qquad\qquad \times \varphi_k(c)^{-2}\varphi_j(\gamma)\varphi_j(c)^{-3}w'E_kww'E_jw \\
&\quad = -2 \sum_{k=r}^{s} \sum_{j=r}^{s} (\lambda_k - \lambda_j)\varphi_k(c)^{-2}\varphi_j(\gamma)\varphi_j(c)^{-3}w'E_kww'E_jw \\
&\quad = -2 \sum_{k=r+1}^{s} \sum_{j=r}^{k-1} \{\varphi_k(c)\varphi_j(\gamma) - \varphi_k(\gamma)\varphi_j(c)\}(\lambda_k - \lambda_j) \\
&\qquad\qquad \times \varphi_k(c)^{-3}\varphi_j(c)^{-3}w'E_kww'E_jw \\
&\quad = -2 \sum_{k=r+1}^{s} \sum_{j=r}^{k-1} (c - \gamma)(\lambda_k - \lambda_j)^2\varphi_k(c)^{-3}\varphi_j(c)^{-3}w'E_kww'E_jw \\
&\quad = 2(\gamma - c) \sum_{k=r+1}^{s} \sum_{j=r}^{k-1} (\lambda_k - \lambda_j)^2\varphi_k(c)^{-3}\varphi_j(c)^{-3}w'E_kww'E_jw
\end{aligned}$$

$> 0$ for $c \in [0,\gamma)$ , $= 0$ for $c = \gamma$ , $< 0$ for $c \in (\gamma,1]$

and is continuous on [0,1], hence so is $\frac{\partial}{\partial c}\mathrm{Eff}(c|\gamma)$. ▮

Lemma 3.17: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. Then the unique maximin-efficient ALUE for $\beta$ is the unique ALUE for $\beta$ whose efficiency when $\mathrm{Cov}(Z) = U$ is equal to its efficiency when $\mathrm{Cov}(Z) = V$.

Proof: Suppose first that w is not an eigenvector of U and V. Let $m(c)$ denote the minimum efficiency of $\hat{\beta}(c)$ over $\gamma \in [0,1]$. Lemma 3.15 implies that $m(c) = \min\{\mathrm{Eff}(c|0), \mathrm{Eff}(c|1)\}$. Lemma 3.16 says that $\mathrm{Eff}(c|0)$ is strictly decreasing and that $\mathrm{Eff}(c|1)$ is strictly increasing on [0,1]. Since $\mathrm{Eff}(c|0)$ and $\mathrm{Eff}(c|1)$ are both continuous functions and $\mathrm{Eff}(0|0) = \mathrm{Eff}(1|1) = 1$, there exists a unique $c^* \in (0,1)$ such that $\mathrm{Eff}(c^*|0) = \mathrm{Eff}(c^*|1)$. $m(c)$ is strictly increasing on $[0,c^*)$ and is strictly decreasing on $(c^*,1]$. Therefore, the unique maximin-efficient ALUE for $\beta$ is $\tilde{\beta} = \hat{\beta}(c^*)$.

Suppose now that w is an eigenvector of both U and V. Then by Lemma 3.14, $\hat{\beta} = (w'w)^{-1}w'Z$ is the unique BLUE, hence the unique ALUE, for $\beta$ and $\hat{\beta}$ has efficiency 1 for $\mathrm{Cov}(Z) = U$ and $\mathrm{Cov}(Z) = V$. ▮

Corollary 3.18: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. Let $W(c) = cU + (1-c)V$. Then there exists a unique maximin-efficient ALUE for $\beta$ given by

$$\tilde{\beta} = \hat{\beta}(c^*) = (w'W(c^*)^+w)w'W(c^*)^+Z$$

where $c^*$ is any value in [0,1] satisfying

$$(w'U^+w)(w'W(c^*)^+UW(c^*)^+w) = (w'V^+w)(w'W(c^*)^+VW(c^*)^+w) \quad . \quad (3.22)$$

Proof: If $w$ is not an eigenvector of $U$ and $V$ then, from the proof of Lemma 3.17, $\tilde{\beta} = \hat{\beta}(c^*)$ where $c^*$ is the unique element in $(0,1)$ satisfying $\mathrm{Eff}(c^*|0) = \mathrm{Eff}(c^*|1)$, which is equivalent to (3.22).

If $w$ is an eigenvector of $U$ and $V$, then $Uw = \lambda w$ and $Vw = \nu w$ for $\lambda, \nu > 0$. Let $c \in [0,1]$. Then $W(c)w = [c\lambda + (1-c)\nu]w \Rightarrow W(c)^+w = [c\lambda + (1-c)\nu]^{-1}w \Rightarrow$ $\hat{\beta}(c) = \hat{\beta} = (w'w)^{-1}w'Z$ and $c^* = c$ satisfies (3.22). ▮

Corollary 3.19: Consider model (3.17) where $w \in \underline{R}(U) \cap \underline{R}(V)$ and $U + V = I$. If $w$ is an eigenvector of $U$ and $V$, then any $c^* \in [0,1]$ satisfies (3.22). Otherwise, $c^* \in [0,1]$ satisfying (3.22) is unique and it also satisfies

$$0 < \min_{(k,j)\in T} \left[1 + \left(\frac{\lambda_k \lambda_j}{\mu_k \mu_j}\right)^{\frac{1}{2}}\right]^{-1} \leq c^* \leq \max_{(k,j)\in T} \left[1 + \left(\frac{\lambda_k \lambda_j}{\mu_k \mu_j}\right)^{\frac{1}{2}}\right]^{-1} < 1,$$

where $T = \{(k,j) : k \neq j \text{ and } w'E_k ww'E_j w > 0\}$.

Proof: The first assertion follows from the proof of Corollary 3.18. The uniqueness of $c^*$ when $w$ is not an eigenvector of $U$ and $V$ also follows from that corollary. In the latter case, equation (3.22) may be re-expressed, using (3.21), as

$$0 = \sum_{k=r}^{s} \sum_{j=r}^{s} \{\varphi_k(0)^{-1}\varphi_j(0) - \varphi_k(1)^{-1}\varphi_j(1)\}\varphi_j(c^*)^{-2} w'E_k ww'E_j w$$

$$= \sum_{k=r}^{s} \sum_{j=r}^{s} (\lambda_k - \lambda_j)\lambda_k^{-1}\mu_k^{-1}\varphi_j(c^*)^{-2} w'E_k ww'E_j w$$

$$= \sum_{k=r+1}^{s} \sum_{j=r}^{k-1} \{\lambda_j\mu_j\varphi_k(c^*)^2 - \lambda_k\mu_k\varphi_j(c^*)^{-2}\}(\lambda_k - \lambda_j)$$
$$\times (\lambda_k\lambda_j\mu_k\mu_j)^{-1}\varphi_k(c^*)^{-2}\varphi_j(c^*)^{-2}w'E_k ww'E_j w$$

$$= \sum_{k=r+1}^{s} \sum_{j=r}^{k-1} \{c^{*2}\lambda_k\lambda_j - (1 - c^*)^2\mu_k\mu_j\}(\lambda_k - \lambda_j)^2(\lambda_k\lambda_j\mu_k\mu_j)^{-1}$$
$$\times \varphi_k(c^*)^{-2}\varphi_j(c^*)^{-2}w'E_k ww'E_j w \quad .$$

If $c^{*2}\lambda_k\lambda_j - (1 - c^*)^2\mu_k\mu_j > 0$ , that is,

$$c^* > [1 + \left(\frac{\lambda_k \lambda_j}{\mu_k \mu_j}\right)^{\frac{1}{2}}]^{-1}$$

for all $(k,j) \in T$ ($T$ is nonempty), then the above sum is positive. If the reverse inequality holds for $(k,j) \in T$, then the sum is negative. ∎

We now apply these results to model (3.2).

Theorem 3.20: Consider model (3.2) where $x \in \underline{R}(V_1) \cap \underline{R}(V_2)$ and $\underline{r}(V_1,V_2) = n$. Let $H = V_1 + V_2$ and for $c \in [0,1]$ let $W(c) = cV_1 + (1 - c)V_2$. Then $\bar{A} = \{\hat{\beta}(c):c \in [0,1]\}$, where

$$\begin{aligned}
\hat{\beta}(c) &= (x'V_2^+x)^{-1}x'H^{-\frac{1}{2}}(H^{-\frac{1}{2}}V_2H^{-\frac{1}{2}})^+H^{-\frac{1}{2}}Y \quad , \quad c = 0 \\
&= (x'W(c)^{-1}x)^{-1}x'W(c)^{-1}Y \quad , \quad c \in (0,1) \\
&= (x'V_1^+x)^{-1}x'H^{-\frac{1}{2}}(H^{-\frac{1}{2}}V_1H^{-\frac{1}{2}})^+H^{-\frac{1}{2}}Y \quad , \quad c = 1 \quad .
\end{aligned}$$

The unique maximin-efficient ALUE for $\beta$ is the unique ALUE for $\beta$ whose efficiency, when $\mathrm{Cov}(Y) = V_1$, is equal to its efficiency when $\mathrm{Cov}(Y) = V_2$.

Proof: As pointed out earlier, the classes of ALUEs and of maximin-efficient ALUEs for $\beta$ are the same for model (3.2) as for model (3.17) when $\underline{r}(V_1,V_2) = n$. By Lemmas 3.14 and 3.17, all that remains is to show that $\hat{\beta}(c)$ as given in Lemma 3.14 is the same as that given here.

Let $W_Z(c)$ denote the $W(c)$ defined in Lemma 3.14. Then for $c \in [0,1]$,

$$\begin{aligned}\hat{\beta}(c) &= (w'W_Z(c)^+w)^{-1}w'W_Z(c)^+Z \\ &= [x'H^{-\frac{1}{2}}(H^{-\frac{1}{2}}W(c)H^{-\frac{1}{2}})^+H^{-\frac{1}{2}}x]^{-1}x'H^{-\frac{1}{2}}(H^{-\frac{1}{2}}W(c)H^{-\frac{1}{2}})^+H^{-\frac{1}{2}}Y \quad .\end{aligned}$$

When $c \in (0,1)$, $W(c)$ is p.d. and $\hat{\beta}(c)$ reduces to the required expression.

Now for any square matrix $M$ and nonsingular matrix $Q$, $Q^{-1}M^+(Q')^{-1}$ is a generalized inverse (g-inverse) of $Q'MQ$. Furthermore, if $y \in \underline{R}(M)$ then $y'M^-y$ is invariant under any choice of g-inverse $M^-$ for $M$. Letting $y = H^{-\frac{1}{2}}x$ and $M = H^{-\frac{1}{2}}W(c)H^{-\frac{1}{2}}$, we have

$$[x'H^{-\frac{1}{2}}(H^{-\frac{1}{2}}W(c)H^{-\frac{1}{2}})^+H^{-\frac{1}{2}}x] = (x'W(c)^+x)$$

for all $c \in [0,1]$. The required expressions for $\hat{\beta}(0)$ and $\hat{\beta}(1)$ then follow. ∎

Corollary 3.21: Under the conditions of Theorem 3.20, the unique maximin-efficient ALUE for $\beta$ is $\hat{\beta}(c^*)$ where $c^*$ is any value in (0,1) satisfying

$$(x'V_1^+x)(x'W(c^*)^{-1}V_1W(c^*)^{-1}x) = (x'V_2^+x)(x'W(c^*)^{-1}V_2W(c^*)^{-1}x). \quad (3.23)$$

If a BLUE exists for $\beta$ then any $c^* \in (0,1)$ satisfies (3.23) and $\hat{\beta}(c^*)$ is the unique BLUE for $\beta$. Otherwise, $c^* \in (0,1)$ satisfying (3.23) is unique.

Proof: When $c^* \in (0,1)$, it is easy to show that (3.23) is equivalent to (3.22) using the same method as in the proof of Theorem 3.20. The first assertion then follows from Corollaries 3.18 and 3.19.

By Lemma 3.14, a BLUE exists for $\beta$ with respect to model (3.17), hence with respect to model (3.2), if and only if w is an eigenvector of both U and V. The rest follows from Corollary 3.19. ∎

We began this chapter by assuming that $\underline{r}(x,V_1,V_2) = n$. We now remove this assumption. Let $r = \underline{r}(x,V_1,V_2)$ in model (3.2). Then $r \leqslant n$. Let G be any n x r matrix such that $\underline{R}(G) = \underline{R}(x,V_1,V_2)$. Set $Y_G = G'Y$. Then

$$E(Y_G) = g\beta \quad , \quad \mathrm{Cov}(Y_G) = \theta_1 G_1 + \theta_2 G_2 \quad , \quad (3.24)$$

where $g = G'x \neq 0$ , $G_1 = G'V_1G$ and $G_2 = G'V_2G$ are n.n.d., and $\underline{r}(g,G_1,G_2) = r$.

Theorem 3.22: Let $\overline{\mathbb{A}}$ denote the class of ALUEs for $\beta$ and let $\overline{\mathcal{E}}$ denote the class of maximin-efficient ALUEs for $\beta$, both with respect to model (3.2). Let $\overline{\mathbb{A}}_G$ and $\overline{\mathcal{E}}_G$ denote these classes with respect to model (3.24). Then

(a) $\bar{\mathbb{A}} = \bar{\mathbb{A}}_G + \bar{\mathbb{A}}_0$ , and

(b) $\bar{\mathcal{E}} = \bar{\mathcal{E}}_G + \bar{\mathbb{A}}_0$ ,

where $\bar{\mathbb{A}}_0 = \{u'Y : u \in R(x,V_1,V_2)^{\perp}\}$ .

Proof: Part (a) follows from Propositions 3.1.2, 3.1.3, and 3.1.7 of Azzam (1981). Furthermore, by Proposition 3.1.4 of Azzam (1981), $\bar{\mathbb{A}}_G$ is essentially complete with respect to model (3.2), so $\bar{\mathcal{E}}_G \subset \bar{\mathcal{E}}$.

If $t'Y_G \in \bar{\mathcal{E}}_G$ and $u'Y \in \bar{\mathbb{A}}_0$, then $b'Y = t'Y_G + u'Y \in \bar{\mathbb{A}}$ and $\mathrm{var}(b'Y|\theta_1,\theta_2) = \mathrm{var}(t'Y_G|\theta_1,\theta_2)$ for all $\theta_1,\theta_2 \geqslant 0$. The minimum efficiency within $\bar{\mathbb{A}}$ of $b'Y$ is therefore equal to the minimum efficiency within $\bar{\mathbb{A}}$ of $t'Y_G$. But $t'Y_G \in \bar{\mathcal{E}}_G \subset \bar{\mathcal{E}}$, so $b'Y \in \bar{\mathcal{E}}$. We conclude that $\bar{\mathcal{E}}_G + \bar{\mathbb{A}}_0 \subset \bar{\mathcal{E}}$.

Suppose $b'Y \in \bar{\mathcal{E}}$. Then $b'Y \in \bar{\mathbb{A}}$ and by part (a) $b'Y = t'Y_G + u'Y$ where $t'Y_G \in \bar{\mathbb{A}}_G \subset \bar{\mathbb{A}}$ and $u'Y \in \bar{\mathbb{A}}_0$. Now $\mathrm{var}(b'Y|\theta_1,\theta_2) = \mathrm{var}(t'Y_G|\theta_1,\theta_2)$ for all $\theta_1,\theta_2 \geqslant 0$, so the minimum efficiency of $b'Y$ within $\bar{\mathbb{A}}$ is the same as the minimum efficiency of $t'Y_G$ within $\bar{\mathbb{A}}$. Since $\bar{\mathbb{A}}_G$ is essentially complete with respect to model (3.2), this is also the minimum efficiency of $t'Y_G$ within $\bar{\mathbb{A}}_G$. Then, since $b'Y \in \bar{\mathcal{E}}$, we have $t'Y_G \in \bar{\mathcal{E}}_G$. Otherwise, there would exist an $s'Y_G \in \bar{\mathbb{A}}_G \subset \bar{\mathbb{A}}$ with greater minimum efficiency within $\bar{\mathbb{A}}_G$ than $t'Y_G$, hence with greater minimum efficiency within $\bar{\mathbb{A}}$ than $b'Y$, a contradiction. Hence $\bar{\mathcal{E}} \subset \bar{\mathcal{E}}_G + \bar{\mathbb{A}}_0$. ∎

Theorem 3.22 says that we can obtain $\overline{\mathcal{E}}$ by applying the results of this chapter to model (3.24) to obtain $\overline{\mathcal{E}}_G$ and then use the relationship $\overline{\mathcal{E}} = \overline{\mathcal{E}}_G + \overline{\mathbb{A}}_0$.

# IV. A VECTOR OF FIXED PARAMETERS

In the previous chapter we obtained the class of maximin-efficient ALUEs for a single fixed parameter in a mixed linear model. We now apply this result to the mixed linear model with a vector of fixed parameters. We shall obtain the class of maximin-efficient ALUEs for any estimable linear parametric function of the fixed parameters.

We start with the model

$$Y = X\beta + B_1 b_1 + B_2 b_2 \tag{4.1}$$

where Y, $B_1$, $b_1$, $B_2$, and $b_2$ are the same as in model (3.1), but where X is a nonzero known n x p matrix and $\beta$ is a p-vector of fixed but unknown parameters. Here $1 \leqslant p \leqslant n$. The first and second moment assumptions are

$$E(Y) = X\beta \quad , \quad Cov(Y) = \theta_1 V_1 + \theta_2 V_2 \quad , \tag{4.2}$$

where $V_1 = B_1 B_1'$ and $V_2 = B_2 B_2'$ are n x n and n.n.d.. We shall denote by $\overline{\mathfrak{L}}_\pi$, $\overline{\mathfrak{A}}_\pi$, and $\overline{\mathfrak{E}}_\pi$ the classes of LUEs, ALUEs, and maximin-efficient ALUEs for an estimable parametric function $\pi'\beta$.

Let $\mathfrak{M}$ denote the expectation space for model (4.2): $\mathfrak{M} = \{E(Y):\beta \in R^p\} = \underline{R}(X)$, and let $m = \dim(\mathfrak{M}) = \underline{r}(X)$. Then $0 < m \leqslant p$. From Theorem 1 of Seely (1970a), m also equals the dimension of $\mathfrak{G}$, the space of estimable linear parametric functions of $\beta$.

Let $\Lambda'\beta$ denote an estimable parametric m-vector having $\pi'\beta$ as its m-th component and satisfying $\underline{r}(\Lambda) = m$. That is, the components of $\Lambda'\beta$ form a basis for $\mathfrak{G}$. Let

$$U = X\Lambda(\Lambda'\Lambda)^{-1} \quad .$$

Then, from Proposition 5.1 of Seely (1979),

$$E(Y) = U\alpha \quad , \quad \alpha \in R^m \tag{4.3}$$

is a full column rank parametrization for $\mathfrak{M}$ such that $\alpha \doteq \Lambda'\beta$, where $\doteq$ denotes the correspondence relation described there. Actually, Seely (1979) assumed the special case $\mathrm{Cov}(Y) = \sigma^2 I$, $\sigma^2 > 0$, but the results we shall use hold here as well.

Now $\alpha \doteq \Lambda'\beta$ implies that the components of $\alpha$ are estimable with respect to parametrization (4.3) and that the classes of LUEs for corresponding components of $\alpha$ and $\Lambda'\beta$ are the same. Hence, so are the classes of ALUEs and maximin-efficient ALUEs.

Estimating $\pi'\beta$ with respect to model (4.2) is, therefore, equivalent to estimating $\alpha_m$ under parametrization (4.3). So, without loss in generality, we can consider the problem of estimating $\beta_p$ in model (4.2).

The case $p = 1$ was discussed in the previous chapter, so we assume that $1 < p \leq n$ and that $\beta_p$ is estimable with respect to model (4.2).

We partition X and $\beta$ according to

$$X = [X_1, x_p] \text{ and } \beta = (\beta_1', \beta_p)' \quad , \tag{4.4}$$

where $X_1$ is an n x (p - 1) matrix with $\underline{r}(X_1) = r_1$, $x_p$ is an n-vector such that $x_p \notin \underline{R}(X_1)$, and $\beta_1$ is a (p - 1)-vector. Then $X\beta = X_1\beta_1 + x_p\beta_p$.

Let H be an n x (n - $r_1$) matrix with columns forming a basis for $\underline{R}(X_1)^{\perp}$. Set $Y_H = H'Y$. Then we have the reduced model

$$Y_H = hx_p + H'B_1b_1 + H'B_2b_2 \tag{4.5}$$

or

$$E(Y_H) = hx_p \quad , \quad Cov\ (Y_H) = \theta_1 H_1 + \theta_2 H_2 \tag{4.6}$$

where $h = H'x_p \neq 0$ and both $H_1 = H'V_1H$ and $H_2 = H'V_2H$ are n.n.d. The reduced models (4.5) and (4.6) have the same form as models (3.1) and (3.2), respectively. Here $Y_H$, h, $\beta_p$, $H_1$, $H_2$, and n - $r_1$ replace Y, x, $\beta$, $V_1$, $V_2$, and n, respectively.

We can apply the results of the previous chapter to obtain the class of maximin-efficient ALUEs for $\beta_p$ with respect to reduced model (4.6). The following lemma says that this will also be the class of maximin-efficient ALUEs for $\beta_p$ with respect to model (4.2).

Lemma 4.1: Consider the model

$$E(Y) = X\beta \quad , \quad Cov(Y) \in \mathcal{V} \quad , \tag{4.7}$$

where $X = [X_1, X_2]$, $\beta = (\beta_1', \beta_2')'$, and $\mathcal{V}$ is a set of n.n.d. matrices. Let H denote a matrix with columns forming a basis for $\underline{R}(X_1)^{\perp}$ and set $Y_H = H'Y$. Then

$$E(Y_H) = H'X\beta = H'X_2\beta_2 \quad , \quad \mathrm{Cov}(Y_H) \in \mathcal{V}_H \qquad , \tag{4.8}$$

where $\mathcal{V}_H = \{H'VH : V \in \mathcal{V}\}$. Let $\bar{\mathfrak{L}}_\pi$, $\bar{\mathfrak{A}}_\pi$, and $\bar{\mathfrak{E}}_\pi$ denote the classes of LUEs, ALUEs, and maximin-efficient ALUEs for the parametric function $\pi'\beta_2$ with respect to model (4.7). Let $\bar{\mathfrak{L}}^*_\pi$, $\bar{\mathfrak{A}}^*_\pi$, and $\bar{\mathfrak{E}}^*_\pi$ denote those classes with respect to reduced model (4.8). Then

(a) $\bar{\mathfrak{L}}_\pi = \bar{\mathfrak{L}}^*_\pi$

(b) $\bar{\mathfrak{A}}_\pi = \bar{\mathfrak{A}}^*_\pi$

(c) $\bar{\mathfrak{E}}_\pi = \bar{\mathfrak{E}}^*_\pi$ .

<u>Proof</u>: $\bar{\mathfrak{L}}_\pi = \{t'Y : E(t'Y) = \pi'\beta_2\}$

$= \{t'Y : X_1't = 0 \quad , \quad X_2't = \pi\}$

$\bar{\mathfrak{L}}^*_\pi = \{s'Y_H : E(s'Y_H) = \pi'\beta_2\}$

$= \{s'Y_H : X_2'Hs = \pi\}$

Suppose $s'Y_H \in \bar{\mathfrak{L}}^*_\pi$. Then $s'Y_H = t'Y$ where $t = Hs$. We have $X_1't = X_1'Hs = 0$ and $X_2't = X_2'Hs = \pi$, so $s'Y_H \in \bar{\mathfrak{L}}_\pi$ and $\bar{\mathfrak{L}}^*_\pi \subset \bar{\mathfrak{L}}_\pi$. Note that this also holds if $\bar{\mathfrak{L}}^*_\pi$ is empty.

Suppose now that $t'Y \in \bar{\mathfrak{L}}_\pi$. Then $X_1't = 0$, so $t \in \underline{N}(X_1') = \underline{R}(X_1)^\perp = \underline{R}(H)$, and $t = Hs$ for some $s$. Furthermore, $\pi = X_2't = X_2'Hs$, so $t'Y = s'Y_H$ where $X_2'Hs = \pi$. Then $t'Y \in \bar{\mathfrak{L}}^*_\pi$ and $\bar{\mathfrak{L}}_\pi \subset \bar{\mathfrak{L}}^*_\pi$. This also holds if $\bar{\mathfrak{L}}_\pi$ is empty. This shows part (a).

If either $\bar{\mathfrak{L}}_\pi = \phi$ or $\bar{\mathfrak{L}}^*_\pi = \phi$, then the other is also empty and $\bar{\mathfrak{A}}_\pi = \bar{\mathfrak{A}}^*_\pi = \bar{\mathfrak{E}}_\pi = \bar{\mathfrak{E}}^*_\pi = \phi$. Let $t'Y = s'Y_H$ be unbiased for $\pi'\beta_2$. Let $\mathrm{Cov}(Y) = V \in \mathcal{V}$. Then $\mathrm{Cov}(Y_H) = H'VH \in \mathcal{V}_H$, and

$$\mathrm{Cov}(s'Y_H) = s'H'VHs = t'Vt = \mathrm{Cov}(t'Y) \quad .$$

Since $\bar{\mathcal{L}}_\pi = \bar{\mathcal{L}}^*_\pi$ and a LUE for $\pi'\beta_2$ has the same variance under both model (4.7) and (4.8), $\bar{\mathbb{A}}_\pi = \bar{\mathbb{A}}^*_\pi$ and $\bar{\mathcal{E}}_\pi = \bar{\mathcal{E}}^*_\pi$. ▮

One approach to obtaining the class of maximin-efficient ALUEs for $\beta_p$ with respect to model (4.2) is, therefore, to transform model (4.2) to model (4.6) and to apply our previous results to that model. The class of maximin-efficient ALUEs for $\beta_p$ with respect to reduced model (4.6) will also be the class of maximin-efficient ALUEs for $\beta_p$ with respect to model (4.2).

We now wish to characterize the class of maximin-efficient ALUEs for $\beta_p$ with respect to model (4.2) without explicitly performing a transformation to model (4.6). We initially assume that $\underline{r}(X,V_1,V_2) = n$ but later drop this assumption.

Lemma 4.2: Let X be partitioned as in (4.4) and let H be defined as for reduced model (4.6). If $\underline{r}(X,V_1,V_2) = n$ then $\underline{r}(h,H_1,H_2) = n - r_1$.

Proof: 
$$\begin{aligned}\underline{r}(h,H_1,H_2) &= \underline{r}(H'x_p,H'V_1H,H'V_2H)\\ &= \underline{r}(H'(X,V_1,V_2))\\ &= \underline{r}(H')\\ &= n - r_1 \quad .\end{aligned}$$ ▮

Consider the artificial model

$$Y = X\beta + B_1\delta_1 + B_2\delta_2 + e \quad , \tag{4.9}$$

where Y, X, $B_1$, and $B_2$ are the same as in model (4.1), but where $\delta_1$ and $\delta_2$ are $q_1$- and $q_2$-vectors of fixed but unknown parameters and e is an n-vector of random errors such that

$$E(e) = 0 \quad , \quad Cov(e) = I_n \quad .$$

Suppose $\beta_p$ is estimable with respect to artificial model (4.9). Then by Theorem 3.3 $x_p \notin \underline{R}(X_1,B_1,B_2)$. Let H be an n x $(n - r_1)$ matrix with orthonormal columns forming a basis for $\underline{R}(X_1)^{\perp}$. Set $Y_H = H'Y$. Then

$$Y_h = h\beta_p + H'B_1\delta_1 + H'B_2\delta_2 + f \tag{4.10}$$

where $h = H'x_p \neq 0$ and $f = H'e$ is a $(n - r_1)$-vector of random errors such that

$$E(f) = 0 \quad , \quad Cov(f) = H'H = I_{n-r_1} \quad .$$

Model (4.10) has the same form as model (3.11). Now $h = H'x_p \in \underline{R}(H'B_1,H'B_2) \Rightarrow H'x_p = H'(B_1,B_2)\alpha$ for some $\alpha \Rightarrow H'(x_p - (B_1,B_2)\alpha) = 0 \Rightarrow x_p - (B_1,B_2)\alpha \in \underline{N}(H') = \underline{R}(X_1)$ $\Rightarrow x_p \in \underline{R}(X_1,B_1,B_2)$. But $x_p \notin \underline{R}(X_1,B_1,B_2)$, so $h \notin \underline{R}(H'B_1,H'B_2)$ and $\beta_p$ is estimable with respect to artificial model (4.10) by Theorem 3.3.

By Lemma 4.2 and Corollary 3.4, the unique BLUE for $\beta_p$ with respect to reduced model (4.6) is the unique BLUE for $\beta_p$ with respect to artificial model (4.10). Applying Lemma 4.1 twice, this estimator is also the unique BLUE for $\beta_p$ with respect to models (4.2) and (4.9). We have shown:

Theorem 4.3: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$. If $\beta_p$ is estimable with respect to artificial model (4.9), then there exists a unique BLUE for $\beta_p$ with respect to that artificial model and it is also the unique BLUE for $\beta_p$ with respect to model (4.2).

We introduce three additional artificial models:

$$E(Y) = X\beta + B_1\delta_1 \quad , \quad \mathrm{Cov}(Y) = V_2 \quad , \tag{4.11}$$

$$E(Y) = X\beta + B_2\delta_2 \quad , \quad \mathrm{Cov}(Y) = V_1 \quad , \text{ and} \tag{4.12}$$

$$E(Y) = X\beta \quad , \quad \mathrm{Cov}(Y) = cV_1 + (1 - c)V_2 \quad , \tag{4.13}$$

where $c \in (0,1)$, $Y$, $X$, $\beta$, $B_1$, and $B_2$ are the same as in model (4.1), $V_1$ and $V_2$ are the same as in model (4.2), but $\delta_1$ and $\delta_2$ are $q_1$- and $q_2$-vectors of fixed but unknown parameters.

Theorem 4.4: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$. If $\beta_p$ is estimable with respect to artificial models (4.11) and (4.12), but is nonestimable with respect to artificial model (4.9), then the class of maximin-efficient ALUEs for $\beta$ and the class of ALUEs for $\beta_p$, both with respect to model (4.2), are the same and are given by

$$\overline{\mathcal{E}} = \overline{\mathcal{A}} = \overline{\mathfrak{B}}(V_1,V_2) \cup \overline{\mathfrak{B}}(V_2,V_1) \cup \bigcup_{0<c<1} \overline{\mathfrak{B}}(cV_1 + (1 - c)V_2),$$

where $\overline{\mathfrak{B}}(V_1,V_2)$, $\overline{\mathfrak{B}}(V_2,V_1)$, and $\overline{\mathfrak{B}}(cV_1 + (1 - c)V_2)$ consist of the unique BLUE for $\beta_p$ with respect to artificial models (4.11), (4.12), and (4.13), respectively.

Proof: Let H be defined as for reduced model (4.6). Then from artificial models (4.11), (4.12), and (4.13) we obtain the reduced artificial models

$$E(Y_H) = h\beta_p + H'B_1\delta_1 \quad , \quad Cov(Y_H) = H_2 \quad , \tag{4.14}$$

$$E(Y_H) = h\beta_p + H'B_2\delta_2 \quad , \quad Cov(Y_H) = H_1 \quad , \text{ and} \tag{4.15}$$

$$E(Y_H) = h\beta_p \quad , \quad Cov(Y_H) = cH_1 + (1 - c)H_2 \; , \tag{4.16}$$

where $Y_H = H'Y$, $h = H'x_p$, $H_1 = H'V_1H$, and $H_2 = H'V_2H$. By Lemma 4.1, the classes of LUEs, ALUEs, and maximin-efficient ALUEs for $\beta_p$ with respect to model (4.2) and artificial models (4.9), (4.11), (4.12), and (4.13) are the same as those for reduced model (4.6) and reduced artificial models (4.10), (4.14), (4.15), and (4.16), respectively.

Theorem 3.8 applies to reduced model (4.6) and reduced artificial models (4.10), (4.14), (4.15), and (4.16) by Lemma 4.2. The statements of this theorem therefore hold with respect to these reduced models. The theorem then follows by repeatedly applying Lemma 4.1. ∎

Theorem 4.5: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$. If $\beta_p$ is estimable with respect to artificial model (4.11) but is nonestimable with respect to artificial model (4.12), then the unique BLUE for $\beta_p$ with respect to artificial model (4.11) is the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2).

Proof: The theorem follows from the proof of Theorem 4.4 by using Theorems 3.12 and 3.13 in place of Theorem 3.8. ∎

Note that the roles of artificial models (4.11) and (4.12) may be reversed in Theorem 4.5.

Theorem 4.6: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$. If $\beta_p$ is nonestimable with respect to artificial models (4.11) and (4.12) then the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2) is the unique BLUE for $\beta_p$ with respect to artificial model (4.13) with $c = c^*$, where $c^*$ is any value in (0,1) such that this LUE has equal efficiencies with respect to model (4.2) when $\mathrm{Cov}(Y) = V_1$ and $\mathrm{Cov}(Y) = V_2$.

Proof: Suppose $\beta_p$ is nonestimable with respect to artificial models (4.11) and (4.12). By Lemma 4.1, $\beta_p$ is nonestimable with respect to reduced artificial models (4.14) and (4.15), so $h \in \underline{R}(H_1) \cap \underline{R}(H_2)$ by Theorem 3.3 and $\underline{r}(H_1,H_2) = n - r_1$ by Lemma 4.2. This theorem follows by applying Theorem 3.20 and Corollary 3.21 to reduced artificial model (4.16) and then applying Lemma 4.1. ∎

Corollary 4.7: Under the conditions of Theorem 4.6, if a BLUE for $\beta_p$ with respect to model (4.2) exists, then it is the unique BLUE for $\beta_p$ with respect to artificial model (4.13) for any $c \in (0,1)$. Otherwise the value of $c^* \in (0,1)$ satisfying Theorem 4.6 is unique.

Proof: Apply Lemmas 4.1 and 4.2 and Corollary 3.21 ▮

We now drop the assumption that $\underline{r}(X,V_1,V_2) = n$. Let $r = \underline{r}(X,V_1,V_2)$ in model (4.2). Then $r \leqslant n$. Let G be any $n \times r$ matrix such that $\underline{R}(G) = \underline{R}(X,V_1,V_2)$. Set $Y_G = G'Y$. Then

$$E(Y_G) = X_G\beta \quad , \quad Cov(Y_G) = \theta_1 G_1 + \theta_2 G_2 \quad ,$$

where $X_G = G'X$ is nonzero, $G_1 = G'V_1G$ and $G_2 = G'V_2G$ are n.n.d., and $\underline{r}(X_G,G_1,G_2) = r$.

Theorem 4.8: Let $\bar{\mathbb{A}}$ denote the class of ALUEs for $\beta_p$ and let $\bar{\mathcal{E}}$ denote the class of maximin-efficient ALUEs for $\beta_p$, both with respect to model (4.2). Let $\bar{\mathbb{A}}_G$ and $\bar{\mathcal{E}}_G$ denote these classes with respect to reduced model (4.17). Then

(a) $\bar{\mathbb{A}} = \bar{\mathbb{A}}_G + \bar{\mathbb{A}}_0$ , and

(b) $\bar{\mathcal{E}} = \bar{\mathcal{E}}_G + \bar{\mathbb{A}}_0$ ,

where $\bar{\mathbb{A}}_0 = \{u'Y : u \in \underline{R}(X,V_1,V_2)^{\perp}\}$.

Proof: The proof is identical to that of Theorem 3.22. ▮

Now suppose $\theta = \theta_1 + \theta_2 > 0$ and it is known that $\gamma = \theta_1/\theta \in [\gamma_L,\gamma_U] \subset [0,1]$ where $\gamma_L \leqslant \gamma_U$. If $\gamma_L = \gamma_U$, then $\gamma = \gamma_L = \gamma_U$ and $(1 - \gamma_L)\theta_1 = \gamma_L\theta_2$. In this case, we can rewrite model (4.2) as

$$E(Y) = X\beta \quad , \quad Cov(Y) = \theta V \quad , \tag{4.18}$$

for some unknown $\theta > 0$ and known n.n.d. matrix V. There then exists a BLUE for every $\pi'\beta$ that is estimable with respect to model (4.18). We therefore suppose that $\gamma_L < \gamma_U$.

When $\gamma_L < \gamma_U$, we have

$$\theta_1 \geqslant \frac{\gamma_L}{1-\gamma_L}\theta_2 \quad \text{and} \quad \theta_2 \geqslant \frac{1-\gamma_U}{\gamma_U}\theta_1 \qquad .$$

Define

$$\psi_1 = \frac{1-\gamma_L}{\gamma_U-\gamma_L}\left[\theta_1 - \frac{\gamma_L}{1-\gamma_L}\theta_2\right] \geqslant 0 \qquad ,$$

$$\psi_2 = \frac{\gamma_U}{\gamma_U-\gamma_L}\left[\theta_2 - \frac{1-\gamma_U}{\gamma_U}\theta_1\right] \geqslant 0 \qquad .$$

Then

$$\theta_1 = \gamma_U\psi_1 + \gamma_L\psi_2 \qquad ,$$

$$\theta_2 = (1 - \gamma_U)\psi_1 + (1 - \gamma_L)\psi_2 \qquad ,$$

$$\theta_1V_1 + \theta_2V_2 = \psi_1W(\gamma_U) + \psi_2W(\gamma_L) \qquad ,$$

where $W(\gamma) = \gamma V_1 + (1 - \gamma)V_2$ .

Note that

$$\theta_1 = \frac{\gamma_L}{1-\gamma_L}\theta_2 + \frac{\gamma_U-\gamma_L}{1-\gamma_L}\psi_1 \qquad ,$$

$$\theta_2 = \frac{1-\gamma_U}{\gamma_U}\theta_1 + \frac{\gamma_U-\gamma_L}{\gamma_U}\psi_2 \qquad ,$$

so $\gamma \in [\gamma_L, \gamma_U]$ for all $\psi_1, \psi_2 \geqslant 0$. Furthermore, $\psi_1 + \psi_2 = 0$ implies

$$\theta_1 = \frac{\gamma_L}{1-\gamma_L}\,\theta_2 = \frac{\gamma_L}{1-\gamma_L}\;\frac{1-\gamma_U}{\gamma_U}\,\theta_1 \qquad ,$$

$$\theta_2 = \frac{1-\gamma_U}{\gamma_U}\,\theta_1 = \frac{1-\gamma_U}{\gamma_U}\;\frac{\gamma_L}{1-\gamma_L}\,\theta_2 \qquad .$$

Since not both $\theta_1 = 0$ and $\theta_2 = 0$, we must have

$$\frac{\gamma_L}{1-\gamma_L} \cdot \frac{1-\gamma_U}{\gamma_U} = 1 \Rightarrow \gamma_L = \gamma_U \quad ,$$

a contradiction. Hence $\psi_1 + \psi_2 > 0$.

We have therefore shown that if $\gamma \in [\gamma_L, \gamma_U] \subset [0,1]$ where $\gamma_L < \gamma_U$, then we can rewrite model (4.2) as

$$E(Y) = X\beta \quad , \quad Cov(Y) = \psi_1 W(\gamma_U) + \psi_2 W(\gamma_L) \tag{4.19}$$

where $\psi_1 \geqslant 0$ and $\psi_2 \geqslant 0$ are unknown, $\psi_1 + \psi_2 > 0$, and both $W(\gamma_U)$ and $W(\gamma_L)$ are known n.n.d. matrices. Therefore this case can be handled using our previous results.

Suppose it is known that $\gamma \in [\gamma_L, 1]$, where $\gamma_L \in (0,1)$. Then

$$W(\gamma_L) = \gamma_L V_1 + (1 - \gamma_L)V_2 \quad \text{and} \quad W(1) = V_1 \quad .$$

So $\underline{R}(W(1)) \subset \underline{R}(W(\gamma_L)) = \underline{R}(V_1,V_2)$, and Theorem 4.4 cannot apply in this case. However, Theorems 4.3, 4.5, and 4.6 may. We restate them below for this case as corollaries:

<u>Corollary 4.9</u>: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$, $\gamma \in [\gamma_L, 1]$, and $\gamma_L \in (0,1)$. If $\beta_p$ is estimable with respect to artificial model (4.9), then the unique BLUE for $\beta_p$ with respect to that artificial model is also the unique BLUE for $\beta_p$ with respect to model (4.2)

<u>Corollary 4.10</u>: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$, $\gamma \in [\gamma_L, 1]$, and $\gamma_L \in (0,1)$. Suppose $\beta_p$ is nonestimable with respect to artificial model (4.9), but $\beta_p$ is estimable with

respect to the artificial model

$$E(Y) = X\beta + B_1\delta_1 \quad , \quad Cov(Y) = W(\gamma_L) \qquad (4.20)$$

where $\delta_1$ is a vector of fixed but unknown parameters and $W(\gamma_L) = \gamma_L V_1 + (1 - \gamma_L)V_2$. Then the unique BLUE for $\beta_p$ with respect to this artificial model is the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2).

Corollary 4.11: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$, $\gamma \in [\gamma_L,1]$, and $\gamma_L \in (0,1)$. If $\beta_p$ is nonestimable with respect to artificial model (4.20), then the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2) when $\gamma \in [\gamma_L,1]$ is the unique BLUE for $\beta_p$ with respect to artificial model (4.13) when $c = c^*$, where $c^*$ is any value in $(\gamma_L,1)$ such that this LUE has equal efficiencies with respect to model (4.2) when $Cov(Y) = W(1) = V_1$ and $Cov(Y) = W(\gamma_L)$.

Corollary 4.11 follows because

$$\begin{aligned}&\{cW(\gamma_U) + (1 - c)W(\gamma_L):c \in (0,1)\}\\ &\qquad = \{dV_1 + (1 - d)V_2:d \in (\gamma_L,\gamma_U)\} \quad .\end{aligned}$$

The roles of $V_1$ and $V_2$ may be reversed in Corollaries 4.9-4.11 to handle the case where $\gamma$ is known to lie in $[0,\gamma_U]$ and $\gamma_U \in (0,1)$.

Suppose, finally, that $\gamma$ is known to lie in $[\gamma_L,\gamma_U]$, a non-empty subset of (0,1). Then $\underline{R}(W(\gamma_U)) = \underline{R}(W(\gamma_L)) = \underline{R}(V_1,V_2)$. Only Theorems 4.3 and 4.6 can apply in this case. Restating these theorems, we obtain two more corollaries:

Corollary 4.12: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$ and $\gamma \in [\gamma_L,\gamma_U] \subset (0,1)$. If $\beta_p$ is estimable with respect to artificial model (4.9), then the unique BLUE for $\beta_p$ with respect to that artificial model is the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2).

Corollary 4.13: Consider model (4.2) where $\underline{r}(X,V_1,V_2) = n$ and $\gamma \in [\gamma_L,\gamma_U] \subset (0,1)$. If $\beta_p$ is nonestimable with respect to artificial model (4.9), then the unique maximin-efficient ALUE for $\beta_p$ with respect to model (4.2) is the unique BLUE for $\beta_p$ with respect to artificial model (4.13) when $c = c^*$, where $c^*$ is any value in $(\gamma_L,\gamma_U)$ such that this LUE has equal efficiencies with respect to model (4.2) when $Cov(Y) = W(\gamma_U)$ and $Cov(Y) = W(\gamma_L)$.

# V. EXAMPLES: RANDOM NESTED WITHIN FIXED

In this chapter we consider the special case of model (1.1),

$$Y = X\beta + B_1b_1 + B_2b_2 \quad , \tag{5.1}$$

where $\underline{R}(X) \subset \underline{R}(B_1) \cap \underline{R}(B_2)$ and $\underline{r}(B_1,B_2) = n$. In the context of the analysis of variance we say that the random effects are nested within the fixed effects. The first and second moment assumptions are given by (1.2),

$$E(Y) = X\beta \quad , \quad Cov(Y) = \theta_1V_1 + \theta_2V_2 \quad , \tag{5.2}$$

where $V_1 = B_1B_1'$ and $V_2 = B_2B_2'$.

We wish to estimate an estimable parametric function $\pi'\beta$. As pointed out in Chapter IV, we can without loss of generality assume that $\pi'\beta = \beta_p$, the p-th component of $\beta$. Then $\underline{r}(V_1,V_2) = n$, $\underline{R}(X) \subset \underline{R}(V_1) \cap \underline{R}(V_2)$, and Theorem 4.6 and Corollary 4.7 apply. Restating them in terms of an estimable $\pi'\beta$ with respect to model (5.2) we have the following two corollaries:

<u>Corollary 5.1</u>: Let $\pi'\beta$ denote an estimable parametric function with respect to model (5.2) where $\underline{R}(X) \subset \underline{R}(V_1) \cap \underline{R}(V_2)$ and $\underline{r}(V_1,V_2) = n$. Then the unique maximin-efficient ALUE for $\pi'\beta$ is the unique BLUE for $\pi'\beta$ when $Cov(Y) = c^*V_1 + (1 - c^*)V_2$, where $c^*$ is any value in $(0,1)$ such that this LUE has equal

efficiencies with respect to model (5.2) when $\mathrm{Cov}(Y) = V_1$ and $\mathrm{Cov}(Y) = V_2$.

Corollary 5.2: Under the conditions of Corollary 5.1, if a BLUE for $\pi'\beta$ exists then any value of $c^*$ in (0,1) will satisfy the condition of equal efficiencies. Otherwise, the value of $c^* \in (0,1)$ satisfying that condition is unique.

Now we apply these results to a few examples. The purpose of these examples is to obtain efficiency curves of selected ALUEs and maximin-efficient ALUEs to illustrate some of the concepts introduced earlier. The numerical results and plotted efficiency curves were obtained using a Hewlett-Packard HP-85 personal computer and accessories.

Example 5.1

Birkes, Seely, and Azzam (1981) discuss the estimation of the mean $\mu$ in the random one-way model

$$Y_{ij} = \mu + a_i + e_{ij} \quad , \quad i = 1,\ldots,t \quad , \quad j = 1,\ldots,n_i \quad ,$$

where the $a_i$ and the $e_{ij}$ are uncorrelated random variables having means zero and variances $\theta_1$ and $\theta_2$, respectively. They show that within the class of LUEs for $\mu$ the class of LUEs of the form $\Sigma c_i \overline{Y}_{i.}$ constitutes a complete class, where $\overline{Y}_{i.}$ denotes the mean for group i, i = 1,...,t.

Let $Y = (\overline{Y}_{1.},\ldots,\overline{Y}_{t.})'$. Then

$$E(Y) = \underline{1}_t\mu \quad , \quad \mathrm{Cov}(Y) = \theta_1 I_t + \theta_2 D \quad , \tag{5.3}$$

where $\underline{1}_t$ is the t-vector of ones, $I_t$ is the t x t identity matrix, and D is the t x t diagonal matrix with diagonal elements $1/n_i$. This is a special case of model (5.2).

Birkes, Seely, and Azzam (1981) characterized the maximin-efficient ALUE for $\mu$ with respect to model (5.3). They then considered an example by Fertig and Mann (1974, p. 221) of groups of strength measurements of varying sizes obtained from 11 different batches of a material. The summary statistics are given in Table 1. They calculated the maximin-efficient ALUE $\hat{\mu}_*$ and compared its efficiency with the efficiencies of three other ALUEs for $\mu$ as a function of the intraclass correlation coefficient $\gamma = \theta_1/\theta$, where $\theta = \theta_1 + \theta_2$.

These calculations were repeated for this thesis using BASIC programs designed for the more general model (5.2). The calculated values for $c^*$, $\hat{\mu}_*$, the overall mean $\hat{\mu}_0(c = 0)$, the unweighted mean of the group means $\hat{\mu}_1(c = 1)$, and their efficiencies at $\gamma = 0,.1,\ldots,1$ all agree with those of Birkes, Seely, and Azzam. These values are summarized in Table 2 and Figure 1.

The minimum efficiency of all ALUEs for $\mu$ is the lesser of

$$\mathrm{Eff}(0|1) = n.^2 \;/\; (t\Sigma n_i^2) \doteq .588$$

and

$$\mathrm{Eff}(1|0) = t^2 \;/\; (n.\Sigma 1/n_i) \doteq .647 \qquad ,$$

where $n. = \Sigma n_i$. The efficiency of $\hat{\mu}_*$ exceeds .888 for all $\gamma \in [0,1]$.

Table 1. Strength Measurements on 11 Batches of Material (Reproduced from Birkes, Seely, and Azzam, 1981).

| Batch Number | Size | Sample Mean | Sample Standard Deviation |
|---|---|---|---|
| 1 | 28 | 193.75 | 4.59 |
| 2 | 18 | 202.22 | 2.61 |
| 3 | 27 | 189.78 | 1.48 |
| 4 | 101 | 186.19 | 2.15 |
| 5 | 7 | 194.14 | 2.91 |
| 6 | 13 | 195.62 | 2.36 |
| 7 | 34 | 201.15 | 2.15 |
| 8 | 21 | 199.64 | 2.89 |
| 9 | 12 | 197.20 | 4.11 |
| 10 | 28 | 197.35 | 4.44 |
| 11 | 19 | 196.07 | 4.83 |

Table 2. ALUEs for $\mu$ in Example 5.1.

| Estimator | c | Value |
|---|---|---|
| $\hat{\mu}_0$ | 0 | 193.42 |
| $\hat{\mu}_*(\gamma \in [0,1])$ | .02873 | 195.01 |
| $\hat{\mu}_*(\gamma \in [.1,1])$ | .22191 | 195.64 |
| $\hat{\mu}_1$ | 1 | 195.74 |

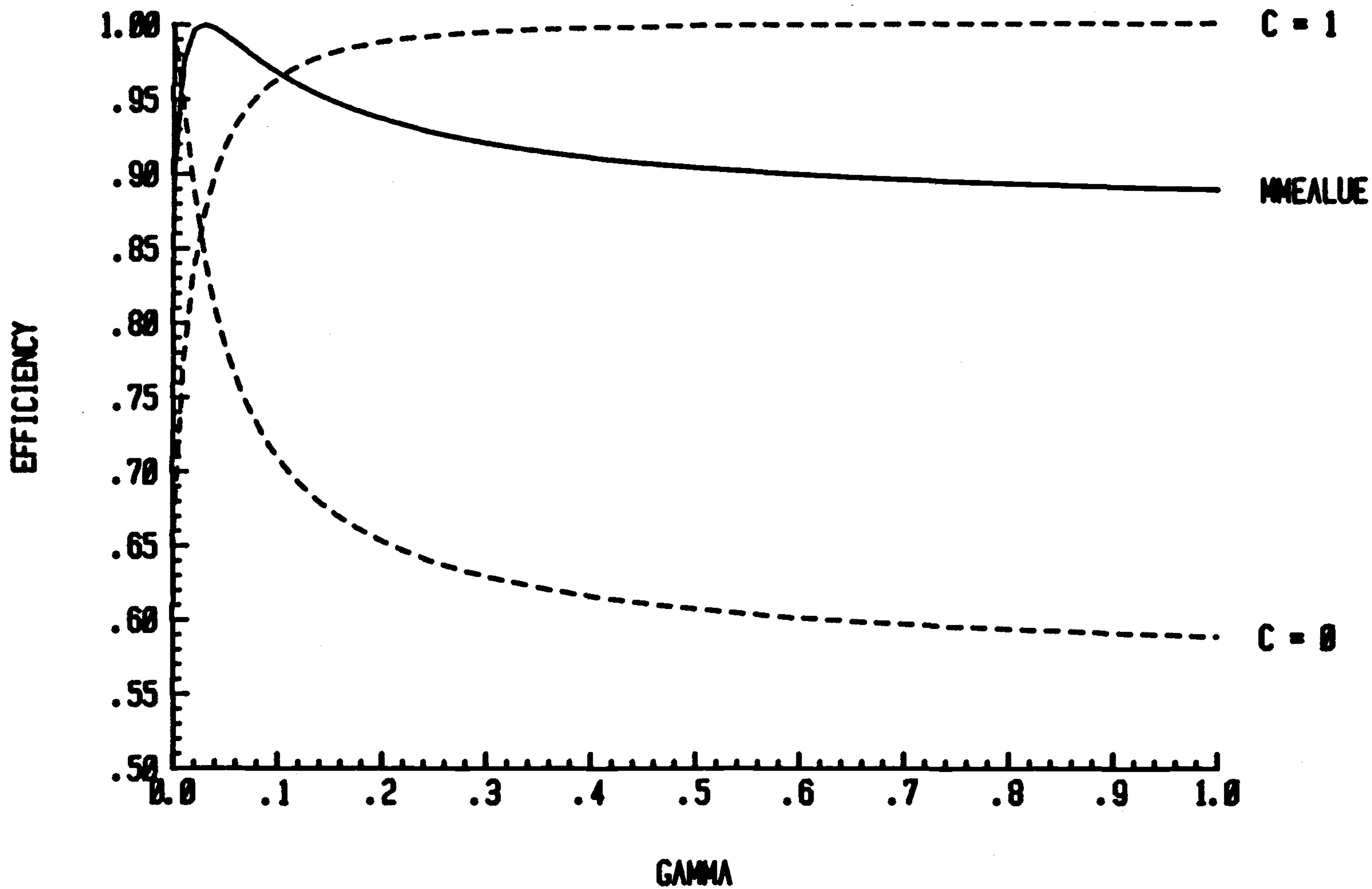


Figure 1. Efficiency curves for selected ALUEs for $\mu$ in example 5.1.

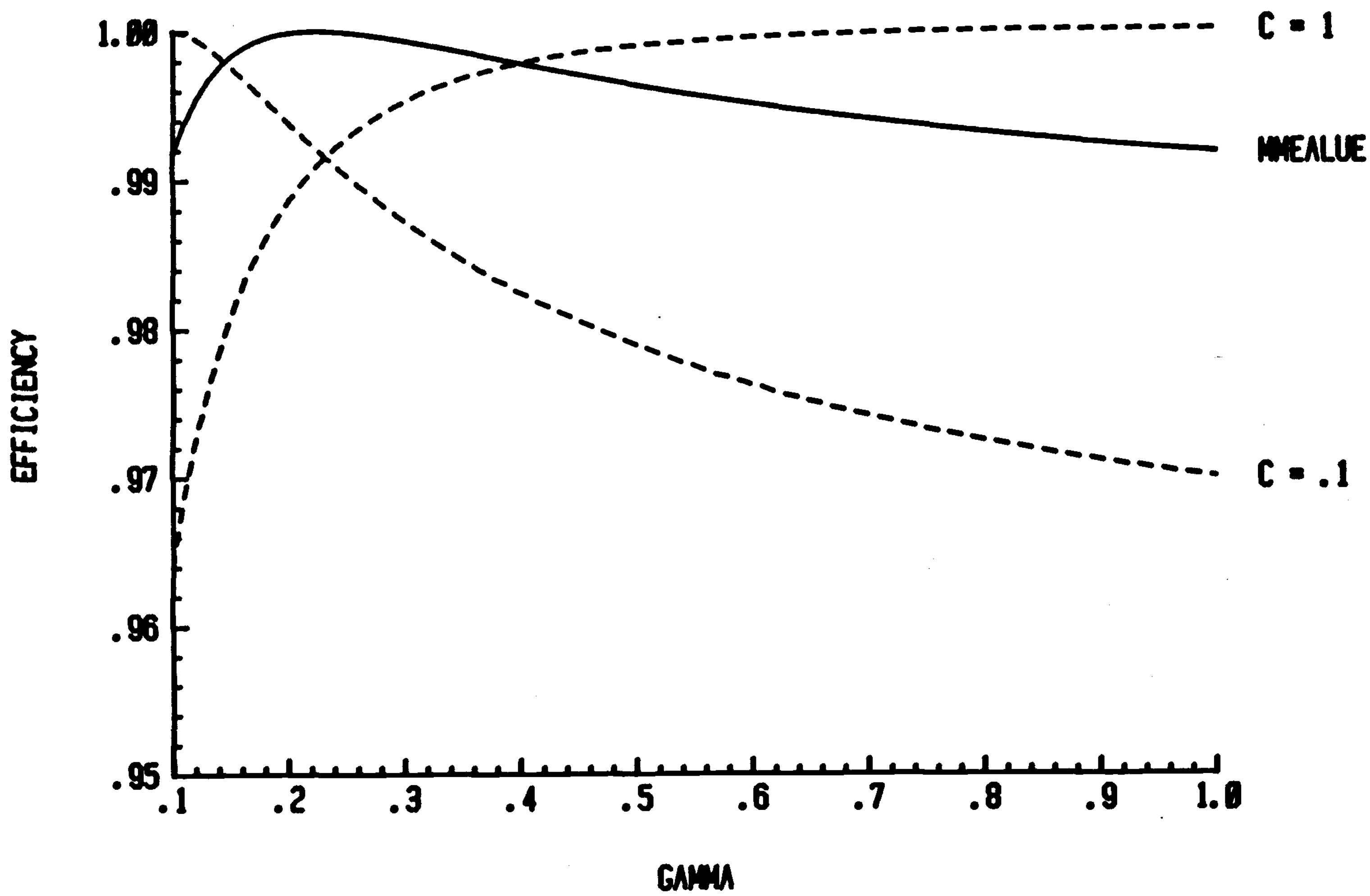


Figure 2. Efficiency curves for selected ALUEs for $\mu$ in example 5.1 when $\gamma \in [0.1,1]$.

If we knew that $\gamma \in [.1,1]$, we might prefer to use $\hat{\mu}_1$. But in this case we could employ Corollary 4.11 to obtain the maximin-efficient ALUE for $\mu$ when $\gamma \in [.1,1]$. This the estimator $\hat{\mu}_*$, where $\mathrm{Eff}(c^*|.1) = \mathrm{Eff}(c^*|1)$. The values of $c^*$ and $\hat{\mu}_*$ and the efficiency of $\hat{\mu}_*$ under this assumption were calculated and are summarized in Table 2 and Figure 2.

Example 5.2

In the previous example we estimated the single fixed parameter in a random one-way model. In this and the following example we consider the mixed two-factor nested model

$$Y_{ijk} = \mu + \alpha_i + b_{ij} + e_{ijk} \quad , \tag{5.4}$$

$i = 1,2,3$ , $j = 1,2,3$ , $k = 1,\ldots,n_{ij}$, where $\mu$ and the $\alpha_i$ are fixed but unknown and the $b_{ij}$ and $e_{ijk}$ are uncorrelated random variables having means zero and variances $\sigma_b^2$ and $\sigma_e^2$, respectively. We apply our results for the nested mixed model to obtain the maximin-efficient ALUEs and their efficiencies for the estimable contrasts $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$.

In this example we use the data given in Table 3. The data are actually taken from an experiment employing a different model, but are used here for illustrative purposes. That experiment is described in Example 6.1 of the next chapter.

In example 1.4 we showed how the mixed two-factor nested model can be put into the form of model (1.1). However, we can show, using a method similar to that used by Birkes, Seely, and

Azzam (1981), that within the class of LUEs for an estimable parametric function of $(\mu,\alpha_1,\alpha_2,\alpha_3)$ the class of LUEs of the form $\Sigma\Sigma\, c_{ij}\overline{Y}_{ij.}$ is complete, where $\overline{Y}_{ij.}$ denotes the mean for cell (i,j). By using this fact we can reduce the computational effort required to obtain the maximin-efficient ALUE for such a parametric function.

Now

$$\overline{Y}_{ij.} = \mu + \alpha_i + b_{ij} + \overline{e}_{ij.} \tag{5.5}$$

where the $b_{ij}$ and the $\overline{e}_{ij.}$ are uncorrelated random variables with means zero and variances $\sigma_b^2$ and $\sigma_e^2/n_{ij}$ respectively. Reparametrizing model (5.5), we obtain

$$E(Y) = X\beta \quad , \quad Cov(Y) = \theta_1 V_1 + \theta_2 V_2 \tag{5.6}$$

where

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad , \qquad Y = \begin{bmatrix} 4.0000 \\ 4.0323 \\ 3.7576 \\ 6.4545 \\ 6.7821 \\ 4.3097 \\ 6.6262 \\ 7.8045 \\ 4.1277 \end{bmatrix} \quad ,$$

$\beta = (\mu + \alpha_1, \alpha_2-\alpha_1, \alpha_3-\alpha_1)'$ , $\theta_1 = 50\sigma_b^2$ , $\theta_2 = \sigma_e^2$ ,
$V_1 = (1/50)I_9$ where $I_9$ is the 9 x 9 identity matrix and $V_2$ is the 9 x 9 diagonal matrix with elements $1/n_{11}, 1/n_{12},\ldots,1/n_{33}$.

Here $\theta_1$ and $V_1$ have been scaled so that $\|V_1\|_2 \doteq \|V_2\|_2$, where $\|V\|_2$ denotes the Frobenius norm (square root of the sum of squared elements) of a matrix V. This scaling has no effect on the maximin-efficient ALUE for a parametric function, but it tends to place c* in the middle of the interval $[\gamma_L, \gamma_U]$. This facilitates the convergence of the routine that finds c* and makes the efficiency graphs more readable.

The values of c* and the estimators $\pi'\hat{\beta}(0)$, $\pi'\hat{\beta}(c^*)$, and $\pi'\hat{\beta}(1)$ are listed in Table 4. The efficiencies of the ALUEs for $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$ are shown in Figures 3, 4, and 5, respectively.

Since for fixed c the LUE $\pi'\hat{\beta}(c)$ is BLUE when Cov(Y) = W(c), any linear combination of such LUEs is the LUE $\pi'\hat{\beta}(c)$ for the same linear combination of corresponding parametric functions. This is illustrated in Table 4 for c = 0,1. However, c* may vary for different parametric functions. As Table 4 illustrates, in this case the maximin-efficient ALUE for a linear combination of parametric functions need not be that same linear combination of maximin-efficient ALUEs for the component parametric functions.

The minimum efficiency of all ALUEs for $\alpha_2 - \alpha_1$ is the lesser of

$$\text{Eff}(0|1) = \frac{2}{3}\,[\Sigma n_{1j}^2/n_{1.}^2 + \Sigma n_{2j}^2/n_{2.}^2]^{-1} \doteq .973$$

and

$$\text{Eff}(1|0) = 9[n_{1.}^{-1} + n_{2.}^{-1}][\Sigma n_{1j}^{-1} + \Sigma n_{2j}^{-1}]^{-1} \doteq .985$$

where $n_{i.} = \sum_j n_{ij}$. The same is true for the minimum efficiencies

Table 3. Cell Numbers and Data for Example 5.2.

| i | j = 1 | 2 | 3 |
|---|---|---|---|
| 1 | $\underline{n}_{11}$ = 34<br>$Y_{11.}$ = 4.0000 | 31<br>4.0323 | 33<br>3.7576 |
| 2 | 66<br>6.4545 | 78<br>6.7821 | 113<br>4.3097 |
| 3 | 107<br>6.6262 | 133<br>7.8045 | 188<br>4.1277 |

Table 4. c* and ALUEs for Example 5.2.

| | Parametric Function | | |
|---|---|---|---|
| | $\alpha_2-\alpha_1$ | $\alpha_3-\alpha_1$ | $\alpha_3-\alpha_2$ |
| c* | .439093 | .371743 | .313869 |
| $\pi'\hat{\beta}(0)$ | 1.68229 | 1.96629 | .28400 |
| $\pi'\hat{\beta}(c^*)$ | 1.82408 | 2.15793 | .34790 |
| $\pi'\hat{\beta}(1)$ | 1.91880 | 2.25617 | .33737 |

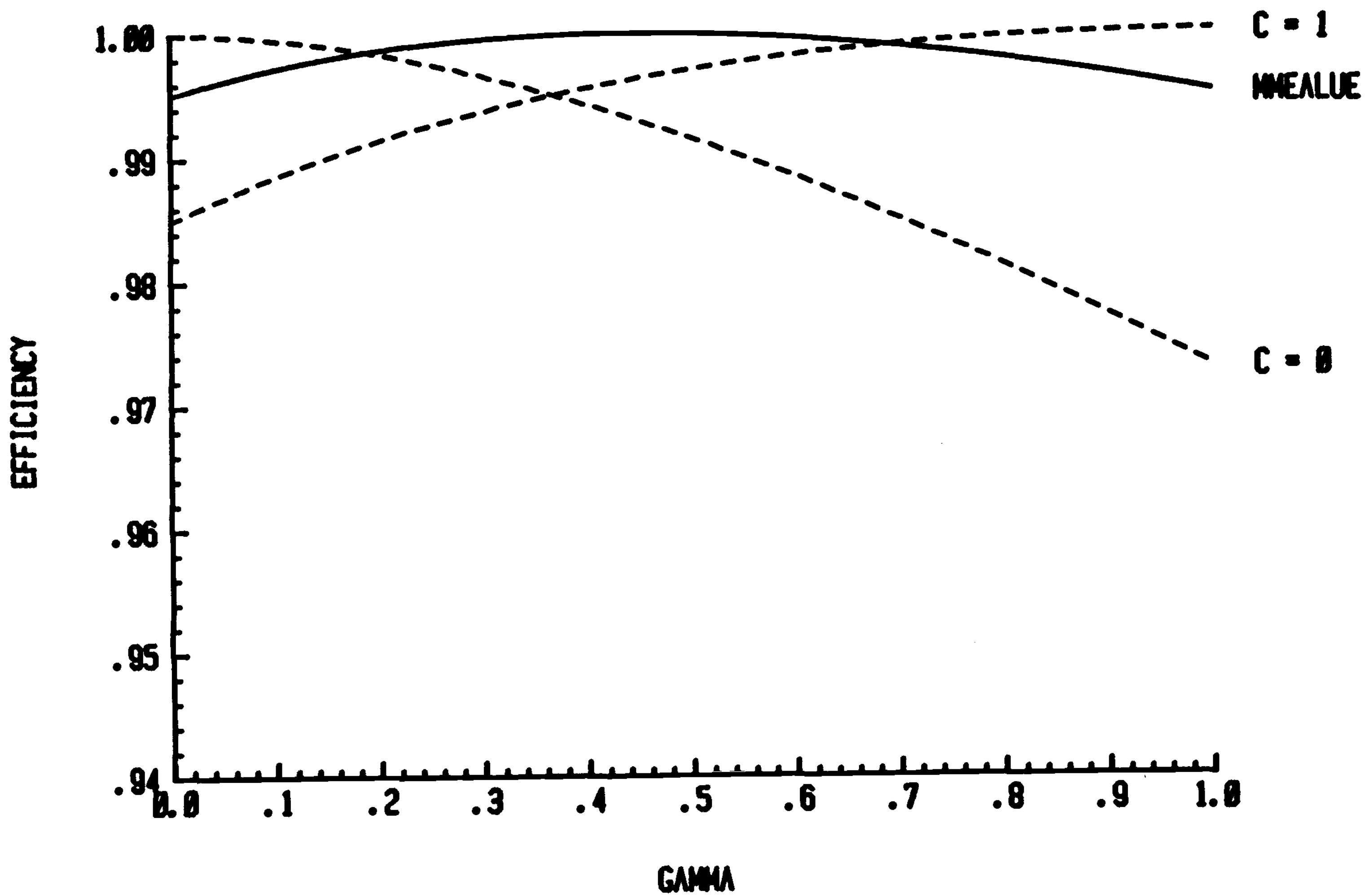


Figure 3. Efficiency curves for selected ALUEs for $\alpha_2 - \alpha_1$ in example 5.2.

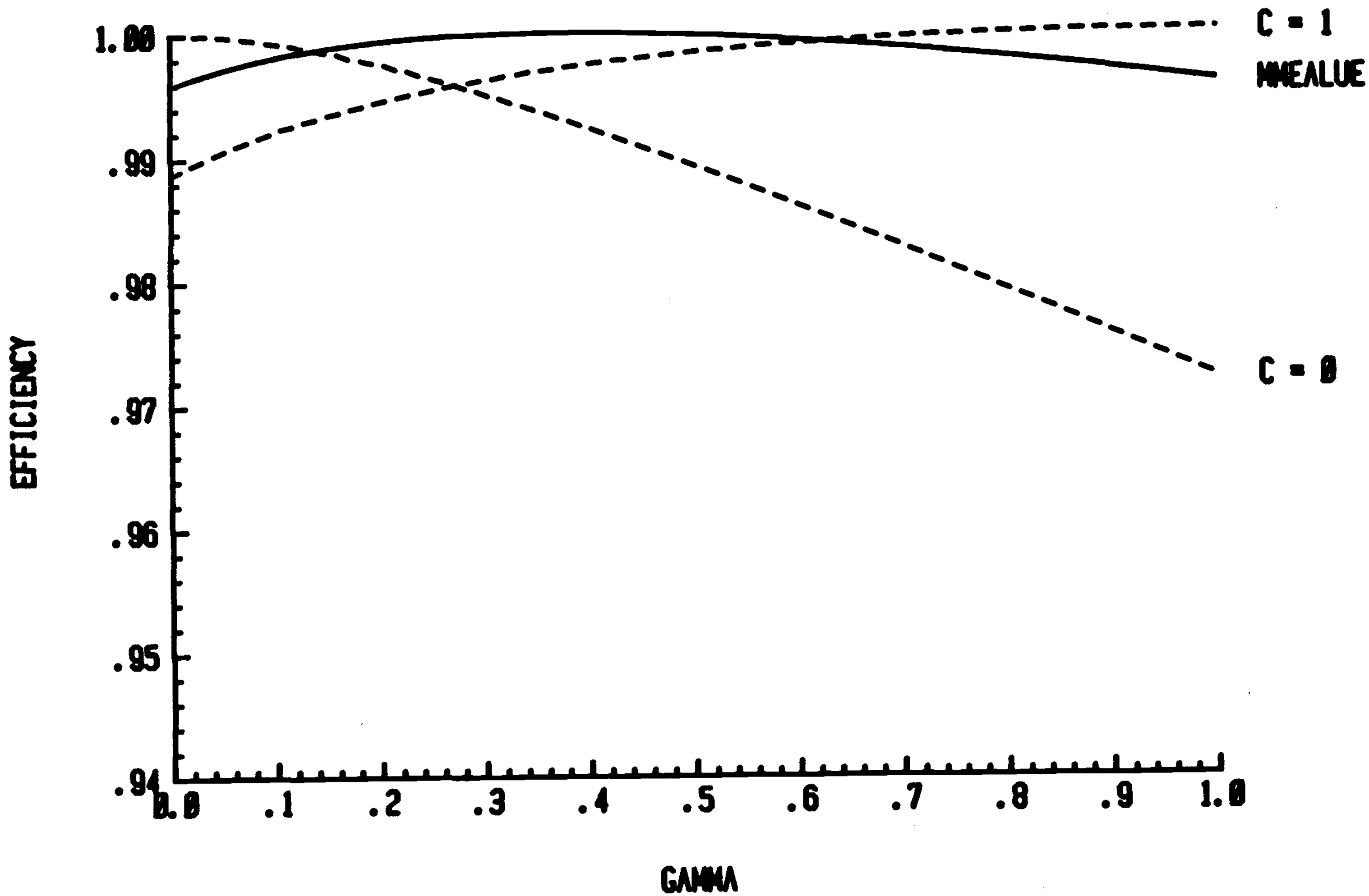


Figure 4. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_1$ in example 5.2.

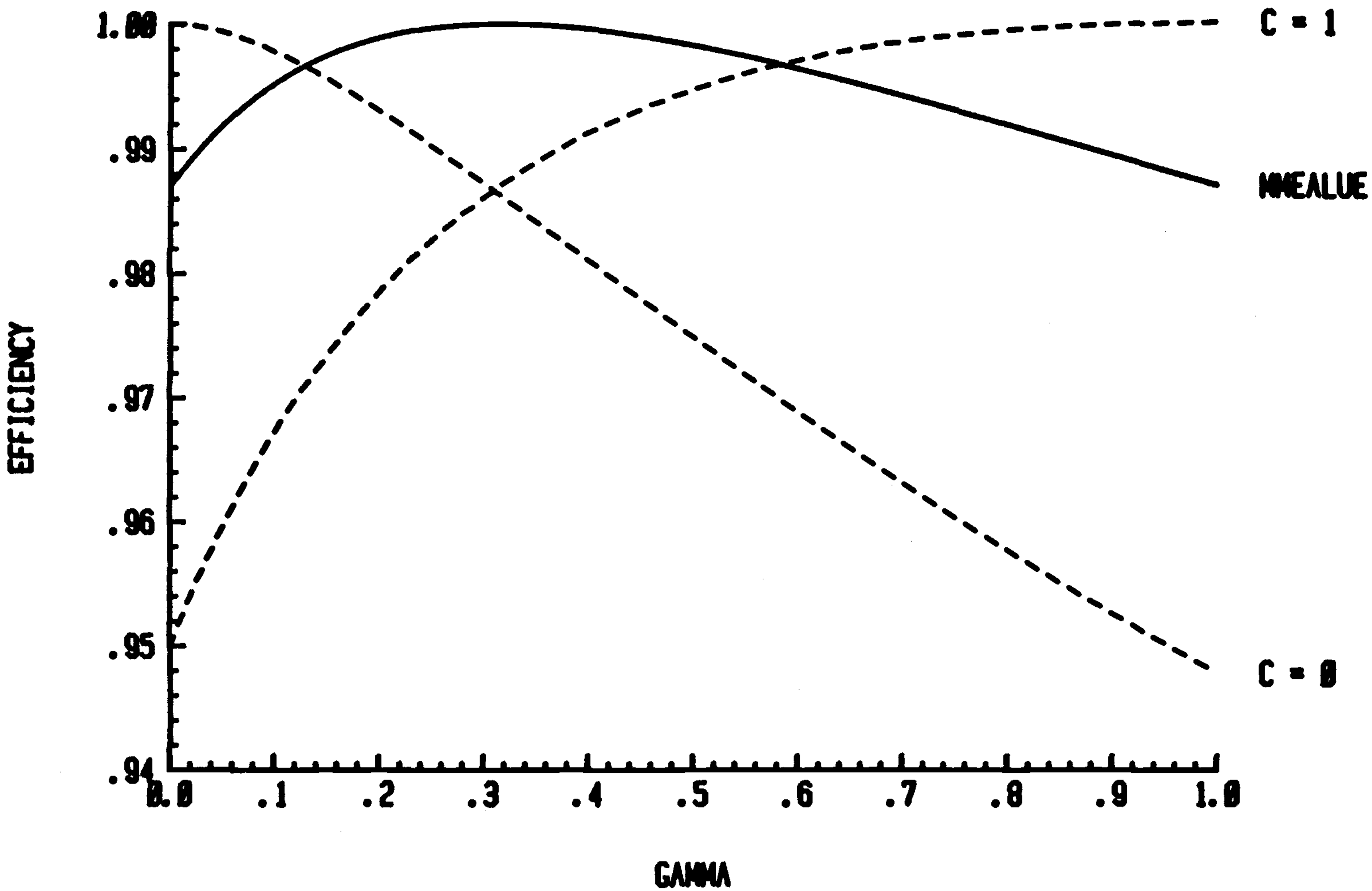


Figure 5. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_2$ in example 5.2.

of ALUEs for $\alpha_3 - \alpha_1$ and $\alpha_3 - \alpha_2$, with similar expressions replacing those above.

As Figures 3, 4, and 5 show, the efficiencies of all ALUEs for $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$ are greater than .947. The reason for the high efficiencies is that the $n_{ij}$'s are roughly equal across the rows in Table 3. If the $n_{ij}$'s were exactly constant across the rows, a BLUE would exist for each of $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$.

Example 5.3

In this example we illustrate how the efficiencies of the ALUEs for $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$ with respect to model (5.6) decrease as the $n_{ij}$'s depart further from constancy across rows.

We use the same model and data as in example 5.2, but transpose the $n_{ij}$ in Table 3. Only $V_2$ changes to reflect the reordering of the $n_{ij}$.

The results are summarized in Table 5 and Figures 6, 7, and 8. The minimum efficiency of all ALUEs occurs for c = 1 when estimating $\alpha_3 - \alpha_2$, and is equal to .661. However, even in this example the maximin-efficient ALUEs for $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$ have efficiencies greater than .91 for all values of $\gamma \in [0,1]$.

Table 5. c* and ALUEs for Example 5.3.

| | Parametric Function | | |
|---|---|---|---|
| | $\alpha_2-\alpha_1$ | $\alpha_3-\alpha_1$ | $\alpha_3-\alpha_2$ |
| c* | .443702 | .423364 | .413697 |
| $\pi'\hat{\beta}(0)$ | 1.49634 | 1.73351 | .23717 |
| $\pi'\hat{\beta}(c^*)$ | 1.73024 | 2.06589 | .34336 |
| $\pi'\hat{\beta}(1)$ | 1.91880 | 2.25617 | .33737 |

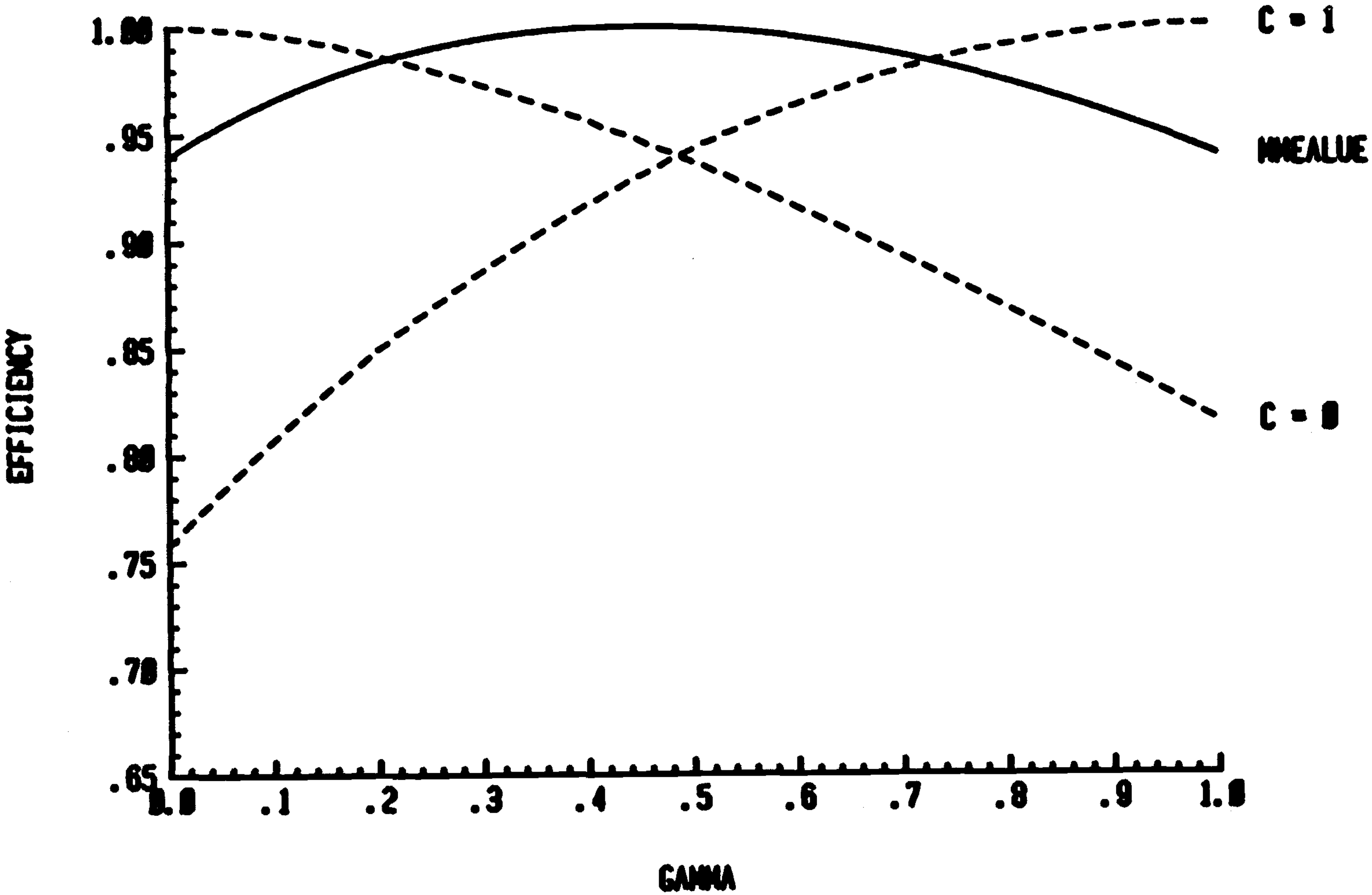


Figure 6. Efficiency curves for selected ALUEs for $\alpha_2 - \alpha_1$ in example 5.3.

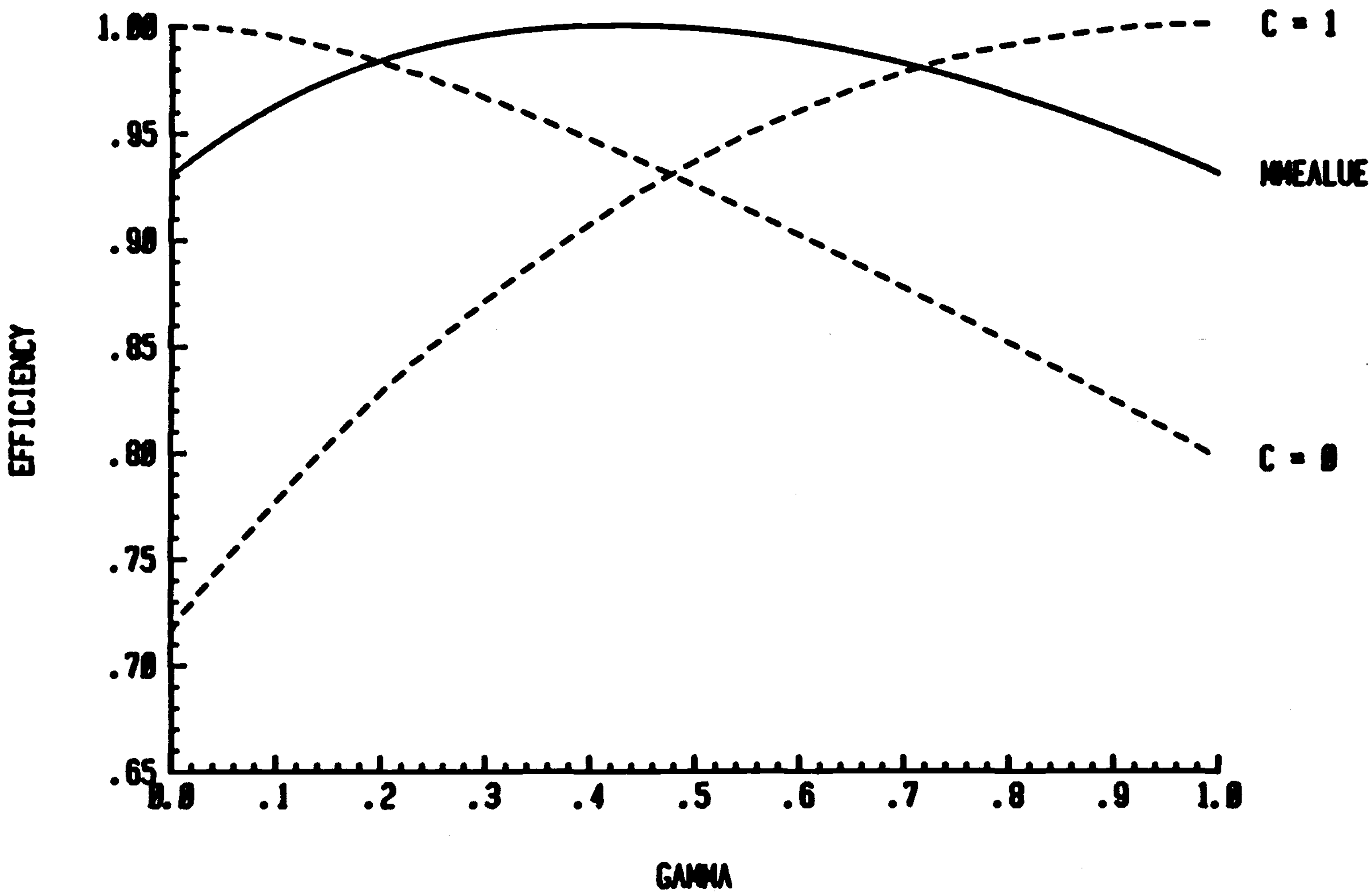


Figure 7. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_1$ in example 5.3.

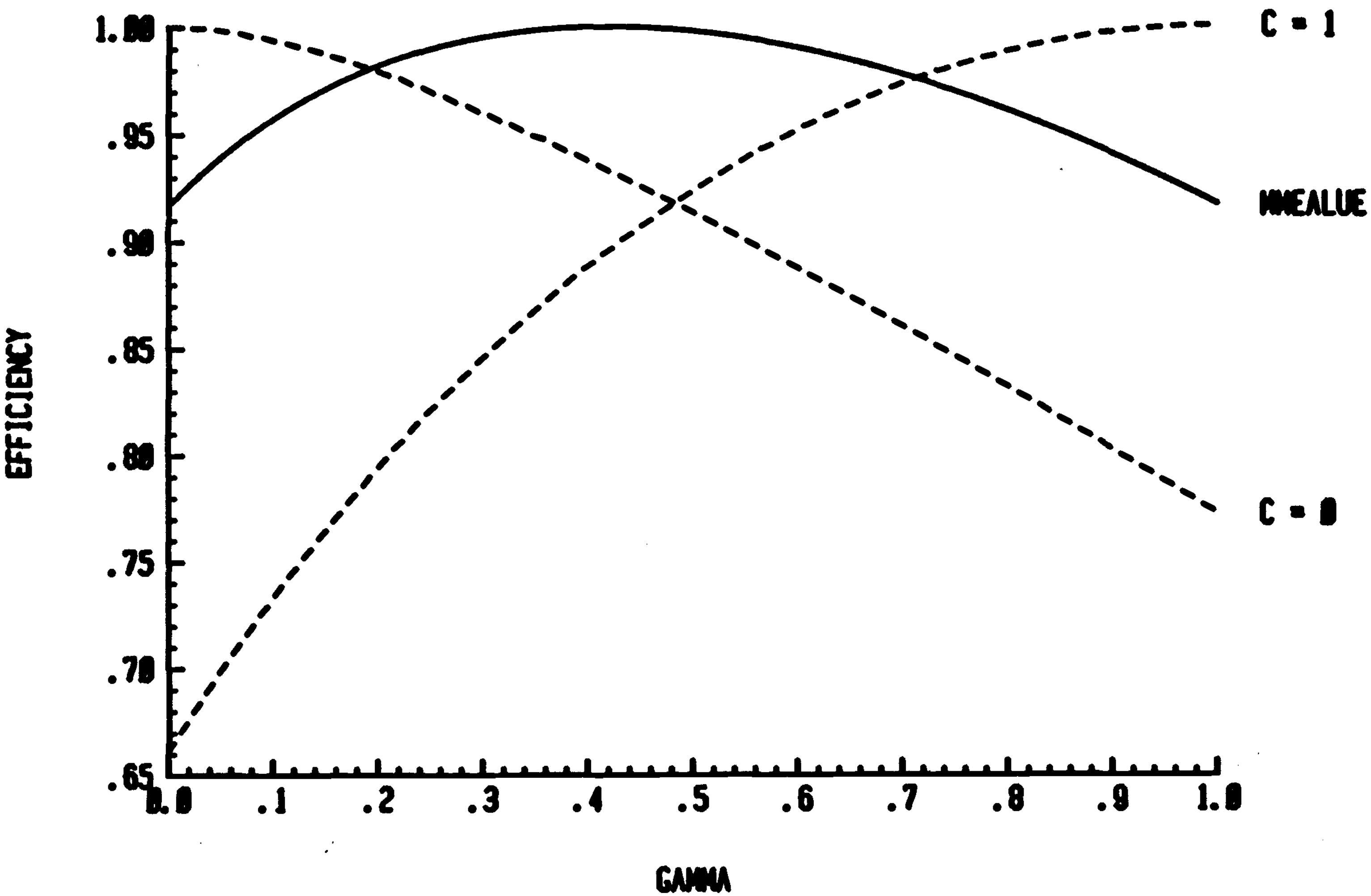


Figure 8. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_2$ in example 5.3.

# VI. EXAMPLES: THE MIXED TWO-WAY ADDITIVE MODEL

We may write this model as

$$Y_{ijk} = \mu + \alpha_i + d_j + e_{ijk} \tag{6.1}$$

where $i = 1,\ldots,I$ , $j = 1,\ldots,J$ , and $k = 1,\ldots,n_{ij}$ . Here $Y_{ijk}$ denotes the k-th observation on the i-th level of $\alpha$ and the j-th level of d. $\mu$ and the $\alpha_i$ are fixed but unknown parameters. The $d_j$ and the $e_{ijk}$ are uncorrelated random variables with means zero and variances $\theta_1$ and $\theta_2$ respectively.

Define

$$n_{i.} = \sum_j n_{ij} \quad , \quad n_{.j} = \sum_i n_{ij} \quad , \quad n_{..} = \sum_i \sum_j n_{ij} \quad .$$

We can assume, without loss of generality, that $n_{i.} > 0$ for $i = 1,\ldots,I$ and $n_{.j} > 0$ for $j = 1,\ldots,J$. We can show, using a method similar to that used in Birkes, Seely, and Azzam (1981), that within the class of LUEs for estimable parametric functions of the fixed parameters, the class of LUEs of the form $\Sigma\,\Sigma c_{ij}\overline{Y}_{ij.}$ is complete, where $\overline{Y}_{ij.}$ denotes the mean of the observations with the i-th level of $\alpha$ and the j-th level of d.

We may write

$$Y = \underline{1}\mu + A\alpha + Bb + e \quad , \tag{6.2}$$

where $Y = (\overline{Y}_{11.},\overline{Y}_{12.},\ldots,\overline{Y}_{IJ.})'$ is the vector of cell means,

$\underline{1}$ is the IJ-vector of ones, $\alpha = (\alpha_1,\ldots,\alpha_I)'$, $b = (d_1,\ldots,d_J)'$, and $e = (\bar{e}_{11.},\bar{e}_{12.},\ldots,\bar{e}_{IJ.})'$. A is an IJ x I matrix of zeroes and ones, each row consisting of I-1 zeroes and one 1 in the column corresponding to the appropriate element of $\alpha$. B is an IJ x J matrix defined similarly.

By setting $X = [\underline{1},A]$ and $\beta = (\mu,\alpha')'$, we obtain a variation of model (4.1),

$$Y = X\beta + Bb + e \quad , \tag{6.3}$$

where $B_1 = B$, $b_1 = b$, $B_2b_2 = e$, and $n = IJ$. Here $V_2$ is an $n \times n$ diagonal matrix with diagonal elements $1/n_{11}, 1/n_{12},\ldots,$ $1/n_{IJ}$. $V_2$ is p.d. and $\underline{r}(X,V_1,V_2) = n$ so Theorems 4.3-4.6 apply.

Let $M_R$ denote model (6.3) or, equivalently,

$$M_R\!:E(Y) = X\beta \quad , \quad Cov(Y) = \theta_1V_1 + \theta_2V_2 \tag{6.4}$$

where $V_1 = BB'$. Let $M_F$ denote the artificial model

$$M_F\!:E(Y) = X\beta + B\delta \quad , \quad Cov(Y) = V_2 \tag{6.5}$$

where $\delta$ is fixed but unknown. Let $\pi'\beta$ be estimable with respect to model $M_R$. Since $V_2$ is p.d., $\pi'\beta$ cannot satisfy the conditions of either Theorems 4.3 or 4.4. If $\pi'\beta$ is estimable with respect to artificial model $M_F$, then by Theorem 4.5 the unique maximin-efficient ALUE for $\pi'\beta$ with respect to model $M_R$ is the same estimator as the unique BLUE for $\pi'\beta$ with respect to artificial model $M_F$.

The class of parametric functions $\pi'\beta$ that are estimable with respect to model $M_R$ is spanned by

$$\{\mu + \alpha_1, \alpha_2 - \alpha_1, \ldots, \alpha_I - \alpha_1\} \quad .$$

By Proposition 4.3 of Birkes, Dodge, and Seely (1976), $\mu + \alpha_1$ cannot be estimable with respect to artificial model $M_F$. Proposition 4.4 of that same paper gives a necessary and sufficient condition for $\alpha_i - \alpha_1$ to be estimable with respect to artificial model $M_F$.

Any parametric function that is a linear combination of those $\alpha_i - \alpha_1$ that are estimable with respect to artificial model $M_F$ will also be estimable with respect to model $M_R$, and its unique maximin-efficient ALUE with respect to model $M_R$ can be found as its unique BLUE with respect to artificial model $M_F$. Since linear combinations of BLUEs are also BLUEs, we see that linear combinations of maximin-efficient ALUEs for such parametric functions with respect to model $M_R$ are also maximin-efficient ALUEs.

If the design is connected, then any $\alpha$-contrast will be estimable with respect to artificial model $M_F$. The maximin-efficient ALUE for any $\alpha$-contrast with respect to model $M_R$ will be the BLUE for that contrast with respect to artificial model $M_F$.

If $\pi'\beta$ is not estimable with respect to artificial model $M_F$ then Theorem 4.6 applies.

Example 6.1

Snedecor and Cochran (1967, section 16.2) present the results of an experiment in which 3 strains of mice were inoculated with 3 isolations (different types) of the mouse typhoid organism. The number of mice and the mean days-to-death are shown for each cell in Table 6.

Snedecor and Cochran assume the original data followed a two-way classification with interaction which we write as

$$Y_{ijk} = \mu + \alpha_i + d_j + (\alpha d)_{ij} + e_{ijk} \quad , \tag{6.4}$$

where $i = 1,2,3$ , $j = 1,2,3$ , and $k = 1,\ldots,n_{ij}$. Here $Y_{ijk}$ represents the days-to-death for mouse k of strain j receiving isolation i. They use an unweighted analysis of cell means assuming the overall mean $\mu$, isolation effect $\alpha_i$, strain effect $d_j$, and interaction $(\alpha d)_{ij}$ are fixed and the $e_{ijk}$ are uncorrelated random variables with means zero and variances $\sigma^2$.

To illustrate the results of this chapter, we suppose that the strain effects $d_j$ are random effects, uncorrelated with themselves and with the $e_{ijk}$, having means zero and variances $\sigma_d^2$, and that no interactions $(\alpha d)_{ij}$ exist.

As in example 5.2, we can show that within the class of LUEs for an estimable parametric function of $(\mu,\alpha_1,\alpha_2,\alpha_3)$, the class of LUEs of the form $\Sigma\Sigma c_{ij}\overline{Y}_{ij.}$ is complete, where (as before) $\overline{Y}_{ij.}$ denotes the mean for cell (i,j). Reparametrizing the resulting model for $\overline{Y}_{ij.}$, we obtain the models

$$M_R: E(Y) = X\beta \quad , \quad Cov(Y) = \theta_1 V_1 + \theta_2 V_2 \quad ,$$

$$M_F: E(Y) = X\beta + B\delta \quad , \quad Cov(Y) = V_2 \quad ,$$

where Y, X, $\beta$, $\theta_2$, and $V_2$ are the same as in model (5.6), but where $\delta = (\delta_1, \delta_2, \delta_3)'$ is fixed but unknown, $\theta_1 = 100\sigma_d^2$, and $V_1 = (1/100)BB'$, where

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

Here $\theta_1$ and $V_1$ have been scaled so that $\|V_1\|_2 \doteq \|V_2\|_2$.

In this example, any $\alpha$-contrast is estimable with respect to both model $M_R$ and artificial model $M_F$. Therefore, to obtain the maximin-efficient ALUE for an $\alpha$-contrast with respect to model $M_R$, we find the BLUE for that contrast with respect to the artificial model $M_F$.

To facilitate computing BLUEs with respect to artificial model $M_F$, we reparametrize that model to

$$M_F': E(Y) = H\eta \quad , \quad Cov(Y) = V_2 \quad ,$$

where $\eta = (\mu + \alpha_1 + \delta_1, \alpha_2 - \alpha_1, \alpha_3 - \alpha_1, \delta_2 - \delta_1, \delta_3 - \delta_1)'$ and

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

The results of this example are summarized in Table 7 and Figures 9, 10, and 11. Table 7 illustrates the remark made above that linear combinations of maximin-efficient ALUEs of $\alpha$-contrasts that are estimable with respect to artificial model $M_F$ are also maximin-efficient ALUEs. In this case, any $\alpha$-contrast can be expressed as a linear combination of $\alpha_2 - \alpha_1$ and $\alpha_3 - \alpha_1$, and its maximin-efficient ALUE is the same linear combination of the maximin-efficient ALUEs for $\alpha_2 - \alpha_1$ and $\alpha_3 - \alpha_1$. For example, the estimate for $\alpha_3 - \alpha_2 = (\alpha_3 - \alpha_1) - (\alpha_2 - \alpha_1)$ is the difference of the estimates for $\alpha_3 - \alpha_1$ and $\alpha_2 - \alpha_1$.

Figures 9, 10, and 11 contain efficiency curves for ALUEs for $\alpha_2 - \alpha_1$, $\alpha_3 - \alpha_1$, and $\alpha_3 - \alpha_2$, respectively. The efficiency curves for those ALUEs with respect to model $M_R$ that are best when $Cov(Y) = cV_1 + (1-c)V_2$ for $c = 0$ and $c = .9$ are included for comparison with the efficiency curve for the maximin-efficient ALUE. As these figures illustrate, except for the maximin-efficient ALUE, all ALUEs of an $\alpha$-contrast have a minimum efficiency of zero. The minimum efficiencies of the maximin-efficient ALUEs for our three $\alpha$-contrasts are given in Table 7.

Table 6. Cell Numbers and Mean Days-to-Death in Three Strains of Mice Inoculated with Three Isolations of the Typhoid Bacillus (Reproduced from Snedecor and Cochran, 1967).

| Isolation | RI | Z | Ba |
|---|---|---|---|
| 9D | 34<br>4.0000 | 31<br>4.0323 | 33<br>3.7576 |
| 11C | 66<br>6.4545 | 78<br>6.7821 | 113<br>4.3097 |
| DSC 1 | 107<br>6.6262 | 133<br>7.8045 | 188<br>7.1277 |

Table 7. Maximin-Efficient ALUEs and Their Minimum Efficiencies for Example 6.1.

| | Parametric Function | | |
|---|---|---|---|
| | $\alpha_2-\alpha_1$ | $\alpha_3-\alpha_1$ | $\alpha_3-\alpha_2$ |
| MMEALUE | 1.91101 | 2.18848 | .27747 |
| min. efficiency | .9949 | .9939 | .9999 |

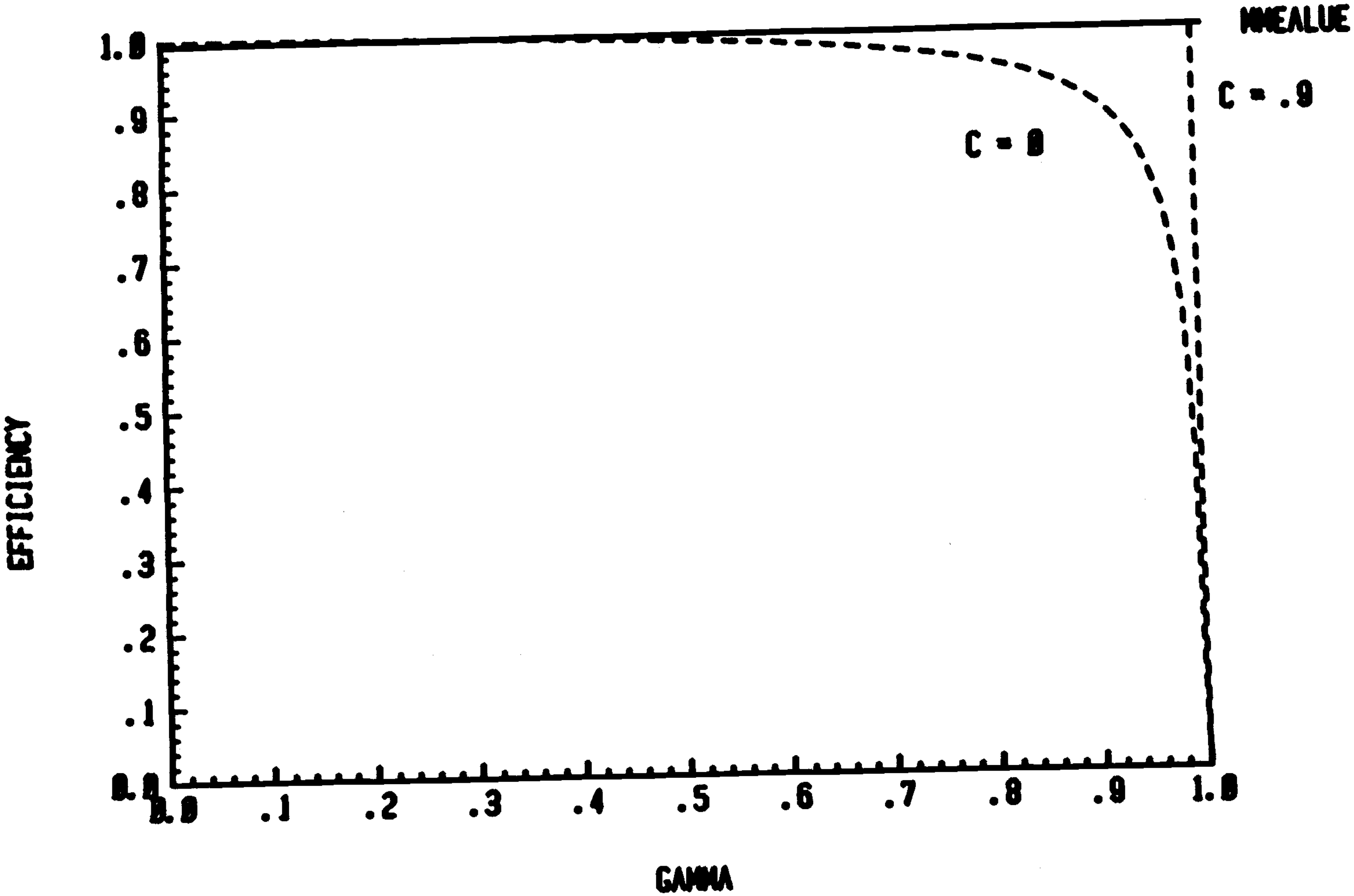


Figure 9. Efficiency curves for selected ALUEs for $\alpha_2 - \alpha_1$ in example 6.1.

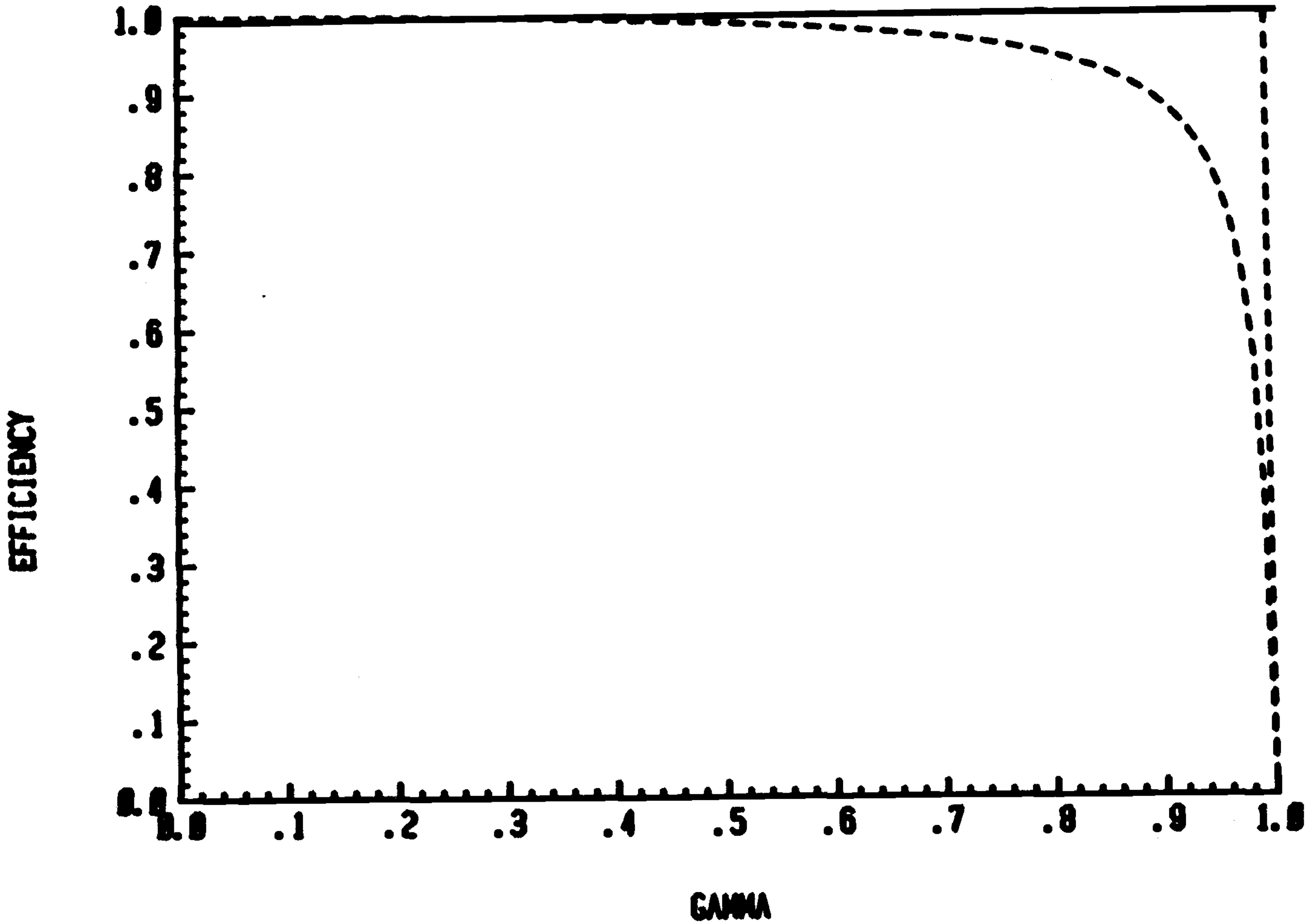


Figure 10. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_1$ in example 6.1.

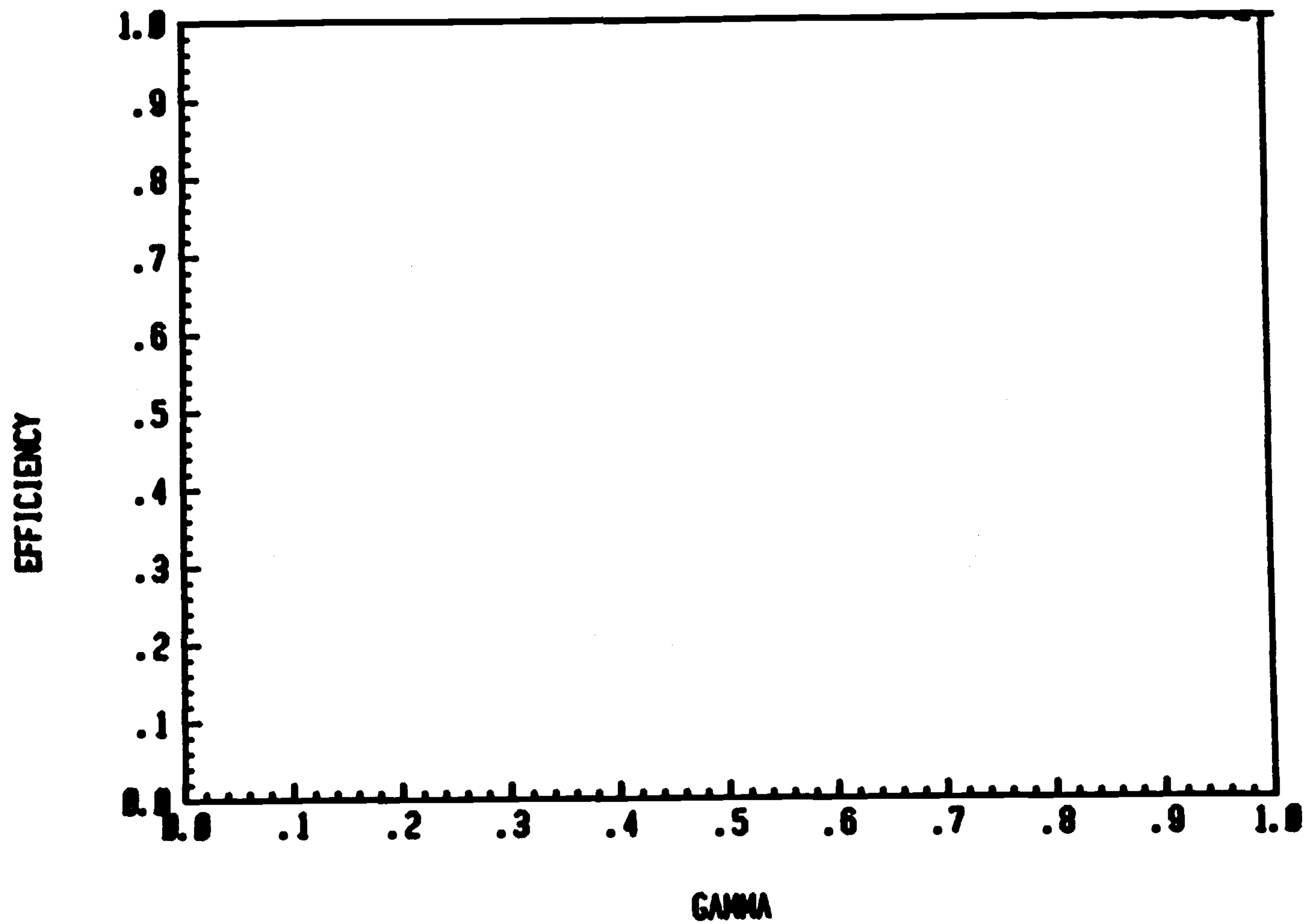


Figure 11. Efficiency curves for selected ALUEs for $\alpha_3 - \alpha_2$ in example 6.1.

Example 6.2

In example 6.1, the design is complete; that is, $n_{ij} > 0$ for all i,j. We consider here an example of a balanced incomplete block design, which is connected but not complete.

We use the mixed two-way additive model as in example 6.1 but have cell numbers $n_{ij}$ given by Table 8. The purpose of this example is to see if the efficiency curves of the maximin-efficient ALUEs and of selected competing ALUEs for $\alpha$-contrasts differ markedly from those of example 6.1.

We have the models

$$M_R : E(Y) = X\beta \quad , \quad \text{Cov}\ (Y) = \theta_1 V_1 + \theta_2 V_2$$
$$M_F : E(Y) = X\beta + B\delta \quad , \quad \text{Cov}\ (Y) = V_2 \quad ,$$

where here $Y = (Y_{11}, Y_{12}, Y_{22}, Y_{23}, Y_{31}, Y_{33})'$,
$\beta = (\mu + \alpha_1,\ \alpha_2 - \alpha_1,\ \alpha_3 - \alpha_1)'$, $\delta = (\delta_1, \delta_2, \delta_3)'$,

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad ,$$

$\theta_1 = \sigma_d^2$, $V_1 = BB'$, $\theta_2 = \sigma_e^2$, and $V_2 = I_6$.

Any $\alpha$-contrast is estimable with respect to both model $M_R$ and artificial model $M_F$. To obtain the maximin-efficient ALUE for an $\alpha$-contrast with respect to model $M_R$, we find the BLUE for that contrast with respect to the artificial model $M_F$.

We again reparametrize artificial model $M_F$ to facilitate computing BLUEs of $\alpha$-contrasts with respect to it:

$$M_F': E(Y) = H\eta \quad , \quad \text{Cov}\ (Y) = V_2 \quad ,$$

where $\eta = (\mu + \alpha_1 + \delta_1, \alpha_2 - \alpha_1, \alpha_3 - \alpha_1, \delta_2 - \delta_1, \delta_3 - \delta_1)'$

and

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

Figure 12 contains efficiency curves for selected ALUEs for $\alpha_2 - \alpha_1$. The efficiency curves for those ALUEs that are best when $\text{Cov}(Y) = cV_1 + (1 - c)V_2$ for $c = 0$ and $c = .9$ are included with the efficiency curve for the maximin-efficient ALUE. Since the design is balanced, the corresponding efficiency curves for the ALUEs for $\alpha_3 - \alpha_1$ and $\alpha_3 - \alpha_2$ are the same as for $\alpha_2 - \alpha_1$. The minimum efficiency of the maximin-efficient ALUE for $\alpha_2 - \alpha_1$ is .7500.

Table 8. Cell Numbers $n_{ij}$ for the Balanced Incomplete Block Design of Example 6.2.

| | Strain of Mice | | |
|---|---|---|---|
| Isolation | RI | Z | Ba |
| 9D | 1 | 1 | 0 |
| 11C | 0 | 1 | 1 |
| DSC 1 | 1 | 0 | 1 |

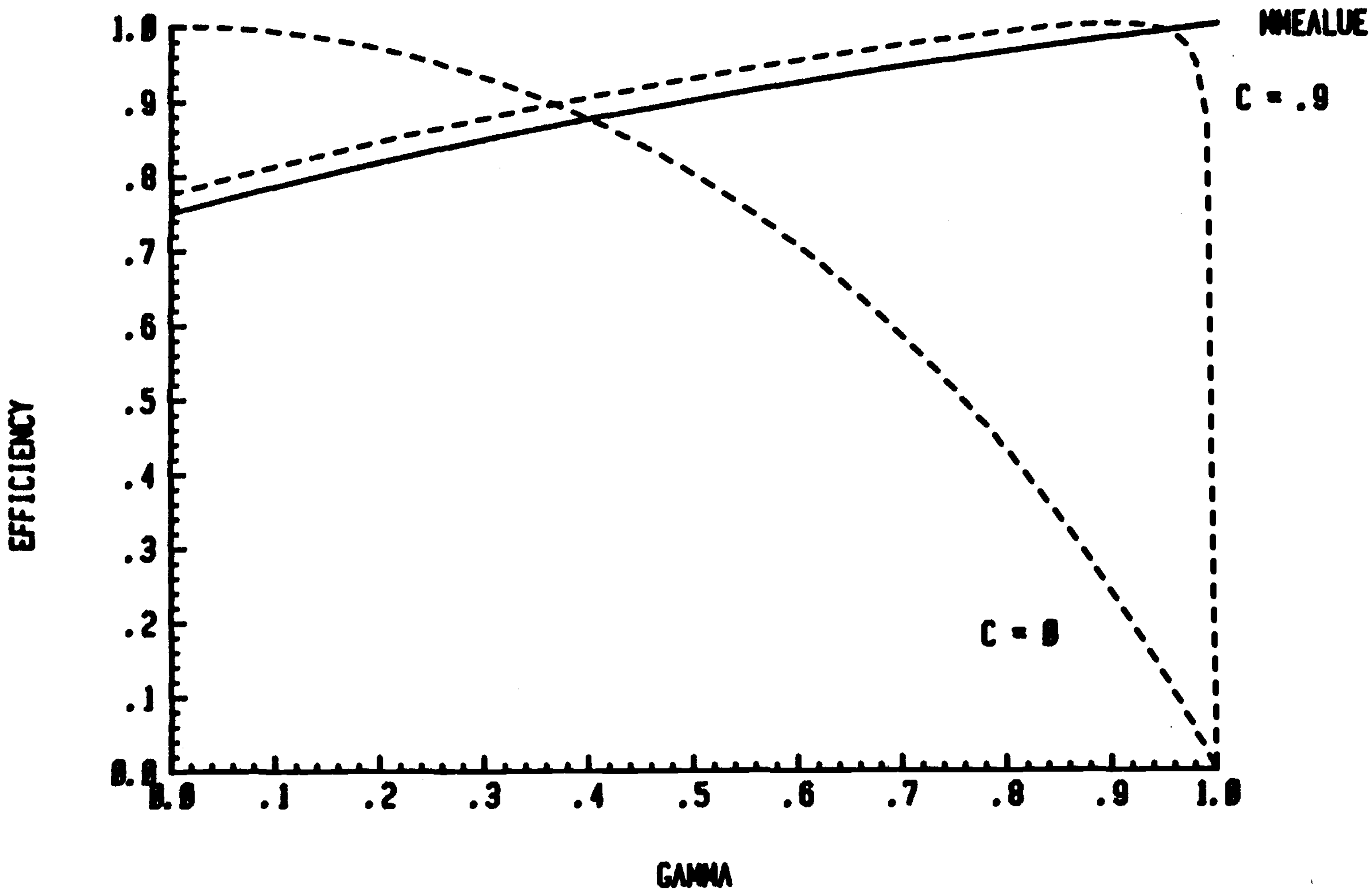


Figure 12. Efficiency curves for selected ALUEs for $\alpha_2 - \alpha_1$ in example 6.2.

# VII. NOTES REGARDING ALTERNATIVE ESTIMATORS

## 7.1 Minimax-Variance ALUE

Birkes, Seely, and Azzam (1981) discuss the estimation of the mean $\mu$ in the random one-way model and characterize the maximin-efficient ALUE for $\mu$. They also point out that the ALUEs for $\mu$ can be compared using their variances, rather than their efficiencies, and suggest one way to do this.

In the course of characterizing the maximin-efficient ALUE for $\mu$, they show that within the class of LUEs for $\mu$ the class of LUEs based upon the $t$ group means $\overline{Y}_{i.}$, $i = 1,\dots,t$, constitutes a complete class. The first and second moments of the group means are summarized by model (5.3),

$$E(Y) = \underline{1}_t\mu \quad , \quad Cov(Y) = \theta_1 I_t + \theta_2 D \quad ,$$

where $Y = (\overline{Y}_{1.},\dots,\overline{Y}_{t.})'$ and $D = diag(1/n_1,\dots,1/n_t)$.

Birkes et al. suggest holding $\theta = \theta_1 + \theta_2$ fixed and selecting the ALUE for $\mu$ with the minimum maximum possible variance. This is the unweighted mean of the group means $\hat{\mu}_1$, which they call the "minimax-variance" LUE for $\mu$.

More generally, one might choose to hold some linear combination $s_1\theta_1 + s_2\theta_2$ fixed and select the ALUE for $\mu$ with the minimum maximum possible variance. We shall show, however, that unless a BLUE for $\mu$ exists, the ALUE selected in this way depends upon the particular linear combination $s_1\theta_1 + s_2\theta_2$ held fixed.

The variance of a linear estimator $t'Y$ with respect to model (5.3) is $\mathrm{var}(t'Y|\theta_1,\theta_2) = \theta_1 t't + \theta_2 t'Dt$ for $\theta_1,\theta_2 \geqslant 0$. If we hold $s_1\theta_1 + s_2\theta_2$ fixed at some value, say K, where $s_1,s_2,K > 0$, then

$$\mathrm{var}(t'Y|\theta_1) = Kt'Dt/s_2 + [t't - s_1 t'Dt/s_2]\theta_1$$

for $\theta_1 \in [0,K/s_1]$. Now, $\mathrm{var}(t'Y|\theta_1)$ is a linear function of $\theta_1$ so the maximum possible variance of $t'Y$, when $s_1\theta_1 + s_2\theta_2 = K$, is

$$\max[\mathrm{var}(t'Y|\theta_1 = 0\ ,\ \theta_2 = K/s_2)\ \ ,\ \mathrm{var}(t'Y|\theta_1 = K/s_1\ ,\ \theta_2 = 0)]$$

$$= \max[K/s_2 \cdot \mathrm{var}(t'Y|\gamma = 0,\ \theta = 1)\ ,\ K/s_1 \cdot \mathrm{var}(t'Y|\gamma = 1,\ \theta = 1)],$$

where $\theta = \theta_1 + \theta_2 > 0$ and $\gamma = \theta_1/\theta$.

Birkes et al. obtain $\bar{A} = \{\hat{\mu}_c : c \in [0,1]\}$, $\mathrm{var}(\hat{\mu}_c|\gamma,\theta)$, and $\mathrm{Eff}(c|\gamma)$. They show that for fixed $\gamma \in [0,1]$, $\mathrm{Eff}(c|\gamma)$ is a continuous positive function of c on [0,1] that is strictly increasing on $(0,\gamma)$ and strictly decreasing on $(\gamma,1)$. So for fixed $\gamma \in [0,1]$ and $\theta > 0$,

$$\mathrm{var}(\hat{\mu}_c|\gamma,\theta) = \mathrm{var}(\hat{\mu}_\gamma|\gamma,\theta)\ /\ \mathrm{Eff}(c|\gamma)$$

is a continuous positive function of c on [0,1] that is strictly decreasing on $(0,\gamma)$ and strictly increasing on $(\gamma,1)$.

We see that $\hat{\mu}_{c'}$ is a "minimax-variance" ALUE for $\mu$ when $s_1\theta_1 + s_2\theta_2 = K$ if $c' \in [0,1]$ and

$$\max[\operatorname{var}(\hat{\mu}_{c'}|\gamma = 0\ ,\ \theta = 1)/s_2\ ,\ \operatorname{var}(\hat{\mu}_{c'}|\gamma = 1\ ,\ \theta = 1)/s_1]$$

$$= \min_{0 \le c \le 1} \max[\operatorname{var}(\hat{\mu}_c|\gamma = 0\ ,\ \theta = 1)/s_2\ ,\ \operatorname{var}(\hat{\mu}_c|\gamma = 1\ ,\ \theta = 1)/s_1].$$

Furthermore, $\operatorname{var}(\hat{\mu}_c|\gamma = 0\ ,\ \theta = 1)$ and $\operatorname{var}(\hat{\mu}_c|\gamma = 1\ ,\ \theta = 1)$ are continuous positive functions of $c$ on $[0,1]$ that are strictly increasing and strictly decreasing on $(0,1)$, respectively.

For sufficiently small $s_1/s_2$, we have

$$\max[\operatorname{var}(\hat{\mu}_c|\gamma = 0\ ,\ \theta = 1)/s_2\ ,\ \operatorname{var}(\hat{\mu}_c|\gamma = 1\ ,\ \theta = 1)/s_1]$$

$$= \operatorname{var}(\hat{\mu}_c|\gamma = 1\ ,\ \theta = 1)/s_1$$

for all $c \in [0,1]$, and $\hat{\mu}_{c'} = \hat{\mu}_1$. Similarly, for sufficiently large $s_1/s_2$ we have $\hat{\mu}_{c'} = \hat{\mu}_0$. If no BLUE for $\mu$ exists, then $\hat{\mu}_0 \neq \hat{\mu}_1$ and the "minimax-variance" ALUE differs for different choices of $s_1$ and $s_2$. In fact, it is clear that by appropriately selecting $s_1$ and $s_2$ we can make any ALUE for $\mu$ a "minimax-variance" ALUE.

### 7.2 Nonlinear Unbiased Estimators

The traditional estimators of a parametric function $\pi'\beta$ in model (1.1) or model (1.2) have the form $\pi'\tilde{\beta}$ where $\tilde{\beta}$ is a nonlinear estimator for $\beta$.

The method of maximum likelihood (ML) simultaneously estimates the fixed parameters $\beta$ and the variance components $\theta_1$ and $\theta_2$ in model (1.2) by maximizing the likelihood of Y under the assumption that Y has a multivariate normal distribution. Under this assumption, the maximum likelihood method can be viewed as a two-stage procedure where the variance components are first estimated by maximum likelihood and $\pi'\beta$ is then estimated by its BLUE under the assumption that $\theta_1$ and $\theta_2$ have these estimated values.

The other traditional methods are also two-stage procedures. The variance components are first estimated by some method. $\pi'\beta$ is then estimated by its BLUE under the assumption that $\theta_1$ and $\theta_2$ have these estimated values.

Kackar and Harville (1981) show that these two-stage procedures give unbiased estimators for $\pi'\beta$ provided the distribution of Y is symmetric about its expected value, the variance component estimators are translation-invariant and are even functions of Y, and the resulting estimator of $\pi'\beta$ has finite expectation.

In that same paper they show, in particular, that Henderson's (1953) methods 1, 2, and 3 yield variance component estimators that are even and translation-invariant. They also show that the maximum likelihood (ML) and the Patterson and Thompson (1974) restricted maximum likelihood (REML) estimators are even and translation-invariant under the assumption that Y has a multivariate normal distribution. Finally, they point out that the locally minimum variance quadratic unbiased translation-invariant estimators (MIVQUEs) introduced by LaMotte (1973) and Rao (1971b and 1972) and the minimum norm quadratic unbiased translation-invariant estimators (MINQUEs) introduced by Rao (1970, 1971a, and 1972) are even, translation-invariant estimators of the variance components by their very definition. The same is true of those estimators considered by Olsen, Seely, and Birkes (1976).

The estimator of $\pi'\beta$ in a two-stage estimation procedure will have finite expectation provided that the ratio $\tilde{\gamma} = \tilde{\theta}_1/(\tilde{\theta}_1 + \tilde{\theta}_2)$ is restricted to be nonnegative. The ML and REML variance component estimators are, by definition, nonnegative. The Henderson, MIVQUE, and MINQUE estimators may be negative, but a truncated version of $\tilde{\gamma}$ may be used to guarantee the finite expectation of the two-stage estimator of $\pi'\beta$.

We conclude that the traditional two-stage estimators of $\pi'\beta$ in model (1.1) or (1.2) are unbiased, provided they are constructed to have finite expectation.

More generally, let $Y$ follow model (1.1) or (1.2) where $Y$ is symmetrically distributed about its mean $X\beta$ and let the random matrix $W(Y)$ and the random variable $t(Y)$ satisfy

$$W(Y + Xh) = W(Y) = W(-Y) \quad , \tag{7.1}$$

$$t(Y + Xh) = t(Y) + \pi'h \quad \text{and} \quad t(Y) = -t(-Y) \quad . \tag{7.2}$$

Then Seely and Hogg (1982) show that $t(Y)$ is symmetrically distributed about $\pi'\beta$. If the expectations exist, $t(Y)$ is unbiased for $\pi'\beta$ and is uncorrelated with each element of $W(Y)$.

If $\Lambda$ is a p x m matrix such that $\Lambda'\beta$ is a vector of estimable parametric functions with respect to model (1.2) and $W(Y)$ satisfies (7.1) and is nonsingular with probability one, then Seely and Hogg (1982) remark that the random vector

$$T(Y) = \Lambda'(X'W(Y)X)^{-}X'W(Y)Y$$

is symmetrically distributed about $\Lambda'\beta$. If the expectations exist, T(Y) is unbiased for $\Lambda'\beta$ and each element of T(Y) is uncorrelated with each element of W(Y).

When the traditional two-stage estimators of an estimable $\pi'\beta$ are used, the data vector Y is usually assumed to have a multivariate normal distribution. We make this assumption throughout the rest of this chapter. The following discussion of efficiencies and admissibility of ALUEs of $\pi'\beta$ within the class of unbiased estimators of $\pi'\beta$ under this assumption parallels that found in Birkes, Seely, and Azzam (1981).

Lemma 7.1: Consider model (1.2) where $\underline{r}(X,V_1,V_2) = n$. Let $\pi'\hat{\beta}(\gamma)$ denote the unique LUE of the parametric function $\pi'\beta$ that is best when $\gamma = \theta_1/(\theta_1 + \theta_2)$ is known. If Y has a multivariate normal distribution, then the variance of $\pi'\hat{\beta}(\gamma)$ is the lower bound for the variances of all unbiased estimators of $\pi'\beta$.

Proof: Suppose $\widetilde{\pi'\beta}$ is an unbiased estimator of $\pi'\beta$. Let $\gamma = \gamma_0 \in [0,1]$ be fixed. Then $\widetilde{\pi'\beta}$ is unbiased with respect to the smaller family of multivariate normal distributions for Y with expectation $X\beta$ and covariance $\theta[\gamma_0 V_1 + (1 - \gamma_0)V_2]$, $\theta \geq 0$. For this smaller family of distributions, $\pi'\hat{\beta}(\gamma_0)$ is the unique minimum variance unbiased estimator of $\pi'\beta$, hence $\mathrm{var}(\widetilde{\pi'\beta}\,|\gamma_0,\theta) \geq \mathrm{var}(\pi'\hat{\beta}(\gamma_0)\,|\gamma_0,\theta)$ with strict inequality unless $\widetilde{\pi'\beta} = \pi'\hat{\beta}(\gamma_0)$ with probability one. Since $\gamma_0$ was chosen arbitrarily within [0,1], the lemma follows. ∎

Throughout this thesis the efficiency $\mathrm{Eff}(c|\gamma)$ of the ALUE $\pi'\hat{\beta}(c)$ has been relative to the variance of the ALUE $\pi'\hat{\beta}(\gamma)$ that is best when $\gamma$ is known. But by Lemma 7.1, this variance is the greatest lower bound of the variances of all unbiased estimators of $\pi'\beta$. Hence the efficiency $\mathrm{Eff}(c|\gamma)$ of $\pi'\hat{\beta}(c)$ is with respect to all unbiased estimators.

Furthermore, since $\pi'\hat{\beta}(c)$ is the unique (with probability one) estimator whose variance attains the lower bound when $\gamma = c$, an ALUE of $\pi'\beta$ is also admissible within the class of unbiased estimators of $\pi'\beta$.

# BIBLIOGRAPHY

Azzam, Abdul-Mordy Hamed. 1981. Admissibility in linear models. Ph.D. dissertation, Oregon State University.

Birkes, David, Yadolah Dodge and Justus Seely. 1976. Spanning sets for estimable contrasts in classification models. Ann. Statist. 4:86-107.

Birkes, David, Justus Seely and Abdul-Mordy Azzam. 1981. An efficient estimator of the mean in a two-stage nested model. Technometrics 23:143-148.

Ferguson, T. 1967. Mathematical Statistics. Academic Press, New York.

Fertig, K. W. and N. R. Mann. 1974. Population percentile estimation for the unbalanced random-effect model: estimation of the strength of a material when sampling from various batches of the material. In Reliability and Biometry, eds. F. Proschan and R. J. Serfling. Society for Industrial and Applied Mathematics, Philadelphia.

Harville, D. A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Amer. Statist. Assoc. 72:320-338.

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

Kackar, R. N. and D. A. Harville. 1981. Unbiasedness of two-stage estimation and prediction procedures for mixed linear models. Commun. Statist. - Theor. Meth. A10(13):1249-1261.

LaMotte, Lynn Roy. 1973. Quadratic estimation of variance components. Biometrics 29:311-330.

LaMotte, Lynn Roy. 1977. On admissibility and completeness of linear unbiased estimators in a general linear model. J. Amer. Statist. Assoc. 72:438-441.

LaMotte, Lynn Roy. 1982. Admissibility in linear estimation. Ann. Statist. 10:245-255.

LaMotte, L. R. and A. McWhorter. 1978. An exact test for the presence of random walk coefficients in a linear regression model. J. Amer. Statist. Assoc. 73:816-820.

Olsen, Anthony Robert. 1973. Quadratic unbiased estimation for two variance components. Ph.D. dissertation, Oregon State University.

Olsen, Anthony, Justus Seely and David Birkes. 1976. Invariant quadratic unbiased estimation for two variance components. Ann. Statist. 4:878-890.

Patterson, H. D. and R. Thompson. 1974. Maximum likelihood estimation of components of variance. Proc. Eighth International Biometric Conf. 197-207.

Rao, C. R. 1970. Estimation of heteroscedastic variances in linear models. J. Amer. Statist. Assoc. 65:161-172.

Rao, C. R. 1971a. Estimation of variance and covariance components - MINQUE theory. J. Mult. Anal. 1:257-275.

Rao, C. R. 1971b. Minimum variance quadratic unbiased estimation of variance components. J. Mult. Anal. 1:445-456.

Rao, C. R. 1972. Estimation of variance and covariance components in linear models. J. Amer. Statist. Assoc. 67:112-115.

Rao, C. R. and S. K. Mitra. 1971. Generalized Inverse of Matrices and Its Applications. Wiley, New York.

Seely, Justus. 1970. Linear spaces and unbiased estimation. Ann. Math. Statist. 41:1725-1734.

Seely, Justus. 1979. Parametrizations and correspondences in linear models. Dept. of Statistics Tech. Report #72, Oregon State University.

Seely, Justus. 1983a. Private communication.

Seely, Justus. 1983b. Some continuity properties of dispersion matrices of estimators in linear models. Dept. of Statistics Tech. Report #88, Oregon State University.

Seely, J. and R. V. Hogg. 1982. Symmetrically distributed and unbiased estimators in linear models. Commun. Statist. - Theor. Meth. A11(7):721-729.

Seely, Justus and George Zyskind. 1971. Linear spaces and minimum variance unbiased estimation. Ann. Math. Statist. 42:691-703.

Snedecor, G. W. and W. G. Cochran. 1967. Statistical Methods. Iowa State University Press, Ames.

Watson, G. S. 1967. Linear least squares regression. Ann. Math. Statist. 38:1679-1699.

Zyskind, George. 1967. On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. Ann. Math. Statist. 38:1092-1109.