Statistics and Applications {**ISSN 2452-7395 (online)**}
Volume 16 No. 1, 2018 (New Series), pp 91-103

Revisiting a General Pattern of Domination Using the Graybill–Deal Estimator

A.S. Hedayat, Bikas.K. Sinha and Keyu Nie
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago

Received: November 24, 2017; Revised: December 10, 2017; Accepted: January 05, 2018

Abstract

This is a review article in a popular area of research which Professor J.N.K. Rao had touched upon - during his distinguished research career. We convey our congratulations and best wishes to JNK Rao on his attaining 80 yrs. We wish him a long, active and peaceful life in the years to come.

Herein we review the problem of unbiased estimation of the common parameter (θ) involved in the linear regression models of the means of two independent normal populations with unequal and unknown variances. We examine the popular Graybill–Deal estimator for the common parameter θ and ask the question: When will the Graybill–Deal estimator possess uniformly smaller variance than that of the individual unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ arising out of the two given models? It turns out that the result depends only on the quantities m_1 and m_2, corresponding to the error degrees of freedom, irrespective of the nature of the linear mean models. We find the same result continues to hold in situations wherein the p ($p \geq 2$) linear regression models involve k ($k > 1$) common estimable parameter(s) in the mean models. In this context, we use the criterion of 'Loewner Order Domination' of information or dispersion matrices.

Key words: Common Parameter Estimation; Graybill-Deal Estimator; Loewner Order Domination; Linear Regression

1 Introduction

The common mean estimation problem was first introduced by Cochran (1937), while he was considering combining a series of similar experiments. The general setting for this kind of problem is: suppose we have p independent groups of normal variables with sample size n_i, for the i-th group, having the sample mean $\bar{x}_i \sim N(\mu, \frac{\sigma_i^2}{n_i})$, where $i = 1, 2, \ldots, p$. The setup presupposes that there is a common unknown mean μ for the p populations and the problem considered is that of efficient unbiased estimation of μ based on the data from the p groups.

For $p = 2$, Cochran (1937) suggested the unbiased estimator

Corresponding Author: Bikas K. Sinha
Email id: bikassinha1946@gmail.com

$$\hat{\mu}_C = \frac{\left(\bar{x}_1 \frac{n_1}{\sigma_1^2} + \bar{x}_2 \frac{n_2}{\sigma_2^2}\right)}{\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2}}. \tag{1.1}$$

This estimator is the best linear unbiased estimator for μ, assuming that the two variances are known. Motivated by Cochran's (1937) work, Graybill and Deal (1959) introduced their estimator $\hat{\mu}_{GD|2}$ by replacing the true variances with their corresponding unbiased estimators:

$$\hat{\mu}_{GD|2} = \frac{\left(\bar{x}_1 \frac{n_1}{s_1^2} + \bar{x}_2 \frac{n_2}{s_2^2}\right)}{\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2}}, \tag{1.2}$$

where $s_i^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$ for $i = 1, 2$. In view of distributional independence of $\bar{x}_i$ and s_i^2, $\hat{\mu}_{GD|2}$ is an unbiased estimator for μ. Furthermore, Graybill and Deal established that $\hat{\mu}_{GD|2}$ is uniformly superior to any single unbiased estimator of μ if and only if the following condition holds:

Either both n_1 and $n_2 > 10$ or $n_1 = 10$ $(n_2 = 10)$ and $n_2 > 18$ (respectively $n_1 > 18$) .

Norwood and Hinkelmann (1977) extended Graybill and Deal's (1959) result to general p groups, and they established that $\hat{\mu}_{GD|p}$ is a uniformly better estimator of μ than each $\bar{x}_i$ if and only if either $n_i > 10$ for $i = 1, 2, \ldots, p$ or $n_i = 10$ for some i, and $n_j > 18$ $(i, j = 1, 2, \ldots, p$ for each $j \neq i)$, where

$$\hat{\mu}_{GD|p} = \sum_{i=1}^{p} \frac{\bar{x}_i \frac{n_i}{s_i^2}}{\sum_{j=1}^{p} \frac{n_j}{s_j^2}}. \tag{1.3}$$

The properties of such estimators have been widely studied in the literature. In particular we would like to emphasize the work of Meier (1953), Cochran and Carroll (1953), Zacks (1966), Rao and Subrahmaniam (1971), Khatri and Shah (1974), Ghosh and Sinha (1981), Sinha (1985), Hartung (1999) and Krishnamoorthy and Moore (2002).

Shinozaki (1978) extended $\mu_{GD|p}$ to a general form:

$$\hat{\mu}_S = \sum_{i=1}^{p} \frac{\bar{x}_i \frac{c_i}{s_i^{2*}}}{\sum_{j=1}^{p} \frac{c_j}{s_j^{2*}}}, \tag{1.4}$$

where $s_i^{2*} = \frac{s_i^2}{n_i}$. By a careful choice of $(c_1, c_2, \ldots, c_p)$, Shinozaki gave a proof of the claim that $\hat{\mu}_{S,p}$ (same as $\hat{\mu}_S$ in 1.4, and similar to $\hat{\mu}_{GD|p}$ in 1.3) combining p groups is a uniformly better estimator of μ than any $\hat{\mu}_{S,q}$ of combining q $(< p)$ components, if and only if $\frac{c_j}{c_i} \leq 2\frac{(n_i-1)(n_j-5)}{(n_i+1)(n_j-1)}$ for any $1 \leq i \neq j \leq p$ and for all $1 \leq q < p$. It is readily verified that when our choice of $(c_1, c_2, \ldots, c_p)$ corresponds to $(c_1 = c_2 = \cdots = c_p)$, the condition above simplifies to what is stated earlier involving only the sample sizes. Chiou and Cohen (1984), Loh (1991), Kubokawa (1989), Tsukuma and Konno (2003) investigated the multivariate normal counterpart of this problem.

With reference to the general framework of common parameter estimation problem, we observe that JNK Rao and Subrahmaniam (1971) examined the status of Graybill-Deal Type Estimators in the above formulations of (i) common mean estimation in normal samples and (ii) common parameters estimation in linear regression models as also (iii) predictions in such models - all with heterogeneous variances, when the variance components are estimated using MINQUE of Rao (1970).

Much later, the problem in the context of linear regression models was revisited by Krishnamoorthy and Moore (2002)and that work is the basis of our present study. We rework on and review the problem in linear regression set-up and discuss some generalizations.

2 Common Parameter Estimation in General linear regression models with Independent Normal Errors

In environmental pollution studies, in order to understand the environmental factors affecting the mean 'contamination/pollution level' of air/water/land, representative samples are sent to different laboratories for statistical analyses. This corresponds to what is technically addressed as 'Meta Analysis' problem. All the studies in different laboratories have a common goal viz., estimation and assessment of global contamination level in the experimental region. At times, linear or quadratic or higher degree regression models are adequate with(without) common intercept term and/or common slope. Of course, the laboratories are likely to have instruments with different precision levels. In such situations, we call for natural application of Graybill-Deal Type estimators. Our purpose in this section is to examine the effectiveness of such estimators.

In this paper we revisit a linear regression set-up with common parameter. We adopt the general approach of formation of Graybill-Deal - type estimators in such set-ups and then examine conditions for their superiority over corresponding estimators based on partial exposure to the entire body of data. So far in the literature related to Graybill-Deal estimator (GDE), the standard set-up of estimation of normal common mean has been investigated. There are exceptions such as Kubokawa (1989).
Consider p independent linear regression models in matrix form, with sample size n_i, $i = 1, 2, \ldots, p$:

$$\boldsymbol{Y}_i = \boldsymbol{X}_{i(n_i \times k)} \boldsymbol{\theta}_{(k \times 1)} + \boldsymbol{Z}_{i(n_i \times t_i)} \boldsymbol{\tau}_{\boldsymbol{i}(t_i \times 1)} + \boldsymbol{\epsilon}_i$$
$$\boldsymbol{X}_{i(n_i \times k)} = \begin{pmatrix} x_{i11} & x_{i21} & \ldots & x_{ik1} \\ x_{i12} & x_{i22} & \ldots & x_{ik2} \\ \ldots & \ldots & \ldots & \ldots \\ x_{i1n_i} & x_{i2n_i} & \ldots & x_{ikn_i} \end{pmatrix}$$
$$\boldsymbol{\epsilon}_i \sim N\left(\mathbf{0}_{(\boldsymbol{n_i} \times \mathbf{1})}, \sigma_i^2 \mathbf{I}_{(n_i \times n_i)}\right).$$

We assume that these p independent linear regression models share a k-dimensional (k-dim) common estimable parameter vector $\boldsymbol{\theta}_{(k \times 1)}$, and an extra t_i-dimensional estimable parameter vector τ_i at i-th model, $i = 1, 2, \ldots, p$. From the normal equations, the OLS unbiased estimator, $\hat{\boldsymbol{\theta}}_{i(k \times 1)}$, of

$\boldsymbol{\theta}_{(k\times 1)}$ based on the data arising out of the i-th model, has the following distribution

$$\hat{\boldsymbol{\theta}}_{i(k\times 1)} \sim N\left(\boldsymbol{\theta}_{(k\times 1)}, \sigma_i^2 \boldsymbol{W}_{i(k\times k)}\right),$$

where

$$\boldsymbol{W}_{i(k\times k)} = \left(\left(\boldsymbol{X}_{i(n_i\times k)} \;\vdots\; \boldsymbol{Z}_{i(n_i\times t_i)}\right)^T \left(\boldsymbol{X}_{i(n_i\times k)} \;\vdots\; \boldsymbol{Z}_{i(n_i\times t_i)}\right)\right)^{-1}_{(k\times k)},$$

$\boldsymbol{W}_{i(k\times k)}$'s are nonsingular matrices.

An unbiased estimator for σ_i^2 is the mean square residuals s_{ri}^2, which is the sum of square residuals divided by the degrees of freedom $\nu_i = n_i - k - t_i$. Further, it is known that $\frac{\nu_i s_{ri}^2}{\sigma_i^2} \sim \chi^2(\nu_i)$.

In the sequel, we shall deal with the estimation of the common $\boldsymbol{\theta}_{(k\times 1)}$ under three different scenarios: in Section 2.1, we shall study the case $k = 1$ for $p \geq 2$ groups; in Section 2.2, we shall study the case $k \geq 2$ for $p = 2$ groups; and in Section 2.3, we shall study the case $k \geq 2$ for $p > 2$ groups. Krishnamoorthy and Moore (2002) had deduced some such results.

2.1 Single Common Parameter Involving p Groups

Suppose there is only one single common parameter, α, among these p independent linear regression models. As mentioned before, under certain conditions, the GDE will be efficient. Here we use a simple example to express our ideas by assuming that common α is the intercept in two linear regression models.

Example 2.1. *Consider two simple independent linear regression models involving unequal unknown variances with common intercept. We have*

$$\begin{aligned}
\mathbf{Y}_1 &= \alpha \boldsymbol{1}_{n_1\times 1} + \beta_1 \mathbf{X}_{n_1\times 1} + \boldsymbol{\epsilon}_1 \\
\mathbf{Y}_2 &= \alpha \boldsymbol{1}_{n_2\times 1} + \beta_2 \mathbf{Z}_{n_2\times 1} + \boldsymbol{\epsilon}_2 \\
\boldsymbol{\epsilon}_1 &\sim N(\boldsymbol{0}_{n_1\times 1}, \sigma_1^2 \mathbf{I}_{n_1\times n_1}) \\
\boldsymbol{\epsilon}_2 &\sim N(\boldsymbol{0}_{n_2\times 1}, \sigma_2^2 \mathbf{I}_{n_2\times n_2}).
\end{aligned}$$

The OLS estimators $(\hat{\alpha}_1, \hat{\beta}_1)$ and $(\hat{\alpha}_2, \hat{\beta}_2)$, respectively for (α, β_1) and (α, β_2), have the following variance-covariance matrices:

$$\mathbf{Var}\binom{\hat{\alpha}_1}{\hat{\beta}_1} = \sigma_1^2 \begin{pmatrix} n_1 & \sum_{i=1}^{n_1} x_i \\ \sum_{i=1}^{n_1} x_i & \sum_{i=1}^{n_1} x_i^2 \end{pmatrix}^{-1} \text{ and } \mathbf{Var}\binom{\hat{\alpha}_2}{\hat{\beta}_2} = \sigma_2^2 \begin{pmatrix} n_2 & \sum_{i=1}^{n_2} z_i \\ \sum_{i=1}^{n_2} z_i & \sum_{i=1}^{n_2} z_i^2 \end{pmatrix}^{-1}.$$

We also have $E(\hat{\alpha}_1) = E(\hat{\alpha}_2) = \alpha$, $Var(\hat{\alpha}_1) = \sigma_1^2(\frac{1}{n_1} + \frac{\bar{x}^2}{SS_x})$ and $Var(\hat{\alpha}_2) = \sigma_2^2(\frac{1}{n_2} + \frac{\bar{z}^2}{SS_z})$ where $SS_x = \sum_{i=1}^{n_1}(x_i - \bar{x})^2$, and $SS_z = \sum_{i=1}^{n_2}(z_i - \bar{z})^2$ respectively.

If A is a matrix of order $p\times q$ and B is another matrix of order $p\times r$, then $(\boldsymbol{A}\vdots\boldsymbol{B})$ represents a matrix of order $p\times(q+r)$, wherein the columns of A are preceded by the columns of B without any change of their relative positions. In the above, $W_{i(k\times k)}$ represents the $k\times k$ upper submatrix of the right hand side (RHS) matrix of order $(k+t_i)\times(k+t_i)$ in each block matrix representation. Note that the RHS matrix involves inversion of a square matrix, assumed to be positive definite.

It is tacitly assumed that all nuisance parameters (τ_i's) are estimable.

Denote $\frac{1}{n_1} + \frac{\bar{x}^2}{SS_x}$ *by* w_1^{-1}*, and* $\frac{1}{n_2} + \frac{\bar{z}^2}{SS_z}$ *by* w_2^{-1}*. The GDE of* α *is*

$$\hat{\alpha}_{GD|2} = \frac{\frac{\hat{\alpha}_1}{s_{r1}^2 w_1^{-1}} + \frac{\hat{\alpha}_2}{s_{r2}^2 w_2^{-1}}}{\frac{1}{s_{r1}^2 w_1^{-1}} + \frac{1}{s_{r2}^2 w_2^{-1}}}.$$

We can easily extend the formula for $\hat{\alpha}_{GD|2}$ to

$$\hat{\alpha}_{GD|p} = \sum_{i=1}^{p} \frac{\hat{\alpha}_i \frac{w_i}{s_{ri}^2}}{\sum\limits_{i=1}^{p} \frac{w_i}{s_{ri}^2}}$$

in case there are p such models to be combined. It is to be noted that $w_1, w_2, \ldots, w_p$ have similar algebraic expressions. At this stage, we will state and prove a general result on the property of $\hat{\alpha}_{GD|p}$. By a simple application of the results, from Norwood and Hinkelmann (1977), and Shinozaki (1978), we have the following Theorem 2.1 for p independent linear regression models sharing one single common intercept parameter α. We may note in passing that ν_i defined above in the beginning of Section 2 assumes the form $\nu_i = n_i - 2$ for the model in Example 2.1 being studied.

Theorem 2.1. *If* p *independent linear regression models share one single common intercept parameter* α*,* $i = 1, 2, \ldots, p$*, then we have the following results:*

1. $\hat{\alpha}_{GD|p}$ *is an unbiased estimator for* α*.*
2. *A necessary and sufficient condition for* $\hat{\alpha}_{GD|p}$ *to have a smaller variance than each* $\hat{\alpha}_i$*, for all values of* σ_i^2*,* $w_i > 0$ *(*$i = 1, \ldots, p$*) is either*
   (a) $\nu_i > 9$*,* $i = 1, 2, \ldots, p$ *or*
   (b) $\nu_i = 9$ *for some i, and* $\nu_j > 17$ *for* $i, j \in \{1, 2, \ldots, p\}$ *and each* $j \neq i$*.*

Moreover, if either condition $(2a)$ *or* $(2b)$ *is satisfied, then* $\hat{\alpha}_{GD|p}$= $\phi_p(\hat{\alpha}_1, \ldots, \hat{\alpha}_p; s_1^2, \ldots, s_p^2)$ *has a smaller variance than any of the* q $(< p)$ *subgroups* $\hat{\alpha}_{GD|q} = \phi_q(\hat{\alpha}_1, \ldots, \hat{\alpha}_q; s_1^2, \ldots, s_q^2)$*.*

Proof. Set $\sigma_i^{2\star} = \frac{\sigma_i^2}{w_i}$ and $s_{ri}^{2\star} = \frac{s_{ri}^2}{w_i}$ for $i = 1, 2, \ldots, p$. We have $\hat{\alpha}_i \sim N(\alpha, \sigma_i^{2\star})$ and $\frac{\nu_i s_{ri}^{2\star}}{\sigma_i^{2\star}} = \frac{\nu_i s_{ri}^2}{\sigma_i^2} \sim \chi^2(\nu_i)$.

Then GDE $\hat{\alpha}_{GD|p} = \sum\limits_{i=1}^{p} \frac{\hat{\alpha}_i \frac{w_i}{s_{ri}^2}}{\sum_{i=1}^{p} \frac{w_i}{s_{ri}^2}} = \sum\limits_{i=1}^{p} \frac{\frac{\hat{\alpha}_i}{s_{ri}^{2\star}}}{\sum_{i=1}^{p} \frac{1}{s_{ri}^{2\star}}}$.

This is the same setting as in Norwood and Hinkelmann (1977). It is easy to show that $\hat{\alpha}_{GD|p}$ is an unbiased estimator for α, and $Var(\hat{\alpha}_{GD|p}) < \sigma_i^{2\star} = \frac{\sigma_i^2}{w_i}$ $\forall i$ if and only if either Condition (2a) or (2b) holds.

Furthermore, we notice that our $\hat{\alpha}_{GD|p}$ is a special case of $\hat{\alpha}_S$ at $c_i = 1$ for $i = 1, \ldots, p$. From Shinozaki(1978)'s results, we know that $Var(\hat{\alpha}_{GD|p}) < Var(\hat{\alpha}_{GD|q})$, $q < p$ if and only if $2\frac{\nu_i(\nu_j-4)}{(\nu_i+2)\nu_j} \geq \frac{c_j}{c_i} = 1$ for any $i \neq j$. This is equivalent to condition (2a) or (2b) as stated above. □

Remark 1. The statistical independence of the $\hat{\alpha}_i$ and s_{ri}^2 guarantees $\hat{\alpha}_{GD|p}$ to be an unbiased estimator of α.

Remark 2. The necessary and sufficient condition in (2a) or (2b) only concerns the degrees of freedom ν_i in group i, which is related with the sample size n_i. Theorem 2.1 tells us that if the sample size is reasonable enough (subject to condition in (2a) or (2b)), then the GDE of a single common parameter utilizing p independent linear regression models always provides a more efficient unbiased estimator than any single group or any $q(< p)$ subgroups.

2.2 k-dim Common Parameter Involving Two Groups

In this section, we consider $p = 2$ independent linear regression models. The k-dim GDE of the common estimable parameter vector $\boldsymbol{\theta}_{(k\times 1)}$ in matrix forms is:

$$\hat{\boldsymbol{\theta}}_{GD|2} = \left((s_{r1}^2\boldsymbol{W_1})^{-1} + (s_{r2}^2\boldsymbol{W_2})^{-1}\right)^{-1}\left((s_{r1}^2\boldsymbol{W_1})^{-1}\hat{\boldsymbol{\theta}}_1 + (s_{r2}^2\boldsymbol{W_2})^{-1}\hat{\boldsymbol{\theta}}_2\right).$$

It is easily determined that this k-dim GDE is an unbiased estimator of $\boldsymbol{\theta}$. The dispersion matrix of $\hat{\boldsymbol{\theta}}_{GD|2}$ is:

$$\begin{aligned}
&\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})\\
&= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{E}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right)\\
&= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{\theta}\right)\\
&= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|2}|s_{r1}^2, s_{r2}^2\right)\right) + 0\\
&= \boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{W_1}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{W_2}^{-1}\right)\left(s_{r1}^{-2}\boldsymbol{W_1}^{-1} + s_{r2}^{-2}\boldsymbol{W_2}^{-1}\right)^{-1}\right).
\end{aligned}$$

In the above, we are conditionally fixing s_{r1}^2 and s_{r2}^2. So at the end, we only need to compute expectation with respect to these variance estimates.

By way of notation, if a dispersion matrix $\boldsymbol{A}$ is non-negative definite (n.n.d.) we write $\boldsymbol{A} \geq 0$, if it is positive definite (p.d.) we write $\boldsymbol{A} > 0$. The Loewner order domination of a dispersion matrix $\boldsymbol{A}$ over $\boldsymbol{B}$ ($\boldsymbol{A} > \boldsymbol{B}$) means $\boldsymbol{A} - \boldsymbol{B} > 0$, the Loewner order of $\boldsymbol{A}$ below $\boldsymbol{B}$ ($\boldsymbol{A} < \boldsymbol{B}$) means $\boldsymbol{A} - \boldsymbol{B} < 0$.

Lemma 2.2. *If $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ are diagonal matrices, then $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2}) < min(\sigma_1^2\boldsymbol{W}_1, \sigma_2^2\boldsymbol{W}_2)$ if and only if either*

1. $\nu_i > 9$, $i = 1, 2$ *or*
2. $\nu_i = 9$ *for some i, and* $\nu_j > 17$ *for the other* $j \neq i$.

Proof. Let $\boldsymbol{W_1} = \boldsymbol{Diag}(w_{1g}^{-1})$, and $\boldsymbol{W_2} = \boldsymbol{Diag}(w_{2g}^{-1})$, where w_{1g}^{-1} and w_{2g}^{-1} are the g-th diagonal entries of matrices $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ respectively, $g = 1, \ldots, k$. To prove this Lemma, it is enough to show the g-th diagonal entry of the dispersion matrix $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})_{gg} \leq min(\frac{\sigma_1^2}{w_{1g}}, \frac{\sigma_2^2}{w_{2g}})$ for all $g = 1, 2, \ldots, k$, and for all values of σ_1^2, σ_2^2; where $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})_{gg} = Var\left(\frac{\frac{\hat{\theta}_{1g}}{s_{r1}^2 w_{1g}^{-1}} + \frac{\hat{\theta}_{2g}}{s_{r2}^2 w_{2g}^{-1}}}{\frac{1}{s_{r1}^2 w_{1g}^{-1}} + \frac{1}{s_{r2}^2 w_{2g}^{-1}}}\right)$.

From Theorem 2.1, it is known that:

$$Var\left(\frac{\frac{\hat{\theta}_{1g}}{s_{r1}^2 w_{1g}^{-1}} + \frac{\hat{\theta}_{2g}}{s_{r2}^2 w_{2g}^{-1}}}{\frac{1}{s_{r1}^2 w_{1g}^{-1}} + \frac{1}{s_{r2}^2 w_{2g}^{-1}}}\right) < min\left(\frac{\sigma_1^2}{w_{1g}}, \frac{\sigma_2^2}{w_{2g}}\right).$$

if and only if condition stated in (1) or (2) of Lemma 2.2 holds. □

Remark 3. When $\boldsymbol{W_1}$ and $\boldsymbol{W_2}$ are k-dim diagonal matrices, we can decompose the k-dim GDE into k simple single GDEs.

It is pertinent to observe that the above result holds even without the two matrices being diagonal matrices. This is established below.

Theorem 2.3. *In two independent linear regression models, the Loewner order of $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|2})$ is below $min(\sigma_1^2\boldsymbol{W}_1, \sigma_2^2\boldsymbol{W}_2)$, for all values of σ_1^2, σ_2^2, if and only if condition stated in (1) or (2) of Lemma 2.2 holds.*

Proof. Notice that ${\boldsymbol{W_1}}^{-1}$ and ${\boldsymbol{W_2}}^{-1}$ are positive definite matrices. We need to show:

$$\begin{aligned}&\boldsymbol{E}\left(\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}{\boldsymbol{W_1}}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}{\boldsymbol{W_2}}^{-1}\right)\right.\\ &\left.\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}\right) < \sigma_1^2\boldsymbol{W}_1,\end{aligned} \tag{2.1}$$

and

$$\begin{aligned}&\boldsymbol{E}\left(\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}{\boldsymbol{W_1}}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}{\boldsymbol{W_2}}^{-1}\right)\right.\\ &\left.\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}\right) < \sigma_2^2\boldsymbol{W}_2.\end{aligned} \tag{2.2}$$

For inequality (2.1), we can obtain

$$\begin{aligned}&\boldsymbol{E}\left({\boldsymbol{W_1}}^{-1/2}\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}{\boldsymbol{W_1}}^{-1/2}{\boldsymbol{W_1}}^{1/2}\left(\frac{\sigma_1^2}{s_{r1}^4}{\boldsymbol{W_1}}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}{\boldsymbol{W_2}}^{-1}\right)\right.\\ &\left.\left({\boldsymbol{W_1}}^{1/2}{\boldsymbol{W_1}}^{-1/2}\left(s_{r1}^{-2}{\boldsymbol{W_1}}^{-1} + s_{r2}^{-2}{\boldsymbol{W_2}}^{-1}\right)^{-1}{\boldsymbol{W_1}}^{-1/2}\right)\right) < \sigma_1^2{\boldsymbol{W_1}}^{-1/2}\boldsymbol{W}_1{\boldsymbol{W_1}}^{-1/2}.\end{aligned} \tag{2.3}$$

Denote ${\boldsymbol{W_1}}^{-1/2}\boldsymbol{W}_2{\boldsymbol{W_1}}^{-1/2}$ by $\boldsymbol{A}$. Then inequality (2.3) can be rewritten as:

$$\boldsymbol{E}\left(\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{A}^{-1}\right)\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\right) < \sigma_1^2\boldsymbol{I}.$$

Due to the fact that $\boldsymbol{A}$ is symmetric, there exists an orthogonal matrix $\boldsymbol{P}$, while $\boldsymbol{P}^T\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}^T = I$, such that $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P} = \boldsymbol{C}$, where $\boldsymbol{C}$ is a diagonal matrix.
Then we have:

$$\boldsymbol{E}\left(\boldsymbol{P}^T\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\boldsymbol{P}\boldsymbol{P}^T\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{A}^{-1}\right)\boldsymbol{P}\boldsymbol{P}^T\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{A}^{-1}\right)^{-1}\boldsymbol{P}\right) < \sigma_1^2\boldsymbol{P}^T\boldsymbol{I}\boldsymbol{P}.$$

Upon simplification:

$$\begin{aligned}\boldsymbol{E}\Bigg(&\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\\ &\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\Bigg) < \sigma_1^2\boldsymbol{I}.\end{aligned} \tag{2.4}$$

Similarly, from (2.2) we obtain:

$$\begin{aligned}\boldsymbol{E}\Bigg(&\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\\ &\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\Bigg) < \sigma_2^2\boldsymbol{C}.\end{aligned} \tag{2.5}$$

By combining (2.4) and (2.5), we have:

$$\begin{aligned}\boldsymbol{E}\Bigg(&\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\left(\frac{\sigma_1^2}{s_{r1}^4}\boldsymbol{I}^{-1} + \frac{\sigma_2^2}{s_{r2}^4}\boldsymbol{C}^{-1}\right)\\ &\left(s_{r1}^{-2}\boldsymbol{I}^{-1} + s_{r2}^{-2}\boldsymbol{C}^{-1}\right)^{-1}\Bigg) < min\left(\sigma_1^2\boldsymbol{I}, \sigma_2^2\boldsymbol{C}\right).\end{aligned} \tag{2.6}$$

Note that (2.6) exhibits a pattern of the comparison of the GDE against individual estimators based on two diagonal matrices viz., identity matrix and the C matrix. This is exactly the same formulation as in Lemma 2 above. Hence the result follows by an application of Lemma 2.2. □

2.3 General k-dim Common Parameter Involving p Independent Groups

Next we consider the general case of p (> 2) independent groups of linear regression models sharing a k-dim common estimable parameter vector $\boldsymbol{\theta}$. The k-dim GDE of p groups is:

$$\hat{\boldsymbol{\theta}}_{GD|p} = \left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\hat{\boldsymbol{\theta}}_i\right).$$

Again, it is easy to show $\hat{\boldsymbol{\theta}}_{GD|p}$ is an unbiased estimator of $\boldsymbol{\theta}$, with the following dispersion matrix:

$$\begin{aligned}\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) &= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) + \boldsymbol{D}\left(\boldsymbol{E}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \\ &= \boldsymbol{E}\left(\boldsymbol{D}\left(\hat{\boldsymbol{\theta}}_{GD|p}|s_{r1}^2, s_{r2}^2, \ldots, s_{rp}^2\right)\right) \\ &= \boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right).\end{aligned}$$

Lemma 2.4. *If $\boldsymbol{W_i}$'s are diagonal matrices $i = 1, 2, \ldots, p$, then $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) < \min_{i\in\{1,\ldots,p\}}(\sigma_i^2\boldsymbol{W_i})$, for all values of σ_i^2, if and only if condition in (2a) or (2b) of Theorem 2.1 holds.*

Moreover, if condition in (2a) or (2b) of Theorem 2.1 is satisfied, the Loewner order of dispersion matrix, $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ is below $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q})$ for any $q(< p)$ subgroups.

Proof. This is the extension of Lemma 2.2. In case that $\boldsymbol{W_i}$ (i=1,2,...,p)'s are diagonal matrices, that can be decomposed into k single GDE of p groups. Our claim follows from an application of Theorem 2.1. □

If not all $\boldsymbol{W_i}$'s are diagonal matrices, $i \in \{1, 2, \ldots, p\}$, we have the following Theorem.

Theorem 2.5. *Suppose there exists a nonsingular matrix $\boldsymbol{P}$, such that $\boldsymbol{W_i^{-1}} = \boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T$, where $\boldsymbol{C_i}$'s are diagonal matrices for all $i = 1, 2, \ldots, p$. Then the Loewner order of $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ is below $\boldsymbol{D}(\boldsymbol{\theta}_{GD|q})$ for any $q(< p)$ subgroups, for all values of σ_i^2, if and only if condition in (2a) or (2b) of Theorem 2.1 holds.*

Proof. We need to examine the validity of

$$\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p}) < \boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q}), q < p.$$

Equivalently,

$$\begin{aligned}&\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right) \\ &< \boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{W_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{W_i}^{-1}\right)^{-1}\right).\end{aligned}$$

Since $\boldsymbol{W_i^{-1}} = \boldsymbol{P}\boldsymbol{C_i^{-1}}\boldsymbol{P}^T$,then

$$\begin{aligned}&\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\right) \\ &< \boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{P}\boldsymbol{C_i}^{-1}\boldsymbol{P}^T\right)^{-1}\right).\end{aligned}$$

Upon simplification:

$$\boldsymbol{E}\left(\boldsymbol{P}^{T^{-1}}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\boldsymbol{P}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\boldsymbol{P}^T\boldsymbol{P}^{T^{-1}}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right)$$
$$< \boldsymbol{E}\left(\boldsymbol{P}^{T^{-1}}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\boldsymbol{P}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\boldsymbol{P}^T\boldsymbol{P}^{T^{-1}}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right).$$

This reduces to:

$$\boldsymbol{E}\left(\boldsymbol{P}^{-1^T}\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right)$$
$$< \boldsymbol{E}\left(\boldsymbol{P}^{-1^T}\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\right),$$

which requires

$$\boldsymbol{E}\left(\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{p}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{p} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\right)$$
$$< \boldsymbol{E}\left(\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\left(\sum_{i=1}^{q}\frac{\sigma_i^2}{s_{ri}^4}\boldsymbol{C_i}^{-1}\right)\left(\sum_{i=1}^{q} s_{ri}^{-2}\boldsymbol{C_i}^{-1}\right)^{-1}\right).$$

Since $\boldsymbol{C_i}'s$ are diagonal matrices, the results follows by an application of Lemma 2.4. □

Remark 4. Generally, the existence of such a nonsingular matrix $\boldsymbol{P}$ that diagonalizes all $\boldsymbol{W_i}$ simultaneously is not guaranteed. However, Corollary 2.6 below provides a special case.

Suppose in the most general representation of the linear regression model described in the beginning of Section 2, $\boldsymbol{Z}_{i(n_i\times t_i)}$ does not exist, which indicates that all these p independent groups of linear regression models are following the same linear regression. In such a case, $\boldsymbol{W}_{i(k\times k)} = (\boldsymbol{X}_i^T\boldsymbol{X}_i)^{-1}$.

Corollary 2.6. *Suppose $\boldsymbol{Z}_{i(n_i\times t_i)}$'s do not exist. For p independent groups of linear regression models sharing a 2-dim common estimable parameter $\boldsymbol{\theta} = \binom{\theta_1}{\theta_2}$, if $\frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)}$=constant, $i = 1, 2, \ldots, p$, then a necessary and sufficient condition for the Loewner order of $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|p})$ to be below $\boldsymbol{D}(\hat{\boldsymbol{\theta}}_{GD|q})$ for any $q(< p)$ subgroups, for all values of σ_i^2, is that the condition in (2a) or (2b) of Theorem 2.1 holds.*

Proof. From the linear regression theory, we know:

$$
\begin{aligned}
\boldsymbol{W_i}^{-1} &= \begin{pmatrix} \sum_{j=1}^{n_i} x_{i1j}^2 & \sum_{j=1}^{n_i} x_{i1j}x_{i2j} \\ \sum_{j=1}^{n_i} x_{i1j}x_{i2j} & \sum_{j=1}^{n_i} x_{i2j}^2 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} & 1 \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{n_i} x_{i1j}^2 & 0 \\ 0 & \sum_{j=1}^{n_i} x_{i2j}^2 - \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)^2}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} \end{pmatrix} \\
&\quad \begin{pmatrix} 1 & \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

If $\frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} = b$, a constant, then let $\boldsymbol{P} = \begin{pmatrix} 1 & 0 \\ \frac{\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right)}{\left(\sum_{j=1}^{n_i} x_{i1j}^2\right)} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$.

Hence from Theorem 2.5, the result follows. □

Remark 5. When this 2-dim common estimable parameter of interest $\boldsymbol{\theta} = \binom{\theta_1}{\theta_2}$ contains intercept, viz. θ_1 , then it asks $\sum_{j=1}^{n_i} x_{i1j}x_{i2j} / \sum_{j=1}^{n_i} x_{i1j}^2 = \bar{x}_{i2\cdot}$, which is the sample mean, to be constant.

Remark 6. The ratio $\left(\sum_{j=1}^{n_i} x_{i1j}x_{i2j}\right) / \left(\sum_{j=1}^{n_i} x_{i1j}^2\right) = \frac{\|\vec{x}_{i2\cdot} \cos\theta\|}{\|\vec{x}_{i1\cdot}\|}$, where θ is the angle between two covariate variable vectors $\vec{x}_{i1\cdot}$ and $\vec{x}_{i2\cdot}$. This suggests that, typically in non-intercept linear models, if we can pre-select our $\vec{x}_{i1\cdot}$ and $\vec{x}_{i2\cdot}$ to make this ratio constant, then we can obtain a more efficient estimator of this common parameter of interest by utilizing all p models, as long as we can collect enough observations for each linear model.

3 Conclusion

We revisited the properties of GDE in two and higher dimensions. We argued that GDE is still an unbiased estimator for the vector parameter of interest. We also found that the condition (2a) or (2b) in Theorem 2.1 (condition (1) or (2) in Lemma 2.2 for groups of two) continued to hold when estimating a k-dim common parameter vector for p independent groups of linear regression models. Consequently the GDE of k-dim common parameters by combining these p groups provides a better and more efficient estimator.

In the linear regression model that we have revisited here, we tacitly assumed that the regression parameters are fixed and unknown. In the literature there are studies on what are called 'random coefficient regression models', such as Carter and Yang (1986) and Liski et al. (1996). We may postulate a model with fixed unknown α parameter (the intercept) but the coefficients are random. The problem of estimation of the common mean α in such scenarios is a rather routine exercise. We propose to examine the domination results in such scenarios in a subsequent communication.

Acknowledgements

Research is supported by the U.S. National Science Foundation Grant DMS-1306394. This work was undertaken during regular visits of Bikas.K. Sinha at University of Illinois, Chicago and completed during the summer of 2017.

References

Carter, R.L., and Yang, M.C.K. (1986). Large sample inference in random coefficient regression models. *Communications in Statistics- Theory and Methods*, **15**, 2507-2525.

Chiou, W-J and Cohen, A. (1984). On estimating a common multivariate normal mean vector. *Annals of the Institute of Statistical Mathematics*, **37**, 499-506.

Cochran, W.G. (1937). Problems arising in the analysis of a series of similar experiments. Supplement to *Journal of the Royal Statistical Society*, **4**, 102-118.

Cochran, W.G. and Carroll, S.P. (1953). A sampling investigation of the efficiency of weighting inversely as the estimated variance. *Biometrics*, **9**, 447-459.

Ghosh, J.K. and Sinha, B.K. (1981). A necessary and sufficient condition for second order admissibility with applications to Berkson's bioassay problem. *The Annals of Statistics*, **9**,1334-1338.

Graybill, F.A. and Deal, R.B. (1959). Combining unbiased estimators. *Biometrics*, **15**, 543-550.

Hartung, J. (1999). A note on combining dependent tests of significance. *Biometrical Journal*, **41**, 849-855.

Khatri, C.G. and Shah K.R. (1974). Estimation of location parameters from two linear models under normality. *Communications in Statistics*, **3**, 647-663.

Krishnamoorthy, K. and Moore, B. C. (2002). Combining information for prediction in linear regression. *Metrika*, **56**, 73-81.

Kubokawa, T. (1989). Estimating common parameters of growth curve models under a quadratic loss. *Communications in Statistics- Theory and Methods*, **46**, 3149-3155.

Liski, E.P., Luoma, F.A., Tampere, and Sinha, B.K. (1996). Optimal designs in random coefficient linear regression models. *Calcutta Statistical Association Bulletin*, **18**, 211-229.

Loh,L.W. (1991). Estimating the common mean of two multivariate normal distributions. *The Annals of Statistics*, **19**, 297-313.

Meier, P. (1953). Variance of the weighted mean. *Biometrics*, **9**, 59-73.

Norwood, T.E. and Hinkelmann,K.Jr. (1977). Estimating the common mean of several normal populations. *The Annals of Statistics*, **5**, 1047-1050.

Rao, C. R. (1970). "Estimation of heteroscedastic variances in linear models. *Journal of the American Statistical Association*, **65**, 161-172.

Rao, J. N.K and Subrahmaniam, K. (1971). Combining independent estimators and estimation in linear regression with unequal variances. *Biometrics*, **27**, 971-990.

Shinozaki, N. (1978). A note on estimating the common mean of k normal distributions and the stein problem. *Communications in Statistics- Theory and Method*, **7**, 1421-1432.

Sinha, B.K. (1985). Unbiased estimation of the variance of the Graybill-Deal estimator of the common mean of several normal populations. *The Canadian Journal of Statistics*, **13**, 243-247.

Tsukuma, H. and Konno, Y. (2003). Modifying the Graybill-Deal estimator of the common regression matrix in two growth curve models." *Research Institute for Mathematical Sciences*, **1334**, 95-111.

Zacks, S. (1966). Unbiased estimation of the common mean. *Journal of the American Statistical Association*, **61**, 467-476.