Alternative Estimation of the Common Mean of Two Normal Populations with Order Restricted Variances

Authors: Adarsha Kumar Jena
– Department of Mathematics
National Institute of Technology Rourkela
Rourkela-769008, India (jadarsha@gmail.com)

Manas Ranjan Tripathy*
– Department of Mathematics
National Institute of Technology Rourkela
Rourkela-769008, India (manas@nitrkl.ac.in)

Nabendu Pal
– Department of Mathematics
University of Louisiana at Lafayette
Lafayette, Louisiana, USA 70504 (nabendu@louisiana.edu)

Abstract:

- The problem of estimating the common mean of two normal populations has been considered when it is known a priori that the variances are ordered. Under order restriction on the variances some new alternative estimators have been proposed including one that uses the maximum likelihood estimator (MLE) numerically. Further, it has been proved that each of these new estimators beats their unrestricted counterparts in terms of stochastic domination as well as Pitman measure of closeness criterion. Sufficient conditions for improving estimators in certain classes of equivariant estimators have been proved, and consequently improved estimators have been obtained under order restriction on the variances. A detailed simulation study has been done in order to evaluate the performances of all the proposed estimators using an existing estimator as a benchmark. From our simulation study, it has been established that the new alternative estimators improve significantly upon their unrestricted counterparts and compete well with an existing estimator under order restriction on the variances.

Key-Words:

- *Common mean; Equivariant estimator; Inadmissibility; Maximum likelihood estimator; Mean squared error; Ordered variances; Pitman measure of closeness; Percentage of relative risk; Stochastic dominance.*

AMS Subject Classification:

- 62F10; 62F30; 62C15.

*Corresponding Author E-mail address: manas@nitrkl.ac.in, manasmath@yahoo.co.in

1. INTRODUCTION

Suppose we have two normal populations with a common mean 'μ' and possibly different variances σ_1^2 and σ_2^2. More specifically, let $\underset{\sim}{X} = (X_1, X_2, \ldots, X_m)$ and $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)$ be independent random samples taken from two normal populations $N(\mu, \sigma_1^2)$ and $N(\mu, \sigma_2^2)$ respectively. The problem is to estimate the common mean 'μ' under the assumption that the variances follow the ordering $\sigma_1^2 \le \sigma_2^2$. In order to evaluate the performance of an estimator the loss functions

$$L_1(d, \mu) = \left(\frac{d-\mu}{\sigma_1}\right)^2, \tag{1.1}$$

$$L_2(d, \mu) = |d - \mu|, \tag{1.2}$$

and

$$L_3(d, \mu) = (d - \mu)^2, \tag{1.3}$$

are typically used, where d is an estimator for estimating 'μ' and $\underset{\sim}{\alpha} = (\mu, \sigma_1^2, \sigma_2^2)$; $\sigma_1^2 \le \sigma_2^2$. Furthermore the risk of an estimator d is defined by

$$R(d, \mu) = E_{\underset{\sim}{\alpha}}\{L_i(d, \mu)\};\ i = 1, 2, 3.$$

The problem of estimating the common mean of two or more normal populations, without considering the order restriction on the variances, is quite popular and has a long history in the literature of statistical inference. In fact, the origin of the problem has been in the recovery of inter-block information in the study of balanced incomplete block designs problem (see Yates [23]). Moreover, the problem has received considerable attention by several pioneer researchers in the last few decades due to its practical applications as well as the theoretical challenges involved in it. This well known problem arises in situations, where two or more measuring devices in a laboratory are used to measure certain quantity, several independent agencies are employed to test the effectiveness of certain new drugs produced by a developer, two or more different methods are used to evaluate certain characteristic etc.. Under these circumstances, if it is assumed that the samples drawn follow normal distributions, then the task boils down to draw inference on the common mean when the variances are unknown and unequal. We refer to some excellent papers by Chang and Pal [4], Lin and Lee [12] and Kelleher [10] for applications as well as examples of such nature. Probably, Graybill and Deal [8] were the first to consider this well-known common mean problem under normality assumption, without taking into account the order restriction on the variances. They proposed a combined estimator by taking convex combination of two sample means with weights as the functions of sample variances. Their combined estimator performs better than the individual sample means in terms of mean squared error when the sample sizes are at least 11. Since then a lot of attention has been paid in this direction by several researchers. In fact, the main goal has been to obtain either some competitors to Graybill-Deal estimator or some alternative estimators which may perform better than both the sample means. Also few attempts have been made to prove the admissibility or inadmissibility of the Graybill-Deal estimator. For a detailed literature review and recent updates on estimating the common mean of two or more normal populations without taking into account the order restriction on the variances, we refer to Khatri and Shah [11], Brown and Cohen [3], Cohen and Sackrowitz [6], Moore and Krishnamoorthy [14], Pal and Sinha [15], Pal et al. [16], Tripathy and Kumar [20, 21] and the references cited therein.

On the other hand, relatively less attention has been paid in estimating the common mean 'μ' when it is known a priori, that the variances follow certain simple ordering, say, $\sigma_1^2 \leq \sigma_2^2$. As an application of the common mean estimation under two ordered variances one can cite the example of evaluating the octane level of a particular grade of gasoline by the state inspectors in the United States. Usually the inspectors evaluating the octane level of gasoline sold at a gasoline station take two types of samples. Multiple samples of gasoline are taken on spot and their octane levels are quickly evaluated by a hand held device which is less precise and hence have high variance. Another batch of gasoline samples is taken and sent to state labs for a detailed, time consuming analysis of the octane level which is more accurate and has a smaller variance. If the spot analysis shows the mean octane level within a certain margin of the declared octane level then the inspector gives the seller a pass. Otherwise, the results from the lab tests are combined with the spot tests to determine the mean octane level. Disciplinary actions against the seller can be taken only if the combined estimate of the mean octane level falls below the declared level by a substantial margin. Probably Elfessi and Pal [7] were the first to consider this model with some justification and proposed an estimator that performs better than the Graybill-Deal estimator. In fact, their proposed estimator performs better than the Graybill-Deal estimator in terms of stochastic domination as well as universal domination. Later on, their results have been extended to the case of $k(\geq 2)$ normal populations by Misra and van der Meulen [13]. Chang et al. [5] also considered the estimation of a common mean under order restricted variances. They proposed a broad class of estimators that includes estimator proposed by Elfessi and Pal [7]. In fact, their proposed estimators stochastically dominate the estimators which do not obey the order restriction on the variances. However, for practical applications purpose, it is essential to have the specific estimators. Moreover, it is also necessary to know the amount of risk reduction after using the prior information regarding the ordering of the variances. Also we note that, the problem of estimation of a common standard deviation of several normal populations when the means are known to follow a simple ordering has been considered by Tripathy et al. [22] from a decision theoretic point of view.

In view of the above, we have proposed certain alternative estimators for the common mean when it is known a priori that the variances are ordered. These new estimators, which utilize the information about variance ordering, are shown to dominate their unrestricted counterparts (proposed by Moore and

Krishnamoorthy [14], Khatri and Shah [11], Brown and Cohen [3], Tripathy and Kumar [20]) stochastically, universally and in terms of Pitman nearness criterion. Moreover we have obtained a plug-in type restricted MLE which beats the unrestricted MLE with respect to a squared error loss function which has been shown numerically. In addition to these, we derive a sufficient condition for improving equivariant estimators using orbit-by-orbit improvement technique of Brewster and Zidek [2]. It is also interesting to see the performance of MLE with respect to other estimators (including the existing one proposed by Elfessi and Pal [7]), under order restriction on the variances, which is lacking in the literature. We also observe that a detailed and in-depth study to compare the performances of all the existing estimators for the common mean under order restricted variances is lacking in the literature. Therefore, we intend to study the performances of all the estimators, - both proposed as well as the existing ones, through a comprehensive simulation study which may fill the knowledge gap and provide useful information to the researchers from an application point of view.

The rest of the work is organized as follows. In Section 2, certain basic results have been discussed and a new plug-in type restricted MLE for the common mean μ has been proposed. In Section 3, some alternative estimators for the common mean μ have been constructed under order restriction on the variances. It is shown that the proposed estimators dominate their old counterparts proposed by Moore and Krishnamoorthy [14], Khatri and Shah [11], Brown and Cohen [3], Tripathy and Kumar [20] in terms of stochastic domination as well as universal domination. Moreover, in Section 4, we have proved that these alternative estimators also dominate their respective unrestricted counterparts in terms of Pitman measure of closeness criterion (see Pitman [18]). Sufficient conditions for improving the estimators which are invariant under affine transformations have been proved in Section 5, and consequently improved estimators have been derived. Interestingly, these improved estimators turned out to be the same as obtained in Section 2. We note that a theoretical comparison of all these proposed estimators seems difficult, and hence a simulation study has been carried out in order to compare numerically the risk functions of all the proposed estimators in Section 6. Moreover, the percentage of risk improvements of all the improved estimators upon their unrestricted counterparts have been noted with respect to all the three loss functions (1.1)-(1.3), which are quite significant. The percentage of relative risk improvements of all the proposed estimators have been obtained with respect to the Graybill-Deal estimator (treated as a benchmark) and recommendations have been made there. Finally we conclude our remarks with some examples to compute the estimates in Section 7.

2. SOME BASIC RESULTS

In this section, we discuss the statistical model and propose some alternative estimators for the common mean μ, when it is known a priori that the variances follow the simple ordering, that is, $\sigma_1^2 \leq \sigma_2^2$.

Let $\underset{\sim}{X} = (X_1, X_2, \ldots, X_m)$ and $\underset{\sim}{Y} = (Y_1, Y_2, \ldots, Y_n)$ be independent random samples taken from two normal populations with a common mean μ and possibly different variances σ_1^2 and σ_2^2 respectively. Let $N(\mu, \sigma_i^2)$ be denote the normal population with mean μ and variance σ_i^2; $i = 1, 2$. The target is to derive certain estimators for μ, when it is known a priori that, the variances are ordered, that is, $\sigma_1^2 \leq \sigma_2^2$ or equivalently $\sigma_1 \leq \sigma_2$. We note, that a minimal sufficient statistics (not complete) for this model exists and is given by $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$ where

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i,\ \bar{Y} = \frac{1}{n}\sum_{j=1}^{n} Y_j,\ S_1^2 = \sum_{i=1}^{m}(X_i - \bar{X})^2,\ S_2^2 = \sum_{j=1}^{n}(Y_j - \bar{Y})^2. \tag{2.1}$$

We further note that, $\bar{X} \sim N(\mu, \sigma_1^2/m)$, $\bar{Y} \sim N(\mu, \sigma_2^2/n)$, $S_1^2/\sigma_1^2 \sim \chi^2_{m-1}$, and $S_2^2/\sigma_2^2 \sim \chi^2_{n-1}$. When there is no order restrictions on the variances, a number of estimators have been proposed by several researchers in the recent past. Let us consider the following well known estimators for the common mean μ when there is no order restriction on the variances.

$$\begin{aligned}
d_{GD} &= \frac{m(m-1)S_2^2\bar{X} + n(n-1)S_1^2\bar{Y}}{m(m-1)S_2^2 + n(n-1)S_1^2} \quad \text{(Graybill and Deal [8])},\\
d_{KS} &= \frac{m(m-3)S_2^2\bar{X} + n(n-3)S_1^2\bar{Y}}{m(m-3)S_2^2 + n(n-3)S_1^2} \quad \text{(Khatri and Shah [11])},\\
d_{MK} &= \frac{\bar{X}\sqrt{m(m-1)}S_2 + \bar{Y}\sqrt{n(n-1)}S_1}{\sqrt{m(m-1)}S_2 + \sqrt{n(n-1)}S_1} \quad \text{(Moore and Krishnamoorthy [14])},\\
d_{TK} &= \frac{\bar{X}\sqrt{m}c_nS_2 + \bar{Y}\sqrt{n}c_mS_1}{\sqrt{m}c_nS_2 + \sqrt{n}c_mS_1} \quad \text{(Tripathy and Kumar [20])},\\
d_{BC1} &= \bar{X} + \left\{\frac{(\bar{Y}-\bar{X})b_1S_1^2/m(m-1)}{S_1^2/m(m-1) + S_2^2/(n(n+2)) + (\bar{Y}-\bar{X})^2/(n+2)}\right\}\\
d_{BC2} &= \bar{X} + (\bar{Y}-\bar{X})\left\{\frac{b_2n(n-1)S_1^2}{n(n-1)S_1^2 + m(m-1)S_2^2}\right\} \quad \text{(Brown and Cohen [3])},\\
d_{GM} &= \frac{m\bar{X} + n\bar{Y}}{m+n} \quad \text{(grand sample mean)},
\end{aligned}$$

where $c_m = \Gamma(\frac{m-1}{2})/(\sqrt{2}\Gamma(\frac{m}{2}))$, $c_n = \Gamma(\frac{n-1}{2})/(\sqrt{2}\Gamma(\frac{n}{2}))$, $0 < b_1 < b_{\max}(m,n)$, $0 < b_2 < b_{\max}(m, n-3)$, and $b_{max}(m,n) = 2(n+2)/nE(\max(1/V, 1/V^2))$. Here V is a random variable having F-distribution with $(n+2)$ and $(m-1)$ degrees of freedom.

Finally we consider the MLE of μ whose closed form does not exist (see Pal et al. [16]). The MLE of μ can be obtained numerically by solving the following system of three equations in three unknowns μ, σ_1^2, and σ_2^2.

$$\mu = \frac{\frac{m}{\sigma_1^2}\bar{x} + \frac{n}{\sigma_2^2}\bar{y}}{\frac{m}{\sigma_1^2} + \frac{n}{\sigma_2^2}}, \tag{2.2}$$

$$\sigma_1^2 = \frac{s_1^2}{m} + \Big(\frac{n\sigma_1^2}{n\sigma_1^2 + m\sigma_2^2}\Big)^2 (\bar{x} - \bar{y})^2, \tag{2.3}$$

$$\sigma_2^2 = \frac{s_2^2}{n} + \Big(\frac{m\sigma_2^2}{n\sigma_1^2 + m\sigma_2^2}\Big)^2 (\bar{x} - \bar{y})^2. \tag{2.4}$$

Here $(\bar{x}, \bar{y}, s_1^2, s_2^2)$ denotes the observed value of $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$. Let the solution of the above system of equations be $\hat{\mu}_{ML}$, $\hat{\sigma}_{1ML}^2$ and $\hat{\sigma}_{2ML}^2$. These are the MLEs of μ, σ_1^2 and σ_2^2 respectively, when there is no order restriction on the variances.

Next, we discuss some results on estimating common mean when it is known a priori that the variances follow the simple ordering $\sigma_1^2 \le \sigma_2^2$. Let $\beta = (n(n-1)S_1^2)/(m(m-1)S_2^2 + n(n-1)S_1^2)$. Under order restriction on the variances, Elfessi and Pal [7] proposed a new estimator, call it $\hat{d}_{EP}$ which is given by

$$\hat{d}_{EP} = \begin{cases} (1-\beta)\bar{X} + \beta\bar{Y}, & \text{if } \frac{S_1^2}{m-1} \le \frac{S_2^2}{n-1} \\ \beta^*\bar{X} + (1-\beta^*)\bar{Y}, & \text{if } \frac{S_1^2}{m-1} > \frac{S_2^2}{n-1}, \end{cases}$$

where

$$\beta^* = \begin{cases} \beta, & \text{if } m = n \\ \frac{m}{m+n}, & \text{if } m \ne n. \end{cases}$$

In the above definition of $\hat{d}_{EP}$ for the case $m = n$, when $\beta^* = \beta$, we mean β as well as the conditions must be simplified for $m = n$.

It is well known that the estimator $\hat{d}_{EP}$ dominates d_{GD} stochastically as well as universally when $\sigma_1^2 \le \sigma_2^2$. Further Misra and van der Meulen [13] extended these dominance results to the case of $k(\ge 2)$ normal populations and also proved that the estimator $\hat{d}_{EP}$ performs better than d_{GD} in terms of Pitman measure of closeness criterion. The MLE of μ has been obtained by solving the system of equations numerically as shown above (see equations (2.2) to (2.4))). When the variances are ordered, using the isotonic version of the MLEs of σ_i^2, we obtain plug-in type restricted MLEs (numerically) of σ_1^2 and σ_2^2 respectively as

$$\hat{\sigma}_{1R}^2 = \begin{cases} \hat{\sigma}_{1ML}^2, & \text{if } \hat{\sigma}_{1ML}^2 \le \hat{\sigma}_{2ML}^2 \\ \frac{1}{2}(\hat{\sigma}_{1ML}^2 + \hat{\sigma}_{2ML}^2), & \text{if } \hat{\sigma}_{1ML}^2 > \hat{\sigma}_{2ML}^2, \end{cases}$$

and

$$\hat{\sigma}_{2R}^2 = \begin{cases} \hat{\sigma}_{2ML}^2, & \text{if } \hat{\sigma}_{1ML}^2 \le \hat{\sigma}_{2ML}^2 \\ \frac{1}{2}(\hat{\sigma}_{1ML}^2 + \hat{\sigma}_{2ML}^2), & \text{if } \hat{\sigma}_{1ML}^2 > \hat{\sigma}_{2ML}^2 \end{cases}$$

(see Barlow et al. [1]). Substituting these estimators in (2.2), we get a plug-in type restricted MLE, (call it d_{RM}) for μ as

$$d_{RM} = \frac{m\hat{\sigma}_{2R}^2\bar{X} + n\hat{\sigma}_{1R}^2\bar{Y}}{m\hat{\sigma}_{2R}^2 + n\hat{\sigma}_{1R}^2}.$$

Further using the grand sample mean of the two populations, one gets another plug-in type restricted MLE of μ, call it $\hat{d}_{RM}$, and is given by

$$\hat{d}_{RM} = \begin{cases} \hat{\mu}_{ML}, & \text{if } \hat{\sigma}^2_{1ML} \leq \hat{\sigma}^2_{2ML} \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \hat{\sigma}^2_{1ML} > \hat{\sigma}^2_{2ML}. \end{cases}$$

Through a simulation study, Tripathy and Kumar [20] concluded that the estimators d_{MK} and d_{TK} compete with each other and perform better than d_{GD} when the variances are not far away from each other. Authors also mentioned that for small values of the ratios of the variances, the estimator d_{KS} compete with d_{GD}. Hence, it is quite evident that one needs to find alternative estimators for μ which may compete with $\hat{d}_{EP}$ when $\sigma_1^2 \leq \sigma_2^2$ or equivalently $\sigma_1 \leq \sigma_2$. In the next sections to follow (Sections 3 and 4), we propose some new estimators which dominate their respective unrestricted counterparts stochastically as well as universally and may compete with $\hat{d}_{EP}$ in terms of risks. Now onwards for convenient we will denote $\hat{d}_{EP}$ as $\hat{d}_{GD}$.

Remark 2.1. One can construct another plug-in type estimator for μ by replacing the estimators $\hat{\sigma}^2_{1R}$ and $\hat{\sigma}^2_{2R}$ in d_{RM} by $\hat{\sigma}^2_{1R} = \min(\hat{\sigma}^2_{1ML}, (m\hat{\sigma}^2_{1ML} + n\hat{\sigma}^2_{2ML})/(m+n))$ and $\hat{\sigma}^2_{2R} = \max(\hat{\sigma}^2_{2ML}, (m\hat{\sigma}^2_{1ML} + n\hat{\sigma}^2_{2ML})/(m+n))$ respectively when $m \neq n$. It has been revealed from our numerical study (Section 6) that it acts as a competitor of d_{RM}.

Remark 2.2. The estimators d_{RM} and $\hat{d}_{RM}$ are seen to perform equally good, which has been checked from our simulation study in Section 6. Hence we include only $\hat{d}_{RM}$ in our numerical comparison in Section 6.

3. STOCHASTIC DOMINATION UNDER ORDER RESTRICTION ON THE VARIANCES

In this section we propose some alternative estimators for the common mean μ under order restriction on the variances that is when it is known a priori that $\sigma_1^2 \leq \sigma_2^2$ or equivalently $\sigma_1 \leq \sigma_2$. Further it will be shown that each of these alternative estimators dominate their unrestricted counterparts proposed by Moore and Krishnamoorthy [14], Tripathy and Kumar [20], Khatri and Shah [11] and Brown and Cohen [3] stochastically under order restriction on the variances.

To start with, let us define

$$
\begin{aligned}
\beta_1 &= \frac{\sqrt{n(n-1)}S_1}{\sqrt{m(m-1)}S_2 + \sqrt{n(n-1)}S_1}, \\
\beta_2 &= \frac{\sqrt{n}c_m S_1}{\sqrt{m}c_n S_2 + \sqrt{n}c_m S_1}, \\
\beta_3 &= \frac{n(n-3)S_1^2}{m(m-3)S_2^2 + n(n-3)S_1^2}, \\
\beta_4 &= \frac{b_2 S_1^2}{S_1^2 + S_2^2}.
\end{aligned}
$$

We propose the following estimators for the common mean μ, when the variances known to follow the simple ordering $\sigma_1^2 \leq \sigma_2^2$.

$$
\hat{d}_{MK} = \begin{cases} (1-\beta_1)\bar{X} + \beta_1\bar{Y}, \text{ if } \frac{\sqrt{n-1}S_1}{\sqrt{m-1}S_2} \leq \sqrt{\frac{n}{m}}, \\ \\ \beta_1^*\bar{X} + (1-\beta_1^*)\bar{Y}, \text{ if } \frac{\sqrt{n-1}S_1}{\sqrt{m-1}S_2} > \sqrt{\frac{n}{m}}, \end{cases}
$$

$$
\hat{d}_{TK} = \begin{cases} (1-\beta_2)\bar{X} + \beta_2\bar{Y}, \text{ if } \frac{S_1}{S_2} \leq \sqrt{\frac{n}{m}}\frac{c_n}{c_m}, \\ \\ \beta_2^*\bar{X} + (1-\beta_2^*)\bar{Y}, \text{ if } \frac{S_1}{S_2} > \sqrt{\frac{n}{m}}\frac{c_n}{c_m}, \end{cases}
$$

$$
\hat{d}_{KS} = \begin{cases} (1-\beta_3)\bar{X} + \beta_3\bar{Y}, \text{ if } \frac{S_1^2}{S_2^2} \leq \frac{m-3}{n-3}, \\ \\ \beta_3^*\bar{X} + (1-\beta_3^*)\bar{Y}, \text{ if } \frac{S_1^2}{S_2^2} > \frac{m-3}{n-3}, \end{cases}
$$

where for $i = 1, 2, 3$ we denote

$$
\beta_i^* = \begin{cases} \beta_i, & \text{if } m = n, \\ \\ \frac{m}{m+n}, & \text{if } m \neq n. \end{cases}
$$

Finally, we propose an estimator for the case of equal sample sizes as

$$
\hat{d}_{BC2} = \begin{cases} (1-\beta_4)\bar{X} + \beta_4\bar{Y}, \text{ if } \frac{S_2^2}{S_1^2} \geq (2b_2 - 1), \\ \\ \beta_4\bar{X} + (1-\beta_4)\bar{Y}, \text{ if } \frac{S_2^2}{S_1^2} < (2b_2 - 1). \end{cases}
$$

In the above definitions of the estimators $\hat{d}_{MK}$, $\hat{d}_{TK}$, $\hat{d}_{KS}$ for the case $m = n$, when $\beta_i^* = \beta_i$; $i = 1, 2, 3$, we mean that both β_i and the corresponding conditions must be simplified by putting $m = n$.

To proceed further we need the following two definitions which will be used in developing the section. Let d_1 and d_2 be any two estimators of the unknown parameter say θ.

Definition 3.1. The estimator d_1 is said to dominate another estimator d_2 stochastically if $P_\theta[(d_2-\mu)^2 \le c] \le P_\theta[(d_1-\mu)^2 \le c], \;\; \forall\, c > 0$.

Definition 3.2. Let the loss function $L(d,\theta)$ in estimating θ by d be a non-decreasing function of the error $|d-\theta|$. An estimator d_1 is said to dominate another estimator d_2 universally if $EL(|d_1-\theta|) \le EL(|d_2-\theta|)$, over the parameter space for all $L(.)$ non-decreasing. Further it was shown by Hwang [9] that d_1 dominates d_2 universally if and only if d_1 dominates d_2 stochastically.

Next, we prove the following results for estimating the common mean μ, under order restriction on the variances, which are immediate.

Theorem 3.1. *Let the loss function $L(.)$ be a non-decreasing function of the error $|d-\mu|$. Further assume that the variances are known to follow the ordering $\sigma_1^2 \le \sigma_2^2$. Then for estimating the common mean μ we have the following dominance results.*

(i) *The estimator $\hat{d}_{MK}$ dominates d_{MK} stochastically and hence universally.*

(ii) *The estimator $\hat{d}_{TK}$ dominates d_{TK} stochastically and hence universally.*

(iii) *The estimator $\hat{d}_{KS}$ dominates d_{KS} stochastically and hence universally.*

(iv) *The estimator $\hat{d}_{BC2}$ dominates d_{BC2} stochastically and hence universally.*

Proof: Please see Appendix. □

4. PITMAN MEASURE OF CLOSENESS

In this section, we prove that the new proposed estimators $\hat{d}_{MK}$, $\hat{d}_{TK}$, $\hat{d}_{KS}$, and $\hat{d}_{BC2}$, perform better than their old counterparts in terms of Pitman measure of closeness criterion when it is known a priori that the variances follow the ordering $\sigma_1^2 \le \sigma_2^2$ or equivalently $\sigma_1 \le \sigma_2$. To prove the main results of this section, we need the following results. Let δ_1 and δ_2 be any two estimators of a real parametric function say $\psi(\theta)$. Pitman [18] proposed a measure of relative closeness to the parametric function $\psi(\theta)$ for comparing two estimators in the following fashions.

Definition 4.1. The estimator δ_1 should be preferred to δ_2 if for every θ, $PMC_\theta(\delta_1,\delta_2) = P_\theta(|\delta_1-\psi(\theta)| < |\delta_2-\psi(\theta)| \,|\, \delta_1 \ne \delta_2) \ge \frac{1}{2}$, and with strict inequality for some θ.

The following lemma will be useful for proving the main results of this section, which was proposed by Peddada and Khatree [17].

Lemma 4.1. *Suppose the random vector (X, Y) has a bivariate normal distribution with $E(X) = E(Y) = 0$ and $E(X^2) < E(Y^2)$. Then $P(|X| < |Y|) > \frac{1}{2}$.*

Let $\underset{\sim}{\alpha} = (\mu, \sigma_1^2, \sigma_2^2)$ and $\Omega_R = \{\underset{\sim}{\alpha} = (\mu, \sigma_1^2, \sigma_2^2) : -\infty < \mu < \infty, 0 < \sigma_1^2 \leq \sigma_2^2 < \infty\}$. We prove the following theorem.

Theorem 4.1. *For estimating the common mean μ of two normal populations, when $\sigma_1^2 \leq \sigma_2^2$, we have the following dominance results.*

(i) $PMC(\hat{d}_{MK}, d_{MK}) > \frac{1}{2}, \ \forall \ \underset{\sim}{\alpha} \in \Omega_R.$

(ii) $PMC(\hat{d}_{TK}, d_{TK}) > \frac{1}{2}, \ \forall \ \underset{\sim}{\alpha} \in \Omega_R.$

(iii) $PMC(\hat{d}_{KS}, d_{KS}) > \frac{1}{2}, \ \forall \ \underset{\sim}{\alpha} \in \Omega_R.$

(iv) $PMC(\hat{d}_{BC2}, d_{BC2}) > \frac{1}{2}, \ \forall \ \underset{\sim}{\alpha} \in \Omega_R.$

Proof: The proof of the theorem is easy after using the Lemma 4.1, and hence has been omitted. □

In the next section we will introduce the concept of invariance to our problem and prove some inadmissibility results in the classes of equivariant estimators for the common mean.

5. INADMISSIBILITY RESULTS UNDER ORDER RESTRICTION ON THE VARIANCES

In this section we introduce the concept of invariance to the problem and derive some inadmissibility results for both affine and location equivariant estimators under order restriction on the variances. As a consequence, estimators dominating some of the existing well known estimators for the common mean have been derived, under order restriction on the variances.

5.1. Affine Class

Let us introduce the concept of invariance to our problem. More specifically, consider the affine group of transformations, $G_A = \{g_{a,b} : g_{a,b}(x) = ax + b, a >$

$0, b \in R\}$. Under the transformation $g_{a,b}$, $X_i \to aX_i + b$; $i = 1, 2, \ldots, m$, $Y_j \to aY_j + b$; $j = 1, 2, \ldots, n$ and consequently the sufficient statistics $\bar{X} \to a\bar{X} + b$, $\bar{Y} \to a\bar{Y} + b$, $S_i^2 \to a^2 S_i^2$, $\mu \to a\mu + b$, $\sigma_i^2 \to a^2\sigma_i^2$ and the family of distributions remains invariant. The problem remains invariant if we choose the loss function as (1.1). The form of an affine equivariant estimator for estimating μ, based on the sufficient statistic $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$ is obtained as

$$d_\Psi = \bar{X} + S_1 \Psi(\underset{\sim}{T}), \tag{5.1}$$

where $\underset{\sim}{T} = (T_1, T_2)$, $T_1 = (\bar{Y} - \bar{X})/S_1$, $T_2 = S_2^2/S_1^2$ and Ψ is any real valued function.

Let us define a new function Ψ_0 for the affine equivariant estimator d_Ψ as

$$\Psi_0(\underset{\sim}{t}) = \begin{cases} \frac{n}{n+m}\min(t_1, 0), & \text{if } \Psi(\underset{\sim}{t}) < \frac{n}{n+m}\min(t_1, 0), \\ \\ \Psi(\underset{\sim}{t}), & \text{if } \frac{n}{n+m}\min(t_1, 0) \le \Psi(\underset{\sim}{t}) \le \frac{n}{n+m}\max(t_1, 0), \\ \\ \frac{n}{n+m}\max(t_1, 0), & \text{if } \Psi(\underset{\sim}{t}) > \frac{n}{n+m}\max(t_1, 0). \end{cases} \tag{5.2}$$

The following theorem gives a sufficient condition for improving estimators in the class of affine equivariant estimators of the form (5.1), under order restriction on the variances.

Theorem 5.1. *Let d_Ψ be an affine equivariant estimator of the form (5.1) for estimating the common mean μ and the loss function be the affine invariant loss (1.1). The estimator d_Ψ is inadmissible and is improved by d_{Ψ_0} if $P(\Psi(\underset{\sim}{T}) \neq \Psi_0(\underset{\sim}{T})) > 0$, for some choices of the parameters $\underset{\sim}{\alpha}$; $\sigma_1 \le \sigma_2$.*

Proof: Please see Appendix. □

Next we will apply Theorem 5.1, to obtain some improved estimators for the common mean μ, under the assumption that $\sigma_1^2 \le \sigma_2^2$. It is easy to observe that all the estimators discussed in Section 2, for the common mean μ without taking into account the order restriction on the variances, fall into the class $d_\Psi = \bar{X} + S_1\Psi(\underset{\sim}{T})$. We apply Theorem 5.1 to get their corresponding improved estimators under the assumption that $\sigma_1^2 \le \sigma_2^2$. Let us first consider the estimator $d_{GD} = \bar{X} + S_1\Psi(\underset{\sim}{T})$, where $\Psi(\underset{\sim}{T}) = (n(n-1)T_1)/(m(m-1)T_2 + n(n-1))$. Note that $\Psi(\underset{\sim}{t}) > (n/(m+n))\max(0, t_1)$, when $S_1^2/(m-1) > S_2^2/(n-1)$. Hence the estimator d_{GD} is inadmissible and is improved by the estimator

$$d_{GD}^a = \begin{cases} \frac{m(m-1)S_2^2\bar{X} + n(n-1)S_1^2\bar{Y}}{m(m-1)S_2^2 + n(n-1)S_1^2}, & \text{if } \frac{S_1^2}{m-1} \le \frac{S_2^2}{n-1} \\ \\ \frac{m\bar{X} + n\bar{Y}}{m+n}, & \text{if } \frac{S_1^2}{m-1} > \frac{S_2^2}{n-1}, \end{cases}$$

under order restriction on the variances.

Similarly one can get the estimators which improve upon d_{KS}, d_{MK}, d_{TK}, d_{BC1}, and d_{BC2} respectively as

$$d_{KS}^a = \begin{cases} \frac{m(m-3)S_2^2\bar{X}+n(n-3)S_1^2\bar{Y}}{m(m-3)S_2^2+n(n-3)S_1^2}, & \text{if } \frac{S_1^2}{m-3} \le \frac{S_2^2}{n-3} \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \frac{S_1^2}{m-3} > \frac{S_2^2}{n-3}, \end{cases}$$

$$d_{MK}^a = \begin{cases} \frac{\sqrt{m(m-1)}S_2\bar{X}+\sqrt{n(n-1)}S_1\bar{Y}}{\sqrt{m(m-1)}S_2+\sqrt{n(n-1)}S_1}, & \text{if } \frac{S_1}{\sqrt{m-1}} \le \frac{S_2}{\sqrt{n-1}}, \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \frac{S_1}{\sqrt{m-1}} > \frac{S_2}{\sqrt{n-1}}, \end{cases}$$

$$d_{TK}^a = \begin{cases} \frac{\sqrt{m}c_nS_2\bar{X}+\sqrt{n}c_mS_1\bar{Y}}{\sqrt{m}c_nS_2+\sqrt{n}c_mS_1}, & \text{if } \frac{S_1}{S_2} \le \sqrt{\frac{n}{m}}\frac{c_n}{c_m}, \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \frac{S_1}{S_2} \le \sqrt{\frac{n}{m}}\frac{c_n}{c_m}, \end{cases}$$

$$d_{BC1}^a = \begin{cases} d_{BC1}, & \text{if } \frac{S_2^2}{S_1^2} + n(\frac{\bar{Y}-\bar{X}}{S_1})^2 > \frac{n+2}{m(m-1)}[b_1(m+n)-n] \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \frac{S_2^2}{S_1^2} + n(\frac{\bar{Y}-\bar{X}}{S_1})^2 \le \frac{n+2}{m(m-1)}[b_1(m+n)-n], \end{cases}$$

and

$$d_{BC2}^a = \begin{cases} d_{BC2}, & \text{if } \frac{m(m-1)S_2^2}{n(n-1)S_1^2} \ge b_2(1+\frac{m}{n})-1 \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } \frac{m(m-1)S_2^2}{n(n-1)S_1^2} < b_2(1+\frac{m}{n})-1. \end{cases}$$

Remark 5.1. It is interesting to note that, for the case of unequal sample sizes, that is for $m \neq n$, the estimators $d_{GD}^a = \hat{d}_{GD}$, $d_{KS}^a = \hat{d}_{KS}$, $d_{MK}^a = \hat{d}_{MK}$, $d_{TK}^a = \hat{d}_{TK}$, $d_{BC2}^a = \hat{d}_{BC2}$. However for equal sample sizes, application of the Theorem 5.1 produces different estimators.

Remark 5.2. We note that, though the MLE of μ can not be obtained in a closed form, however from (2.2) it is easy to write $\hat{\mu}_{ML} = \bar{X} + S_1\Psi_{ML}(\underset{\sim}{T})$, where $\Psi_{ML}(\underset{\sim}{T}) = T_1 n\hat{\sigma}_{1ML}^2/(m\hat{\sigma}_{2ML}^2 + n\hat{\sigma}_{1ML}^2)$, and $\hat{\sigma}_{1ML}^2$, $\hat{\sigma}_{2ML}^2$ are to be found by solving (2.3) and (2.4). Though $\Psi_{ML}(\underset{\sim}{T})$ does not have a closed form, for a given dataset(sample values), we can find the value of $\Psi_{ML}(\underset{\sim}{t})$. Therefor, we can find $\Psi_0^{ML}(\underset{\sim}{t})$ by using (5.2). Hence we can apply Theorem 5.1 and find the value of the improved estimator $d_{ML}^a = \bar{X} + S_1\Psi_0^{ML}(\underset{\sim}{t})$ which does not have a closed form. It has been observed in our simulation study that the improved version of the MLE appears to have the identical risk as the estimator d_{RM}.

5.2. Location Class

A larger class of estimators than the class considered above is the class of location equivariant estimators. Let $G_L = \{g_c : g_c(x) = x + c, -\infty < c < \infty\}$ be the location group of transformations. Under the transformation g_c, we observe that, $\bar{X} \to \bar{X} + c$, $\bar{Y} \to \bar{Y} + c$, $S_1^2 \to S_2^2$, $S_2^2 \to S_2^2$, and the parameters $\mu \to \mu + c$, $\sigma_1 \to \sigma_1$. The family of probability distributions is invariant and consequently the estimation problem is also invariant under the loss (1.1). Based on the minimal sufficient statistics $(\bar{X}, \bar{Y}, S_1^2, S_2^2)$ the form of a location equivariant estimator for estimating the common mean μ is thus obtained as

$$d_\psi = \bar{X} + \psi(\underset{\sim}{U}), \tag{5.3}$$

where $\underset{\sim}{U} = (T, S_1^2, S_2^2)$, $T = \bar{Y} - \bar{X}$, and ψ is a real valued function. Let us define a function ψ_0 for the location equivariant estimator d_ψ as

$$\psi_0(\underset{\sim}{t}) = \begin{cases} \frac{n}{n+m}\min\{t,0\}, & \text{if } \psi(\underset{\sim}{u}) < \frac{n}{m+n}\min\{t,0\}, \\ \psi(\underset{\sim}{u}), & \text{if } \frac{n}{n+m}\min\{t,0\} \le \psi(\underset{\sim}{u}) \le \frac{n}{n+m}\max\{t,0\}, \\ \frac{n}{n+m}\max\{t,0\}, & \text{if } \psi(\underset{\sim}{u}) > \frac{n}{n+m}\max\{t,0\}. \end{cases} \tag{5.4}$$

The following theorem gives a sufficient condition for improving location equivariant estimators under the condition that the variances follow the ordering $\sigma_1^2 \le \sigma_2^2$.

Theorem 5.2. *Let d_ψ be a location equivariant estimator for estimating the common mean μ and the loss function be (1.1). Let the function $\psi_0(\underset{\sim}{u})$ be as defined in (5.6). The estimator d_ψ is inadmissible and is improved by d_{ψ_0} if $P_{\underset{\sim}{\alpha}}(\psi(\underset{\sim}{U}) \neq \psi_0(\underset{\sim}{U})) > 0$ for some choices of the parameters $\underset{\sim}{\alpha} = (\mu, \sigma_1^2, \sigma_2^2)$; $\sigma_1^2 \le \sigma_2^2$.*

Proof: The proof is similar to the proof of the Theorem 5.1, and hence has been omitted for brevity. □

Remark 5.3. We also observe that all the estimators proposed in Section 2, including the MLE (whose closed form does not exist) belong to the class $d_\psi(\underset{\sim}{U}) = \bar{X} + \psi(\underset{\sim}{U})$. Hence as an application of Theorem 5.2, produces improved estimators. Further we note that, though location class produces larger class of estimators, the sufficient conditions in Theorem 5.2, does not help to obtain different improved estimators than those obtained by applying Theorem 5.1, under order restriction on the variances. In fact, the sufficient conditions in Theorem 5.1 and 5.2 for improving equivariant estimators produces the same improved estimators under order restricted variances.

Remark 5.4. The performances of all the improved estimators which has been proposed in Section 2 as well as in this section by applying Theorem 5.1, will be evaluated in Section 6, using the affine invariant loss function L_1. Further the percentage of risk improvements upon their respective old counterparts has been noted.

Remark 5.5. We note that the estimator d_{GM}, also belongs to the classes given in (5.1) and (5.5). However, the conditions in Theorem 5.1 and Theorem 5.2 for improving it, do not satisfy. Hence the estimator d_{GM} could not be improved by applying either Theorem 5.1 or Theorem 5.2, under the condition that the variances are ordered that is, $\sigma_1^2 \le \sigma_2^2$.

6. A SIMULATION STUDY

It should be noted that, in Section 2 we have constructed the plug-in type restricted MLE $\hat{d}_{RM}$ for the common mean μ, taking into account the order restriction on the variances. Moreover, in Sections 3 and 4 we have also constructed some alternative estimators such as $\hat{d}_{MK}$, $\hat{d}_{TK}$, $\hat{d}_{KS}$, and $\hat{d}_{BC2}$ and proved that each of these estimators dominate their old unrestricted counterparts in terms of stochastic domination as well as Pitman measure of closeness criterion. Furthermore in Section 5, we have proposed some improved estimators namely d^a_{GD}, d^a_{KS}, d^a_{MK}, d^a_{TK}, d^a_{BC1}, d^a_{BC2} by an application of Theorem 5.1 and 5.2. In addition to all these estimators, we have also considered the improved estimator $\hat{d}_{GD}$ proposed by Elfessi and Pal [7]. In order to know the performances of all these improved estimators, one needs to compare the risk functions. We observe that an analytical comparison of all these estimators seems quite impossible, hence in this section we compare the risk functions of all the improved estimators numerically through Monte-Carlo simulation method. For this purpose we have generated $20,000$ random samples of sizes m and n respectively from $N(\mu, \sigma_1^2)$ and $N(\mu, \sigma_2^2)$, with the condition that $\sigma_1^2 \le \sigma_2^2$. The accuracy of the simulation has been checked and the error has been checked which is seen up to 10^{-3}. To proceed further, we define the percentage of risk improvements of all the improved estimators upon each of their unrestricted counterparts as follows.

$$P1 = \left(1 - \frac{R(\hat{d}_{GD}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad P2 = \left(1 - \frac{R(\hat{d}_{KS}, \mu)}{R(d_{KS}, \mu)}\right) \times 100,$$

$$P3 = \left(1 - \frac{R(\hat{d}_{MK}, \mu)}{R(d_{MK}, \mu)}\right) \times 100, \quad P4 = \left(1 - \frac{R(\hat{d}_{TK}, \mu)}{R(d_{TK}, \mu)}\right) \times 100,$$

$$P5 = \left(1 - \frac{R(d^a_{GD}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad P6 = \left(1 - \frac{R(d^a_{KS}, \mu)}{R(d_{KS}, \mu)}\right) \times 100,$$

$$P7 = \left(1 - \frac{R(d^a_{MK}, \mu)}{R(d_{MK}, \mu)}\right) \times 100, \quad P8 = \left(1 - \frac{R(d^a_{TK}, \mu)}{R(d_{TK}, \mu)}\right) \times 100,$$

$$P9 = \left(1 - \frac{R(\hat{d}_{RM}, \mu)}{R(d_{ML}, \mu)}\right) \times 100.$$

In order to compare the performances of all the improved estimators among themselves we use the affine loss function (1.1). It is better to compare the risk functions of all the improved estimators with respect to a benchmark estimator which can be the Graybill-Deal (see Graybill and Deal [8]) estimator. We define the percentage of relative risk performances of all the improved estimators with respect to the benchmark estimator d_{GD} as follows.

$$R1 = \left(1 - \frac{R(\hat{d}_{GD}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad R2 = \left(1 - \frac{R(\hat{d}_{KS}, \mu)}{R(d_{GD}, \mu)}\right) \times 100,$$

$$R3 = \left(1 - \frac{R(\hat{d}_{MK}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad R4 = \left(1 - \frac{R(\hat{d}_{TK}, \mu)}{R(d_{GD}, \mu)}\right) \times 100,$$

$$R5 = \left(1 - \frac{R(d^a_{GD}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad R6 = \left(1 - \frac{R(d^a_{MK}, \mu)}{R(d_{GD}, \mu)}\right) \times 100,$$

$$R7 = \left(1 - \frac{R(d^a_{BC1}, \mu)}{R(d_{GD}, \mu)}\right) \times 100, \quad R8 = \left(1 - \frac{R(d^a_{BC2}, \mu)}{R(d_{GD}, \mu)}\right) \times 100,$$

$$R9 = \left(1 - \frac{R(\hat{d}_{RM}, \mu)}{R(d_{GD}, \mu)}\right) \times 100.$$

It is easy to observe that the risks of all the estimators are functions of τ with respect to the loss function L_1 as given in (1.1), where we denote $\tau = \sigma_1^2/\sigma_2^2$, $0 < \tau \leq 1$. We note that, when the sample sizes are unequal (that is when $m \neq n$) $\hat{d}_{GD} = d^a_{GD}$, $\hat{d}_{KS} = d^a_{KS}$, $\hat{d}_{MK} = d^a_{MK}$, and $\hat{d}_{TK} = d^a_{TK}$. Further we notice that for equal sample sizes (that is when $m = n$) $\hat{d}_{GD} = \hat{d}_{KS}$ and $\hat{d}_{MK} = \hat{TK}$. In our simulation study we have chosen $b_1 = \frac{1}{2}b_{\max}(m, n)$ and $b_2 = \frac{1}{2}b_{\max}(m, n-3)$, where the values of $b_{\max}(m, n)$ have been taken from the table given in Brown and Cohen [3]. Moreover we observe that for $b_2 = 1$, the estimator $d_{BC2} = d_{GD}$ also when $b_2 = 0$, it reduces to $\bar{X}$. The percentage of risk improvements of d^a_{BC1}, d^a_{BC2} and $\hat{d}_{BC2}$ upon their unrestricted counterparts are seen to be very marginal and hence have not been tabulated. The simulation study has been carried out for various combinations of sample sizes while the parameter $\tau \in (0, 1]$. For illustration purpose we have presented the percentage of risk improvements as well as the percentage of relative risk improvements of all the estimators for some choices of sample sizes in Tables 1 - 7. In Tables 1 and 2 we have presented the percentage of risk improvements of all the improved estimators upon their unrestricted counterparts for equal and unequal sample sizes respectively with respect to the loss function (1.1). Particularly, in Table 1, the percentage of risk improvements of all the improved estimators have been presented for the sample sizes $(5, 5)$, $(12, 12)$, $(20, 20)$ and $(30, 30)$. The first and the seventh column represent the values of τ and the rest of the columns represent the percentage of risk improvements of each of the improved estimators. The table consists of several cells. In each cell, corresponding to one choice of τ, there correspond four values of percentage of risk values for the sample sizes $(5,5)$, $(12,12)$, $(20,20)$ and $(30,30)$. Table 2, is divided into two parts, specifically the first half (column second to sixth) represents the percentage of risk performances for all the estimators with sample sizes $(5,10)$, and $(12,20)$. The second part (column seventh to eleventh) represents the percentage of risk improvements for the sample sizes $(10,5)$ and $(20,12)$. In this table the first column also represents the values of τ and the columns second onwards represent the percentage of risk improvements of all the estimators upon their unrestricted counterparts. In this table each cell contains two values of percentage of risk improvements. These two values correspond to one value of τ. In a very similar fashion the percentage of risk improvements of all the estimators have been presented in Tables 3 to 5 for equal and unequal sample sizes with respect to the loss functions (1.2) and (1.3).

The percentage of relative risk improvements of all the improved estimators with respect to the benchmark estimator d_{GD} (denoted as Ri; $i=1,2,7$) have been presented in Tables 6 and 7 for equal and unequal sample sizes respectively. Specifically, in Table 6 we have presented the percentage of relative risk performances of all the improved estimators for the sample sizes $(5,5)$, $(12,12)$ and $(20,20)$. The Table 6 consists of eight columns and each column have several cells. Corresponding to each value of τ there correspond three values of percentage of relative risks. These three values correspond to three sample sizes $(5,5)$, $(12,12)$ and $(20,20)$ respectively. In a very similar way we have presented the percentage of relative risk improvements of all the improved estimators for the unequal sample sizes $(5,10)$, $(12,20)$, $(10,5)$ and $(20,12)$ in Table 7. Moreover, we have also plotted the risk values of all the improved estimators with respect to the loss function (1.1), against the choices of τ in Figure 1. Specifically, Figure 1 (a)-(b) presents the graph for equal sample sizes whereas Figure 1 (c)-(f) presents for unequal sample sizes. We note that the estimators $\hat{d}_{GD}$, $\hat{d}_{KS}$, $\hat{d}_{MK}$, $\hat{d}_{TK}$, $\hat{d}_{RM}$, d^a_{GD}, d^a_{MK}, d^a_{BC1}, d^a_{BC2} have been denoted by GDI, KSI, MKI, TKI, RML, GDA, MKA, BC1A and BC2A respectively in Figure 1 (a)-(f).

The following observations have been made during our simulation study as well as from the tables, which we discuss separately for equal and unequal sample sizes.

Case I: $m=n$.

1. The percentage of risk improvements as well as the risk values of all the new estimators upon their respective unrestricted counterparts decreases as the sample sizes increase for fixed values of the parameter, with respect to the loss functions L_1, L_2 and L_3.
2. Let the loss function be L_1. The percentage of risk improvement of $\hat{d}_{GD}$ (see $P1$) is seen maximum up to 12%. The percentage of risk improvement of d^a_{GD} (see $P5$) is seen maximum up to 10%. The percentage of risk improvement of $\hat{d}_{MK}$ (see $P3$) is seen maximum up to 7%. The percentage of

risk improvement of d^a_{MK} (see $P6$) is seen maximum up to 6%, where the percentage of risk improvement of $\hat{d}_{RM}$ (see $P9$) is seen maximum up to 20%.

3. Let the loss function be L_2. The maximum percentage of risk improvement of $\hat{d}_{GD}$, d^a_{GD}, $\hat{d}_{MK}$, d^a_{MK}, and $\hat{d}_{RM}$ upon their respective unrestricted counterparts are seen near to 6%, 5%, 4%, 3% and 7% respectively. The maximum percentage of risk improvement is seen in the case of $\hat{d}_{RM}$ for small sample sizes and when σ_1^2 and σ_2^2 are close to each other.

4. Let the loss function be L_3. The maximum percentage of risk improvement of $\hat{d}_{GD}$, d^a_{GD}, $\hat{d}_{MK}$, d^a_{MK}, and $\hat{d}_{RM}$ upon their respective unrestricted counterparts are seen respectively as 11%, 10%, 7%, 5% and 20%. The maximum percentage of risk improvement of each of the estimators has been noticed for small sample sizes and when the variances are close to each other.

5. Here we note that, the percentage of risk improvements of all the new estimators upon their respective unrestricted counterparts are approximated values only which have been obtained numerically and hence it may vary with sample sizes.

6. The above numerical results $(2)-(4)$ validates the theoretical findings in Sections 3, 4, and 5.

7. The risk values of all the estimators such as $\hat{d}_{GD}$, d^a_{GD}, $\hat{d}_{MK}$, d^a_{MK}, d^a_{BC1}, d^a_{BC2}, and $\hat{d}_{RM}$, decrease as the sample sizes increase. Further for fixed sample sizes, as the values of τ varies from 0 to 1, the risk values of all the estimators decrease. It has been noticed that, for small values of τ (say $0 < \tau < 0.25$), the percentage of relative risk improvement of ${d^a}_{BC1}$ is maximum and seen up to 15%. For the values of τ near to 1, (say for the range $0.50 < \tau < 1$) the estimators $\hat{d}_{MK}$ and d^a_{MK} have almost same percentage of relative risk improvements. For moderate values of τ (say $0.50 < \tau < 0.75$), the estimators $\hat{d}_{MK}$ and d^a_{MK} perform equally well, however as the sample sizes increases from moderate to large, the performance of these two estimators decrease and compete well with $\hat{d}_{GD}$. In fact, the dominance regions of $\hat{d}_{MK}$ and d^a_{MK} upon $\hat{d}_{GD}$ decrease. It has also been noticed that the estimators $\hat{d}_{GD}$, $\hat{d}_{MK}$, and d^a_{BC1} compete with d^a_{GD}, d^a_{MK} and d^a_{BC2} respectively.

Case II: $m \neq n$.

1. The percentage of risk improvements of all the improved estimators decrease as the sample sizes increase for fixed values of σ_1^2 and σ_2^2 with respect to the loss functions L_1, L_2 and L_3.

2. Let us first consider the loss function L_1. The percentage of risk improvement of $\hat{d}_{GD}$ upon d_{GD} (denoted as $P1$) is seen maximum up to 16%, the maximum percentage of risk improvement of $\hat{d}_{KS}$ upon d_{KS} (denoted as

$P2$) is seen near to 8%. The maximum percentage of risk improvement of $\hat{d}_{MK}$ and $\hat{d}_{TK}$ over their corresponding unrestricted counterparts are seen near to 14% and 13% respectively. The maximum percentage of risk improvement of $\hat{d}_{RM}$ upon d_{ML} is seen up to 15%. We also note that, these maximum risk improvements have been noticed when $m > n$ for all the estimators.

3. Let us consider the loss function L_2. The maximum percentage of risk improvement of $\hat{d}_{GD}$ upon d_{GD} is seen up to 7%. The maximum percentage of risk improvement of $\hat{d}_{KS}$ over d_{KS} is seen near to 4%. The maximum percentage of risk improvement of $\hat{d}_{MK}$ upon d_{MK} is seen near to 7%. The maximum percentage of risk improvement of $\hat{d}_{TK}$ upon d_{TK} is seen near to 7%. The maximum percentage of risk improvement of $\hat{d}_{RM}$ upon d_{ML} is seen near to 13%.

4. Consider the loss function L_3. The maximum percentage of risk improvement of $\hat{d}_{GD}$ upon d_{GD} is seen up to 13%. The maximum percentage of risk improvement of $\hat{d}_{KS}$ upon d_{KS} is seen near to 8%. The maximum percentage of risk improvement of $\hat{d}_{MK}$ upon d_{MK} is seen near to 13%. The maximum percentage of risk improvement of $\hat{d}_{TK}$ upon d_{TK} is seen near to 13%. The maximum percentage of risk improvement of $\hat{d}_{RM}$ upon d_{ML} is seen near to 36%.

5. Here we note that, the percentage of risk improvements of all the improved estimators upon their respective unrestricted counterparts are approximated values only which have been obtained numerically and hence it may vary with sample sizes, however the trends remain the same.

6. The above numerical results $(2)-(4)$ also validates the theoretical findings in Sections 3, 4, and 5.

7. The risk values of all the estimators, such as $\hat{d}_{GD}$, $\hat{d}_{KS}$, $\hat{d}_{MK}$, $\hat{d}_{TK}$, d^a_{BC1}, d^a_{BC2}, and $\hat{d}_{RM}$, decrease as the sample sizes increase. It has been noticed that for small values of τ (say $0 < \tau < 0.15$), the percentage of relative risk improvements of d^a_{BC1} and d^a_{BC2} are maximum and seen up to 12%. For the values of τ near to 1, (say $0.75 < \tau < 1$) the estimator $\hat{d}_{KS}$ (for $m < n$) and $\hat{d}_{MK}$, $\hat{d}_{TK}$ (when $m > n$) has maximum percentage of relative risk improvements. For moderate values of τ, the estimators $\hat{d}_{MK}$ and $\hat{d}_{TK}$ perform equally well, however as the sample sizes increase from moderate to large the performance of these two estimators decrease and in this case the estimators $\hat{d}_{GD}$ and $\hat{d}_{KS}$ perform better.

From the above discussions and also from our simulation study the following conclusions can be drawn regarding the use of the proposed estimators in practice.

1. Let us consider that the sample sizes are equal, that is $m = n$. When the variance of the first population is much smaller compare to the second, we recommend to use d^a_{BC1}. When the variance of both the populations are

close to each other, we recommend to use either $\hat{d}_{MK}$ or d^a_{MK}, as they compete with each other. In other cases, that is neither the variances differ too much nor close to each other, the estimators $\hat{d}_{MK}$ and d^a_{MK} can be used for small sample sizes (say $m, n \leq 10$), and $\hat{d}_{MK}$ or $\hat{d}_{GD}$ for moderate to large sample sizes.

2. Consider that the sample sizes are unequal, that is $m \neq n$. When the variance of the first population is much smaller than the second, we recommend to use either the estimator d^a_{BC1} or d^a_{BC2}. When the variances of both the populations are close to each other, the estimators $\hat{d}_{KS}$ or $\hat{d}_{TK}$ (for $m < n$) and $\hat{d}_{TK}$ or $\hat{d}_{MK}$ (for $m > n$) can be recommended for use. However for moderate ranges of τ, the estimators $\hat{d}_{MK}$ or $\hat{d}_{TK}$ (for $m < n$) and the estimators $\hat{d}_{KS}$, $\hat{d}_{GD}$, $\hat{d}_{RM}$ or $\hat{d}_{TK}$ (for $m > n$) can be preferred as they all perform equally well.

7. CONCLUDING REMARKS AND EXAMPLES

In this paper, we have re-investigated the problem of estimating common mean of two normal populations when the variances are known to follow the ordering $\sigma_1^2 \leq \sigma_2^2$. It should be noted that, Elfessi and Pal [7] considered this model and obtained an estimator which dominates the well known Graybill-Deal (see Graybill and Deal [8]) estimator in terms of stochastic domination as well as Pitman measure of closeness criterion. We have proposed some new estimators for the common mean under order restricted variances which beat their unrestricted counterparts (previously proposed by Khatri and Saha [11], Moore and Krishnamoorthy [14], Tripathy and Kumar [20], Brown and Cohen [3]) stochastically, universally and in terms of Pitman measure of closeness criterion and compete well with the estimator proposed by Elfessi and Pal [7]. Moreover, we have obtained a plug-in type restricted MLE which beats the unrestricted MLE with respect to a squared error loss function. In addition to these, we have derived a sufficient condition for improving equivariant estimators using orbit-by-orbit improvement technique of Brewster and Zidek [2]. To the best of our knowledge, the performance of the MLE of the common mean has not been discussed under order restricted variances in the literature which was also lacking and we have tried to answer up to some extent. We have carried out a detailed and in-depth simulation study in order to compare the performances of both proposed as well as existing estimators with that of the plug-in type restricted MLE which was lacking in the literature. Under order restriction on the variances we have recommended estimators that can be used in practice. We hope that the current study may fill the knowledge gap and provide useful information to the researchers from an application point of view.

Example 7.1. (Simulated Data): The following two data sets each of size 10 from normal distributions have been generated using the software R, with

$\tau\downarrow$	$P1$	$P5$	$P3$	$P6$	$P9$	$\tau\downarrow$	$P1$	$P5$	$P3$	$P6$	$P9$
	1.93	1.17	0.72	0.41	17.78		9.76	8.91	5.93	4.43	11.82
0.05	0.00	0.00	0.00	0.00	1.63	0.55	3.85	2.77	2.06	1.27	3.53
	0.00	0.00	0.00	0.00	0.00		1.42	0.94	0.73	0.42	1.04
	0.00	0.00	0.00	0.00	0.00		0.43	0.29	0.22	0.13	0.31
	5.66	3.53	2.56	1.46	13.90		8.93	9.13	5.42	4.42	11.77
0.10	0.00	0.00	0.00	0.00	0.31	0.60	3.95	2.92	2.11	1.33	3.55
	0.00	0.00	0.00	0.00	0.00		1.50	1.09	0.77	0.48	1.23
	0.00	0.00	0.00	0.00	0.00		0.83	0.53	0.42	0.24	0.57
	6.52	4.19	3.19	1.87	11.42		8.46	9.08	5.14	4.34	11.39
0.15	0.16	0.09	0.07	0.04	0.46	0.65	3.32	2.92	1.79	1.26	3.61
	0.00	0.00	0.00	0.00	0.00		1.77	1.38	0.92	0.59	1.55
	0.00	0.00	0.00	0.00	0.00		1.01	0.70	0.52	0.31	0.75
	7.88	5.24	4.09	2.44	10.55		7.43	9.14	4.62	4.32	11.39
0.20	0.51	0.30	0.23	0.13	0.67	0.70	4.07	3.59	2.23	1.57	4.32
	0.02	0.01	0.01	0.01	0.01		2.12	1.70	1.10	0.73	1.92
	0.00	0.00	0.00	0.00	0.00		1.30	0.93	0.67	0.41	0.99
	8.30	5.65	4.49	2.73	10.63		6.29	9.04	3.80	4.14	11.61
0.25	0.49	0.30	0.23	0.13	0.47	0.75	3.22	3.41	1.76	1.41	4.22
	0.06	0.04	0.03	0.02	0.04		2.09	1.88	1.10	0.78	2.11
	0.01	0.00	0.00	0.00	0.00		1.04	0.98	0.53	0.39	1.05
	9.23	6.50	5.14	3.18	10.28		6.64	9.82	4.14	4.55	11.69
0.30	0.88	0.53	0.43	0.24	0.80	0.80	2.96	3.81	1.61	1.49	4.80
	0.15	0.09	0.07	0.04	0.11		2.30	2.27	1.22	0.92	2.57
	0.07	0.04	0.03	0.02	0.04		1.51	1.31	0.78	0.54	1.4
	10.02	7.35	5.70	3.63	11.26		3.39	8.52	2.11	3.65	10.86
0.35	1.57	0.99	0.79	0.45	1.48	0.85	2.86	4.15	1.61	1.61	5.05
	0.26	0.15	0.13	0.07	0.19		2.21	2.57	1.17	0.99	2.92
	0.05	0.03	0.02	0.01	0.03		1.67	1.67	0.87	0.66	1.81
	9.70	7.42	5.60	3.68	11.00		3.38	9.01	2.07	3.80	11.72
0.40	2.11	1.35	1.07	0.62	1.81	0.90	1.47	3.65	0.80	1.26	4.67
	0.52	0.32	0.26	0.14	0.37		1.35	2.31	0.72	0.81	2.64
	0.04	0.02	0.02	0.01	0.02		1.25	1.66	0.65	0.60	1.79
	10.91	8.78	6.49	4.41	12.00		1.82	8.51	1.19	3.53	10.70
0.45	2.35	1.53	1.22	0.70	1.99	0.95	0.81	3.62	0.47	1.20	4.48
	0.55	0.36	0.27	0.16	0.41		1.27	2.60	0.69	0.89	2.94
	0.19	0.11	0.09	0.05	0.12		0.05	1.32	0.02	0.36	1.46
	10.17	8.62	6.02	4.25	11.23		0.68	7.72	0.49	2.83	9.87
0.50	3.47	2.36	1.82	1.08	3.06	1.00	0.55	3.34	0.31	0.93	4.21
	1.11	0.73	0.56	0.33	0.85		0.21	2.23	0.11	0.59	2.55
	0.32	0.21	0.16	0.09	0.23		0.78	2.09	0.42	0.67	2.29

Table 1: Percentage of Risk Improvements of all the Proposed Estimators Using the Loss L_1 for the Sample Sizes $(m,n) = (5,5),(12,12),(20,20),(30,30)$

a common mean $\mu = 25$ and with the condition that $\sigma_1^2 \le \sigma_2^2$. Data Set A: 24.28, 25.94, 25.76, 29.14, 28.39, 23.51, 23.43, 22.60, 22.29, 28.26. Data Set B: 24.61, 23.70, 26.25, 29.11, 26.13, 23.52, 25.57, 22.34, 26.19, 23.04. The sufficient statistics can be computed as $\bar{x} = 25.36$, $\bar{y} = 25.04$, $s_1^2 = 57.69$, $s_2^2 = 36.46$. Based on the summery data it is seen that $s_1^2 > s_2^2$. This is a case where the improved estimators can be obtained. The various estimators are computed as $d_{GD} = 25.17$, $\hat{d}_{GD} = 25.24$, $d^a_{GD} = 25.20$, $d_{KS} = 25.17$, $\hat{d}_{KS} = 25.24$, $d^a_{KS} = 25.20$, $d_{MK} = 25.18$, $\hat{d}_{MK} = 25.22$, $d^a_{MK} = 25.20$, $d_{TK} = 25.18$, $\hat{d}_{TK} = 25.22$, $d^a_{TK} = 25.20$,

$\tau \downarrow$	$(m,n) = (5,10),(12,20)$					$(m,n) = (10,5),(20,12)$				
	$P1$	$P2$	$P3$	$P4$	$P9$	$P1$	$P2$	$P3$	$P4$	$P9$
0.05	0.00	0.01	0.00	0.00	0.02	2.18	0.89	2.19	1.98	43.23
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.86
0.10	0.12	0.47	0.00	0.00	0.26	4.21	1.94	4.54	4.12	39.23
	0.00	0.00	0.00	0.00	0.05	0.02	0.00	0.10	0.09	0.83
0.15	0.18	0.92	0.01	0.01	0.68	5.79	2.89	6.20	5.65	32.09
	0.00	0.01	0.00	0.00	0.03	0.02	0.01	0.19	0.18	0.73
0.20	0.61	1.80	0.06	0.07	1.26	6.32	3.12	7.25	6.60	26.89
	0.01	0.03	0.00	0.00	0.05	0.25	0.18	0.53	0.50	0.34
0.25	0.72	2.34	0.06	0.07	1.45	7.70	3.77	8.79	8.05	25.31
	0.04	0.07	0.00	0.00	0.11	0.43	0.30	1.06	1.00	0.80
0.30	1.06	3.18	0.08	0.10	2.17	8.27	4.10	9.36	8.56	19.80
	0.12	0.19	0.00	0.00	0.22	0.56	0.40	1.35	1.27	0.44
0.35	1.47	4.01	0.09	0.11	2.59	9.66	4.86	10.33	9.50	20.65
	0.22	0.34	0.00	0.00	0.41	1.01	0.76	2.03	1.93	0.97
0.40	1.74	4.60	0.13	0.16	2.83	10.28	5.06	11.21	10.32	19.58
	0.39	0.58	0.01	0.01	0.72	1.34	0.99	2.65	2.52	1.37
0.45	2.06	5.09	0.18	0.22	3.41	10.6	5.31	11.32	10.44	17.64
	0.55	0.81	0.01	0.02	0.98	1.77	1.30	3.30	3.15	1.51
0.50	2.56	6.05	0.21	0.26	4.08	11.26	5.93	11.62	10.75	17.39
	0.77	1.12	0.03	0.03	1.15	1.91	1.41	3.41	3.26	1.61
0.55	2.66	6.18	0.25	0.31	4.02	11.91	6.15	11.91	11.04	18.19
	1.08	1.53	0.05	0.06	1.66	2.31	1.71	3.82	3.65	1.97
0.60	2.69	6.22	0.25	0.30	4.07	11.77	6.11	11.73	10.88	17.98
	1.26	1.75	0.04	0.05	1.84	2.39	1.78	3.91	3.74	2.12
0.65	2.98	6.52	0.31	0.37	4.49	12.84	6.70	12.26	11.39	17.61
	1.36	1.88	0.06	0.07	2.01	3.24	2.49	4.56	4.38	2.95
0.70	2.86	6.28	0.31	0.36	4.21	12.64	6.65	11.67	10.83	16.90
	1.78	2.35	0.11	0.13	2.62	3.55	2.72	4.79	4.61	3.25
0.75	3.21	6.75	0.30	0.36	4.72	12.37	6.56	11.15	10.37	16.05
	1.93	2.53	0.12	0.13	2.66	3.41	2.63	4.37	4.20	3.27
0.80	3.75	7.44	0.46	0.54	5.39	12.47	6.46	11.25	10.45	15.33
	2.06	2.73	0.10	0.10	2.79	3.97	3.10	4.58	4.42	3.80
0.85	3.31	6.95	0.32	0.38	4.81	11.99	6.54	9.93	9.23	15.25
	2.43	3.09	0.18	0.19	3.30	4.00	3.10	4.51	4.34	3.83
0.90	3.41	6.87	0.39	0.45	4.89	11.40	5.77	9.40	8.70	14.64
	2.33	2.96	0.18	0.19	3.26	4.16	3.21	4.46	4.29	4.01
0.95	3.33	6.78	0.29	0.35	4.96	11.83	6.45	9.25	8.58	14.60
	2.26	2.86	0.15	0.16	3.02	3.95	3.05	3.90	3.76	3.91
1.00	2.99	6.27	0.24	0.28	4.32	11.40	6.12	8.47	7.85	13.82
	1.94	2.50	0.07	0.08	2.83	4.09	3.23	3.52	3.38	4.09

Table 2: Percentage of Risk Improvements of all the Proposed Estimators Using the Loss L_1 for Unequal Sample Sizes

$d_{BC1} = 25.25$, $d^a_{BC1} = 25.25$, $d_{BC2} = 25.18$, $\hat{d}_{BC2} = 25.20$, $d^a_{BC2} = 25.23$, $d_{ML} = 25.17$, $\hat{d}_{RM} = 25.20$, $d^a_{ML} = 25.20$. Depending upon the variance ratios, the improved estimators can be used.

Example 7.2. Rohatgi and Saleh [19], (p.515) discussed one example regarding the mean life time (in hours) of light bulbs. Suppose a random sample of 9 bulbs has sample mean 1309 hours with standard deviation of 420 hours. A second sample of 16 bulbs chosen from a different batch has sample mean

$(m, n) \downarrow$	$(\sigma_1^2, \sigma_2^2) \downarrow$	$L_2 -$ Loss					$L_3 -$ Loss				
		$P1$	$P5$	$P3$	$P6$	$P9$	$P1$	$P5$	$P3$	$P6$	$P9$
	(0.05, 0.10)	4.62	3.96	2.81	1.99	5.37	10.73	9.21	6.55	4.69	11.72
	(0.05, 0.30)	2.27	1.33	1.20	0.65	3.28	6.48	4.25	3.23	1.91	11.12
	(0.05, 0.50)	1.26	0.70	0.63	0.33	3.06	3.94	2.41	1.72	0.97	13.05
	(0.05, 0.70)	0.77	0.42	0.36	0.19	3.54	3.59	2.22	1.48	0.83	16.24
	(0.05, 1.00)	0.61	0.32	0.27	0.14	3.27	2.79	1.75	1.07	0.61	19.27
	(1.00, 1.10)	1.19	4.25	0.76	1.78	5.45	1.99	8.39	1.21	3.43	10.50
	(1.00, 1.50)	3.71	4.17	2.27	1.96	5.45	8.26	9.29	5.10	4.45	11.27
(5, 5)	(1.00, 2.00)	4.53	3.81	2.74	1.89	5.15	10.33	8.67	6.21	4.37	11.48
	(1.00, 2.50)	4.71	3.43	2.76	1.74	4.91	9.80	7.52	5.66	3.72	10.99
	(1.00, 3.00)	4.31	2.91	2.48	1.48	4.35	9.40	6.78	5.28	3.34	10.71
	(2.00, 2.10)	0.21	3.94	0.18	1.43	5.03	1.61	8.57	0.98	3.49	10.69
	(2.00, 2.30)	1.43	4.41	0.92	1.86	5.64	3.37	8.83	2.06	3.75	11.11
	(2.00, 2.50)	3.16	4.87	2.06	2.25	6.19	6.11	9.37	3.86	4.33	12.05
	(2.00, 2.70)	3.06	4.37	1.87	1.99	5.64	6.34	8.90	3.96	4.13	11.08
	(2.00, 3.00)	4.26	4.39	2.67	2.14	5.74	8.38	9.18	5.21	4.46	11.75
	(0.05, 0.10)	1.56	1.04	0.84	0.49	1.24	2.61	1.90	1.37	0.85	2.49
	(0.05, 0.30)	0.04	0.02	0.02	0.01	0.09	0.17	0.09	0.07	0.04	0.26
	(0.05, 0.50)	0.01	0.00	0.00	0.00	0.04	0.01	0.00	0.00	0.00	0.48
	(0.05, 0.70)	0.00	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.29
	(0.05, 1.00)	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.38
	(1.00, 1.10)	1.39	2.19	0.83	0.88	2.65	0.98	3.51	0.53	1.16	4.47
	(1.00, 1.50)	1.99	1.68	1.12	0.76	2.01	3.94	3.35	2.13	1.47	4.21
(12, 12)	(1.00, 2.00)	1.41	0.98	0.78	0.46	1.19	2.77	1.98	1.46	0.89	2.59
	(1.00, 2.50)	0.89	0.59	0.46	0.27	0.75	1.73	1.16	0.88	0.52	1.50
	(1.00, 3.00)	0.61	0.36	0.31	0.17	0.47	1.35	0.82	0.67	0.37	1.18
	(2.00, 2.10)	0.26	1.84	0.15	0.57	2.31	1.37	3.98	0.76	1.37	4.99
	(2.00, 2.30)	0.92	1.80	0.46	0.63	2.22	1.80	3.61	1.00	1.31	4.45
	(2.00, 2.50)	1.18	1.74	0.64	0.66	2.13	3.37	4.06	1.86	1.64	4.95
	(2.00, 2.70)	2.08	1.93	1.15	0.83	2.33	4.04	3.90	2.21	1.66	4.77
	(2.00, 3.00)	1.85	1.55	1.00	0.69	1.87	3.78	3.30	2.04	1.43	4.16
	(0.05, 0.10)	0.56	0.34	0.29	0.16	0.38	1.12	0.72	0.57	0.33	0.81
	(0.05, 0.30)	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.01
	(0.05, 0.50)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.05, 0.70)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.05, 1.00)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(1.00, 1.10)	0.34	1.00	0.20	0.34	1.14	1.23	2.37	0.65	0.81	2.74
	(1.00, 1.50)	0.95	0.72	0.48	0.31	0.81	1.97	1.51	1.03	0.65	1.67
(20, 20)	(1.00, 2.00)	0.62	0.37	0.33	0.18	0.41	1.10	0.69	0.56	0.32	0.79
	(1.00, 2.50)	0.21	0.12	0.10	0.05	0.14	0.42	0.25	0.21	0.11	0.28
	(1.00, 3.00)	0.12	0.06	0.06	0.03	0.07	0.28	0.16	0.14	0.07	0.20
	(2.00, 2.10)	0.23	1.14	0.15	0.37	1.30	0.40	2.21	0.21	0.67	2.59
	(2.00, 2.30)	0.95	1.22	0.47	0.44	1.40	1.52	2.21	0.80	0.80	2.52
	(2.00, 2.50)	1.09	1.11	0.59	0.46	1.23	2.01	2.05	1.06	0.81	2.29
	(2.00, 2.70)	1.03	0.92	0.53	0.37	1.03	2.45	2.07	1.29	0.88	2.31
	(2.00, 3.00)	0.80	0.65	0.41	0.27	0.74	1.79	1.39	0.93	0.60	1.55

Table 3: Percentage of Risk Improvements of all the Proposed Estimators Using the Loss L_2 and L_3 loss

1205 hours and standard deviation 390 hours. A two sample t-test fails to reject the hypothesis that the means are equal. Suppose it is known a priori that the variance of first sample is smaller than the second one. This is a situation where our model will be useful. On the basis of these samples, we have $m = 9$,

$(m,n)\downarrow$	$(\sigma_1^2,\sigma_2^2)\downarrow$	L_2- Loss					L_3- Loss				
		$P1$	$P2$	$P3$	$P4$	$P9$	$P1$	$P2$	$P3$	$P4$	$P9$
(5, 10)	(0.05, 0.10)	1.27	2.84	0.12	0.14	1.87	2.46	5.73	0.18	0.22	3.86
	(0.05, 0.30)	0.11	0.42	0.01	0.01	0.25	0.29	1.10	0.02	0.02	0.72
	(0.05, 0.50)	0.03	0.13	0.00	0.00	0.08	0.10	0.39	0.01	0.01	0.28
	(0.05, 0.70)	0.01	0.06	0.00	0.00	0.04	0.04	0.15	0.00	0.00	0.08
	(0.05, 1.00)	0.00	0.02	0.00	0.00	0.01	0.00	0.04	0.00	0.00	0.05
	(1.00, 1.10)	1.44	2.97	0.13	0.15	2.15	3.52	6.91	0.38	0.45	4.94
	(1.00, 1.50)	1.59	3.45	0.17	0.20	2.37	3.21	6.94	0.32	0.39	4.86
	(1.00, 2.00)	1.21	2.89	0.09	0.11	1.92	2.56	6.03	0.21	0.26	4.17
	(1.00, 2.50)	0.81	2.21	0.06	0.07	1.42	1.92	4.95	0.15	0.19	3.29
	(1.00, 3.00)	0.59	1.71	0.04	0.05	1.04	1.22	3.55	0.08	0.10	2.43
	(2.00, 2.10)	1.73	3.37	0.21	0.24	2.52	3.00	6.39	0.20	0.25	4.43
	(2.00, 2.30)	1.75	3.56	0.17	0.20	2.52	3.50	7.07	0.36	0.42	4.90
	(2.00, 2.50)	1.67	3.47	0.17	0.20	2.47	3.36	7.03	0.32	0.39	5.13
	(2.00, 2.70)	1.62	3.50	0.15	0.18	2.46	3.20	6.96	0.32	0.38	4.73
	(2.00, 3.00)	1.68	3.60	0.18	0.21	2.51	3.15	6.83	0.32	0.38	4.98
(12,20)	(0.05, 0.10)	0.45	0.64	0.02	0.03	0.68	0.77	1.13	0.01	0.01	1.24
	(0.05, 0.30)	0.00	0.01	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.02
	(0.05, 0.50)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.05, 0.70)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.05, 1.00)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(1.00, 1.10)	0.96	1.26	0.04	0.04	1.32	2.28	2.93	0.16	0.17	3.24
	(1.00, 1.50)	0.77	1.05	0.03	0.04	1.10	1.66	2.25	0.10	0.10	2.36
	(1.00, 2.00)	0.37	0.53	0.01	0.02	0.56	0.74	1.05	0.03	0.03	1.12
	(1.00, 2.50)	0.18	0.27	0.00	0.00	0.28	0.35	0.53	0.00	0.01	0.61
	(1.00, 3.00)	0.08	0.13	0.00	0.00	0.17	0.17	0.29	0.00	0.00	0.34
	(2.00, 2.10)	1.14	1.45	0.07	0.07	1.51	2.13	2.69	0.16	0.17	2.86
	(2.00, 2.30)	1.19	1.52	0.08	0.08	1.60	2.42	3.08	0.21	0.22	3.21
	(2.00, 2.50)	0.88	1.18	0.05	0.05	1.20	1.78	2.37	0.12	0.13	2.58
	(2.00, 2.70)	0.82	1.08	0.06	0.06	1.14	2.06	2.69	0.16	0.17	2.88
	(2.00, 3.00)	0.76	1.01	0.05	0.05	1.06	1.59	2.12	0.10	0.11	2.27

Table 4: Percentage of Risk Improvements of all the Proposed Estimators Using the Loss L_2 and L_3.

$n = 16$, $\bar{x} = 1309$, $\bar{y} = 1205$, $s_1 = (\sqrt{m-1})420$, $s_2 = (\sqrt{n-1})390$. The various estimators for the common mean are obtained as $d_{GD} = 1269.27$, $\hat{d}_{GD} = 1271.24$, $d_{KS} = 1267.44$, $\hat{d}_{KS} = 1271.24$, $d_{MK} = 1274.22$, $\hat{d}_{MK} = 1274.22$, $d_{TK} = 1274.01$, $\hat{d}_{TK} = 1274.01$, $d_{BC1} = 1284.37$, $d^a_{BC1} = 1284.37$, $d_{BC2} = 1287.32$, $d^a_{BC2} = 1287.32$, $d_{ML} = 1269.87$, $\hat{d}_{RM} = 1271.24$, and $d^a_{ML} = 1271.24$. In this situation we recommend to use either of the estimators $\hat{d}_{GD}$, $\hat{d}_{KS}$, or d^a_{ML}.

8. Appendix

Proof of Theorem 3.1

Proof: (i) First we prove the dominance result for equal sample sizes

$(m,n)\downarrow$	$(\sigma_1^2,\sigma_2^2)\downarrow$	L_2- Loss					L_3- Loss				
		$P1$	$P2$	$P3$	$P4$	$P9$	$P1$	$P2$	$P3$	$P4$	$P9$
(10, 5)	(0.05, 0.10)	4.70	2.27	5.51	5.07	6.73	10.89	5.43	11.52	10.63	17.75
	(0.05, 0.30)	1.79	0.81	2.49	2.24	6.69	5.34	2.37	6.24	5.64	28.01
	(0.05, 0.50)	1.08	0.53	1.40	1.26	7.27	3.51	1.53	3.80	3.43	31.23
	(0.05, 0.70)	0.71	0.30	0.96	0.86	7.45	3.48	1.58	3.29	2.99	40.10
	(0.05, 1.00)	0.43	0.21	0.60	0.54	6.71	1.83	0.63	1.97	1.77	43.66
	(1.00, 1.10)	5.78	3.04	4.89	4.51	7.06	12.93	7.12	10.35	9.65	16.21
	(1.00, 1.50)	5.57	2.92	5.52	5.10	7.09	12.55	6.61	11.90	11.04	16.55
	(1.00, 2.00)	4.88	2.44	5.66	5.21	6.60	11.60	5.89	12.07	11.16	19.34
	(1.00, 2.50)	4.30	2.07	5.24	4.80	7.21	10.41	5.32	11.19	10.31	19.67
	(1.00, 3.00)	3.42	1.58	4.52	4.11	6.99	9.87	5.14	10.47	9.67	20.79
	(2.00, 2.10)	6.24	3.44	5.05	4.69	7.48	11.78	6.31	9.13	8.44	15.06
	(2.00, 2.30)	6.06	3.21	5.19	4.83	7.05	12.13	6.47	10.05	9.32	15.78
	(2.00, 2.50)	6.11	3.30	5.59	5.19	7.23	12.55	6.60	10.90	10.13	15.61
	(2.00, 2.70)	6.13	3.24	5.82	5.41	7.43	13.47	7.47	11.94	11.14	16.27
	(2.00, 3.00)	5.94	3.16	5.96	5.52	7.40	12.78	6.70	12.05	11.19	17.28
(20,12)	(0.05, 0.10)	0.88	0.64	1.71	1.63	0.74	1.91	1.42	3.46	3.29	1.71
	(0.05, 0.30)	0.03	0.02	0.10	0.10	0.11	0.03	0.22	0.20	0.00	1.16
	(0.05, 0.50)	0.00	0.00	0.02	0.02	0.06	0.00	0.00	0.03	0.03	2.40
	(0.05, 0.70)	0.00	0.00	0.00	0.00	0.04	0.01	0.00	0.02	0.02	1.01
	(0.05, 1.00)	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	1.66
	(1.00, 1.10)	2.08	1.64	2.19	2.12	2.04	4.26	3.38	4.20	4.06	4.16
	(1.00, 1.50)	1.71	1.31	2.47	2.38	1.55	3.02	2.30	4.33	4.16	2.72
	(1.00, 2.00)	0.91	0.68	1.73	1.65	0.74	2.03	1.51	3.62	3.46	1.75
	(1.00, 2.50)	0.53	0.39	1.18	1.12	0.43	1.27	0.93	2.61	2.48	1.13
	(1.00, 3.00)	0.34	0.24	0.83	0.78	0.28	0.64	0.43	1.67	1.58	0.50
	(2.00, 2.10)	1.71	1.26	1.72	1.65	1.65	3.90	3.06	3.70	3.56	3.89
	(2.00, 2.30)	1.80	1.40	1.98	1.91	1.74	3.83	2.97	4.11	3.96	3.83
	(2.00, 2.50)	1.80	1.40	2.20	2.12	1.66	3.52	2.69	4.35	4.18	3.27
	(2.00, 2.70)	1.74	1.33	2.38	2.29	1.62	3.63	2.80	4.68	4.51	3.47
	(2.00, 3.00)	1.64	1.25	2.39	2.30	1.45	3.53	2.73	4.80	4.62	3.34

Table 5: Percentage of Risk Improvements of all the Proposed Estimators Using the Loss L_2 and L_3

that is $m=n$. Consider the estimator $\hat{d}_{MK}$ which is given by

$$\hat{d}_{MK} = \begin{cases} (1-\beta_1)\bar{X}+\beta_1\bar{Y}, & \text{if } S_1 \le S_2, \\ \beta_1\bar{X}+(1-\beta_1)\bar{Y}, & \text{if } S_1 > S_2. \end{cases}$$

Our target is to show that,

$$P[(\hat{d}_{MK}-\mu)^2 \le c] \le P[(d_{MK}-\mu)^2 \le c], \quad \forall\, c>0. \tag{8.1}$$

Which is equivalent to show that

$$P[(d_{MK}-\mu)^2 \le c | S_1 > S_2] \le P[(\hat{d}_{MK}-\mu)^2 \le c | S_1 > S_2], \quad \forall\, c>0.$$

Denoting $X_1^* = (1-\beta_1)\bar{X}+\beta_1\bar{Y}$ and $X_2^* = \beta_1\bar{X}+(1-\beta_1)\bar{Y}$, we observe that $X_1^*-\mu \sim N(0,\sigma^2)$, $X_2^*-\mu \sim N(0,\sigma_*^2)$, where $\sigma^2 = (1-\beta_1)^2\frac{\sigma_1^2}{m}+\beta_1^2\frac{\sigma_2^2}{m}$ and $\sigma_*^2 = \beta_1^2\frac{\sigma_1^2}{m}+(1-\beta_1)^2\frac{\sigma_2^2}{m}$. Thus incorporating all these information the above inequality reduces to,

$$\Phi\Big(\frac{\sqrt{c}}{\sigma}\Big) \le \Phi\Big(\frac{\sqrt{c}}{\sigma_*}\Big), \ \forall\, c>0 \text{ and } S_1 > S_2, \tag{8.2}$$

$\tau\downarrow$	$(m,n)=(5,5),(12,12),(20,20)$						
	$R1$	$R5$	$R3$	$R6$	$R7$	$R8$	$R9$
	1.93	1.17	-38.84	-39.28	13.40	12.32	-18.31
0.05	0.00	0.00	-44.28	-44.28	1.63	1.52	-0.52
	0.00	0.00	-41.76	-41.76	0.31	0.29	0.26
	6.52	4.19	-5.86	-7.31	11.43	7.96	-0.74
0.15	0.16	0.09	-17.8	-17.84	2.18	1.85	0.76
	0.00	0.00	-19.5	-19.5	0.80	0.65	0.55
	8.30	5.65	4.14	2.37	6.41	0.82	2.75
0.25	0.49	0.30	-6.49	-6.60	-0.05	-0.59	0.26
	0.06	0.04	-8.94	-8.96	0.00	-0.12	0.15
	10.02	7.35	10.34	8.37	-0.03	-8.24	4.55
0.35	1.57	0.99	-1.19	-1.54	-0.14	-0.64	0.46
	0.26	0.15	-3.54	-3.59	-1.15	-1.15	0.07
	10.91	8.78	13.07	11.15	-3.63	-13.69	6.22
0.45	2.35	1.53	2.36	1.85	-3.04	-3.62	0.75
	0.55	0.36	-0.72	-0.83	-2.12	-1.98	0.06
	9.76	8.91	14.54	13.18	-8.86	-20.58	6.29
0.55	3.85	2.77	5.09	4.32	-4.05	-4.93	1.80
	1.42	0.94	1.64	1.34	-2.43	-2.08	0.53
	8.46	9.08	14.31	13.59	-13.9	-27.3	6.72
0.65	3.32	2.92	6.02	5.51	-7.72	-8.41	1.77
	1.77	1.38	2.88	2.56	-3.81	-3.38	0.92
	6.29	9.04	13.72	14.02	-21.68	-37.8	6.07
0.75	3.22	3.41	6.31	5.98	-9.55	-9.94	2.31
	2.09	1.88	3.85	3.53	-5.45	-4.59	1.40
	3.39	8.52	12.52	13.89	-26.94	-44.56	5.66
0.85	2.86	4.15	6.75	6.76	-11.21	-11.56	3.07
	2.21	2.57	4.70	4.53	-6.40	-5.48	2.09
	1.82	8.51	12.45	14.52	-31.13	-50.45	5.42
0.95	0.81	3.62	6.04	6.73	-15.23	-15.35	2.32
	1.27	2.60	4.20	4.40	-8.28	-7.02	2.20
	0.68	7.72	10.26	13.22	-35.40	-55.92	4.94
1.00	0.55	3.34	4.97	6.14	-17.04	-16.80	2.14
	0.21	2.23	3.34	4.02	-8.87	-7.50	1.83

Table 6: Percentage of Relative Risk Improvements of all the Proposed Estimators Using the Loss L_1 for Equal Sample Sizes

where $\Phi(.)$ is the cumulative distribution function of a standard normal random variable.

The inequality (8.2) is equivalent to show that, $\sigma^2 > \sigma_*^2$, when $S_1 > S_2$. This is true as, $\sigma^2 - \sigma_*^2 > 0$ when $\sigma_1^2 \leq \sigma_2^2$ and $S_1 > S_2$. This proves the case of equal sample sizes.

Next we prove the result for the case of unequal sample sizes that is $m \neq n$. Denoting $V_1 = \sqrt{\frac{m}{m-1}}S_1$ and $V_2 = \sqrt{\frac{n}{n-1}}S_2$, the estimator $\hat{d}_{MK}$ can be written as,

$$\hat{d}_{MK} = \begin{cases} d_{MK}, & \text{if } V_1 \leq V_2 \\ \frac{m\bar{X}+n\bar{Y}}{m+n}, & \text{if } V_1 > V_2. \end{cases}$$

$\tau \downarrow$	$(m,n) = (5,10), (12,20), (10,5), (20,12)$						
	$R1$	$R2$	$R3$	$R4$	$R7$	$R8$	$R9$
0.05	0.00	-8.4	-23.32	-25.50	3.46	3.08	-1.69
	0.00	-0.53	-30.31	-31.00	0.79	0.67	0.54
	2.18	6.68	-63.88	-59.61	10.25	9.64	-4.87
	0.00	0.34	-57.97	-56.84	1.18	0.97	0.68
0.15	0.18	-6.53	-0.40	-1.37	0.01	-0.93	-2.03
	0.00	-0.71	-8.10	-8.45	0.13	0.00	-0.13
	5.79	9.87	-22.3	-20.38	13.2	11.27	5.33
	0.02	0.57	-30.61	-29.93	1.57	1.27	1.01
0.25	0.72	-2.63	5.58	5.09	-6.48	-7.87	-2.44
	0.04	-0.43	-0.79	-0.99	-2.23	-2.29	-0.66
	7.70	10.26	-6.39	-5.35	11.3	7.81	8.14
	0.43	0.98	-17.05	-16.61	1.79	1.42	1.25
0.35	1.47	1.49	8.63	8.48	-13.83	-15.54	-1.73
	0.22	0.01	2.81	2.70	-4.38	-4.29	-0.84
	9.66	10.74	2.29	2.88	9.38	4.71	9.32
	1.01	1.36	-8.73	-8.46	0.84	0.29	1.45
0.45	2.06	3.77	9.03	9.09	-22.4	-24.25	-1.3
	0.55	0.59	4.21	4.15	-6.82	-6.24	-0.50
	10.6	10.59	7.92	8.23	7.92	1.97	10.32
	1.77	2.04	-3.95	-3.77	1.31	0.78	1.98
0.55	2.66	5.76	9.55	9.74	-27.42	-29.41	-0.41
	1.08	1.52	5.02	5.02	-10.3	-9.21	0.26
	11.91	11.00	11.17	11.35	6.33	-0.93	11.03
	2.31	2.24	0.27	0.35	-0.60	-1.21	1.99
0.65	2.98	6.84	9.14	9.47	-35.76	-37.23	-0.34
	1.36	1.97	4.65	4.69	-13.33	-12.04	0.67
	12.84	11.37	13.59	13.68	4.43	-3.84	11.46
	3.24	3.00	3.15	3.20	-1.05	-1.75	2.65
0.75	3.21	7.98	8.47	8.93	-43.11	-44.95	-0.14
	1.93	2.67	4.58	4.65	-15.73	-13.78	1.42
	12.37	10.08	14.27	14.29	0.59	-9.52	10.97
	3.41	3.00	4.61	4.63	-2.65	-3.39	2.65
0.85	3.31	8.48	7.91	8.45	-51.29	-52.66	-0.84
	2.43	3.20	3.95	4.04	-17.53	-15.15	2.18
	11.99	9.02	14.76	14.74	-3.30	-14.75	10.02
	4.00	3.47	5.95	5.94	-3.31	-3.93	3.14
0.95	3.33	9.10	7.54	8.17	-56.73	-58.04	-0.47
	2.26	3.14	3.25	3.38	-21.16	-18.37	2.10
	11.83	8.84	14.60	14.56	-4.47	-16.63	10.04
	3.95	3.31	6.44	6.42	-4.8	-5.49	2.92
1.00	2.99	9.03	6.59	7.28	-62.71	-63.43	-1.96
	1.94	2.84	2.52	2.65	-24.06	-21.01	1.91
	11.4	8.07	14.64	14.59	-5.9	-19.2	9.47
	4.09	3.55	5.81	5.81	-4.87	-5.47	3.18

Table 7: Percentage of Relative Risk Improvements of all the Proposed Estimators Using the Loss L_1 for Unequal Sample Sizes

Proceeding as before, one needs to show that,

$$P[(d_{MK}-\mu)^2 \le c|V_1 > V_2] \le P[(\hat{d}_{MK}-\mu)^2 \le c|V_1 > V_2], \quad \forall\ c > 0.$$

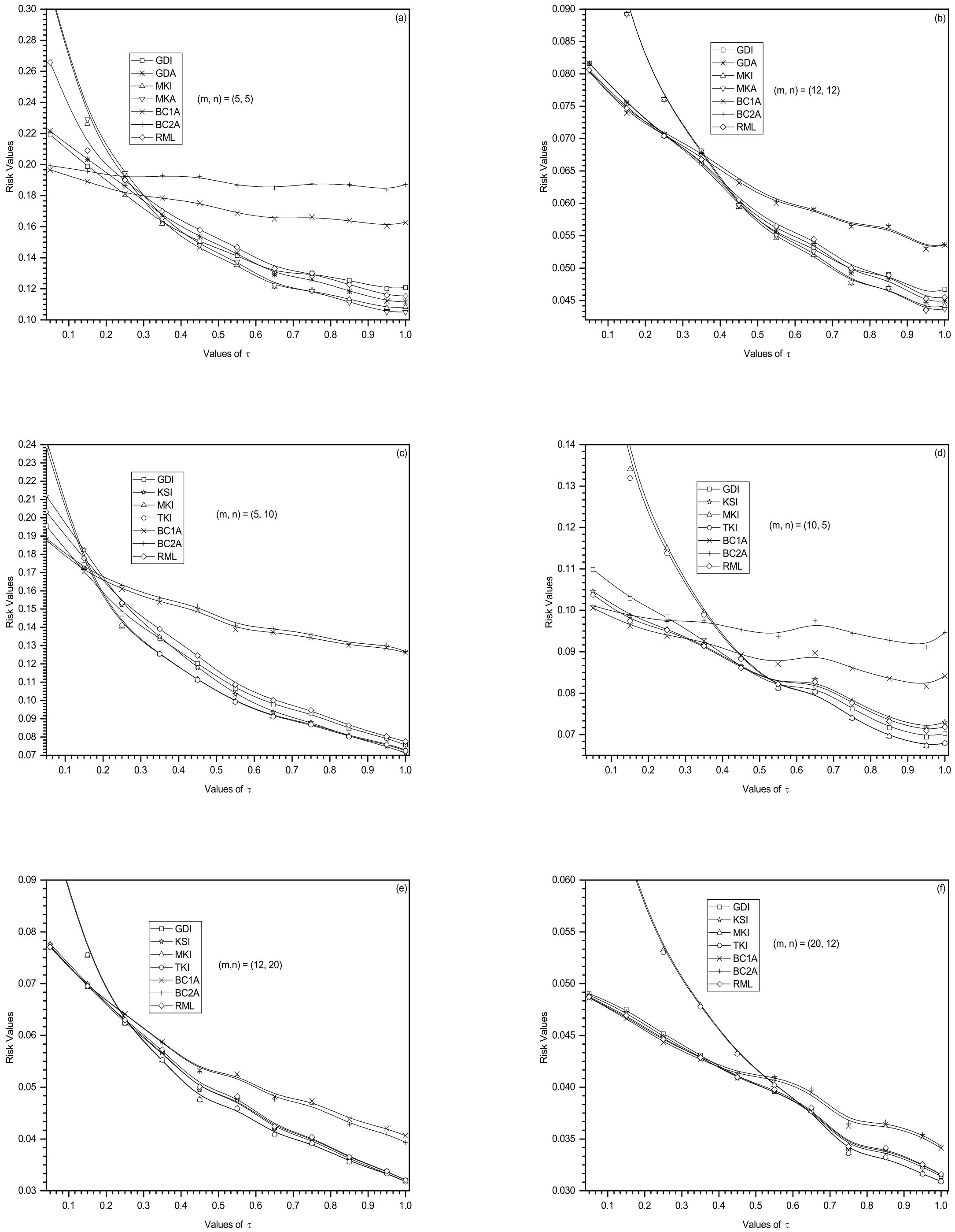

Figure 1: $(a) - (f)$ Comparison of risk values of several estimators for common mean μ using the loss L_1 for sample sizes $(5,5)$, $(12,12)$, $(5,10)$, $(10,5)$, $(12,20)$ and $(20,12)$ respectively.

Which is further equivalent to show that

$$\Phi\Big(\frac{\sqrt{c}}{\nu}\Big) \leq \Phi\Big(\frac{\sqrt{c}}{\nu_*}\Big), \ \forall\ c > 0,$$

where $\nu^2 = \frac{(m-1)s_2^2\sigma_1^2+(n-1)s_1^2\sigma_2^2}{(\sqrt{m(m-1)}s_2+\sqrt{n(n-1)}s_1)^2}$ and $\nu_*^2 = \frac{m\sigma_1^2+n\sigma_2^2}{(m+n)^2}$. This is further equivalent to show that $\nu^2 > \nu_*^2$, $\forall\ c > 0$ when $V_1 > V_2$.

This is equivalent to show that,

$$\frac{\sigma_1^2 + \sigma_2^2\lambda^2}{(\sqrt{m}+\sqrt{n}\lambda)^2} > \frac{m\sigma_1^2 + n\sigma_2^2}{(m+n)^2}, \ \forall\ c > 0, \sigma_1 \leq \sigma_2,$$

where $\lambda = \frac{\sqrt{n-1}S_1}{\sqrt{m-1}S_2}$. Let $h(\lambda) = \frac{\sigma_1^2+\sigma_2^2\lambda^2}{(\sqrt{m}+\sqrt{n}\lambda)^2}$. To show that $h(\lambda) > h(\sqrt{\frac{n}{m}})$, for $\lambda > \sqrt{\frac{n}{m}}$. We observe that $\frac{dh}{d\lambda} \leq 0$, if $\lambda \leq \sqrt{\frac{n}{m}}\frac{\sigma_1^2}{\sigma_2^2} \leq \sqrt{\frac{n}{m}}$, as $\sigma_1^2/\sigma_2^2 \leq 1$. Further $\frac{dh}{d\lambda} > 0$, when $\lambda > \sqrt{\frac{n}{m}}$. Hence $h(\lambda)$ is increasing in the interval $[\sqrt{\frac{n}{m}}, \infty)$. Universal domination follows from Definition 3.2. This proves (i). The proofs of $(ii)-(iv)$ are very much similar to the proof of (i) and hence have been omitted. This completes the proof of the Theorem 3.1. $\square$

Proof of Theorem 5.1

Proof: The theorem can be proved by using a well known technique for improving equivariant estimators proposed by Brewster and Zidek [2]. To proceed, let us consider the conditional risk function of d_Ψ given $\underset{\sim}{T} = \underset{\sim}{t}$:

$$R(\underset{\sim}{\alpha}, d_\Psi|\underset{\sim}{t}) = \frac{1}{\sigma_1^2}E\{(\bar{X} + S_1\Psi(\underset{\sim}{T}) - \mu)^2|\underset{\sim}{T} = \underset{\sim}{t}\}.$$

The above risk function is convex in $\Psi(\underset{\sim}{t})$ and attains minimum at

$$\Psi(\underset{\sim}{t}, \underset{\sim}{\alpha}) = \frac{E\{(\mu - \bar{X})S_1|\underset{\sim}{T} = \underset{\sim}{t}\}}{E\{S_1^2|\underset{\sim}{T} = \underset{\sim}{t}\}}. \tag{8.3}$$

To evaluate the conditional expectations involved in the above expression, we use the following transformations. Let us define $V_1 = (\sqrt{m}(\bar{X} - \mu))/\sigma_1$, $V_2 = (\sqrt{m}(\bar{Y} - \mu))/\sigma_1$, $W_1 = S_1^2/\sigma_1^2$ and $W_2 = S_2^2/\sigma_2^2$. With these substitution the expression for $\Psi(\underset{\sim}{t}, \underset{\sim}{\alpha})$ then reduces to,

$$\Psi(\underset{\sim}{t}, \underset{\sim}{\alpha}) = -\frac{E(V_1W_1^{\frac{1}{2}}|\underset{\sim}{T} = \underset{\sim}{t})}{\sqrt{m}E(W_1|\underset{\sim}{T} = \underset{\sim}{t})}. \tag{8.4}$$

These conditional expectations have been evaluated in [20] and are given by,

$$E(W_1|\underset{\sim}{T} = \underset{\sim}{t}) = \frac{m+n-1}{\lambda},$$

and

$$E(V_1W_1^{\frac{1}{2}}|\underset{\sim}{T} = \underset{\sim}{t}) = -\frac{n\sqrt{m}(m+n-1)t_1}{(n+m\rho)\lambda},$$

where

$$\lambda = \frac{mnt_1^2}{n+m\rho} + \frac{t_2}{\rho} + 1, \quad \rho = \frac{\sigma_2^2}{\sigma_1^2} \geq 1.$$

Substituting these expressions in (8.4), we get the minimimizing choice of $\Psi(\underset{\sim}{t}, \underset{\sim}{\alpha})$ as

$$\hat{\Psi}(\underset{\sim}{t}, \rho) = \frac{nt_1}{n+m\rho}.$$

In order to derive the inadmissibility condition of the theorem, we need the supremum and infimum values of $\hat{\Psi}(\underset{\sim}{t}, \rho)$ with respect to $\rho \in [1, \infty)$ for fixed values of $\underset{\sim}{T} = \underset{\sim}{t}$. We consider the following two cases to obtain the supremum and infimum of $\hat{\Psi}(\underset{\sim}{t}, \rho)$.

Case-I: Let $t_1 \geq 0$. Now the function $\hat{\Psi}(\underset{\sim}{t}, \rho)$ is decreasing with respect to $\rho \in [1, \infty)$. Hence we obtain

$$\inf_{\rho \geq 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \lim_{\rho \to \infty} \hat{\Psi}(\underset{\sim}{t}, \rho) = 0 \text{ and } \sup_{\rho \geq 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \lim_{\rho \to 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \frac{nt_1}{n+m}.$$

Case-II: Let $t_1 < 0$. The function $\hat{\Psi}(\underset{\sim}{t}, \rho)$ is an increasing function of ρ. So for this case we obtain

$$\inf_{\rho \geq 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \lim_{\rho \to 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \frac{nt_1}{n+m} \text{ and } \sup_{\rho \geq 1} \hat{\Psi}(\underset{\sim}{t}, \rho) = \lim_{\rho \to \infty} \hat{\Psi}(\underset{\sim}{t}, \rho) = 0.$$

Combining the Case-I and Case-II, it is easy to define the function $\Psi_0(\underset{\sim}{t})$ as given in (5.2). Utilizing the function $\Psi_0(\underset{\sim}{t})$ and as an application of Theorem 3.1 (in Brewster and Zidek [2]), we get $R(d_{\Psi_0}, \underset{\sim}{\alpha}) \leq R(d_{\Psi}, \underset{\sim}{\alpha})$, when $\sigma_1 \leq \sigma_2$. This completes the proof of the theorem. $\square$

A Sample Program Code of the Simulation Study: As suggested by an anonymous reviewer, below we present a sample program code of the simulation study for equal sample sizes.

```
library(MASS)
library(nleqslv)
M=20000
n1=30
n2=30
b=1.5434/2.0
c=1.5045/2.0
sd2=1.0
mu=0.0
cm=gamma((n1-1)/2)/(sqrt(2)*gamma(n1/2))
cn=gamma((n2-1)/2)/(sqrt(2)*gamma(n2/2))
for(sd1 in seq(0.05,1.0,0.05))
{x1=matrix(0,n1,M)
x2=matrix(0,n2,M)
m1=array(0,M)
m2=array(0,M)
s1=array(0,M)
s2=array(0,M)
d=array(0,M)
a1=array(0,M)
a2=array(0,M)
a3=array(0,M)
a4=array(0,M)
```

```
a5=array(0,M)
a6=array(0,M)
a7=array(0,M)
a8=array(0,M)
a9=array(0,M)
a10=array(0,M)
a11=array(0,M)
a12=array(0,M)
a13=array(0,M)
a14=array(0,M)
a15=array(0,M)
a16=array(0,M)
b1=array(0,M)
b2=array(0,M)
b3=array(0,M)
b4=array(0,M)
b5=array(0,M)
b6=array(0,M)
b7=array(0,M)
b8=array(0,M)
b9=array(0,M)
b10=array(0,M)
b11=array(0,M)
b12=array(0,M)
b13=array(0,M)
b14=array(0,M)
b15=array(0,M)
b16=array(0,M)
c1=array(0,M)
c2=array(0,M)
c3=array(0,M)
c4=array(0,M)
c5=array(0,M)
c6=array(0,M)
c7=array(0,M)
c8=array(0,M)
c9=array(0,M)
c10=array(0,M)
c11=array(0,M)
c12=array(0,M)
c13=array(0,M)
c14=array(0,M)
c15=array(0,M)
c16=array(0,M)
c17=array(0,M)
c18=array(0,M)
c19=array(0,M)
c20=array(0,M)
e1=array(0,M)
e2=array(0,M)
GD=array(0,M)
GDI=array(0,M)
GDA=array(0,M)
PsiGD=array(0,M)
Psi1=array(0,M)
Psi2=array(0,M)
PsiKS=array(0,M)
PsiMK=array(0,M)
PsiTK=array(0,M)
KS=array(0,M)
KSI=array(0,M)
KSA=array(0,M)
MK=array(0,M)
MKI=array(0,M)
MKA=array(0,M)
TK=array(0,M)
TKI=array(0,M)
TKA=array(0,M)
ML=array(0,M)
T1=array(0,M)
T2=array(0,M)
T3=array(0,M)
T4=array(0,M)
V1R=array(0,M)
V2R=array(0,M)
MLR=array(0,M)
PsiBC1=array(0,M)
BC1A=array(0,M)
BC1=array(0,M)
BC2=array(0,M)
PsiBC2=array(0,M)
BC2A=array(0,M)
g1=array(0,M)
g2=array(0,M)
g3=array(0,M)
g4=array(0,M)
RML=array(0,M)
beta1=array(0,M)
beta2=array(0,M)
beta3=array(0,M)
```

```
beta4=array(0,M)
beta5=array(0,M)
BC2I=array(0,M)
PsiML=array(0,M)
t1=array(0,M)
MLA=array(0,M)
for(j in 1:M)
{x1[,j] = rnorm(n1, mean = mu, sd = sqrt(sd1))
x2[,j] = rnorm(n2, mean = mu, sd = sqrt(sd2))
m1[j] = mean(x1[,j])
m2[j] = mean(x2[,j])
d[j] = m2[j] - m1[j]
s1[j] = sum((x1[,j] - m1[j])^2)
s2[j] = sum((x2[,j] - m2[j])^2)
t1[j] = d[j]/sqrt(s1[j])
beta1[j] = (n2 * (n2 - 1) * s1[j])/((n2 * (n2 - 1) * s1[j]) + (n1 * (n1 - 1) * s2[j]))
a1[j] = ((1 - beta1[j]) * m1[j]) + (beta1[j] * m2[j])
GD[j] = ((a1[j] - mu)/sqrt(sd1))^2
a2[j] = s1[j]/(n1 - 1)
a3[j] = s2[j]/(n2 - 1)
if(a2[j] <= a3[j])
{a4[j] = a1[j]}
else{a4[j] = (beta1[j] * m1[j]) + ((1 - beta1[j]) * m2[j])}
GDI[j] = ((a4[j] - mu)/sqrt(sd1))^2
a5[j] = (n2 * (n2 - 1)) * (d[j]/sqrt(s1[j]))
a6[j] = (n1 * (n1 - 1) * (s2[j]/s1[j])) + (n2 * (n2 - 1))
PsiGD[j] = a5[j]/a6[j]
Psi1[j] = (n2/(n1 + n2)) * min((d[j]/sqrt(s1[j])), 0)
Psi2[j] = (n2/(n1 + n2)) * max((d[j]/sqrt(s1[j])), 0)
if(PsiGD[j] < Psi1[j])
{a7[j] = Psi1[j]}
elseif(PsiGD[j] > Psi2[j])
{a7[j] = Psi2[j]}
else{a7[j] = PsiGD[j]}
a8[j] = m1[j] + (sqrt(s1[j]) * a7[j]) GDA[j] = ((a8[j] - mu)/sqrt(sd1))^2
beta2[j] = (n2 * (n2 - 3) * s1[j])/((n2 * (n2 - 3) * s1[j]) + (n1 * (n1 - 3) * s2[j]))
a9[j] = (beta2[j] * m2[j]) + ((1 - beta2[j]) * m1[j])
KS[j] = ((a9[j] - mu)/sqrt(sd1))^2
a10[j] = s1[j]/(n1 - 3)
a11[j] = s2[j]/(n2 - 3)
if(a10[j] <= a11[j])
{a12[j] = a9[j]}
else{a12[j] = (beta2[j] * m1[j]) + ((1 - beta2[j]) * m2[j])}
KSI[j] = ((a12[j] - mu)/sqrt(sd1))^2
a13[j] = (n2 * (n2 - 3)) * (d[j]/sqrt(s1[j]))
a14[j] = (n1 * (n1 - 3) * (s2[j]/s1[j])) + (n2 * (n2 - 3))
PsiKS[j] = a13[j]/a14[j]
if(PsiKS[j] < Psi1[j])
{a15[j] = Psi1[j]}
elseif(PsiKS[j] > Psi2[j])
{a15[j] = Psi2[j]}
else{a15[j] = PsiKS[j]}
a16[j] = m1[j] + (sqrt(s1[j]) * a15[j])
KSA[j] = ((a16[j] - mu)/sqrt(sd1))^2
beta3[j] = sqrt(n2 * (n2 - 1) * s1[j])/(sqrt(n2 * (n2 - 1) * s1[j]) + sqrt(n1 * (n1 - 1) * s2[j]))
b1[j] = ((1 - beta3[j]) * m1[j]) + (beta3[j] * m2[j])
MK[j] = ((b1[j] - mu)/sqrt(sd1))^2
b2[j] = sqrt(n1/(n1 - 1)) * sqrt(s1[j])
b3[j] = sqrt(n2/(n2 - 1)) * sqrt(s2[j])
if(b2[j] <= b3[j])
{b4[j] = b1[j]}
else{b4[j] = (beta3[j] * m1[j]) + ((1 - beta3[j]) * m2[j])}
MKI[j] = ((b4[j] - mu)/sqrt(sd1))^2
b5[j] = sqrt(n2 * (n2 - 1)) * (d[j]/sqrt(s1[j]))
b6[j] = (sqrt(n1 * (n1 - 1)) * sqrt(s2[j]/s1[j])) + (sqrt(n2 * (n2 - 1)))
PsiMK[j] = b5[j]/b6[j]
if(PsiMK[j] < Psi1[j])
{b7[j] = Psi1[j]}
elseif(PsiMK[j] > Psi2[j])
{b7[j] = Psi2[j]}
else{b7[j] = PsiMK[j]}
b8[j] = m1[j] + (sqrt(s1[j]) * b7[j])
MKA[j] = ((b8[j] - mu)/sqrt(sd1))^2
beta4[j] = (sqrt(n2 * s1[j]) * cm)/((sqrt(n2 * s1[j]) * cm) + (sqrt(n1 * s2[j]) * cn))
b9[j] = ((1 - beta4[j]) * m1[j]) + (beta4[j] * m2[j])
TK[j] = ((b9[j] - mu)/sqrt(sd1))^2
b10[j] = cm * sqrt(s1[j]) * sqrt(n1)
b11[j] = cn * sqrt(s2[j]) * sqrt(n2)
if(b10[j] <= b11[j])
{b12[j] = b9[j]}
else{b12[j] = (beta4[j] * m1[j]) + ((1 - beta4[j]) * m2[j])}
TKI[j] = ((b12[j] - mu)/sqrt(sd1))^2
b13[j] = sqrt(n2) * cm * (d[j]/sqrt(s1[j]))
b14[j] = (sqrt(n1) * cn * sqrt(s2[j]/s1[j])) + (sqrt(n2) * cm)
PsiTK[j] = b13[j]/b14[j]
if(PsiTK[j] < Psi1[j])
{b15[j] = Psi1[j]}
elseif(PsiTK[j] > Psi2[j])
```

```
{b15[j] = Psi2[j]}
else{b15[j] = PsiTK[j]}
b16[j] = m1[j] + (sqrt(s1[j]) * b15[j])
TKA[j] = ((b16[j] - mu)/sqrt(sd1))^2
c1[j] = (d[j] * b * s1[j])/(n1 * (n1 - 1))
c2[j] = (s1[j]/(n1 * (n1 - 1))) + (s2[j]/(n2 * (n2 + 2))) + ((d[j] * d[j])/(n2 + 2))
c3[j] = c1[j]/c2[j]
c4[j] = m1[j] + c3[j]
BC1[j] = ((c4[j] - mu)/sqrt(sd1))^2
c5[j] = (d[j] * b)/(n1 * (n1 - 1) * sqrt(s1[j]))
c6[j] = 1/(n1 * (n1 - 1))
c7[j] = s2[j]/(s1[j] * n2 * (n2 + 2))
c8[j] = ((d[j]/sqrt(s1[j]))^2)/(n2 + 2)
PsiBC1[j] = c5[j]/(c6[j] + c7[j] + c8[j])
c9[j] = m1[j] + (sqrt(s1[j]) * PsiBC1[j])
if(PsiBC1[j] < Psi1[j])
{c10[j] = Psi1[j]}
elseif(PsiBC1[j] > Psi2[j])
{c10[j] = Psi2[j]}
else{c10[j] = PsiBC1[j]}
c11[j] = m1[j] + (sqrt(s1[j]) * c10[j])
BC1A[j] = ((c11[j] - mu)/sqrt(sd1))^2
c12[j] = d[j] * c * s1[j] * (n2 * (n2 - 1))
c13[j] = (n2 * (n2 - 1) * s1[j]) + (n1 * (n1 - 1) * s2[j])
c14[j] = c12[j]/c13[j]
c15[j] = m1[j] + c14[j]
BC2[j] = ((c15[j] - mu)/sqrt(sd1))^2
c16[j] = (d[j]/sqrt(s1[j])) * c * n2 * (n2 - 1)
c17[j] = (n2 * (n2 - 1)) + (n1 * (n1 - 1) * (s2[j]/s1[j]))
PsiBC2[j] = c16[j]/c17[j]
if(PsiBC2[j] < Psi1[j])
{c18[j] = Psi1[j]}
elseif(PsiBC2[j] > Psi2[j])
{c18[j] = Psi2[j]}
else{c18[j] = PsiBC2[j]}
c19[j] = m1[j] + (sqrt(s1[j]) * c18[j])
BC2A[j] = ((c19[j] - mu)/sqrt(sd1))^2
beta5[j] = (c * n2 * (n2 - 1) * s1[j])/((n2 * (n2 - 1) * s1[j]) + (n1 * (n1 - 1) * s2[j]))
if((2 * c) <= (1 + (s2[j]/s1[j])))
{c20[j] = ((1 - beta5[j]) * m1[j]) + (beta5[j] * m2[j])}
else{c20[j] = (beta5[j] * m1[j]) + ((1 - beta5[j]) * m2[j])}
BC2I[j] = ((c20[j] - mu)/sqrt(sd1))^2
fnewton < -function(x)
{y < -numeric(3)
d11 = n2 * x[1] * d[j]
d22 = n1 * d[j] * x[2]
d33 = (n2 * x[1]) + (n1 * x[2])
d44 = (n1 * m1[j]/x[1]) + (n2 * m2[j]/x[2])
d55 = (n1/x[1]) + (n2/x[2])
y[1] < -x[1] - (s1[j]/n1) - (d11/d33)^2
y[2] < -x[2] - (s2[j]/n2) - (d22/d33)^2
y[3] < -x[3] - (d44/d55)
y}
xstart < -c(s1[j]/(n1 - 1), s2[j]/(n2 - 1), m1[j])
T1[j] = nleqslv(xstart, fnewton, control = list(btol = 0.0001), method = "Newton")x[1]
T2[j] = nleqslv(xstart, fnewton, control = list(btol = 0.0001), method = "Newton")x[2]
T3[j] = nleqslv(xstart, fnewton, control = list(btol = 0.0001), method = "Newton")x[3]
ML[j] = ((T3[j] - mu)/sqrt(sd1))^2
V1R[j] = min(T1[j], ((n1 * T1[j]) + (n2 * T2[j]))/(n1 + n2))
V2R[j] = max(T2[j], ((n1 * T1[j]) + (n2 * T2[j]))/(n1 + n2))
g1[j] = (n1 * V2R[j] * m1[j]) + (n2 * V1R[j] * m2[j])
g2[j] = (n1 * V2R[j]) + (n2 * V1R[j])
g3[j] = g1[j]/g2[j]
MLR[j] = ((g3[j] - mu)/sqrt(sd1))^2
if(T1[j] <= T2[j])
{g4[j] = T3[j]}
else{g4[j] = ((n1 * m1[j]) + (n2 * m2[j]))/(n1 + n2)}
RML[j] = ((g4[j] - mu)/sqrt(sd1))^2
PsiML[j] = (n2 * T1[j] * t1[j])/((n1 * T2[j]) + (n2 * T1[j]))
if(PsiML[j] < Psi1[j])
{e1[j] = Psi1[j]}
elseif(PsiML[j] > Psi2[j])
{e1[j] = Psi2[j]}
else{e1[j] = PsiML[j]}
e2[j] = m1[j] + (sqrt(s1[j]) * e1[j])
MLA[j] = ((e2[j] - mu)/sqrt(sd1))^2
} tau = sd1/sd2
R1 = sum(GD)/M
R2 = sum(GDI)/M
R3 = sum(GDA)/M
R4 = sum(KS)/M
R5=sum(KSI)/M
R6=sum(KSA)/M
R7=sum(MK)/M
R8=sum(MKI)/M
R9=sum(MKA)/M
R10=sum(TK)/M
R11=sum(TKI)/M
```

```
R12=sum(TKA)/M
R13=sum(BC1)/M
R14=sum(BC1A)/M
R15=sum(BC2)/M
R16=sum(BC2A)/M
R222=sum(BC2I)/M
R17=sum(ML)/M
R18=sum(MLR)/M
R19=sum(RML)/M
R20=sum(MLA)/M
P1=round(((R1-R2)/R1)*100,2)
P2=round(((R1-R3)/R1)*100,2)
P3=round(((R4-R5)/R4)*100,2)
P4=round(((R4-R6)/R4)*100,2)
P5=round(((R7-R8)/R7)*100,2)
P6=round(((R7-R9)/R7)*100,2)
P7=round(((R10-R11)/R10)*100,2)
P8=round(((R10-R12)/R10)*100,2)
P9=round(((R13-R14)/R13)*100,2)
P10=round(((R15-R16)/R15)*100,2)
P11=round(((R17-R18)/R17)*100,2)
P12=round(((R17-R19)/R17)*100,2)
P13=round(((R15-R222)/R15)*100,2)
P14=round(((R17-R20)/R17)*100,2)
PR1=round(((R1-R2)/R1)*100,2)
PR2=round(((R1-R3)/R1)*100,2)
PR3=round(((R1-R8)/R1)*100,2)
PR4=round(((R1-R9)/R1)*100,2)
PR5=round(((R1-R14)/R1)*100,2)
PR6=round(((R1-R16)/R1)*100,2)
PR7=round(((R1-R222)/R1)*100,2)
PR8=round(((R1-R19)/R1)*100,2)
PR9=round(((R1-R20)/R1)*100,2)
}}
```

ACKNOWLEDGMENTS

The Second author (Manas Ranjan Tripathy) would like to thank Department of Science and Technology (SERB), [EMR/2017/003078] New Delhi, India for providing some financial support. The authors would like to thank an anonymous reviewer and the editor for their valuable suggestions and comments which have helped in improving the presentation of the paper.

REFERENCES

[1] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, M. D.(1972). *Statistical Inference Under Order Restrictions,* **Wiley, New York**.

[2] Brewster, J. F., Zidek, J. V.(1974). Improving on equivariant estimators. *Annals of Statistics,* **2,** 21–38.

[3] Brown, L. D. and Cohen, A.(1974) Point and confidence estimation of a common mean and recovery of inter-block information. *Annals of Statistics,* **2(5),** 963–976.

[4] Chang, C. and Pal, N.(2008). Testing on the common mean of several normal distributions. *Computational Statistics and Data Analysis,* **53**, 321–333.

[5] Chang, Y. T., Oono, Y. and Shinozaki, N.(2012). Improved estimators for the common mean and ordered means of two normal distributions with ordered variances. *Journal of Statistical Planning and Inference,* **142(9),** 2619–2628.

[6] Cohen, A. and Sackrowitz, H.(1974). On estimating the common mean of two normal distributions. *Annals of Statistics,* **2(6),** 1274–1282.

[7] Elfessi, A. and Pal, N.(1992). A note on the common mean of two normal populations with order restricted variances. *Communications in Statistics- Theory and Methods,* **21(11),** 3177–3184.

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

[9] Hwang, J. T.(1985). Universal domination and stochastic domination: Estimation simultaneously under a broad class of loss functions. *Annals of Statistics,***13(1)**, 295–314.

[10] Kelleher, T.(1996). A Bayes method for estimating the common mean of two normal populations. *Communications in Statistics- Theory and Methods,* **25(9)**, 2141–2157.

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

[12] Lin, S. and Lee, J. C.(2005). Generalized inferences on the common mean of several normal populations. *Journal of Statistical Planning and Inference,* **134**, 568–582.

[13] Misra, N. and van der Meulen, E. C.(1997). On estimation of the common mean of $k(\geq 2)$ normal populations with order restricted variances. *Statistics and Probability Letters,* **36,** 261–267.

[14] Moore, B. and Krishnamoorthy, K.(1997). Combining independent normal sample means by weighting with their standard errors. *Journal of Statistical Compution and Simulation,* **58,** 145–153.

[15] Pal, N. and Sinha, B. K.(1996). Estimation of a common mean of several normal populations : A review. *Far East Journal of Mathematical Sciences, Special Volume, Part I*, 97–110.

[16] Pal, N., Lin, J. J., Chang, C. H. and Kumar, S.(2007). A revisit to the common mean problem: Comparing the maximum likelihood estimator with the Graybill-Deal estimator. *Computational Statistics and Data Analysis,* **51(12)**, 5673–5681.

[17] Peddada, S. D. and Khatree, R.(1986). On Pitman nearness and variance of estimators. *Communications in Statistics- Theory and Methods,* **15**, 3005–3018.

[18] Pitman, E. J. G.(1937). The closest estimates of statistical parameters. *Mathematical Proceedings of the Cambridge Philosophical Society,* **33(2)**, 212–222.

[19] Rohatgi, V. K. and Saleh, A. K. Md. E.(2003). An Introduction to Probability and Statistics, **John Wiley,** $2nd$ **Edition, New Work**.

[20] Tripathy, M. R. and Kumar, S.(2010). Estimating a common mean of two normal populations. *Journal of Statistical Theory and Applications,* **9(2),** 197–215.

[21] Tripathy, M. R. and Kumar, S.(2015). Equivariant estimation of common mean of several normal populations. *Journal of Statistical Computation and Simulation,* **85(18),** 3679–3699.

[22] Tripathy, M. R., Kumar, S. and Pal, N.(2013). Estimating common standard deviation of two normal populations with ordered means. *Statistical Methods and Applications,* **22(3),** 305–318.

[23] Yates, F.(1940). The recovery of inter-block information in balanced incomplete block designs. *Annals of Eugenics,* **10**, 317–325.