SIMULTANEOUS ESTIMATION OF SEVERAL INTRACLASS CORRELATION COEFFICIENTS

S.E. Ahmed
Department of Mathematics and Statistics
University of Regina
Regina, Saskatchewan, Canada, S4S 0A2
e-mail: ahmed@albiruni.math.uregina.ca

A. K. Gupta
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, Ohio 43403, USA

S. M. Khan
Department of Mathematics and Statistics
University of Regina
Regina, Saskatchewan, Canada, S4S 0A2

C. J. Nicol
Department of Economics
University of Regina
Regina, Saskatchewan, Canada, S4S 0A2

October 20, 1999

ABSTRACT

Based on shrinkage and preliminary test rules, various estimators are proposed for estimation of several intraclass correlation coefficients when independent samples are drawn from multivariate normal populations. It is demonstrated that the James-Stein type estimators are asymptotically superior to the usual estimators. Furthermore, it is also indicated through asymptotic results that none of the preliminary test and shrinkage estimators dominate each other, though they perform relatively well as compared to the classical estimator. The relative dominance picture of the estimators is presented. A Monte Carlo study is performed to appraise the properties of the proposed estimators for small samples.

Key words and phrases: *paired estimation, intraclass correlation coefficients, asymptotic distributional risk, shrinkage estimators and preliminary test, local alternatives, Monte Carlo, maximum likelihood estimation.*

1. INTRODUCTION

Let $\mathbf{X}^{(l)}_{(p\times 1)}$ be a random vector, which has multivariate normal distribution with mean vector $\boldsymbol{\mu}^{(l)}_{(p\times 1)} = (\mu_l, \cdots, \mu_l)'$ and covariance matrix

$$\boldsymbol{\Sigma}_{l_{(p\times p)}} = \sigma_l^2[(1-\rho_l)\mathbf{I}_l + \rho_l \mathbf{1}_l \mathbf{1}_l']$$

where $\mathbf{I}_{l_{(p\times p)}}$ is the identity matrix and $\mathbf{1}_{l_{(p\times 1)}} = (1, \cdots, 1)'$, $\quad l = 1, \cdots, q$. Then the covariance matrix $\boldsymbol{\Sigma}_l$ is said to have an intraclass correlation structure. The problem of interest here is to estimate ρ_l when we have the uncertain prior information in the form of the null hypothesis

$$H_o : \rho_1 = \rho_2 = \cdots = \rho_q = \rho, \tag{1.1}$$

where ρ is unspecified.

The intraclass correlation coefficient is used to measure the degree of resemblance between siblings concerning a certain characteristic, such as blood pressure, stature, body weight or lung capacity and has been studied by a host of researchers. Kapata (1993) has studied hypothesis testing concerning a constant intraclass correlation for families of varying size. Donner (1986) has provided a comprehensive review for the inference procedure in the one-way random effect model.

Thus, the problem of estimating the correlation parameter frequently arises in many medical and bio-statistical applications. Suppose that the researcher has collected data from different (say q) research stations of similar conditions. The researcher is interested in estimating the ρ_l simultaneously on the basis of q random samples and has reason to believe that all the population correlation values may be equal. In the present investigation, we propose and examine the properties of the *unrestricted estimator* (*UE*), the *pooled estimator* (*PE*), the *preliminary-test estimator* (*PTE*) and the *James-Stein rule* or *shrinkage estimator* (*SE*) using the asymptotic distributional quadratic risk (*ADR*) measure.

Useful discussions of some of the implications of the estimators' parametric theory are given by Judge and Bock (1978) and Ahmed (1992a), for example. More generally, these estimators abound in a wide range of statistical applications, as evidenced by the bibliographies of Bancroft and Han (1977) and Han *et al.* (1988). For asymptotic results on the subject see to Ahmed (1991, 1992b) and Gupta *et al.* (1989). For an excellent review of the shrinkage estimators, readers are referred to Stigler (1990).

2. PROPOSED ESTIMATORS

Let $\mathbf{X}^{(l)}_i = \left(\mathbf{X}^{(l)}_{1i}, \cdots, \mathbf{X}^{(l)}_{pi}\right)$, $\quad i = 1, 2, \cdots, n_l$ be a random sample of size n_l from a p-variate normal distribution with mean vector $\boldsymbol{\mu}_l = (\mu_l, \cdots, \mu_l)'$ and covariance matrix $\boldsymbol{\Sigma}_l$, where $\boldsymbol{\Sigma}_l$ has an intraclass correlation structure. Then for the full model, the *UE* of ρ, is defined as

$$(2.1) \qquad r_l = \frac{\sum_{i=1}^{n_l}\sum_{j\neq k}^{p}(x_{ji}^{(l)} - \bar{x}_l)(x_{ki}^{(l)} - \bar{x}_l)}{(p-1)\sum_{i=1}^{n_l}\sum_{j=1}^{p}(x_{ji}^{(l)} - \bar{x}_l)^2}, \quad l = 1, 2, \cdots, q,$$

where $\bar{x}_l = \frac{1}{pn_l}\sum_{i=1}^{n_l}\sum_{j=1}^{p} x_{ji}^{(l)}$. Further, we introduce the following well-known transformation:

$$(2.2) \qquad \rho_l^* = \sqrt{\frac{(p-1)}{2p}} ln\left(\frac{1+(p-1)\rho_l}{1-\rho_l}\right),$$

where *ln* means *logarithm to the base e*. To construct the *UE* of $\boldsymbol{\rho}^* = (\rho_1^*, \rho_2^*, \cdots, \rho_q^*)'$, we replace unknown ρ_l with its empirical estimate. This yields, $\hat{\boldsymbol{\rho}}^* = (\hat{\rho}_1^*, \hat{\rho}_2^*, \cdots, \hat{\rho}_q^*)'$,

$$(2.3) \qquad \hat{\rho}_l^* = \sqrt{\frac{(p-1)}{2p}} ln\left(\frac{1+(p-1)r_l}{1-r_l}\right), \quad l = 1, 2, \cdots, q,$$

where r_l is the UE of ρ_l given in (2.1). Fisher (1958, Chapter 7) showed that the transformation (2.3) is a variance stabilizing transformation for any value of p and follows an approximately normal distribution with mean ρ_l^* and variance $\frac{1}{(n_l-2)}$. However, the normal approximation becomes poorer as the value of p increases. To overcome this difficulty Konishi (1985) proposed the following transformation

$$(2.4) \qquad \theta_l = \rho_l^* - \frac{(7-5p)}{n_l\sqrt{18p(p-1)}}.$$

To construct the *UE* of $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_q)'$, we replace unknown parameters with their empirical estimates in (2.4) and denote them by $\hat{\boldsymbol{\theta}}_n^U = (\hat{\theta}_1^U, \hat{\theta}_2^U, \cdots, \hat{\theta}_q^U)'$. This yields,

$$(2.5) \qquad \hat{\theta}_l^U = \hat{\rho}_l^* - \frac{(7-5p)}{n_l\sqrt{18p(p-1)}},$$

which follows an approximately normal distribution with mean $\theta_l = \rho_l^* - \frac{(7-5p)}{n_l\sqrt{18p(p-1)}}$ and variance $\frac{1}{(n_l-2)}$. In this case, the bias of θ_l is of order $\frac{1}{n_l}$. So if this is unimportant and negligible, the test for the equality of q intraclass correlations under consideration is equivalent to a test of the equality of the values of θ_l. The hypothesis $H_o : \rho_1 = \cdots = \rho_l$ is therefore equivalent to

$$(2.6) \qquad H_o : \theta_1 = \cdots = \theta_l.$$

In this case a reasonable estimate of the common intraclass correlation coefficient is

$$(2.7) \qquad \hat{\boldsymbol{\theta}}_n^P = (\hat{\theta}_n^P \cdots, \hat{\theta}_n^P)', \quad \hat{\theta}_n^P = \frac{1}{n}\sum_{l=1}^{q}(n_l-2)\hat{\theta}_l^U, \quad n = \sum_{l=1}^{q}(n_l-2),$$

which we call the *PE* (Elston, 1975). Generally speaking, the *PE* yields smaller asymptotic risk at and near the null hypothesis at the expense of poorer performance in the rest of the parameter space, where its risk is unbounded. In order to avoid this undesirable property of the *PE*, it is natural to develop an estimator which is a combination of $\hat{\boldsymbol{\theta}}_n^U$ and $\hat{\boldsymbol{\theta}}_n^P$ by performing a preliminary test on the null hypothesis. This methodology was proposed by Bancroft (1944). Thus, we define the *PTE*, $\hat{\boldsymbol{\theta}}_n^{PT} = \left(\hat{\theta}_1^{PT}, \cdots, \hat{\theta}_q^{PT}\right)'$, such that

$$\hat{\boldsymbol{\theta}}_n^{PT} = \hat{\boldsymbol{\theta}}_n^U I(D_n \geq d_{n,\alpha}) + \hat{\boldsymbol{\theta}}_n^P I(D_n < d_{n,\alpha}), \tag{2.8}$$

where $I(A)$ is an indicator function of a set A and

$$D_n = n(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P)'\boldsymbol{\Lambda}(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P), \tag{2.9}$$

with

$$\boldsymbol{\Lambda} = \text{Diag}(\lambda_l), \quad \lambda_{n_l} = \frac{(n_l - 2)}{n}. \tag{2.10}$$

Thus, for a given level of significance α, $(0 < \alpha < 1)$, let $d_{n,\alpha}$ be the upper $(100\alpha)\%$ critical value using the distribution of D_n under H_o. Furthermore, under the null hypothesis, the distance statistic D_n follows the central chi-square distribution with $(q-1)$ degrees of freedom as $n \to \infty$ in such a way that $\lambda_{n_l} \to \lambda_l \in (0,1)$ (Konishi and Gupta (1989)).

The *PTE* (conditional on the value of α) is a convex combination of the *UE* and *PE*, formed using a test-statistic of the H_o, (2.6), and has bounded quadratic risk. Although, the *PTE* has bounded risk, it is sensitive to departures from H_o, (irrespective of the value of α), so may not be adequate for all values of $\boldsymbol{\theta}$. To overcome this shortcoming, we combine the James-Stein (1961) approach with pre-test rules, and propose the *SE*. The *SE* resembles the James and Stein rule estimator (James and Stein, 1961), which performs well over the entire parameter space $\boldsymbol{\theta} \in \Omega$ relative to $\hat{\boldsymbol{\theta}}_n^U$. We propose the *SE*, $\hat{\boldsymbol{\theta}}_n^S = (\hat{\theta}_1^S, \cdots \hat{\theta}_q^S)'$, as follows:

$$\hat{\boldsymbol{\theta}}_n^S = \hat{\boldsymbol{\theta}}_n^P + \{1 - (q-3)D_n^{-1}\}(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P), \quad q \geq 4. \tag{2.11}$$

In the next section, we obtain expressions for the asymptotic biases and risks of these estimators (*UE, PE, PTE* and *SE*).

3. MAIN RESULTS

In this article, we shall study the properties of the proposed estimators in an asymptotic setting, using a quadratic loss function. Let $\boldsymbol{\theta}_n^*$ be an estimator of $\boldsymbol{\theta}$ and $\mathbf{W}$ be a positive semi-definite matrix. Consider the quadratic loss function

$$L(\boldsymbol{\theta}_n^*, \boldsymbol{\theta}) = n(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})'\mathbf{W}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}). \tag{3.1}$$

Then, the *quadratic risk* for $\boldsymbol{\theta}_n^*$ is given by

$$(3.2) \qquad \begin{aligned} R(\boldsymbol{\theta}_n^*, \boldsymbol{\theta}) =& nE\{(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})'\mathbf{W}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})\} \\ =& n\,\mathrm{trace}\,[\mathbf{W}\{E(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})'\}]\,. \end{aligned}$$

Further, $\boldsymbol{\theta}_n^*$ will be termed an *inadmissible estimator* of $\boldsymbol{\theta}$ if there exists an alternative estimator $\boldsymbol{\theta}_n^o$ such that

$$(3.3) \qquad \mathcal{R}(\boldsymbol{\theta}_n^o, \boldsymbol{\theta}) \le \mathcal{R}(\boldsymbol{\theta}_n^*, \boldsymbol{\theta}) \quad \text{for all} \quad \boldsymbol{\theta},$$

with strict inequality for some $\boldsymbol{\theta}$. If, instead of (3.3) holding for every n, we have

$$(3.4) \qquad \lim_{n\to\infty} \mathcal{R}(\boldsymbol{\theta}_n^o, \boldsymbol{\theta}) \le \lim_{n\to\infty} \mathcal{R}(\boldsymbol{\theta}_n^*, \boldsymbol{\theta}) \quad \text{for all} \quad \boldsymbol{\theta},$$

with strict inequality for some $\boldsymbol{\theta}$, then $\boldsymbol{\theta}_n^*$ is termed an *asymptotically inadmissible estimator* of $\boldsymbol{\theta}$. However, the expression in (3.4) is usually difficult to obtain, hence we consider the *asymptotic distributional risk* (*ADR*) for a sequence $\{K_{(n)}\}$ of local alternatives

$$(3.5) \qquad K_{(n)} : \boldsymbol{\theta} = \boldsymbol{\theta}_n, \quad \text{where} \quad \boldsymbol{\theta}_n = \theta\mathbf{1}_q + \frac{\boldsymbol{\delta}}{\sqrt{n}},$$

where $\boldsymbol{\delta}$ is a fixed real vector and $\mathbf{1}_q = (1, \cdots, 1)'$. Note that $\boldsymbol{\delta} = \mathbf{0}$ implies $\boldsymbol{\theta}_n = \rho\mathbf{1}_q$, so (2.6) is a particular case of $\{K_{(n)}\}$. The *asymptotic distribution function* (*ADF*) of $\{\sqrt{n}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta})\}$ is given by

$$(3.6) \qquad G(\mathbf{y}) = \lim_{n\to\infty} P\{\sqrt{n}(\boldsymbol{\theta}_n^* - \boldsymbol{\theta}) \le \mathbf{y}\},$$

where $\boldsymbol{\theta}_n^*$ is any estimator of $\boldsymbol{\theta}$ for which the limit in (3.6) exists. Also, let

$$(3.7) \qquad \mathbf{Q} = \int\int \cdots \int \mathbf{y}\mathbf{y}'dG(\mathbf{y}).$$

Then, the *ADR* is defined by

$$(3.8) \qquad ADR(\boldsymbol{\theta}_n^*; \boldsymbol{\theta}) = \mathrm{trace}(\mathbf{W}\mathbf{Q}).$$

In an effort to compute the *ADR* of *UE, PE, PTE* and *SE*, we first note that (2.8) can be rewritten in the following form

$$\hat{\boldsymbol{\theta}}_n^{PT} = \hat{\boldsymbol{\theta}}_n^{U} - (\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})I(D_n < d_{n,\alpha}).$$

Thus, in the case of *fixed* alternatives,

$$(3.9) \qquad \begin{aligned} n(\hat{\boldsymbol{\theta}}_n^{PT} - \hat{\boldsymbol{\theta}}_n^{U})'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^{PT} - \hat{\boldsymbol{\theta}}_n^{U}) =& n(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})I(D_n < d_{n,\alpha}) \\ =& D_n I(D_n < d_{n,\alpha}) \frac{n(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})}{n(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})'\boldsymbol{\Lambda}(\hat{\boldsymbol{\theta}}_n^{U} - \hat{\boldsymbol{\theta}}_n^{P})} \\ \le& \{D_n I(D_n < d_{n,\alpha})\}ch_{\max}(\mathbf{W}\boldsymbol{\Lambda}^{-1}) \\ \le& \{D_n I(D_n < d_{n,\alpha})\}\mathrm{trace}(\mathbf{W}\boldsymbol{\Lambda}^{-1}) \end{aligned}$$

where $ch_{\max}(\mathbf{A})$ is the largest characteristic root of the matrix $\mathbf{A}$. Also, for $\boldsymbol{\theta} \notin H_o$, $E\{D_n I(D_n < d_{n,\alpha})\} \leq d_{n,\alpha}\{P(D_n < d_{n,\alpha})\}$. But the test statistic D_n is consistent, hence $E\{D_n I(D_n < d_{n,\alpha})\} \to 0 \quad \text{as} \quad n \to \infty$. Thus, for fixed $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_n^U$ and $\hat{\boldsymbol{\theta}}_n^{PT}$ have the same (bounded) risk, asymptotically.

For $\hat{\boldsymbol{\theta}}_n^S$ we note that

$$
\begin{aligned}
n(\hat{\boldsymbol{\theta}}_n^S - \hat{\boldsymbol{\theta}}_n^U)'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^S - \hat{\boldsymbol{\theta}}_n^U) =&(q-3)^2 D_n^{-2}\{n(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P)'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P)\} \\
&\leq (q-3)^2\{n(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P)'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P)\}^{-1} \\
&\{ch_{\max}(\mathbf{W}\boldsymbol{\Lambda}^{-1})\}^2.
\end{aligned} \tag{3.10}
$$

In addition, on the set $\{D_n = 0\}$, we have $\hat{\boldsymbol{\theta}}_n^S = \hat{\boldsymbol{\theta}}_n^U = \hat{\boldsymbol{\theta}}_n^{PT}$. For $\boldsymbol{\theta} \notin H_o$,

$$E\{D_n^{-1} I(D_n > 0)\} \to 0 \quad \text{as} \quad n \to \infty.$$

In other words, $\hat{\boldsymbol{\theta}}_n^S$ and $\hat{\boldsymbol{\theta}}_n^U$ become asymptotically risk equivalent for every $\boldsymbol{\theta}$ not in H_o. The arguments are similar to Gupta *et al.* (1989), hence we omit the details.

Finally, for any $\boldsymbol{\theta} \notin H_o$, $(\hat{\boldsymbol{\theta}}_n^P - \boldsymbol{\theta}) \overset{a.s.}{\to} \zeta(\neq 0)$, and

$$n(\hat{\boldsymbol{\theta}}_n^P - \boldsymbol{\theta})'\mathbf{W}(\hat{\boldsymbol{\theta}}_n^P - \boldsymbol{\theta}) \overset{p}{\to} +\infty, \quad \text{as} \quad n \to \infty.$$

The asymptotic risk of $\hat{\boldsymbol{\theta}}_n^P$, for any $\boldsymbol{\theta} \notin H_o$, approaches $+\infty$. However, the asymptotic risk of $\hat{\boldsymbol{\theta}}_n^U$ is bounded for every $\boldsymbol{\theta} \in \Omega$. The following theorem summarizes the results.

Theorem 3.1: When $\boldsymbol{\theta} \notin H_o, \hat{\boldsymbol{\theta}}_n^P$ has asymptotic risk of $+\infty$, while $\hat{\boldsymbol{\theta}}_n^S$, $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^U$ have the same finite asymptotic risk.

To study the *ADB* and *ADR* of the estimators, we consider the Pitman alternatives

$$K_{(n)} : \boldsymbol{\theta}_n = \theta \mathbf{1}_q + \frac{\boldsymbol{\delta}}{\sqrt{n}}, \tag{3.11}$$

Given the following lemmas:

Lemma 3.1: $\mathbf{X}_n = \sqrt{n}(\hat{\boldsymbol{\theta}}_n^U - \boldsymbol{\theta}) \sim N_q(\boldsymbol{\delta}, \boldsymbol{\Lambda}^{-1})$.

Lemma 3.2: $\mathbf{Y}_n = \sqrt{n}(\hat{\boldsymbol{\theta}}_n^U - \hat{\boldsymbol{\theta}}_n^P), \sim N_q(\boldsymbol{\beta}, \boldsymbol{\Lambda}^{-1}\mathbf{H}')$, where

$$\boldsymbol{\beta} = \mathbf{H}\boldsymbol{\delta}, \quad \mathbf{H} = \mathbf{I}_q - \mathbf{J}\boldsymbol{\Lambda}, \quad \mathbf{J} = \mathbf{1}_q\mathbf{1}_q'.$$

Lemma 3.3: $\mathbf{Z}_n = \sqrt{n}(\hat{\boldsymbol{\theta}}_n^P - \boldsymbol{\theta}) \sim N_q(\mathbf{0}, \mathbf{J})$, here we assume that $\boldsymbol{\lambda}'\boldsymbol{\delta} = 0$, where $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_q)$,

where we note in passing that $\mathbf{Y}_n = \sqrt{n}\mathbf{H}\hat{\boldsymbol{\theta}}_n^U$, the foregoing results can be summarised in the following theorem.

Theorem 3.2:

$$\begin{pmatrix}\mathbf{X}_n\\ \mathbf{Y}_n\end{pmatrix} \sim N_{2q}\left\{\begin{pmatrix}\boldsymbol{\delta}\\ \boldsymbol{\beta}\end{pmatrix}, \begin{pmatrix}\boldsymbol{\Lambda}^{-1} & \boldsymbol{\Lambda}^{-1}\mathbf{H}'\\ \mathbf{H}\boldsymbol{\Lambda}^{-1} & \boldsymbol{\Lambda}^{-1}\mathbf{H}'\end{pmatrix}\right\}, \tag{3.12}$$

$$\begin{pmatrix}\mathbf{Z}_n\\ \mathbf{Y}_\mathbf{n}\end{pmatrix} \sim N_{2q}\left\{\begin{pmatrix}\mathbf{0}\\ \boldsymbol{\beta}\end{pmatrix}, \begin{pmatrix}\mathbf{J} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Lambda}^{-1}\mathbf{H}'\end{pmatrix}\right\}. \tag{3.13}$$

Under the conditions in Theorem 3.2, we can state the following theorem, describing the asymptotic behaviour of the test statistic, D_n,

Theorem 3.3: The test statistic, $D_n = n(\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^P)'\boldsymbol{\Lambda}(\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^P)$, is distributed asymptotically as a non-central chi-square random variable with $(q-1)$ degrees of freedom and non-centrality parameter $\Delta = \boldsymbol{\beta}'\boldsymbol{\Lambda}\boldsymbol{\beta}$. Thus, under the null hypothesis and for large n, D_n closely follows the central chi-square distribution with $(q-1)$ degrees of freedom. For given α, the critical value of D_n may be approximated by $\chi^2_{q-1,\alpha}$, the upper $(100\alpha)\%$ point of the chi-square distribution with $(q-1)$ degrees of freedom.

Under local alternatives, and by virtue of the results in Theorems 3.2–3.3, we arrive at expressions for the *ADB* and *ADR* of *UE, PE, PTE* and *SE*. These are summarised in Theorems 3.4 and 3.5 respectively.

Theorem 3.4:

$$ADB(\hat{\boldsymbol{\theta}}_n^U) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\boldsymbol{\theta}}_n^U - \boldsymbol{\theta})\} = \mathbf{0}, \tag{3.14}$$

$$ADB(\hat{\boldsymbol{\theta}}_n^P) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\boldsymbol{\theta}}_n^P - \boldsymbol{\theta})\} = -\boldsymbol{\beta}, \tag{3.15}$$

$$ADB(\hat{\boldsymbol{\theta}}_n^{PT}) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\boldsymbol{\theta}}_n^{PT} - \boldsymbol{\theta})\} = -\boldsymbol{\beta}H_{q+1}(\chi^2_{q-1,\alpha};\Delta), \tag{3.16}$$

$$ADB(\hat{\boldsymbol{\theta}}_n^S) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\boldsymbol{\theta}}_n^S - \boldsymbol{\theta})\} = -(q-3)\boldsymbol{\beta}E(\chi^{-2}_{q+1}(\Delta)). \tag{3.17}$$

Proof. (3.16) and (3.17) are derived using straightforward computations following the same arguments in Section 4.3 of Judge and Bock (1978), (and the details are therefore omitted here). Moreover, the derivation of these formulae are similar to the case treated in Gupta et. al. (1989).

The equations in Theorem 3.4 reveal that all estimators are biased except $\hat{\boldsymbol{\theta}}_n^U$ which is asymptotically unbiased. However, the biases of $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^S$ are bounded in Δ. In addition, they are asymptotically unbiased (in sense of Δ) but $\hat{\boldsymbol{\theta}}_n^P$ is not.

Equations (3.18)–(3.21) below provide expressions for the *ADR* of the estimators:

Theorem 3.5:

$$ADR(\hat{\boldsymbol{\theta}}_n^U, \boldsymbol{\theta}) = \text{trace}(\mathbf{W}\boldsymbol{\Lambda}^{-1}), \tag{3.18}$$

$$ADR(\hat{\boldsymbol{\theta}}_n^P, \boldsymbol{\theta}) = \text{trace}(\mathbf{W}\mathbf{J}) + \Delta_w\ , \quad \Delta_w = \boldsymbol{\beta}'\mathbf{W}\boldsymbol{\beta}, \tag{3.19}$$

$$ADR(\hat{\boldsymbol{\theta}}^{PT}, \boldsymbol{\theta}) = \text{trace}(\mathbf{W}\boldsymbol{\Lambda}^{-1}) - \text{trace}(\mathbf{WC}) H_{q+1}(\chi^2_{q-1,\alpha}; \Delta) + \Delta_w \{2H_{q+1}(\chi^2_{q-1,\alpha}; \Delta) - H_{q+3}(\chi^2_{q-1,\alpha}; \Delta)\}, \tag{3.20}$$

where $\mathbf{C} = \boldsymbol{\Lambda}^{-1} - \mathbf{J}$,

$$ADR(\hat{\boldsymbol{\theta}}_n^S, \boldsymbol{\theta}) = \text{trace}(\mathbf{W}\boldsymbol{\Lambda}^{-1}) + \Delta_w (q-3)(q+1) E(\chi^{-4}_{q+3}(\Delta)) - (q-3)\text{trace}(\mathbf{WC})\{2E(\chi^{-2}_{q+1}(\Delta)) - (q-3)E(\chi^{-4}_{q+1}(\Delta))\}, \tag{3.21}$$

Proof. (3.18) and (3.19) are straightforward. The representations given by (3.20) and (3.21) are obtained by using the same arguments and corresponding computations as detailed in Section 4.3 of Judge and Bock (1978).

In order to make risk analysis more meaningful, we consider the special case where $\mathbf{W} = \boldsymbol{\Lambda}$, and hence $\Delta_w = \Delta$. The expressions in (3.18)–(3.21) then reduce to:

Corollary 3.1:

$$ADR(\hat{\boldsymbol{\theta}}_n^U, \boldsymbol{\theta}) = q \tag{3.22}$$

$$ADR(\hat{\boldsymbol{\theta}}_n^P, \boldsymbol{\theta}) = 1 + \Delta \tag{3.23}$$

$$ADR(\hat{\boldsymbol{\theta}}_n^{PT}, \boldsymbol{\theta}) = q - (q-1) H_{q+1}(\chi^2_{q-1,\alpha}; \Delta) + \Delta \{2H_{q+1}(\chi^2_{q-1,\alpha}; \Delta) - H_{q+3}(\chi^2_{q-1,\alpha}; \Delta)\}, \tag{3.24}$$

$$ADR(\hat{\boldsymbol{\theta}}_n^S, \boldsymbol{\theta}) = q + \Delta (q-3)(q+1) E(\chi^{-4}_{q+3}(\Delta)) - q(q-3)\{2E(\chi^{-2}_{q+1}(\Delta)) - (q-3)E(\chi^{-4}_{q+1}(\Delta))\}. \tag{3.25}$$

In the next section, we investigate the *relative ADR* properties of the proposed estimators.

4. RISK ANALYSIS FOR VARIOUS ESTIMATORS

In this section, we consider the special case where $\mathbf{W} = \boldsymbol{\Lambda}$. In this framework, we investigate the risk functions of the various estimators, to determine their dominance characteristics.

First, we note that $\hat{\boldsymbol{\theta}}_n^U$ has a constant risk while the risk of $\hat{\boldsymbol{\theta}}_n^P$ becomes unbounded as Δ moves away from $\mathbf{0}$ crossing the risk of $\hat{\boldsymbol{\theta}}_n^U$. Furthermore, we note that

$$ADR(\hat{\boldsymbol{\theta}}_n^P; \boldsymbol{\theta}) \leq ADR(\hat{\boldsymbol{\theta}}_n^U; \boldsymbol{\theta}) \quad \text{if} \quad \Delta \leq q-1. \tag{4.1}$$

Thus, $\hat{\boldsymbol{\theta}}_n^P$ dominates $\hat{\boldsymbol{\theta}}_n^U$ in the interval $[0, q-1]$. Clearly, when Δ moves away from H_o beyond the value $(q-1)$, the risk of $\hat{\boldsymbol{\theta}}_n^P$ increases and becomes unbounded.

Combining $\hat{\boldsymbol{\theta}}_n^U$ and $\hat{\boldsymbol{\theta}}_n^P$ yields $\hat{\boldsymbol{\theta}}_n^{PT}$, as described earlier. The risk of $\hat{\boldsymbol{\theta}}_n^{PT}$ is bounded in Δ, beginning at an initial value of $[q - (q-1)H_{q+1}(\chi^2_{q-1,\alpha}; 0)]$. Then, as Δ deviates from 0, the risk function of $\hat{\boldsymbol{\theta}}_n^{PT}$ monotonically approaches the risk of $\hat{\boldsymbol{\theta}}_n^U$ after first crossing the risk function of $\hat{\boldsymbol{\theta}}_n^U$ and achieving a maximum value. The specific behaviour of $\hat{\boldsymbol{\theta}}_n^{PT}$ in these respects depends on the value of α.

Using (3.23) and (3.24), the ADR expressions for $\hat{\boldsymbol{\theta}}_n^P$ and $\hat{\boldsymbol{\theta}}_n^{PT}$ respectively, it can be shown that

$$\frac{ADR(\hat{\boldsymbol{\theta}}^{PT};\boldsymbol{\theta})}{ADR(\hat{\boldsymbol{\theta}}_n^P,\boldsymbol{\theta})} \leq 1 \quad \text{if}$$

$$\Delta \leq \frac{(q-1)\{1-H_{q+1}(\chi^2_{q-1,\alpha};\Delta)\}}{1-2H_{q+1}(\chi^2_{q-1,\alpha};\Delta)+H_{q+3}(\chi^2_{q-1,\alpha};\Delta)}.$$

Thus, $\hat{\boldsymbol{\theta}}_n^{PT}$ dominates $\hat{\boldsymbol{\theta}}_n^P$ if

$$\Delta \in \left[0, \frac{(q-1)\{1-H_{q+1}(\chi^2_{q-1,\alpha};\Delta)\}}{1-2H_{q+1}(\chi^2_{q-1,\alpha};\Delta)+H_{q+3}(\chi^2_{q-1,\alpha};\Delta)}\right),$$

Using (3.22) and (3.24), the ADR expressions for $\hat{\boldsymbol{\theta}}_n^U$ and $\hat{\boldsymbol{\theta}}_n^{PT}$ respectively, it can be seen that

$$ADR(\hat{\boldsymbol{\theta}}_n^{PT};\boldsymbol{\theta}) \leq ADR(\hat{\boldsymbol{\theta}}_n^U;\boldsymbol{\theta}) \quad \text{if}$$

$$\Delta \leq \frac{(q-1)H_{q+1}(\chi^2_{q-1,\alpha};\Delta)}{2H_{q+1}(\chi^2_{q-1,\alpha};\Delta)-H_{q+3}(\chi^2_{q-1,\alpha};\Delta)}. \tag{4.2}$$

Thus, $\hat{\boldsymbol{\theta}}_n^{PT}$ dominates $\hat{\boldsymbol{\theta}}_n^U$ for some values of Δ, but the reverse is true for other values of Δ. As a partial check, when $\alpha \to 0$, then $\hat{\boldsymbol{\theta}}^{PT}$ dominates $\hat{\boldsymbol{\theta}}_n^U$ in the interval $[0,(q-1))$.

The above discussions, allows us to conclude that none of the three estimators, $\hat{\boldsymbol{\theta}}_n^P$, $\hat{\boldsymbol{\theta}}_n^U$ and $\hat{\boldsymbol{\theta}}_n^{PT}$, asymptotically dominates the other two. Finally, therefore, we consider ADR comparisons of $\hat{\boldsymbol{\theta}}_n^S$ with the above three estimators.

For $\hat{\boldsymbol{\theta}}_n^S$ and $\hat{\boldsymbol{\theta}}_n^U$, it can be seen from the expressions (3.22) and (3.25) that

$$\frac{ADR(\hat{\boldsymbol{\theta}}_n^S;\boldsymbol{\theta})}{ADR(\hat{\boldsymbol{\theta}}_n^U;\boldsymbol{\theta})} \leq 1, \quad \text{for all} \quad \Delta, \tag{4.3}$$

with strict inequality holding for some Δ. The largest gain in risk is achieved near the null hypothesis. Therefore, the risk of $\hat{\boldsymbol{\theta}}_n^S$ is smaller than the risk of $\hat{\boldsymbol{\theta}}_n^U$ in the entire parameter space, and the upper limit for the former is attained when $\Delta \to \infty$. This clearly indicates the asymptotic inferiority of $\hat{\boldsymbol{\theta}}_n^U$ to $\hat{\boldsymbol{\theta}}_n^S$ under local alternatives. More specifically, the risk of $\hat{\boldsymbol{\theta}}_n^S$ begins at an initial value of 3, and increases monotonically towards q as Δ moves away from 0. Once again, note that for the dominance of $\hat{\boldsymbol{\theta}}_n^S$ over $\hat{\boldsymbol{\theta}}_n^U$, we require $q \geq 4$.

We next consider a comparison of the ADR performances of $\hat{\boldsymbol{\theta}}_n^S$ versus $\hat{\boldsymbol{\theta}}_n^P$ under H_o. Using the respective expressions for the ADR of these estimators in (3.23) and (3.25), after simplification we have

$$ADR(\hat{\boldsymbol{\theta}}_n^S, \boldsymbol{\theta}) - ADR(\hat{\boldsymbol{\theta}}_n^P, \boldsymbol{\theta}) = 2. \tag{4.4}$$

Thus, the risk of $\hat{\boldsymbol{\theta}}_n^P$ is substantially smaller than the risk of $\hat{\boldsymbol{\theta}}_n^S$ when the null hypothesis is true. As Δ increases, then $E(\chi_{q+1}^{-4}(\Delta))$ decreases, so the opposite conclusion holds. In general, $\hat{\boldsymbol{\theta}}_n^S$ does not dominate $\hat{\boldsymbol{\theta}}_n^P$ for Δ in the interval $[0, \Delta^+)$, where

$$\Delta^+ = \frac{(q-1)\big[1-(q-3)\{E(\chi_{q+1}^{-2}(\Delta)) + \Delta E(\chi_{q+3}^{-4}(\Delta))\}\big]}{\{1-(q-3)(q+1)E(\chi_{q+3}^{-4}(\Delta)\}}. \tag{4.5}$$

Alternatively, when Δ deviates from the null hypothesis beyond Δ^+, then $\hat{\boldsymbol{\theta}}_n^S$ dominates $\hat{\boldsymbol{\theta}}_n^P$ in the rest of the parameter space. Hence, neither $\hat{\boldsymbol{\theta}}_n^S$ nor $\hat{\boldsymbol{\theta}}_n^P$ asymptotically dominates the other under local alternatives.

From a practical point of view, it is of interest to investigate numerically potential values of Δ^+, and hence the size of the interval $[\Delta^+, \infty)$. This will provide a motivation for using $\hat{\boldsymbol{\theta}}_n^S$ over the pooled estimator when conditions are appropriate. We first observe that Δ^+ is function of q and Δ. Thus, for given values of q, we can calculate the implied values of Δ^+. Table 1 provides the values of Δ^+ for various values of q.

Table 1: *The values of* Δ^+ at selected values of q.

q	4	5	6	7	8	9	10	11	12	13	14
Δ^+	2.60	3.04	3.52	3.95	4.25	4.76	4.85	5.21	5.40	5.74	6.02
q	15	16	17	18	19	20	25	30	35	40	50
Δ^+	6.30	6.56	6.72	6.96	7.15	7.43	8.33	9.02	10.2	10.6	10.9

The foregoing Table therefore indicates the appropriate estimator a researcher might employ on the basis of its ADR properties, depending on the magnitude of anticipated deviations from H_o. To motivate this idea further, we provide the following discussion, based on the use of actual data, collected in a series of Family Expenditure Surveys by Statistics Canada.

Illustrative Example

We consider the case of expenditures on food by households in Canada, and how these are intra-correlated across different regions of the country. Family Expenditure Surveys (FAMEX) have been carried out by Statistics Canada for the years 1969, 1974, 1978, 1982, 1984, 1986, 1990, 1992 and 1996. Researchers might be interested in how the intra-correllation for food in the country varies across time. This can be explored using data from the different FAMEX survey years.

Clearly, expenditures on food are influenced by a number of factors, including household size, housing tenure status (renters versus home owners), amongst other things. There is, however, less likely to be variation in household food expenditures across regions. In a typical FAMEX data set, Statistics Canada divides the country into five regions: Atlantic Provinces (Prince Edward Island, New Brunswick, Newfoundland and Nova Scotia), Québec, Ontario, Prairie Provinces (Alberta, Manitoba and Saskatchewan) and British Columbia. In the context of the notation in this paper, we thus let $p = 5$ be the number of regions. We can then think of random drawings of households of particular types from a multivariate Normal distribution, with $p = 5$.

For the purposes of this example, we confine our attention to married-couple households living in owned accommodation. There is substantial empirical evidence that spending patterns by household size and housing tenure are statistically significantly different. See Barnes and Gillingham (1984) and Nicol (1989), for example. Given households of this type, the aim is then to determine the intraclass correlation coefficients in different years for which data are available. We focus on five survey years: 1978, 1982, 1986, 1990 and 1992, thus implying $q = 5$ in the notation above.

Food expenditures by households of the type described above were drawn at random from each of the five surveys. The total sample size in each FAMEX survey is different, hence we obtain sub-samples of different sizes for each year as follows: 1978, $n_1 = 45$; 1982, $n_2 = 65$; 1986, $n_3 = 55$; 1990, $n_4 = 20$; and 1992, $n_5 = 30$.

Using the foregoing data, $\hat{\boldsymbol{\theta}}^U$ and $\hat{\boldsymbol{\theta}}^P$ were computed with the following results

$$\hat{\boldsymbol{\theta}}^U = [0.199, -0.060, 0.114, -0.018, -0.365]' \tag{4.6}$$

$$\hat{\boldsymbol{\theta}}^P = [0.001, 0.001, 0.001, 0.001, 0.001]' \tag{4.7}$$

Based on these estimates, one can calculate the test statistic, (2.9), which is asymptotically distributed with a central chi-square distribution with $q - 1 = 4$ degrees of freedom under the H_o given in (2.6). The $(100\alpha)\%$ critical value of such a random variable is 9.488, when $\alpha = 0.05$. For this particular application, the realised test statistic is 6.360, which has an upper-tail probability value of 0.174, given a distribution based on the null hypothesis. Consequently, at $\alpha = 0.05$, $\hat{\boldsymbol{\theta}}^P = \hat{\boldsymbol{\theta}}^{PT}$. However, using the realised value of the test statistic, we compute $\hat{\boldsymbol{\theta}}^S$ based on (2.11) as

$$\hat{\boldsymbol{\theta}}^S = [0.137, -0.041, 0.078, -0.012, -0.250]' \tag{4.8}$$

In practical applications, the true nature of the deviation (if any) from H_o is unknown. On the basis of the analytical results presented earlier, a conservative approach to estimator choice would appear to favour $\hat{\boldsymbol{\theta}}^S$, on the basis of the relative *ADR* of the various estimators. This example serves to illustrate that, indeed, this approach has merit. In particular, as we will see in the Monte Carlo simulation presented in the next section, $\hat{\boldsymbol{\theta}}^{PT}$ and $\hat{\boldsymbol{\theta}}^S$ can dominate $\hat{\boldsymbol{\theta}}^U$ and $\hat{\boldsymbol{\theta}}^P$, even for fairly small deviations, $\Delta^\bullet = \sum_{i=1}^q (\boldsymbol{\theta} - \boldsymbol{\theta}_o)^2$.

5. MONTE CARLO SIMULATION

In this section a simulation study is carried out to investigate the properties of the proposed estimators for small samples. We have numerically calculated the risks of $\hat{\boldsymbol{\theta}}^U(R_1)$, $\hat{\boldsymbol{\theta}}^P(R_2)$, $\hat{\boldsymbol{\theta}}^{PT}(R_3)$, and $\hat{\boldsymbol{\theta}}^S(R_4)$ by simulation. Using such simulated data, it is also possible to compute a *simulated maximum likelihood estimator* (*SMLE*), which we denote $\hat{\boldsymbol{\theta}}^{ML}$. Then, the simulated risk of $\hat{\boldsymbol{\theta}}^{ML}(R_5)$ can also be computed. We describe how $\hat{\boldsymbol{\theta}}^{ML}$ is calculated in what follows.

Using the foregoing simulated risks, we define the notion of the *simulated relative efficiency* (*RE*) of an estimator, $\boldsymbol{\theta}^\star$, compared to another estimator $\boldsymbol{\theta}^\diamond$ by

$$RE(\boldsymbol{\theta}^\star : \boldsymbol{\theta}^\diamond) = 100\frac{R(\boldsymbol{\theta}^\diamond)}{R(\boldsymbol{\theta}^\star)}, \tag{5.1}$$

where $R(\boldsymbol{\theta}^\diamond)$ and $R(\boldsymbol{\theta}^\star)$ are the simulated risks of the estimators $\boldsymbol{\theta}^\star$ and $\boldsymbol{\theta}^\diamond$ respectively. Keep in mind that a value of RE greater than 100 indicates the degree of superiority of $\boldsymbol{\theta}^\star$ over $\boldsymbol{\theta}^\diamond$. Thus, the simulated efficiency of the various proposed estimators (and of the *SMLE*), relative to $\hat{\boldsymbol{\theta}}^U$, are given by:

$$RE_{k-1} = 100\frac{R_1}{R_k}, \quad k = 2, \ldots, 5. \tag{5.2}$$

Thus, RE_1, RE_2, RE_3 and RE_4 are the RE of $\hat{\boldsymbol{\theta}}^P$, $\hat{\boldsymbol{\theta}}^{PT}$, $\hat{\boldsymbol{\theta}}^S$ and $\hat{\boldsymbol{\theta}}^{ML}$ respectively, relative to $\hat{\boldsymbol{\theta}}^U_n$.

We assume that the populations from which we wish to draw the simulated data have multivariate normal distributions with mean vectors $\boldsymbol{\mu} = \mathbf{1}$, $\sigma^2 = 1$. For convenience, we set n_l equal for all $l = 1, \ldots, q$, where $q = 4$. We also set $p = 4$. To determine the role of sample size, we carry out three simulation experiments, with $n_l = 20, 30$ and 50. For each of the three experiments, we use five-hundred replications. This number was chosen since it is sufficiently large to identify the patterns we are seeking to examine, but small enough that computational time is not a major consideration. However, we note in passing that the computations for R_5, the simulated risks for $\hat{\boldsymbol{\theta}}^{ML}$, took three calendar days on a Sun Enterprise 4000 six-processor server, for $n_l = 50$.

Random numbers for the simulations were generated using *IMSL* sub-routine *DRNMVN* for $p = 4$, $q = 4$ and n_l as indicated above. From these data, five-hundred estimates each of $\hat{\boldsymbol{\theta}}^U$ and $\hat{\boldsymbol{\theta}}^P$ were computed, and used to simulate the distribution of D_n under $H_o : \theta_1 = \theta_2 = \theta_3 = \theta_4$. The cut-off points of the simulated distribution of D_n were then obtained, yielding simulated critical values. These can be compared with the true critical values of D_n under H_o, which is asymptotically distributed as $\chi^2(q-1)$.

The estimators $\hat{\boldsymbol{\theta}}^{PT}$ and $\hat{\boldsymbol{\theta}}^S$ were computed using the simulated distribution of D_n, and associated realisations of this test statistic for each replication. With respect to $\hat{\boldsymbol{\theta}}^{PT}$, the simulated critical value for a given α level was used.

The computation of $\hat{\boldsymbol{\theta}}^{ML}$ is a little more complicated. Since there is no closed-form solution for the maximum likelihood estimator in this case, we *simulate* computation of such an estimator using the generated data. The contribution of one observation to the log-likelihood function for this model can be represented by the following transformation of such contributions (see, for example, Donner and Koval, 1980):

$$
\begin{aligned}
-2\ln L =& \kappa + (p-1)\ln(1-\rho) + \ln W + \\
(5.3) \qquad & \left[\left(\frac{W-\rho}{W}\right)\sum_{j=1}^{p}[X_{ij}-\mu]^2 - \rho\sum_{j=1}^{p}\sum_{l\neq j}^{p}\left(\frac{(X_{ij}-\mu)(X_{il}-\mu)}{W}\right)\right]/(1-\rho),
\end{aligned}
$$

for $i = 1,\ldots,q$, where κ is a constant and $W = [1+(p-1)\rho]$. This expression depends on the value of ρ. Thus, while the data are generated conditional on a specific value for ρ, the simulated estimator, $\hat{\boldsymbol{\theta}}^{ML}$, is obtained by evaluating the sum of all such contributions at a grid of values for ρ, then selecting that value of the total which yields the lowest value of the sum over all observations of (5.3). This procedure is repeated for each of the five-hundred replications, yielding five-hundred estimates of $\hat{\boldsymbol{\theta}}^{ML}$. These estimates are then used to compute the simulated risk, R_5, for $\hat{\boldsymbol{\theta}}_n^{ML}$.

The above discussion deals with computation of the various estimators under H_o, their numerical risks, $R_1,\ldots,R_5$, and the implied RE referred to earlier. Recalling the definition of $\Delta^\bullet = \sum_{i=1}^{q}(\boldsymbol{\theta}-\boldsymbol{\theta}_o)^2$ from the *Illustrative Example* of Section 4, we can study the RE, $RE_1,\ldots,RE_4$. Tables 2–4 give these results, from which it can be seen that all estimators attain maximum efficiency relative to $\hat{\boldsymbol{\theta}}^{U}$ when $\Delta^\bullet = 0$.

In order to investigate the behavior of the estimators for $\Delta^\bullet > 0$, additional samples were generated from multivariate normal populations, assuming a shift to the right by an amount $\Delta^\bullet = (\boldsymbol{\theta}-\boldsymbol{\theta}_o)^2$ when $\boldsymbol{\theta} \neq \boldsymbol{\theta}_o$. The efficiencies of the various estimators were calculated based on five-hundred replications, again for $p = 4$, $q = 4$, $\alpha = 0.05$ with $n_l = 20, 30, 50$, and for $\alpha = 0.35$ with $n_l = 30$. Tables 2–4 provide the estimated relative efficiencies for the various estimates over $\hat{\boldsymbol{\theta}}^{U}$ for these respective sample sizes. Figures 1–3 contain the same information in a graphical form.

It is apparent from the tables and figures that $\hat{\boldsymbol{\theta}}^{P}$ dominates the other four estimators near the null hypothesis. Alternatively, as the hypothesis error grows, the performance of $\hat{\boldsymbol{\theta}}^{P}$ becomes the worst. Hence it is not a desirable strategy to choose $\hat{\boldsymbol{\theta}}^{P}$ as a general approach. On the other hand, the performance of $\hat{\boldsymbol{\theta}}^{PT}$ is stable for such departures. That is, it achieves maximum efficiency at $\Delta^\bullet = 0$, then drops to a minimum, thereafter tending to the risk of $\hat{\boldsymbol{\theta}}^{U}$.

The relative efficiency of $\hat{\boldsymbol{\theta}}^{PT}$ is higher than that of $\hat{\boldsymbol{\theta}}^{S}$ near the null hypothesis. However, for larger values of $\Delta^\bullet$, the opposite conclusion holds. More importantly, $\hat{\boldsymbol{\theta}}^{S}$ is superior to $\hat{\boldsymbol{\theta}}^{U}$ for *all* values of $\Delta^\bullet$.

The comparisons of $\hat{\boldsymbol{\theta}}^{P}$, $\hat{\boldsymbol{\theta}}^{PT}$ and $\hat{\boldsymbol{\theta}}^{S}$ to $\hat{\boldsymbol{\theta}}^{U}$ in these tables can also be considered relative to their performance versus $\hat{\boldsymbol{\theta}}^{ML}$ in the same context. In practical situations, it is not possible to obtain $\hat{\boldsymbol{\theta}}^{ML}$, so it is of interest to consider how our proposed estimators perform relative to an estimator with the normally extremely desirable properties inherent in the maximum likelihood approach. From Tables 2–4 and Figures 1–3, it is evident that $\hat{\boldsymbol{\theta}}^{PT}$ or $\hat{\boldsymbol{\theta}}^{S}$ exhibit superior RE performances than $\hat{\boldsymbol{\theta}}^{ML}$ in virtually every situation where $\Delta^{\bullet} > 0$.

In short, Tables 2–4 reveal that, for $\Delta^{\bullet}$ close to 0, all the proposed estimators are highly efficient relative to $\hat{\boldsymbol{\theta}}^{U}$. Further, for larger values of $\Delta^{\bullet}$, the performance of the estimators is similar to the analysis of asymptotics provided in Section 4.

We have also assessed the performance of $\hat{\boldsymbol{\theta}}^{S}$ relative to $\hat{\boldsymbol{\theta}}^{PT}$ at a larger size, α, of the test. The SE out-performs the PTE for larger values of α for all $\Delta^{\bullet}$. These results are given in the column labelled RE_{2a} in Table 3, where $n_l = 30$. It is seen that for $\alpha \geq 0.35$, the proposed estimator $\hat{\boldsymbol{\theta}}^{S}$ also dominates $\hat{\boldsymbol{\theta}}^{PT}$ for all $\Delta^{\bullet}$.

To conclude, based on this simulation, we find that the properties of the various estimators are in accordance with the asymptotic results presented above. Simulations for other choices of q were also carried out, yielding similar conclusions. However, for large values of q the relative efficiency of $\hat{\boldsymbol{\theta}}^{S}$ over $\hat{\boldsymbol{\theta}}^{U}$ is substantial.

6. CONCLUDING REMARKS

We have discussed various estimators of $\boldsymbol{\theta}$, a function of the intraclass correlation coefficient, when q samples are available to increase the efficiency of these estimators. We use information obtained via preliminary test, and by incorporating the information provided by test statistics in the estimation process to obtain $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^{S}$. The asymptotic distribution theory of $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^{S}$, and of their risks depend on the asymptotic normality of $\hat{\boldsymbol{\theta}}_n^{P}$ and $\hat{\boldsymbol{\theta}}_n^{U}$, as well as on the asymptotic non-central χ^2 distribution of the test statistic. We conclude that $\hat{\boldsymbol{\theta}}_n^{S}$ is more efficient than $\hat{\boldsymbol{\theta}}_n^{U}$ in the whole parameter space, while the performances of $\hat{\boldsymbol{\theta}}_n^{P}$ and $\hat{\boldsymbol{\theta}}_n^{PT}$ depend on the value of Δ.

The decision whether to use $\hat{\boldsymbol{\theta}}_n^{P}$, $\hat{\boldsymbol{\theta}}_n^{PT}$, or $\hat{\boldsymbol{\theta}}_n^{S}$ rests with the user. We recommend that, if the hypothesis is true or $\Delta \in [0, (q-1)]$, then select $\hat{\boldsymbol{\theta}}_n^{P}$ simply because it has the lowest risk as compared to the other estimators. However, if the experimenter has no knowledge about Δ, which is generally the case, then $\hat{\boldsymbol{\theta}}_n^{S}$ should be used because it dominates $\hat{\boldsymbol{\theta}}_n^{U}$ for all values of Δ. Further, $\hat{\boldsymbol{\theta}}_n^{S}$ dominates $\hat{\boldsymbol{\theta}}_n^{PT}$ for a range of α. An empirical example serves to support this claim.

From the point of view of robust-efficiency, both $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^{S}$ may be advocated, leaning more towards $\hat{\boldsymbol{\theta}}_n^{S}$ since the size Δ is generally unknown and unlikely to be very small. In any event, $\hat{\boldsymbol{\theta}}_n^{S}$ performs better than $\hat{\boldsymbol{\theta}}_n^{P}$ for Δ in the interval $[\Delta^{+}, \infty)$, where Δ^{+} is given in equation (4.5). It should be kept in mind that $\hat{\boldsymbol{\theta}}_n^{S}$ can only be used for $q > 3$, since otherwise it is undefined. With respect to $\hat{\boldsymbol{\theta}}_n^{ML}$, on the other hand, its performance relative to $\hat{\boldsymbol{\theta}}_n^{PT}$ and $\hat{\boldsymbol{\theta}}_n^{S}$ can be superior for some sample sizes, and some values of ρ. However, for anything more than minor values of Δ, this ceases to be the case. Furthermore, in practice, a true maximum likelihood estimator is not available in this estimation environment. Consequently, the apparent superiority of the *simulated* maximum likelihood estimator, $\hat{\boldsymbol{\theta}}_n^{ML}$, is not a realistic option available to the applied researcher.

ACKNOWLEDGMENTS

The authors would like to thank the referees for their helpful comments and providing some references. The work of first author was supported by grants from the Natural Sciences and Engineering Research Council of Canada and a part of this investigation has been done while he was visiting Bowling Green State University, Ohio, USA. The work of the fourth author was supported, in part, by the Social Sciences and Humanities Research of Canada. Some of the results in this paper are based on the 1978, 1982, 1986, 1990 and 1992 public-use tapes for the Canadian Family Expenditure Surveys. The use and interpretation of these data are solely the responsibility of the authors.

REFERENCES

Ahmed, S.E. (1991). To pool or not to pool: The discrete data. Probability and Statistics Letters **11**, 233-237.

Ahmed, S. E.(1992a). Shrinkage preliminary test estimation in multivariate normal distributions. Journal of Statistical Computation and Simulation **43**, 177-195.

Ahmed, S. E.(1992b). Large-sample pooling procedure for correlation. The Statistician **41**, 425-438.

Bancroft, T.A. (1944). On biases in estimation due to use of preliminary test of significance. Annals of Mathematical Statistics, **15**, 190-204.

Bancroft, T.A. and C.P. Han (1977). Inference based on conditional specification: A note and a bibliography. International Statistical Review, **45**, 117-127.

Barnes, R. and R. Gillingham (1984). Demographic effects in demand analysis: estimation of the quadratic expenditure system. Review of Economics and Statistics, **66**, 591-601.

Donner, A. and J. J. Koval (1980). The estimation of intraclass correlation in the analysis of family data. Biometrics, **36**, 19-25.

Donner, A. (1986). A review of inference procedure for the intra class correlation coefficient in the one-way random effect model. International Statistical Review, **54**, 67-82.

Elston, R.C.(1975). Correlation between correlations. Biometrika, **62**, 133-148.

Fisher, R. A. (1958). Statistical methods for research workers. Edinburgh:Oliver & Boyd.

James W. and C. Stein (1961). Estimation with quadratic loss. Proceeding of the Fourth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. 361-379.

Gupta, A. K., A.K.M.E. Saleh and P.K. Sen (1989). Improved estimation in a contingency table: independence structure. Journal of American Statistical Association **84**, 525-532.

Han, C.P., C.V. Rao and J. Ravichandran (1988). Inference based on conditional specification: A second bibliography. Communications in Statistics–Theory and Methods **17**, 1945-1964.

Judge, G. G. and M. E. Bock (1978). The statistical implication of pre-test and Stein-rule estimators in econometrics. Amsterdam:North Holland.

Katapa, R. S. (1993). A test of hypothesis on familial correlations. Biometrics, **49**, 569-576.

Konishi, S. (1985). Normalizing and variance stabilizing transformations for intraclass correlations. Ann. Inst. Statist. Math. **37**, 87-94.

Konishi, S. and A. K. Gupta (1989). Testing the equality of several intraclass correlation coefficients. Journal of Statistical Planning and Inference **21**, 93-105.

Nicol, C. J. (1989). Testing a theory of exact aggregation. Journal of Business & Economic Statistics, **7**, 259-265.

Stigler, S. M. (1990). The 1988 Neyman Memorial Lecture: A Galtonian perspective on shrinkage estimators. Statistical Science **5**, 147-155.

Survey of Family Expenditures Microdata Files,(1978, 1982, 1986, 1990 and 1992). Family expenditure surveys section, Statistics Canada, Ottawa, Canada.

Table 2
Simulated Efficiency of the Estimators
Over $\hat{\boldsymbol{\theta}}^{\boldsymbol{U}}$ for $q = 4, \alpha = .05, n = 20$

$\Delta^{\bullet}$	RE_1	RE_2	RE_3	RE_4
0.000	315.091	226.024	127.189	118.079
0.010	284.877	208.644	132.148	115.106
0.040	224.738	171.583	126.429	109.959
0.086	168.522	136.789	114.960	106.019
0.149	125.979	106.433	110.570	97.022
0.229	95.578	86.749	106.974	91.996
0.325	73.887	74.461	105.948	86.079
0.441	58.125	68.973	104.690	77.732
0.578	46.391	67.408	103.387	74.494
0.741	37.436	69.258	102.062	69.354
0.934	30.436	76.101	101.010	66.633
1.167	24.840	82.102	100.443	62.130
1.452	20.269	90.928	100.171	58.421
1.810	16.456	94.958	100.031	55.315
2.279	13.201	99.052	99.954	52.306
2.932	10.347	100.000	99.912	49.745

The simulated critical value of the test statistic is 7.9332 The tabulated critical value is 7.8147.

Figure 1: Comparisons of SRE for PE, PTE, SE and SMLE

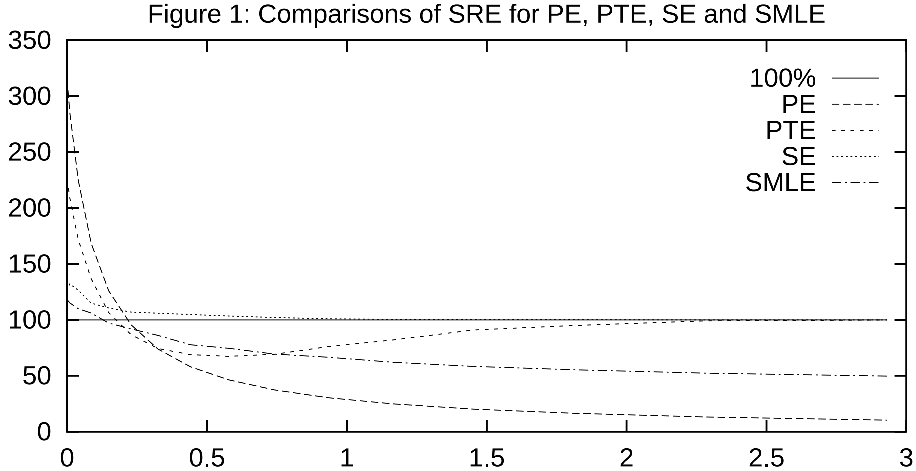

Table 3: Simulated Efficiency of the Estimators Over $\hat{\boldsymbol{\theta}}^{\boldsymbol{U}}$ for $q = 4, \alpha = .05$ and $.35, n = 30$

$\Delta^{\bullet}$	RE_1	RE_2	RE_{2a}	RE_3	RE_4
0.000	337.399	237.593	130.563	137.873	154.050
0.010	286.510	207.340	123.708	115.277	150.001
0.040	201.627	152.068	110.197	120.840	131.678
0.086	137.362	109.750	99.691	115.782	119.114
0.149	96.038	85.512	93.365	104.169	106.176
0.229	69.629	71.925	91.742	106.971	94.662
0.325	52.158	67.081	93.175	105.120	86.615
0.441	40.112	67.506	95.738	103.166	76.756
0.578	31.474	71.337	98.839	102.280	69.916
0.741	25.065	82.376	99.688	101.672	64.078
0.934	20.163	90.554	99.289	101.159	57.379
1.167	16.314	98.799	99.604	100.860	50.485
1.452	13.217	99.295	100.000	100.690	49.372
1.810	10.668	98.904	100.000	100.568	44.680
2.279	8.518	100.000	100.000	100.468	42.705
2.932	6.653	100.000	100.000	100.375	40.682

The simulated critical value of the test statistic is 8.2297 at $\alpha = 0.05$ and 3.1196 at $\alpha = 0.35$. The tabulated critical values are 7.8147 and 3.2831 respectively

Figure 2: Comparisons of SRE for PE, PTE(0.05), PTE(0.35),SE and SMLE

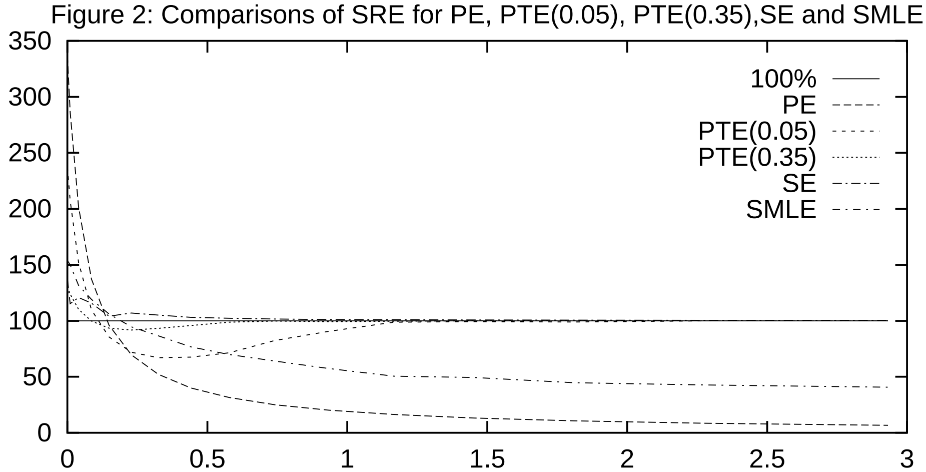

Table 4
Simulated Efficiency of the Estimators
Over $\hat{\boldsymbol{\theta}}^{\boldsymbol{U}}$ for $q = 4, \alpha = .05, n = 50$

$\Delta^{\bullet}$	RE_1	RE_2	RE_3	RE_4
0.000	374.581	249.910	132.858	130.701
0.010	280.509	192.422	125.059	117.191
0.040	164.354	127.695	118.609	101.061
0.086	99.225	86.739	107.779	90.837
0.149	64.599	68.471	104.662	78.699
0.229	44.856	62.999	102.889	66.670
0.325	32.696	67.367	101.815	59.462
0.441	24.698	78.022	101.244	53.035
0.578	19.149	88.784	100.881	47.333
0.741	15.126	96.769	100.631	40.669
0.934	12.102	99.131	100.458	37.820
1.167	9.756	100.000	100.333	34.656
1.452	7.885	100.000	100.243	30.769
1.810	6.353	100.000	100.177	29.461
2.279	5.065	100.000	100.125	26.668
2.932	3.951	100.000	100.088	25.343

The simulated critical value of the test statistic is 8.8637 The tabulated critical value is 7.8147.

Figure 3: Comparisons of SRE for PE, PTE, SE and SMLE