Double shrink methodologies to determine the sample size via covariance structures[★]

Kazuyoshi Yata^a, Makoto Aoshima^{b,*}

^a*Graduate School of Pure and Applied Sciences, University of Tsukuba, Ibaraki, Japan*

^b*Institute of Mathematics, University of Tsukuba, Ibaraki, Japan*

Abstract

We consider fixed-size estimation for a linear function of mean vectors from $\pi_i : N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, \dots, k$, when every $\boldsymbol{\Sigma}_i$ has some structure. The goal of inference is to construct a fixed-span confidence region with required accuracy. We find a sample size for each π_i with the help of the ‘double shrink methodology’, that is introduced by this paper, via covariance structures of $\boldsymbol{\Sigma}_i$, $i = 1, \dots, k$. We estimate the sample size in a two-stage sampling and give a fixed-span confidence region that has the coverage probability approximately second-order consistent with the required accuracy. Some simulations are carried out to see moderate sample size performances of the proposed methodologies.

Key words: Asymptotic uniformity; Intraclass correlation model; Optimal sample size; Two-stage sampling

1 Introduction

Suppose that we have $\pi_i : N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, $i = 1, \dots, k$, independent, normally distributed populations, having unknown mean vector $\boldsymbol{\mu}_i$ and unknown covariance matrix $\boldsymbol{\Sigma}_i$. We assume that $\boldsymbol{\Sigma}_i$ for every i has the following structure

[★] Research of the second author was partially supported by Grant-in-Aid for Scientific Research (B), Japan Society for the Promotion of Science (JSPS), under Contract Number 18300092.

^{*} Corresponding author. Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan; Fax: +81-298-53-6501.

Email address: aoshima@math.tsukuba.ac.jp (Makoto Aoshima).

with unknown positive scalars σ_{ij} , $j = 1, \dots, \ell$:

$$\mathbf{\Sigma}_i = \sigma_{i1}^2 \mathbf{A}_1 + \dots + \sigma_{i\ell}^2 \mathbf{A}_\ell \quad (1)$$

with a fixed ℓ ($1 \leq \ell \leq p$), where $\mathbf{A}_j$ is a $p \times p$ known symmetric matrix with rank r_j having $1 \leq r_1 \leq \dots \leq r_\ell$, $\sum_{j=1}^\ell r_j = p$ and $\sum_{j=1}^\ell \mathbf{A}_j = \mathbf{I}_p$. A special case of the structure is the intraclass correlation model defined by $\mathbf{\Sigma}_i = \sigma_i^2 \{(1 - \rho_i) \mathbf{I}_p + \rho_i \mathbf{1}\mathbf{1}^T\}$, $\mathbf{1} = (1, \dots, 1)^T$: $\mathbf{A}_1 = p^{-1} \mathbf{1}\mathbf{1}^T$ ($r_1 = 1$), $\mathbf{A}_2 = \mathbf{I}_p - \mathbf{A}_1$ ($r_2 = p - 1$), $\sigma_{i1}^2 = \sigma_i^2(p\rho_i + 1 - \rho_i)$, $\sigma_{i2}^2 = \sigma_i^2(1 - \rho_i)$ and $\ell = 2$.

Let $\mathbf{X}_{i1}, \mathbf{X}_{i2}, \dots$ be a sequence of independent and identically distributed (i.i.d.) random vectors from each π_i . Having recorded $\mathbf{X}_{i1}, \dots, \mathbf{X}_{in_i}$ for each π_i , let us write $\bar{\mathbf{X}}_{in_i} = \sum_{s=1}^{n_i} \mathbf{X}_{is} / n_i$ and $\mathbf{n} = (n_1, \dots, n_k)$. We are interested in estimating the linear function $\boldsymbol{\mu} = \sum_{i=1}^k b_i \boldsymbol{\mu}_i$, where b_i 's are known and nonzero scalars. Let $\mathbf{T}_\mathbf{n} = \sum_{i=1}^k b_i \bar{\mathbf{X}}_{in_i}$. Given $d > 0$, we define

$$R_\mathbf{n} = \{\boldsymbol{\mu} \in R^p : \|\mathbf{T}_\mathbf{n} - \boldsymbol{\mu}\| \leq d\}, \quad (2)$$

where $\|\cdot\|$ is the Euclidean norm. Now, with some given $\alpha \in (0, 1)$, our goal is to construct a fixed-span confidence region $R_\mathbf{n}$ such that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_\mathbf{n}) \geq 1 - \alpha \quad \text{for all } \boldsymbol{\theta}, \quad (3)$$

where $\boldsymbol{\theta} = (\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_k, \sigma_{11}, \dots, \sigma_{1\ell}, \dots, \sigma_{k1}, \dots, \sigma_{k\ell})$.

From (1), there exists a $p \times p$ orthogonal matrix $\mathbf{H}$ such that

$$\mathbf{H}^T \mathbf{\Sigma}_i \mathbf{H} = \begin{pmatrix} \sigma_{i1}^2 \mathbf{I}_{r_1} & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & \sigma_{i\ell}^2 \mathbf{I}_{r_\ell} \end{pmatrix}$$

for all π_i , $i = 1, \dots, k$. Then, $\mathbf{H}^T(\mathbf{T}_\mathbf{n} - \boldsymbol{\mu})$ is distributed as $N_p(\mathbf{O}, \sum_{i=1}^k n_i^{-1} b_i^2 \mathbf{H}^T \mathbf{\Sigma}_i \mathbf{H})$. Let V_j ($j = 1, \dots, \ell$) be a mutually independent chi-square random variable with r_j ($j = 1, \dots, \ell$) degrees of freedom (d.f.). We note that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_\mathbf{n}) = P_{\boldsymbol{\theta}} \left(\sum_{j=1}^\ell \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{n_i} V_j \leq d^2 \right). \quad (4)$$

The purpose of this paper is to find the sample size for each π_i in order to have the coverage probability, given by (4), satisfying requirement (3). There are many literatures related to this inference problem. For instance, see Aoshima (2005) and Ghosh et al. (1997) for a review. We, however, emphasize on a previous work given by Aoshima and Takada (2006) for the present problem. They gave a method to find the sample size for each π_i and constructed a region $R_\mathbf{n}$ satisfying requirement (3) in a Stein's (1945) two-stage sampling scheme. They proposed an approximation to the sample size in order to overcome the complexity of its calculation and showed that the approximation

gives a region $R_{\mathbf{n}}$ with accuracy approximately equal to $1 - \alpha$. In this paper, we propose a different method, the ‘double shrink methodology’, to find the sample size for each π_i . It is shown that our findings in this paper improve both the sample size and the approximation to the coverage probability.

In Section 2, we introduce a new methodology, the ‘double shrink methodology’, to find the sample size for each π_i via covariance structures. In Section 3, we give an approximation to the coverage probability along with an expansion formula of the sample size to satisfy requirement (3). In Section 4, we numerically show that the double shrink methodology, proposed in this paper, improves both the sample size and the approximation to the coverage probability in several simulation studies.

2 Double shrink methodology

In this section, we introduce a new methodology to find the sample sizes satisfying requirement (3) via covariance structures. We name it the double shrink methodology. Let us explain how it finds the sample sizes to satisfy requirement (3). We briefly write n_i as a real number for a while.

Let $F_p(\cdot)$ denote the cumulative distribution function of a chi-square random variable with p d.f. Let a be the constant such that $F_p(a) = 1 - \alpha$. We tentatively consider $n_i = (a/d^2)|b_i|\sigma_{i1} \sum_{i'=1}^k |b_{i'}|\sigma_{i'1}$ that yields the smallest sum $\sum_{i=1}^k n_i$ to hold that $\sum_{i=1}^k b_i^2 \sigma_{i1}^2 / n_i = d^2/a$. Let $\delta_j = (\sum_{i'=1}^k |b_{i'}|\sigma_{i'1})^{-1} \sum_{i=1}^k |b_i| \sigma_{ij}^2 / \sigma_{i1}$, $j = 2, \dots, \ell$. Let $\delta = \max_{j=2, \dots, \ell} \delta_j$. If $\delta \leq 1$, we have from (4) that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{n}}) = P_{\boldsymbol{\theta}}\left(V_1 + \sum_{j=2}^{\ell} \delta_j V_j \leq a\right) \geq P_{\boldsymbol{\theta}}\left(\sum_{j=1}^{\ell} V_j \leq a\right) = F_p(a) = 1 - \alpha.$$

If $\delta > 1$, we modify the sample size as $n_i = (a\delta/d^2)|b_i|\sigma_{i1} \sum_{i'=1}^k |b_{i'}|\sigma_{i'1}$ in order to have that

$$P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{n}}) = P_{\boldsymbol{\theta}}\left(\frac{1}{\delta}V_1 + \sum_{j=2}^{\ell} \frac{\delta_j}{\delta} V_j \leq a\right) \geq P_{\boldsymbol{\theta}}\left(\sum_{j=1}^{\ell} V_j \leq a\right) = F_p(a) = 1 - \alpha.$$

When we consider both the cases of δ simultaneously, the sample size required to satisfy requirement (3) is given by $n_i = (a/d^2)|b_i|\sigma_{i1} \max_{1 \leq j \leq \ell} \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}^2 / \sigma_{i'1}$.

Figures 1–3 summarize the idea of the double shrink methodology in the case that $p = 2$, $\ell = 2$ and $\delta (= \delta_2) > 1$. In Figure 1, the ellipse indicates a $1 - \alpha$ confidence region for $\boldsymbol{\mu}$ and the shaded circle indicates the fixed region with the radius of d from the centre $\boldsymbol{\mu}$. When the sample is taken up to size $n_i = (a/d^2)|b_i|\sigma_{i1} \sum_{i'=1}^k |b_{i'}|\sigma_{i'1}$ for each π_i , the confidence region is shrunk up to the circle along the eigenvector having axis number $l = 1$ as seen in Figure 2. Next, when the sample is additionally taken until the total sample

for each π_i becomes of size $n_i = (a/d^2)|b_i|\sigma_{i1} \max_{1 \leq j \leq \ell} \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}^2 / \sigma_{i'1}$, the confidence region is shrunk up to the circle along the eigenvector having axis number $l = 2$ as seen in Figure 3.

On the other hand, Aoshima and Takada (2006) considered the sample size given by

$$n_i \geq \frac{a}{d^2} \max_{1 \leq j \leq \ell} (|b_i|\sigma_{ij} \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}) \quad (= \tilde{n}_i, \text{ say}) \quad (5)$$

for each π_i . As seen in Figure 4, the confidence region of such a size is included in the circle and it may not come in contact with the circle.

Comparing Figure 4 with Figure 3, it should be noted that the double shrink methodology improves the approximation to the coverage probability in (4). One can expect that the double shrink methodology successfully prevents the confidence region from both meeting requirement (3) excessively and oversampling too much.

Generally, for a fixed l ($1 \leq l \leq \ell$), the double shrink methodology finds the sample size for each π_i as

$$n_i \geq \frac{a}{d^2} |b_i|\sigma_{il} \max_{1 \leq j \leq \ell} \sum_{i'=1}^k |b_{i'}| \frac{\sigma_{i'j}^2}{\sigma_{i'l}} \quad (= C_i, \text{ say}). \quad (6)$$

Then, the region $R_{\mathbf{n}}$ given by (2) satisfies requirement (3). Since σ_{ij} 's are unknown, it is necessary to estimate C_i 's in (6) with some pilot samples. We consider a two-stage estimation methodology to determine the sample sizes $\mathbf{n}$ under the following assumptions: For fixed l ($1 \leq l \leq \ell$),

(A1) There exists some j_l such that $\sum_{i=1}^k |b_i|\sigma_{ij_l}^2 / \sigma_{il} > \sum_{i=1}^k |b_i|\sigma_{ij}^2 / \sigma_{il}$ for all j ($1 \leq j \neq j_l \leq \ell$);

(A2) There exists a known and positive lower bound $\sigma_{il\star}$ for σ_{il} such that

$$\sigma_{il} > \sigma_{il\star}, \quad i = 1, \dots, k$$

(cf. Mukhopadhyay and Duggan, 1999).

1. Having m_0 (≥ 4) fixed, define

$$m = \max \left\{ m_0, \left\lceil \frac{a}{d^2} \min_{1 \leq i \leq k} |b_i|\sigma_{il\star} \sum_{i'=1}^k |b_{i'}|\sigma_{i'l\star} \right\rceil + 1 \right\}, \quad (7)$$

where $[x]$ denotes the largest integer less than x . According to (7), take a pilot sample $\mathbf{X}_{is}$, $s = 1, \dots, m$ of size m and calculate $\mathbf{S}_i = \nu^{-1} \sum_{s=1}^m (\mathbf{X}_{is} - \bar{\mathbf{X}}_{im})(\mathbf{X}_{is} - \bar{\mathbf{X}}_{im})^T$ for all π_i , $i = 1, \dots, k$, where $\bar{\mathbf{X}}_{im} = \sum_{s=1}^m \mathbf{X}_{is} / m$ and $\nu = m - 1$. Then, define

$$S_{ij}^2 = r_j^{-1} \text{tr}(\mathbf{A}_j \mathbf{S}_i) \quad (8)$$

as an unbiased estimate of σ_{ij}^2 for all $i = 1, \dots, k$; $j = 1, \dots, \ell$. Define the total

sample size for all π_i , $i = 1, \dots, k$, by

$$N_i = \max \left\{ m, \left\lceil \frac{u_l}{d^2} |b_i| S_{il} \max_{1 \leq j \leq \ell} \sum_{i'=1}^k |b_{i'}| \frac{S_{i'j}^2}{S_{i'l}} \right\rceil + 1 \right\}, \quad (9)$$

where u_l (> 0) is determined later. Let $\mathbf{N} = (N_1, \dots, N_k)$.

2. Take an additional sample $\mathbf{X}_{is}$, $s = m + 1, \dots, N_i$ of size $N_i - m$ for each π_i . By combining the initial sample and the additional sample, calculate $\bar{\mathbf{X}}_{iN_i} = N_i^{-1} \sum_{s=1}^{N_i} \mathbf{X}_{is}$ for each π_i . Then, define the region $R_{\mathbf{N}}$ as in (2) with $\mathbf{T}_{\mathbf{N}} = \sum_{i=1}^k b_i \bar{\mathbf{X}}_{iN_i}$.

3 Second-order approximations

For fixed l ($1 \leq l \leq \ell$), let j_l and j_{lm} be the indices such that

$$\begin{aligned} \max_{1 \leq j \leq \ell} \sum_{i=1}^k |b_i| \frac{\sigma_{ij}^2}{\sigma_{il}} &= \sum_{i=1}^k |b_i| \frac{\sigma_{ij_l}^2}{\sigma_{il}} \quad \text{in (6),} \\ \max_{1 \leq j \leq \ell} \sum_{i=1}^k |b_i| \frac{S_{ij}^2}{S_{il}} &= \sum_{i=1}^k |b_i| \frac{S_{ij_{lm}}^2}{S_{il}} \quad \text{in (9),} \end{aligned} \quad (10)$$

respectively. Throughout, we write that

$$\begin{aligned} \tau_{l\star} &= \min_{1 \leq i \leq k} |b_i| \sigma_{il\star} \sum_{i'=1}^k |b_{i'}| \sigma_{i'l\star}, \\ f_i &= |b_i| (\sigma_{ij_l}^2 / \sigma_{il}) \left(\sum_{i'=1}^k |b_{i'}| \sigma_{i'j_l}^2 / \sigma_{i'l} \right)^{-1} \quad (i = 1, \dots, k). \end{aligned}$$

Theorem 3.1. Assume that $a \geq p - 2$. Choose u_l in (9) as $u_l = a(1 + \nu^{-1} \hat{s}_l)$ where

$$\hat{s}_l = \begin{cases} \frac{1}{r_l} + \frac{(a-p) \sum_{i=1}^k b_i^2 S_{il}^2 - r_l k \tau_{l\star}}{2r_l (\sum_{i=1}^k |b_i| S_{il})^2} & (j_{lm} = l), \\ \left(\sum_{i=1}^k |b_i| \frac{S_{ij_{lm}}^2}{S_{il}} \right)^{-2} \left(\frac{a+2-p}{2r_{j_{lm}}} \sum_{i=1}^k b_i^2 \frac{S_{ij_{lm}}^4}{S_{il}^2} - \frac{\tau_{l\star}}{2} \sum_{i=1}^k \frac{S_{ij_{lm}}^2}{S_{il}^2} \right) & (j_{lm} \neq l) \end{cases} \quad (11)$$

with S_{ij}^2 's given by (8). Then, the region $R_{\mathbf{N}}$ is asymptotically second-order consistent as $d \rightarrow 0$ in the sense that

$$\begin{aligned} P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{N}}) &\geq E_{\boldsymbol{\theta}} \left\{ F_p \left(d^2 \left(\max_{1 \leq j \leq \ell} \left(\sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{N_i} \right) \right)^{-1} \right) \right\} \\ &= 1 - \alpha + o(d^2) \quad \text{for all } \boldsymbol{\theta}. \end{aligned} \quad (12)$$

Proof. For fixed l ($1 \leq l \leq \ell$), let j_{l_N} be the index such that

$$\max_{1 \leq j \leq \ell} \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{N_i} = \sum_{i=1}^k \frac{b_i^2 \sigma_{ij_{l_N}}^2}{N_i}.$$

We have from (4) that

$$\begin{aligned} & P_{\boldsymbol{\theta}}(\boldsymbol{\mu} \in R_{\mathbf{N}}) \\ & \geq E_{\boldsymbol{\theta}} \left\{ F_p \left(d^2 \left(\sum_{i=1}^k \frac{b_i^2 \sigma_{ij_{l_N}}^2}{N_i} \right)^{-1} \right) \right\} \\ & = E_{\boldsymbol{\theta}} \left\{ F_p \left(a \left(\sum_{i=1}^k f_i \frac{C_i}{N_i} \right)^{-1} \right) \right\} + E_{\boldsymbol{\theta}} \left\{ I_{\{j_l \neq j_{l_N}\}} F_p \left(d^2 \left(\sum_{i=1}^k \frac{b_i^2 \sigma_{ij_{l_N}}^2}{N_i} \right)^{-1} \right) \right\}, \end{aligned} \quad (13)$$

where $I_{\{j_l \neq j_{l_N}\}}$ is the indicator function. Now, let us define a new function as follows.

$$g(u_1, \dots, u_k) = F_p(av^{-1}), \quad v = f_1 u_1^{-1} + \dots + f_k u_k^{-1} \quad \text{for } u_i > 0, \quad i = 1, \dots, k.$$

Denoting $F'_p(w)$, $F''_p(w)$ for the first and second derivatives of $F_p(w)$ respectively, one can verify the following expressions of the partial derivatives of $g(u_1, \dots, u_k)$. For all $1 \leq i \neq j \leq k$, we have that

$$\begin{aligned} \frac{\partial g}{\partial u_i} &= a F'_p(a/v) f_i v^{-2} u_i^{-2}, \\ \frac{\partial^2 g}{\partial u_i^2} &= a \{ a F''_p(a/v) f_i^2 v^{-4} u_i^{-4} + 2 F'_p(a/v) f_i^2 v^{-3} u_i^{-4} - 2 F'_p(a/v) f_i v^{-2} u_i^{-3} \}, \\ \frac{\partial^2 g}{\partial u_i \partial u_j} &= a \{ a F''_p(a/v) f_i f_j v^{-4} u_i^{-2} u_j^{-2} + 2 F'_p(a/v) f_i f_j v^{-3} u_i^{-2} u_j^{-2} \}. \end{aligned}$$

For the first term in (13), we use the Taylor expansion to claim that

$$\begin{aligned} & E_{\boldsymbol{\theta}} \left\{ F_p \left(a \left(\sum_{i=1}^k f_i \frac{C_i}{N_i} \right)^{-1} \right) \right\} = E_{\boldsymbol{\theta}} \left\{ g \left(\frac{N_1}{C_1}, \dots, \frac{N_k}{C_k} \right) \right\} \\ & = 1 - \alpha + a F'_p(a) \sum_{i=1}^k f_i E_{\boldsymbol{\theta}} \left(\frac{N_i - C_i}{C_i} \right) \\ & \quad + \frac{a}{2} \sum_{i=1}^k (a F''_p(a) f_i^2 + 2 F'_p(a) f_i^2 - 2 F'_p(a) f_i) E_{\boldsymbol{\theta}} \left\{ \left(\frac{N_i - C_i}{C_i} \right)^2 \right\} \\ & \quad + \frac{a}{2} \sum_{i \neq j} (a F''_p(a) f_i f_j + 2 F'_p(a) f_i f_j) E_{\boldsymbol{\theta}} \left\{ \left(\frac{N_i - C_i}{C_i} \right) \left(\frac{N_j - C_j}{C_j} \right) \right\} \\ & \quad + E_{\boldsymbol{\theta}}(\mathfrak{R}), \end{aligned} \quad (14)$$

where

$$E_{\boldsymbol{\theta}}(\mathfrak{R}) = \frac{1}{6} \sum_{i,j,\ell} E_{\boldsymbol{\theta}} \left\{ \frac{\partial^3 g}{\partial u_i \partial u_j \partial u_\ell} \Big|_{\mathbf{u}=\boldsymbol{\xi}} \left(\frac{N_i - C_i}{C_i} \right) \left(\frac{N_j - C_j}{C_j} \right) \left(\frac{N_\ell - C_\ell}{C_\ell} \right) \right\} \quad (15)$$

with suitable random variables ξ_i 's between 1 and N_i/C_i , $i = 1, \dots, k$, $\mathbf{u} = (u_1, \dots, u_k)$ and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)$. With the help of Lemmas 5 and 7 in Appendix, we obtain the following expansions from (14):

$$\begin{aligned} & E_{\boldsymbol{\theta}} \left\{ g \left(\frac{N_1}{C_1}, \dots, \frac{N_k}{C_k} \right) \right\} \\ &= \begin{cases} 1 - \alpha + \frac{aF'_p(a)}{\nu_l} \left(s_l r_l - 1 + \frac{1}{2} \sum_{i=1}^k f_i B_i + \sum_{i=1}^k f_i^2 + a \frac{F''_p(a)}{F'_p(a)} \sum_{i=1}^k f_i^2 \right) \\ \quad + o(\nu^{-1}) & (j_l = l), \\ 1 - \alpha + \frac{aF'_p(a)}{\nu_l} \left(s_l r_l + \frac{1}{2} \sum_{i=1}^k f_i B_i + a \frac{r_l F''_p(a)}{r_{j_l} F'_p(a)} \sum_{i=1}^k f_i^2 \right) + o(\nu^{-1}) & (j_l \neq l), \end{cases} \end{aligned} \quad (16)$$

where $\nu_l = \nu r_l$, $B_i = C_i^{-1} \nu_l$ and s_l is a constant such that $E_{\boldsymbol{\theta}}(\hat{s}_l) = s_l + o(1)$. Noting that $P_{\boldsymbol{\theta}}(j_{l_m} = j_l) = 1 + O(\nu^{-3})$ from Lemma 1, one yields for (11) that

$$s_l = \begin{cases} \frac{1}{r_l} + \frac{(a-p) \sum_{i=1}^k b_i^2 \sigma_{il}^2 - r_l k \tau_{l*}}{2r_l (\sum_{i=1}^k |b_i| \sigma_{il})^2} & (j_l = l), \\ \left(\sum_{i=1}^k |b_i| \frac{\sigma_{ij_l}^2}{\sigma_{il}^2} \right)^{-2} \left(\frac{(a+2-p)}{2r_{j_l}} \sum_{i=1}^k b_i^2 \frac{\sigma_{ij_l}^4}{\sigma_{il}^2} - \frac{\tau_{l*}}{2} \sum_{i=1}^k \frac{\sigma_{ij_l}^2}{\sigma_{il}^2} \right) & (j_l \neq l). \end{cases} \quad (17)$$

Combining (17) and the results that $aF''_p(a)/F'_p(a) = (p - a - 2)/2$ and $\sum_{i=1}^k f_i B_i = r_l \tau_{l*} (\sum_{i=1}^k \sigma_{ij_l}^2 / \sigma_{il}^2) (\sum_{i=1}^k |b_i| \sigma_{ij_l}^2 / \sigma_{il})^{-2} + O(d^2)$ with (16), we have that

$$E_{\boldsymbol{\theta}} \left\{ F_p \left(a \left(\sum_{i=1}^k f_i \frac{C_i}{N_i} \right)^{-1} \right) \right\} = 1 - \alpha + o(d^2). \quad (18)$$

For the second term in (13), with the help of Lemma 6, we evaluate as $d \rightarrow 0$ that

$$E_{\boldsymbol{\theta}} \left\{ I_{\{j_l \neq j_{l_N}\}} F_p \left(d^2 \left(\sum_{i=1}^k \frac{b_i^2 \sigma_{ij_{l_N}}^2}{N_i} \right)^{-1} \right) \right\} \leq P_{\boldsymbol{\theta}}(j_l \neq j_{l_N}) = O(d^6). \quad (19)$$

Combining (18) and (19) with (13), the result can be obtained. $\square$

Theorem 3.2. *The two-stage procedure (7)–(9) with (11) has as $d \rightarrow 0$:*

$$E_{\boldsymbol{\theta}}(N_i - C_i) = \begin{cases} (2\tau_{l\star}r_l)^{-1} \{ |b_i|\sigma_{il} \sum_{i'=1}^k |b_{i'}|\sigma_{i'l} + (a-p)f_i \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l}^2 + b_i^2 \sigma_{il}^2 \} \\ + \frac{1}{2}(1 - kf_i) + o(1) \quad \text{for } i = 1, \dots, k & (j_l = l), \\ \\ (2\tau_{l\star}r_{j_l})^{-1} \left\{ (a+2-p) \frac{\sigma_{il}^2}{\sigma_{ij_l}^2} f_i \sum_{i'=1}^k b_{i'}^2 \frac{\sigma_{i'j_l}^4}{\sigma_{i'l}^2} + \frac{r_{j_l}}{r_l} \left(|b_i|\sigma_{il} \sum_{i' \neq i} |b_{i'}| \frac{\sigma_{i'j_l}^2}{\sigma_{i'l}} \right) \right\} \\ + \frac{1}{2} \left(1 - \frac{\sigma_{il}^2}{\sigma_{ij_l}^2} f_i \sum_{i'=1}^k \frac{\sigma_{i'j_l}^2}{\sigma_{i'l}^2} \right) + o(1) \quad \text{for } i = 1, \dots, k & (j_l \neq l). \end{cases}$$

Proof. The results are obtained by Lemma 5 in Appendix straightforwardly. $\square$

Now, in the two-stage procedure (7)–(9) with (11), one may determine the axis number l ($1 \leq l \leq \ell$) so as to minimize the sum $\sum_{i=1}^k N_i$ as follows: Let l_o denote the axis number l that minimizes the sum $\sum_{i=1}^k C_i$ in (6). We assume that l_o is determined uniquely. For fixed l_o , let j_o be the index such that

$$\max_{1 \leq j \leq \ell} \sum_{i=1}^k |b_i| \frac{\sigma_{ij}^2}{\sigma_{il_o}} = \sum_{i=1}^k |b_i| \frac{\sigma_{ij_o}^2}{\sigma_{il_o}}.$$

Let $\tau_{ijl} = |b_i|\sigma_{il} \sum_{i'=1}^k |b_{i'}|\sigma_{i'l}^2 / \sigma_{i'l}$ ($i = 1, \dots, k; j = 1, \dots, \ell; l = 1, \dots, \ell$). We assume for (A1)–(A2) that

(A1') There exists some j_l for each l such that $\tau_{ij_l l} > \tau_{ijl}$ for all j ($1 \leq j \neq j_l \leq \ell$);

(A2') There exists a known and positive lower bound $\sigma_{ij\star}$ for σ_{ij} such that

$$\sigma_{ij} > \sigma_{ij\star}, \quad i = 1, \dots, k; \quad j = 1, \dots, \ell.$$

Having m_0 (≥ 4) fixed, we start with a pilot sample of size m as

$$m = \max \left\{ m_0, \left\lceil \frac{a}{d^2} \min_{\substack{1 \leq i \leq k \\ 1 \leq j \leq \ell}} |b_i| \sigma_{ij\star} \sum_{i'=1}^k |b_{i'}| \sigma_{i'j\star} \right\rceil + 1 \right\} \quad (20)$$

on behalf of (7). We carry out the first-stage sampling of the two-stage procedure and calculate the sum of the required total sample sizes, $\sum_{i=1}^k N_i$, for each l ($1 \leq l \leq \ell$), using (9) after replacing $\tau_{l\star}$ with

$$\tau_{o\star} = \min_{\substack{1 \leq i \leq k \\ 1 \leq j \leq \ell}} |b_i| \sigma_{ij\star} \sum_{i'=1}^k |b_{i'}| \sigma_{i'j\star}$$

in (11). Now, find the number l ($= l_{om}$, say) that gives the minimum sum $\sum_{i=1}^k N_i$ and fix l_{om} as the axis number hereafter. One may utilize the two-stage estimation methodology stated in Section 2 after replacing both (7) and

(9) with (20) and

$$N_i = \max \left\{ m, \left\lceil \frac{u_{l_{om}}}{d^2} |b_i| S_{il_{om}} \max_{1 \leq j \leq \ell} \sum_{i'=1}^k |b_{i'}| \frac{S_{i'j}^2}{S_{i'l_{om}}} \right\rceil + 1 \right\}, \quad (21)$$

respectively.

Theorem 3.3. Assume that $a \geq p - 2$. For the axis number l_{om} , let j_{om} be the index such that

$$\max_{1 \leq j \leq \ell} \sum_{i=1}^k |b_i| \frac{S_{ij}^2}{S_{il_{om}}} = \sum_{i=1}^k |b_i| \frac{S_{ij_{om}}^2}{S_{il_{om}}}.$$

Choose $u_{l_{om}}$ in (21) as $u_{l_{om}} = a(1 + \nu^{-1} \hat{s}_{l_{om}})$ where

$$\hat{s}_{l_{om}} = \begin{cases} \frac{1}{r_{l_{om}}} + \frac{(a-p) \sum_{i=1}^k b_i^2 S_{il_{om}}^2 - r_{l_{om}} k \tau_{o*}}{2r_{l_{om}} (\sum_{i=1}^k |b_i| S_{il_{om}})^2} & (j_{om} = l_{om}), \\ \left(\sum_{i=1}^k |b_i| \frac{S_{ij_{om}}^2}{S_{il_{om}}} \right)^{-2} \left(\frac{a+2-p}{2r_{j_{om}}} \sum_{i=1}^k b_i^2 \frac{S_{ij_{om}}^4}{S_{il_{om}}^2} - \frac{\tau_{o*}}{2} \sum_{i=1}^k \frac{S_{ij_{om}}^2}{S_{il_{om}}^2} \right) & (j_{om} \neq l_{om}) \end{cases} \quad (22)$$

with S_{ij}^2 's given by (8). For the two-stage procedure (20)–(21), the region R_N is asymptotically second-order consistent as $d \rightarrow 0$ as stated in (12).

Proof. With the help of Lemma 8 in Appendix, the result can be obtained similarly to Theorem 3.1. $\square$

Theorem 3.4. The two-stage procedure (20)–(21) with (22) has as $d \rightarrow 0$:

$$E_{\theta}(N_i - C_i) = \begin{cases} (2\tau_{o*} r_{l_o})^{-1} \{ |b_i| \sigma_{il_o} \sum_{i'=1}^k |b_{i'}| \sigma_{i'l_o} + (a-p) f_{oi} \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l_o}^2 + b_i^2 \sigma_{il_o}^2 \} \\ + \frac{1}{2} (1 - k f_{oi}) + o(1) \quad \text{for } i = 1, \dots, k & (j_o = l_o), \\ (2\tau_{o*} r_{j_o})^{-1} \left\{ (a+2-p) \frac{\sigma_{il_o}^2}{\sigma_{ij_o}^2} f_{oi} \sum_{i'=1}^k b_{i'}^2 \frac{\sigma_{i'j_o}^4}{\sigma_{i'l_o}^2} + \frac{r_{j_o}}{r_{l_o}} \left(|b_i| \sigma_{il_o} \sum_{i' \neq i} |b_{i'}| \frac{\sigma_{i'j_o}^2}{\sigma_{i'l_o}} \right) \right\} \\ + \frac{1}{2} \left(1 - \frac{\sigma_{il_o}^2}{\sigma_{ij_o}^2} f_{oi} \sum_{i'=1}^k \frac{\sigma_{i'j_o}^2}{\sigma_{i'l_o}^2} \right) + o(1) \quad \text{for } i = 1, \dots, k & (j_o \neq l_o), \end{cases}$$

where

$$f_{oi} = |b_i| (\sigma_{ij_o}^2 / \sigma_{il_o}) \left(\sum_{i'=1}^k |b_{i'}| \sigma_{i'j_o}^2 / \sigma_{i'l_o} \right)^{-1} \quad (i = 1, \dots, k).$$

Proof. The results are obtained by Lemmas 5 and 8 in Appendix straightforwardly. $\square$

Remark 1. If there exists a number l such that $j_l = l$ in (10), we have that $\tilde{n}_i \geq C_i$, where $\tilde{n}_i$ is given by (5) and C_i is given by (6) with such l . It follows for $\tilde{n}_i$'s that

$$F_p \left(d^2 \left(\max_{1 \leq j \leq \ell} \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{\tilde{n}_i} \right)^{-1} \right) \geq 1 - \alpha.$$

Therefore, a region $R_{\mathbf{N}}$ with a sample size estimating $\tilde{n}_i$ for each π_i can no longer satisfy (12). If $\tilde{n}_i$ is modified by

$$\tilde{n}_i^* = \frac{a}{d^2} \left(\max_{1 \leq j \leq \ell} \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{\max_{1 \leq j' \leq \ell} (|b_i| \sigma_{ij'} \sum_{i'=1}^k |b_{i'}| \sigma_{i'j'})} \right) \max_{1 \leq j \leq \ell} (|b_i| \sigma_{ij} \sum_{i'=1}^k |b_{i'}| \sigma_{i'j}),$$

one has that

$$F_p \left(d^2 \left(\max_{1 \leq j \leq \ell} \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{\tilde{n}_i^*} \right)^{-1} \right) = 1 - \alpha.$$

Therefore, a region $R_{\mathbf{N}}$ with a sample size estimating $\tilde{n}_i^*$ for each π_i should satisfy (12). However, it would be much complicated to estimate $\tilde{n}_i^*$.

Remark 2. Let us consider the case that there exists a common and known number j_1 for all π_i such that $\sigma_{ij_1} > \sigma_{ij}$ for all j ($1 \leq j \neq j_1 \leq \ell$) in (1). One can reduce the two-stage procedure (7)–(9) with (11) as follows: We have for any fixed j_2 ($1 \leq j_2 \neq j_1 \leq \ell$) that

$$\begin{aligned} & \sum_{i=1}^k |b_i| \sigma_{ij_2} \max_{1 \leq j \leq \ell} \left(\sum_{i'=1}^k |b_{i'}| \frac{\sigma_{i'j}^2}{\sigma_{i'j_2}} \right) = \sum_{i=1}^k |b_i| \sigma_{ij_2} \sum_{i'=1}^k |b_{i'}| \frac{\sigma_{i'j_1}^2}{\sigma_{i'j_2}} \\ &= \sum_{i=1}^k b_i^2 \sigma_{ij_1}^2 + \sum_{i < i'} |b_i| |b_{i'}| \sigma_{ij_1} \sigma_{i'j_1} \left(\frac{\sigma_{ij_1} \sigma_{i'j_2}}{\sigma_{ij_2} \sigma_{i'j_1}} + \frac{\sigma_{ij_2} \sigma_{i'j_1}}{\sigma_{ij_1} \sigma_{i'j_2}} \right) \\ &\geq \sum_{i=1}^k b_i^2 \sigma_{ij_1}^2 + 2 \sum_{i < i'} |b_i| |b_{i'}| \sigma_{ij_1} \sigma_{i'j_1} = \left(\sum_{i=1}^k |b_i| \sigma_{ij_1} \right)^2 \\ &= \sum_{i=1}^k |b_i| \sigma_{ij_1} \max_{1 \leq j \leq \ell} \left(\sum_{i'=1}^k |b_{i'}| \frac{\sigma_{i'j}^2}{\sigma_{i'j_1}} \right). \end{aligned}$$

So, one naturally chooses j_1 as the axis number l in (6) in order to have the minimum sum $\sum_{i=1}^k C_i$. Under (A1)–(A2), with j_1 for l , we carry out the first-stage sampling with (7). Now, define the total sample size of each π_i as

$$N_i = \max \left\{ m, \left\lceil \frac{u_1}{d^2} |b_i| S_{ij_1} \sum_{i'=1}^k |b_{i'}| S_{i'j_1} \right\rceil + 1 \right\}, \quad (23)$$

where u_1 is given by

$$u_1 = a \left(1 + \frac{1}{\nu} \left(\frac{1}{r_{j_1}} + \frac{(a-p) \sum_{i=1}^k b_i^2 S_{ij_1}^2 - r_{j_1} k \tau_{j_1\star}}{2r_{j_1} \left(\sum_{i=1}^k |b_i| S_{ij_1} \right)^2} \right) \right) \quad (24)$$

with $\tau_{j_1\star} = \min_{1 \leq i \leq k} |b_i| \sigma_{ij_1\star} \sum_{i'=1}^k |b_{i'}| \sigma_{i'j_1\star}$. We carry out the second-stage sampling with (23)–(24). Then, this reduced procedure holds the asymptotic characteristic stated in Theorem 3.1. On the other hand, the total sample size due to Aoshima and Takada (2006) is naturally reduced to (23) with

$$u_1 = a + \frac{a}{2p\nu} \left(\frac{p+2\ell}{p+2} a + 2k\ell - p \right). \quad (25)$$

Remark 3. When the structure (1) specifies an intraclass correlation model as $\Sigma_i = \sigma_i^2 \{(1 - \rho_i) \mathbf{I}_p + \rho_i \mathbf{1}\mathbf{1}^T\}$, the parameters in (1) are given by $\sigma_{i1}^2 = \sigma_i^2(p\rho_i + 1 - \rho_i)$ and $\sigma_{i2}^2 = \sigma_i^2(1 - \rho_i)$ with $r_1 = 1$, $r_2 = p - 1$ and $\ell = 2$. Let us consider the case where $\rho_i > 0$ for all π_i , such as a large dimensional case. Since $\sigma_{i1} > \sigma_{i2}$ for all π_i , one may follow the two-stage procedure (23)–(24) in Remark 2. Then, a candidate for $\sigma_{i1\star}$ appearing in (24) is given by $\sigma_{i1\star} = \sigma_{i\star}$ with $\sigma_{i\star}$ a known and positive lower bound for π_i 's standard deviation.

Remark 4. One may apply Theorems 3.1 and 3.2 to constructing a fixed-size ellipsoidal confidence region for regression parameters. Let us consider a linear regression model as follows:

$$Y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, 2, \dots,$$

where $\mathbf{x}_i$'s are known p -vectors, $\boldsymbol{\beta}$ is an unknown p -vector of regression parameters, and ε_i 's are i.i.d. $N(0, \sigma^2)$ with unknown σ (> 0). Let us denote $\mathbf{y}_n^T = (Y_1, \dots, Y_n)$ and $\mathbf{X}_n^T = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, and assume that $\text{rank}(\mathbf{X}_n) = p$ ($< n$). Having recorded $(\mathbf{x}_i, Y_i)$, $i = 1, \dots, n$ with $n > p$, we estimate $\boldsymbol{\beta}$ by the least square estimator $\hat{\boldsymbol{\beta}}_n = (\mathbf{X}_n^T \mathbf{X}_n)^{-1} \mathbf{X}_n^T \mathbf{y}_n$. Along the lines of Ghosh et al. (pp.198–200, 1997), a fixed-size ellipsoidal confidence region for $\boldsymbol{\beta}$ is defined by

$$R_{\boldsymbol{\beta}_n} = \{\boldsymbol{\beta} \in R^p : (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})^T (n^{-1} \mathbf{X}_n^T \mathbf{X}_n) (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \leq d^2\}$$

for given d (> 0), and our goal is to hold $P_{\boldsymbol{\beta}, \sigma^2}(\boldsymbol{\beta} \in R_{\boldsymbol{\beta}_n}) \geq 1 - \alpha$ for all $(\boldsymbol{\beta}, \sigma^2)$ for given $\alpha \in (0, 1)$. It is easy to see that $P_{\boldsymbol{\beta}, \sigma^2}(\boldsymbol{\beta} \in R_{\boldsymbol{\beta}_n}) = F_p(nd^2/\sigma^2)$, so that n has to be $n \geq a\sigma^2/d^2$ ($= C$, say). We assume that there exists a known and positive lower bound $\sigma_\star$ for σ such that $\sigma > \sigma_\star$. One may estimate σ^2 by $S^2 = (m - p)^{-1}(\mathbf{y}_m - \mathbf{X}_m \hat{\boldsymbol{\beta}}_m)^T (\mathbf{y}_m - \mathbf{X}_m \hat{\boldsymbol{\beta}}_m)$ with a pilot sample of size $m = \max\{m_0, \lceil a\sigma_\star^2/d^2 \rceil + 1\}$ where $m_0 \geq \max\{4, p + 1\}$. Then, the total sample size of the two-stage estimation methodology is defined by $N = \max\{m, \lceil uS^2/d^2 \rceil + 1\}$, where

$$u = a(1 + \nu^{-1} \hat{s}) \quad \text{with} \quad \nu = m - p \quad \text{and} \quad \hat{s} = (a - p + 2)/2 - 0.5(\sigma_\star^2/S^2).$$

Then, from Theorems 3.1 and 3.2 (with $k = 1$, $\ell = 1$ and $r_1 = 1$), it holds as $d \rightarrow 0$ that

$$\begin{aligned} P_{\beta, \sigma^2}(\beta \in R_{\beta_N}) &= 1 - \alpha + o(d^2) \quad \text{for all } (\beta, \sigma^2), \\ E_{\beta, \sigma^2}(N - C) &= (2\sigma_\star^2)^{-1}\sigma^2(a - p + 2) + o(1). \end{aligned}$$

4 Moderate sample performances

In order to study the performance of the two-stage procedure, we take resort to computer simulations. We fix $p = 4$ and $b_i = 1$, $i = 1, \dots, k$. Our goal is to construct 90% fixed-span confidence regions for $\mu = \mu_1 + \dots + \mu_k$. In other words, we have $\alpha = 0.1$ (that is, $a = 7.779$). Independent pseudorandom normal observations from $\pi_i : N_4(\mu_i, \Sigma_i)$, $i = 1, \dots, k$, were generated where Σ_i 's were fixed as $\Sigma_i = \sigma_i^2\{(1 - \rho_i)\mathbf{I}_4 + \rho_i\mathbf{1}\mathbf{1}^T\}$. Then, $\ell = 2$, $(r_1, r_2) = (1, 3)$, $\sigma_{i1}^2 = \sigma_i^2(p\rho_i + 1 - \rho_i)$ and $\sigma_{i2}^2 = \sigma_i^2(1 - \rho_i)$, $i = 1, \dots, k$. We choose $\sigma_i = 1$, $i = 1, \dots, k$. We set $\sigma_{ij\star}^2 = 0.7\sigma_{ij}^2$, $i = 1, \dots, k$; $j = 1, 2$, with ρ_i fixed in each table.

In Table 1, pretending that $\rho_i > 0$, $i = 1, \dots, k$, we compare the performance of the two-stage procedure in Remark 2 with (24), which is given in the first block, with that with (25) due to Aoshima and Takada (2006), which is given in the second block. Let $C = \sum_{i=1}^k C_i$. We consider three cases that (i) $k = 2$, $(\rho_1, \rho_2) = (1/10, 1/5)$ and $C = 90$; (ii) $k = 3$, $(\rho_1, \rho_2, \rho_3) = (1/10, 3/20, 1/5)$ and $C = 135$; (iii) $k = 4$, $(\rho_1, \rho_2, \rho_3, \rho_4) = (1/10, 2/15, 1/6, 1/5)$ and $C = 180$, whereas with a fixed value of C one easily obtains from (6) that $d = 0.707$, 0.866 and 1.001 , respectively. Then, note that $m = 30$ obtained from (7) is same as the one given by Aoshima and Takada (2006).

In Tables 2 and 3, we examine the two-stage procedure (20)–(21) with (22) in the first block, the two-stage procedure (7)–(9) with (11) having axis number $l = 2$ in the second block, and the two-stage procedure due to Aoshima and Takada (2006) in the third block. In Table 2, we consider two cases that (i) $k = 2$, $(\rho_1, \rho_2) = (-0.2, 0.1)$ and $d = 0.458$; (ii) $k = 4$, $(\rho_1, \rho_2, \rho_3, \rho_4) = (-0.2, -0.2, 0.1, 0.1)$ and $d = 0.648$. In Table 3, we consider two cases that (i) $k = 2$, $(\rho_1, \rho_2) = (-0.1, 0.2)$ and $d = 0.567$; (ii) $k = 4$, $(\rho_1, \rho_2, \rho_3, \rho_4) = (-0.1, -0.1, 0.2, 0.2)$ and $d = 0.801$. We calculated the value of m from (20) in the first block, from (7) with $l = 2$ in the second block, and from (2.6) given in Aoshima and Takada (2006) in the third block, respectively, in each table.

In Tables 1–3, the findings obtained by averaging the outcomes from 10,000 ($= R$, say) replications are summarized in each situation. Under a fixed scenario, suppose that the r th replication ends with $N_i = n_{ir}$ ($i = 1, \dots, k$) observations and the corresponding fixed-span confidence region $R_{\mathbf{n}_r}$ based on $\mathbf{n}_r = (n_{1r}, \dots, n_{kr})$ for $r = 1, \dots, R$. Now, $\bar{n}_i = R^{-1} \sum_{r=1}^R n_{ir}$ which estimates C_i

with its estimated standard error $s(\bar{n}_i)$, where $s^2(\bar{n}_i) = (R^2 - R)^{-1} \sum_{r=1}^R (n_{ir} - \bar{n}_i)^2$, $i = 1, \dots, k$. Then, $\bar{n}$ ($= \bar{n}_1 + \dots + \bar{n}_k$) estimates the total fixed sample size C with its estimated standard error $s(\bar{n})$, computed analogously. In the end of the r th replication, we also check whether $\boldsymbol{\mu}$ belongs to the constructed confidence region $R_{\mathbf{n}_r}$ and define $p_r = 1$ (or 0) accordingly as $\boldsymbol{\mu}$ does (or does not) belong to $R_{\mathbf{n}_r}$, $r = 1, \dots, R$. Let $\bar{p} = R^{-1} \sum_{r=1}^R p_r$, which estimates the target coverage probability, having its estimated standard error $s(\bar{p})$ where $s^2(\bar{p}) = R^{-1} \bar{p}(1 - \bar{p})$. For the two-stage procedure (7)–(9) with (11) (or (20)–(21) with (22) or Remark 2 with (24)), the value of u is given as the average number of the outcomes from 10,000 replications. For the two-stage procedure due to Aoshima and Takada (2006), that is given by (25). At the last column, we gave the approximate value of $E_{\boldsymbol{\theta}}(N_i - C_i)$, which was obtained from Theorem 3.2 (or Theorem 3.4) in Section 3 and from Theorem 2.2 in Aoshima and Takada (2006).

Let us explain, for example, the entries from the first block for the case when $k = 2$ in Table 1. We consider $C = 90$, and hence $d = 0.707$, $C_1 = 42.67$, $C_2 = 47.33$ from (6) with $l = 1$. One obtains $m = 30$ using (7) (having $m_0 = 4$, for example). From 10,000 independent simulations, we observed $u = 8.211$, $\bar{n}_1 = 45.18$, $s(\bar{n}_1) = 0.091$, $\bar{n}_2 = 50.03$, $s(\bar{n}_2) = 0.104$, and $\bar{n} = 95.21$, $s(\bar{n}) = 0.175$. Also, we had $\bar{p} = 0.9712$, $s(\bar{p}) = 0.00167$, and $\bar{n}_1 - C_1 = 2.51$, $\bar{n}_2 - C_2 = 2.70$, $\bar{n} - C = 5.21$. At the last column, we had $E(N_1 - C_1) = 2.43$, $E(N_2 - C_2) = 2.68$, $E(N - C) = 5.12$ where $N = \sum_{i=1}^k N_i$. Theorem 3.2 with $j_l = l = 1$ indicates that one may expect $\bar{n}_i - C_i$ to fall in the vicinity of the value of $E(N_i - C_i)$, $i = 1, 2$. One will observe that the values of $E_{\boldsymbol{\theta}}(N_i - C_i)$ are approximated fairly well by these asymptotic values for small d .

Throughout, the two-stage procedure (7)–(9) with (11) or (20)–(21) with (22) or Remark 2 with (24) reduces the sample size required in the two-stage procedure due to Aoshima and Takada (2006). It is obvious specially when the number of populations, k , becomes large. In Table 2, one will observe that the two-stage procedure (7)–(9) with (11) having axis number $l = 2$ reduces the sample size required in the modified procedure (20)–(21) with (22). It is because the former can start with preferable m for the specified axis number, $l = 2$, which is the same number as l_o to minimize the sum $\sum_{i=1}^k C_i$ for the parameter configuration in Table 2. In contrast, in Table 3, one will observe the converse. It is because the axis number $l = 2$ does not coincide with l_o for the parameter configuration in Table 3.

Acknowledgements

The authors would like to thank an anonymous referee and the Associate Editor for their valuable comments.

A Appendix

Throughout this section, for fixed l ($1 \leq l \leq \ell$), we write that

$$\begin{aligned}\tau_{ijl} &= |b_i| \sigma_{il} \sum_{i'=1}^k |b_{i'}| \frac{\sigma_{i'j}^2}{\sigma_{i'l}}, \quad \tau_{jl} = \sum_{i=1}^k |b_i| \frac{\sigma_{ij}^2}{\sigma_{il}}, \\ Y_{ijl} &= |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| \frac{S_{i'j}^2}{S_{i'l}}, \quad Y_{jl} = \sum_{i=1}^k |b_i| \frac{S_{ij}^2}{S_{il}}, \\ \tilde{Y}_{il} &= \max(Y_{i1l}, \dots, Y_{i\ell l}), \quad \tilde{Y}_l = \max(Y_{1l}, \dots, Y_{\ell l})\end{aligned}$$

for $i = 1, \dots, k$; $j = 1, \dots, \ell$. Then, note that $\tilde{Y}_{il} = |b_i| S_{il} \tilde{Y}_l$. From (6), we write that $C_i = a\tau_{ijl}/d^2$. Let d (> 0) go to zero through a sequence such that $a\tau_{l*}/d^2$ always remains an integer. Then, from (7), we may write that $m = a\tau_{l*}/d^2$.

Let $\nu_j = \nu r_j$, $j = 1, \dots, \ell$. For S_{ij}^2 , $j = 1, \dots, \ell$ ($i = 1, \dots, k$), defined by (8), we note that $\nu_j S_{ij}^2 / \sigma_{ij}^2$, $j = 1, \dots, \ell$, are independently distributed as the chi-square distribution with ν_j d.f., $j = 1, \dots, \ell$, respectively. For each i , let W_{ij} , $j = 1, \dots, \ell$, denote the random variables such that $\nu_j W_{ij}$, $j = 1, \dots, \ell$, are independently distributed as the chi-square distribution with ν_j d.f., $j = 1, \dots, \ell$, respectively. Let $\delta_{ij} = W_{ij} - 1$. Then, we have that $S_{ij}^2 = \sigma_{ij}^2(1 + \delta_{ij})$, and $E(\delta_{ij}) = 0$, $E(\delta_{ij}^2) = 2\nu_j^{-1}$, $E(\delta_{ij}^{2t-1}) = O(\nu^{-t})$, $E(\delta_{ij}^{2t}) = O(\nu^{-t})$ and $E((1 + \delta_{ij})^t) = O(1)$, $t = 1, 2, \dots$

Lemma 1 *For fixed l ($1 \leq l \leq \ell$), if there exists some j_l such that $\tau_{jl} > \tau_{j'l}$ for all j' ($1 \leq j' \neq j_l \leq \ell$), we have as $\nu \rightarrow \infty$ that*

$$E_{\boldsymbol{\theta}}(\tilde{Y}_{il}) = E_{\boldsymbol{\theta}}(Y_{ijl}) + O(\nu^{-3/2}) \quad \text{and} \quad E_{\boldsymbol{\theta}}(|Y_{ijl} - \tau_{ijl}|^t) = O(\nu^{-t/2}) \quad (t \geq 2).$$

Proof. We write that

$$\frac{S_{ij}^2}{S_{il}} - \frac{\sigma_{ij}^2}{\sigma_{il}} = \frac{\sigma_{ij}^2}{\sigma_{il}} \left\{ (1 + \delta_{ij}) \left(\frac{1}{\sqrt{1 + \delta_{il}}} - 1 \right) + \delta_{ij} \right\}$$

and

$$\begin{aligned}\frac{S_{il} S_{i'j}^2}{S_{i'l}} - \frac{\sigma_{il} \sigma_{i'j}^2}{\sigma_{i'l}} &= \frac{\sigma_{il} \sigma_{i'j}^2}{\sigma_{i'l}} \left\{ (1 + \delta_{i'j}) \left((\sqrt{1 + \delta_{il}} - 1) \left(\frac{1}{\sqrt{1 + \delta_{i'l}}} - 1 \right) \right. \right. \\ &\quad \left. \left. + (\sqrt{1 + \delta_{il}} - 1) + \left(\frac{1}{\sqrt{1 + \delta_{i'l}}} - 1 \right) \right) + \delta_{i'j} \right\}.\end{aligned}$$

By noting that $E_{\boldsymbol{\theta}}(|(1 + \delta_{il})^{1/2} - 1|^t) = O(\nu^{-t/2})$ and $E_{\boldsymbol{\theta}}(|(1 + \delta_{il})^{-1/2} - 1|^t) = O(\nu^{-t/2})$ ($t \geq 2$), we have for $t \geq 2$ that

$$E_{\boldsymbol{\theta}} \left(\left| \frac{S_{ij}^2}{S_{il}} - \frac{\sigma_{ij}^2}{\sigma_{il}} \right|^t \right) = O(\nu^{-t/2}) \quad \text{and} \quad E_{\boldsymbol{\theta}} \left(\left| \frac{S_{il} S_{i'j}^2}{S_{i'l}} - \frac{\sigma_{il} \sigma_{i'j}^2}{\sigma_{i'l}} \right|^t \right) = O(\nu^{-t/2})$$

and hence that

$$E_{\boldsymbol{\theta}}(|Y_{jl} - \tau_{jl}|^t) = E_{\boldsymbol{\theta}} \left(\left| \sum_{i=1}^k |b_i| \left(\frac{S_{ij}^2}{S_{il}} - \frac{\sigma_{ij}^2}{\sigma_{il}} \right) \right|^t \right) = O(\nu^{-t/2}), \quad (\text{A.1})$$

$$E_{\boldsymbol{\theta}}(|Y_{ijl} - \tau_{ijl}|^t) = E_{\boldsymbol{\theta}} \left(\left| \sum_{i'=1}^k |b_i| |b_{i'}| \left(\frac{S_{il} S_{i'j}^2}{S_{i'l}} - \frac{\sigma_{il} \sigma_{i'l}^2}{\sigma_{i'l}} \right) \right|^t \right) = O(\nu^{-t/2}).$$

Next, there would exist some $\varepsilon_{A_l} (> 0)$ such that $\tau_{jl} > \tau_{j'l} + \varepsilon_{A_l}$ for all $j' (\neq j_l)$. Define $A_{jl} = \{|Y_{jl} - \tau_{jl}| < \varepsilon_{A_l}/2\}$, $j = 1, \dots, \ell$, and $A_l = \cap_{j=1}^{\ell} A_{jl}$. Let $\mathbf{Y}_l = (Y_{1l}, \dots, Y_{\ell l})$. Note that

$$\tilde{Y}_l = Y_{j_l l} \text{ and } \tilde{Y}_{il} = Y_{ij_l l} \text{ when } \mathbf{Y}_l \in A_l. \quad (\text{A.2})$$

Otherwise, from (A.1), we evaluate that

$$\begin{aligned} P_{\boldsymbol{\theta}}(A_l^c) &\leq \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}}(A_{jl}^c) = \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}}(|Y_{jl} - \tau_{jl}| \geq \varepsilon_{A_l}/2) \\ &\leq (2/\varepsilon_{A_l})^6 \sum_{j=1}^{\ell} E_{\boldsymbol{\theta}}(|Y_{jl} - \tau_{jl}|^6) = O(\nu^{-3}). \end{aligned} \quad (\text{A.3})$$

Let I_{A_l} be the indicator function. Since $|E_{\boldsymbol{\theta}}(Y_{ijl} I_{A_l^c})| \leq \sqrt{E_{\boldsymbol{\theta}}(Y_{ijl}^2) P_{\boldsymbol{\theta}}(A_l^c)} = O(\nu^{-3/2})$, we have that

$$E_{\boldsymbol{\theta}}(\tilde{Y}_{il} I_{A_l^c}) = O(\nu^{-3/2}). \quad (\text{A.4})$$

From (A.2) and (A.4), we obtain that

$$E_{\boldsymbol{\theta}}(\tilde{Y}_{il}) = E_{\boldsymbol{\theta}}(\tilde{Y}_{il} I_{A_l}) + E_{\boldsymbol{\theta}}(\tilde{Y}_{il} I_{A_l^c}) = E_{\boldsymbol{\theta}}(Y_{ij_l l}) + O(\nu^{-3/2}).$$

The proof is completed. $\square$

Lemma 2. For the two-stage procedure (7)–(9) with (11), we have as $d \rightarrow 0$ that

$$E_{\boldsymbol{\theta}} \left(N_i - \left\lfloor \frac{u_l}{d^2} \tilde{Y}_{il} \right\rfloor - 1 \right) = O(d).$$

Proof. Let $I_{\{N_i=m\}}$ be the indicator function. Then, we have that

$$\begin{aligned} E_{\boldsymbol{\theta}} \left(N_i - \left\lfloor \frac{u_l}{d^2} \tilde{Y}_{il} \right\rfloor - 1 \right) &= E_{\boldsymbol{\theta}} \left\{ I_{\{N_i=m\}} \left(m - \left\lfloor \frac{u_l}{d^2} \tilde{Y}_{il} \right\rfloor - 1 \right) \right\} \\ &\leq \sqrt{P_{\boldsymbol{\theta}}(N_i = m) E_{\boldsymbol{\theta}} \left\{ \left(m - \left\lfloor \frac{u_l}{d^2} \tilde{Y}_{il} \right\rfloor - 1 \right)^2 \right\}}. \end{aligned} \quad (\text{A.5})$$

From (A.2), we note that $j_{l_m} = j_l$ when $(Y_{1l}, \dots, Y_{\ell l}) \in A_l$. From assumptions (A1)–(A2), we have that $\tau_{ijl} > \tau_{l*}$. Then, it follows that

$$\begin{aligned}
P_{\boldsymbol{\theta}}(N_i = m) &= P_{\boldsymbol{\theta}}\left(\frac{u_l \tilde{Y}_{il}}{d^2} \leq m\right) = P_{\boldsymbol{\theta}}\left(\frac{u_l \tilde{Y}_{il}}{d^2 C_i} - \frac{C_i + 1}{C_i} \leq \frac{m - (C_i + 1)}{C_i}\right) \\
&\leq P_{\boldsymbol{\theta}}\left(\frac{u_l \tilde{Y}_{il}}{a \tau_{ijl}} - 1 - \frac{1}{C_i} \leq \frac{\tau_{l*} - \tau_{ijl}}{\tau_{ijl}}\right) \\
&\leq P_{\boldsymbol{\theta}}\left(\left(\frac{u_l \tilde{Y}_{il}}{a \tau_{ijl}} - 1 - \frac{1}{C_i} \leq \frac{\tau_{l*} - \tau_{ijl}}{\tau_{ijl}}\right) \cap A_l\right) + P_{\boldsymbol{\theta}}(A_l^c) \\
&\leq P_{\boldsymbol{\theta}}\left(\left|\frac{u_{l(jl)} Y_{ijl}}{a \tau_{ijl}} - 1\right| + C_i^{-1} \geq \frac{\tau_{ijl} - \tau_{l*}}{\tau_{ijl}}\right) + P_{\boldsymbol{\theta}}(A_l^c) \quad (\text{A.6}) \\
&\leq \left(\frac{\tau_{ijl} - \tau_{l*}}{\tau_{ijl}}\right)^{-6} E_{\boldsymbol{\theta}}\left\{\left(\left|\frac{u_{l(jl)} Y_{ijl}}{a \tau_{ijl}} - 1\right| + C_i^{-1}\right)^6\right\} + P_{\boldsymbol{\theta}}(A_l^c), \quad (\text{A.7})
\end{aligned}$$

where $u_{l(jl)} (= a(1 + \nu^{-1} \hat{s}_{l(jl)}))$, say) in (A.6) denotes the value of u_l when $(Y_{1l}, \dots, Y_{\ell l}) \in A_l$. Now, one yields that

$$\begin{aligned}
E_{\boldsymbol{\theta}}\left\{\left|\frac{u_{l(jl)} Y_{ijl}}{a \tau_{ijl}} - 1\right|^t\right\} &\leq E_{\boldsymbol{\theta}}\left\{\left(\frac{1}{\tau_{ijl}} \left(|Y_{ijl} - \tau_{ijl}| + \left|\frac{\hat{s}_{l(jl)} Y_{ijl}}{\nu}\right|\right)\right)^t\right\} \\
&= O(\nu^{-t/2}) \quad (t \geq 2). \quad (\text{A.8})
\end{aligned}$$

Here, (A.8) follows from the result that for any $x (\geq 0)$ and $y (\geq 0)$ such that $x + y = t (\geq 2)$, we have from Lemma 1 that

$$\begin{aligned}
E_{\boldsymbol{\theta}}\left(|Y_{ijl} - \tau_{ijl}|^x \left|\nu^{-1} \hat{s}_{l(jl)} Y_{ijl}\right|^y\right) &\leq \sqrt{E_{\boldsymbol{\theta}}(|Y_{ijl} - \tau_{ijl}|^{2x}) E_{\boldsymbol{\theta}}(|\nu^{-1} \hat{s}_{l(jl)} Y_{ijl}|^{2y})} \\
&= O\left(\nu^{-(x/2+y)}\right) = O\left(\nu^{-(t/2+y/2)}\right).
\end{aligned}$$

By combining (A.8) and (A.3) with (A.7), we evaluate that

$$P_{\boldsymbol{\theta}}(N_i = m) = O(d^6). \quad (\text{A.9})$$

We obtain the result in view of (A.5) and (A.9). $\square$

The following result was given by Aoshima and Yata (2007).

Lemma 3. *Let $q (> 0)$ and $h (\geq 0)$ be constants. For a fixed $b (\geq 1)$, let $X_{b\nu}$ denote a chi-square random variable with $b\nu$ d.f. Let $U = qX_{b\nu} - h - [qX_{b\nu} - h]$, where $[x]$ denotes the largest integer less than x . Then, we have that*

$$x - \frac{x(1 + h + x_i)}{q} \sup_z F'_{b\nu}(z) \leq P(U \leq x) \leq x + \frac{x}{q} \sup_z F'_{b\nu}(z),$$

where $x \in (0, 1)$ and $x_i \in (0, x)$, and $F'_{b\nu}(\cdot)$ denotes the first derivative of $F_{b\nu}(\cdot)$.

Lemma 4. For the two-stage procedure (7)–(9) with (11), we have as $d \rightarrow 0$ that

$$E_{\boldsymbol{\theta}} \left\{ \frac{u_l}{d^2} |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| \frac{S_{i'j_{lm}}}{S_{i'l}} - \left[\frac{u_l}{d^2} |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| \frac{S_{i'j_{lm}}}{S_{i'l}} \right] \right\} = \frac{1}{2} + O(d).$$

Proof. Let $U_{\{ijl\}} = (u_{l(j)}/d^2) |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| S_{i'l}^2 / S_{i'l}$ with $u_{l(j)} = a(1 + \nu^{-1} \hat{s}_{l(j)})$. Then, from (A.2)–(A.3), we have as $d \rightarrow 0$ that

$$\begin{aligned} & E_{\boldsymbol{\theta}} \{U_{\{ijl_m\}} - [U_{\{ijl_m\}}]\} \\ &= E_{\boldsymbol{\theta}} \{ (U_{\{ijl_m\}} - [U_{\{ijl_m\}}]) I_{A_l} \} + E_{\boldsymbol{\theta}} \{ (U_{\{ijl_m\}} - [U_{\{ijl_m\}}]) I_{A_l^c} \} \\ &= E_{\boldsymbol{\theta}} \{ U_{\{ijl\}} - [U_{\{ijl\}}] \} + O(d). \end{aligned} \quad (\text{A.10})$$

Let W_{ij} , $j = 1, \dots, \ell$ ($i = 1, \dots, k$), be independently distributed as the chi-square distribution with ν_j d.f., $j = 1, \dots, \ell$, respectively. Let $X_j = \sum_{i=1}^k W_{ij}$ and $V_{ij} = W_{ij}/X_j$, $i = 1, \dots, k$; $j = 1, \dots, \ell$. Then, X_j is distributed as the chi-square distribution with $k\nu_j$ d.f., V_{ij} is distributed as the beta distribution with parameters $\nu_j/2$ and $(k-1)\nu_j/2$, and X_j and $\tilde{V}_l = (V_{1l}, \dots, V_{kl})$ are independent.

When $j_l = l$, we write $\hat{s}_{l(j_l)}$ as

$$\hat{s}_{l(j_l)} = \frac{1}{r_l} + \frac{(a-p)b_i^2 \sigma_{il}^2 V_{il} \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l}^2 V_{i'l}}{2r_l Z_i^2} - \nu_l \frac{b_i^2 \sigma_{il}^2 V_{il} k \tau_{l\star}}{2X_l Z_i^2},$$

where $Z_i = |b_i| \sigma_{il} \sqrt{V_{il} \sum_{i'=1}^k |b_{i'}| \sigma_{i'l} \sqrt{V_{i'l}}}$. Then, we have that $(u_{l(j_l)}/d^2) |b_i| S_{il} \sum_{i'=1}^k |b_{i'}| S_{i'l} = (u_{l(j_l)}/d^2 \nu_l) X_l Z_i$ where

$$\begin{aligned} \frac{u_{l(j_l)} X_l Z_i}{d^2 \nu_l} &= X_l \frac{a Z_i}{d^2 \nu_l} \left(1 + \frac{1}{\nu_l} \left(1 + \frac{(a-p)b_i^2 \sigma_{il}^2 V_{il} \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l}^2 V_{i'l}}{2Z_i^2} \right) \right) \\ &\quad - \frac{ab_i^2 \sigma_{il}^2 V_{il} k \tau_{l\star}}{2d^2 Z_i \nu}. \end{aligned}$$

Let us denote

$$\begin{aligned} Q &= \frac{a Z_i}{d^2 \nu_l} \left(1 + \frac{1}{\nu_l} \left(1 + \frac{(a-p)b_i^2 \sigma_{il}^2 V_{il} \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l}^2 V_{i'l}}{2Z_i^2} \right) \right), \\ H &= \frac{ab_i^2 \sigma_{il}^2 V_{il} k \tau_{l\star}}{2d^2 Z_i \nu}. \end{aligned}$$

Then, we have that $(u_{l(j_l)}/d^2 \nu_l) X_l Z_i = QX_l - H$. Let $U = (QX_l - H) - [QX_l - H]$. From Lemma 3, the conditional distribution of U , given $\tilde{V}_l =$

$\tilde{v}_l$ ($H = h$, $Q = q$), is given for $x \in (0, 1)$ that

$$x - \frac{x(1+h+x_i)}{q} \sup_z F'_{k\nu_l}(z) \leq P_{\boldsymbol{\theta}}(U \leq x | \tilde{V}_l = \tilde{v}_l) \leq x + \frac{x}{q} \sup_z F'_{k\nu_l}(z),$$

where $x_i \in (0, x)$. If $a \geq p-2$, we evaluate that $1/Q \leq r_l \tau_{l\star}/Z_i \leq r_l \tau_{l\star}/(b_i^2 \sigma_{il}^2 V_{il})$ and $H/Q \leq r_l k \tau_{l\star}/(2 \sum_{i'=1}^k b_{i'}^2 \sigma_{i'l}^2 V_{i'l}) \leq r_l k \tau_{l\star}/(2 \min_{1 \leq i \leq k} b_i^2 \sigma_{il}^2)$ ($= \gamma$). Then, we have $E_{\boldsymbol{\theta}}(1/Q) \leq (r_l \tau_{l\star}/b_i^2 \sigma_{il}^2)(k\nu_l - 2)/(\nu_l - 2)$. Here, H/Q is uniformly integrable since $|H/Q| \leq \gamma$, and $1/Q$ is uniformly integrable since $|1/Q| \leq r_l \tau_{l\star}/(b_i^2 \sigma_{il}^2 V_{il})$ with $r_l \tau_{l\star}/(b_i^2 \sigma_{il}^2 V_{il})$ being uniformly integrable. From the result that $\sup_z F'_{k\nu_l}(z) = O(\nu^{-1/2})$ as $\nu \rightarrow \infty$, one can yield that

$$E_{\boldsymbol{\theta}} \left\{ \frac{x}{Q} \sup_z F'_{k\nu_l}(z) \right\} \leq E_{\boldsymbol{\theta}} \left\{ \frac{x(1+H+x_i)}{Q} \sup_z F'_{k\nu_l}(z) \right\} = O(d).$$

From the fact that $E_{\boldsymbol{\theta}}\{P_{\boldsymbol{\theta}}(U \leq x | \tilde{V} = \tilde{v})\} = P_{\boldsymbol{\theta}}(U \leq x)$, we obtain that

$$P_{\boldsymbol{\theta}}(U \leq x) = x + O(d) \quad \text{as } d \rightarrow 0. \quad (\text{A.11})$$

Next, when $j_l \neq l$, let us denote

$$Q = \frac{aZ_i}{d^2\nu_{j_l}} \left(1 + \frac{(a+2-p)b_i^2\sigma_{il}^2V_{il}}{2\nu_{j_l}Z_i^2} \sum_{i'=1}^k b_i^2 \frac{\sigma_{i'j_l}^4 V_{i'j_l}^2}{\sigma_{i'l}^2 V_{i'l}} \right),$$

$$H = \frac{a\tau_{l\star}b_i^2\sigma_{il}^2V_{il}}{2\nu d^2Z_i} \sum_{i'=1}^k \frac{\sigma_{i'j_l}^2 V_{i'j_l}}{\sigma_{i'l}^2 V_{i'l}},$$

where $Z_i = |b_i|\sigma_{il}\sqrt{V_{il}} \sum_{i'=1}^k |b_{i'}|(\sigma_{i'j_l}^2 V_{i'j_l}/\sigma_{i'l}\sqrt{V_{i'l}})$. Then, $(u_{l(j_l)}/d^2)|b_i|S_{il} \sum_{i'=1}^k |b_{i'}|S_{i'j_l}^2/S_{i'l} = QX_{j_l} - H$. Since $a \geq p-2$, we evaluate that

$$E_{\boldsymbol{\theta}} \left\{ \frac{H}{Q} \right\} \leq \frac{r_{j_l} \tau_{l\star} \max_{1 \leq i \leq k} \sigma_{ij_l}^2/\sigma_{il}^2}{2 \min_{1 \leq i \leq k} b_i^2 \sigma_{ij_l}^4/\sigma_{il}^2} E_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^k \frac{1}{V_{il}} \right\}$$

$$= \frac{kr_{j_l} \tau_{l\star} \max_{1 \leq i \leq k} \sigma_{ij_l}^2/\sigma_{il}^2}{2 \min_{1 \leq i \leq k} b_i^2 \sigma_{ij_l}^4/\sigma_{il}^2} \left(\frac{k\nu_{j_l} - 2}{\nu_l - 2} \right),$$

$$E_{\boldsymbol{\theta}} \left\{ \frac{1}{Q} \right\} \leq \left(\frac{r_{j_l} \tau_{l\star}}{b_i^2 \sigma_{ij_l}^2} \right) \left(\frac{k\nu_{j_l} - 2}{\nu_{j_l} - 2} \right).$$

Then, we claim (A.11) similarly to the case when $j_l = l$. Hence, U is asymptotically uniform on $(0, 1)$ as $d \rightarrow 0$. In view of (A.10), the proof is completed. $\square$

Remark 5. When the design value is defined as a constant, the asymptotic uniformity of $P(U \leq x)$ was studied by several authors. See Hall (1981) for $k = 1$ and Takada (2004) for $k \geq 2$. When u_l is not a constant as in this paper, it is much more complicated.

Lemma 5. When $j_l = l$, the two-stage procedure (7)–(9) with (11) has as $d \rightarrow 0$:

$$\begin{aligned} (i) \quad E_{\boldsymbol{\theta}}\{C_i^{-1}(N_i - C_i)\} &= (2\nu_l)^{-1}(2s_l r_l - 1 + f_i + B_i) + O(d^3), \\ (ii) \quad E_{\boldsymbol{\theta}}\{C_i^{-2}(N_i - C_i)^2\} &= (2\nu_l)^{-1}(1 + 2f_i + \sum_{i'=1}^k f_{i'}^2) + O(d^3), \\ (iii) \quad E_{\boldsymbol{\theta}}\{C_i^{-1}(N_i - C_i)C_j^{-1}(N_j - C_j)\} &= (2\nu_l)^{-1}(f_i + f_j + \sum_{i'=1}^k f_{i'}^2) + O(d^3) \\ &\quad (i \neq j); \end{aligned}$$

When $j_l \neq l$, it has as $d \rightarrow 0$:

$$\begin{aligned} (i) \quad E_{\boldsymbol{\theta}}\{C_i^{-1}(N_i - C_i)\} &= (2\nu_l)^{-1}(2s_l r_l + 1 - f_i + B_i) + O(d^3), \\ (ii) \quad E_{\boldsymbol{\theta}}\{C_i^{-2}(N_i - C_i)^2\} &= (2\nu_l)^{-1}(1 - 2f_i + (4r_{j_l}^{-1}r_l + 1) \sum_{i'=1}^k f_{i'}^2) + O(d^3), \\ (iii) \quad E_{\boldsymbol{\theta}}\{C_i^{-1}(N_i - C_i)C_j^{-1}(N_j - C_j)\} &= (2\nu_l)^{-1}((4r_{j_l}^{-1}r_l + 1) \sum_{i'=1}^k f_{i'}^2 - f_i - f_j) \\ &\quad + O(d^3) \quad (i \neq j), \end{aligned}$$

where $B_i = \nu_l/C_i$ and s_l is given by (17).

Proof. Let us write that

$$N_i = rC_i T_i + (1 + [rC_i T_i] - rC_i T_i) + (N_i - [rC_i T_i] - 1),$$

where $r = u_l/a = 1 + \nu^{-1}\hat{s}_l$ and $T_i = \tau_{ij_l l}^{-1}\tilde{Y}_{il}$. Here, from Lemma 4, $U_i = 1 + [rC_i T_i] - rC_i T_i$ is asymptotically distributed as $U(0, 1)$. Let $D_i = N_i - [rC_i T_i] - 1$. From Lemma 2, it follows that $E\{(D_i/\nu)^c\} = O(\nu^{-3/2})$ as $d \rightarrow 0$, where $c (\geq 1)$ is fixed. Then, we have that

$$C_i^{-1}(N_i - C_i) = (rT_i - 1) + \nu^{-1}B_i U_i + C_i^{-1}D_i. \quad (\text{A.12})$$

We have from Lemma 1 that $E_{\boldsymbol{\theta}}(\hat{s}_l) = s_l + o(1)$. When $j_l = l$, we obtain the following results:

$$\begin{aligned} E_{\boldsymbol{\theta}}(rT_i - 1) &= (2\nu_l)^{-1}(2s_l r_l - 1 + f_i) + O(d^3), \\ E_{\boldsymbol{\theta}}\{(rT_i - 1)^2\} &= (2\nu_l)^{-1}(1 + 2f_i + \sum_{i'=1}^k f_{i'}^2) + O(d^3), \\ E_{\boldsymbol{\theta}}\{(rT_i - 1)(rT_j - 1)\} &= (2\nu_l)^{-1}(f_i + f_j + \sum_{i'=1}^k f_{i'}^2) + O(d^3) \quad (i \neq j). \end{aligned} \quad (\text{A.13})$$

When $j_l \neq l$, we obtain the following results:

$$\begin{aligned} E_{\boldsymbol{\theta}}(rT_i - 1) &= (2\nu_l)^{-1}(2s_l r_l + 1 - f_i) + O(d^3), \\ E_{\boldsymbol{\theta}}\{(rT_i - 1)^2\} &= (2\nu_l)^{-1}(1 - 2f_i + (4r_{j_l}^{-1}r_l + 1) \sum_{i'=1}^k f_{i'}^2) + O(d^3), \\ E_{\boldsymbol{\theta}}\{(rT_i - 1)(rT_j - 1)\} &= (2\nu_l)^{-1}(-f_i - f_j + (4r_{j_l}^{-1}r_l + 1) \sum_{i'=1}^k f_{i'}^2) + O(d^3) \\ &\quad (i \neq j). \end{aligned} \quad (\text{A.14})$$

Let us combine these results with the expectations of (A.12). Let $U_i = U_0 + \varepsilon_i$, where U_0 is a $U(0, 1)$ random variable and ε_i is the remainder term. Then, note that $(E\{(rT_i - 1)\nu^{-1}\varepsilon_j\})^2 \leq E\{(rT_i - 1)^2\}E(\nu^{-2}\varepsilon_j^2) = o(\nu^{-3})$ so that $E\{(rT_i - 1)\nu^{-1}\varepsilon_j\} = o(\nu^{-3/2})$. The results are obtained straightforwardly. $\square$

Lemma 6. *For the two-stage procedure (7)–(9) with (11), let j_{l_N} be the index such that $\max_{1 \leq j \leq \ell} (\sum_{i=1}^k b_i^2 \sigma_{ij}^2 / N_i) = \sum_{i=1}^k b_i^2 \sigma_{ij_{l_N}}^2 / N_i$. Then, we have as $d \rightarrow 0$ that*

$$P_{\boldsymbol{\theta}}(j_{l_N} \neq j_l) = O(d^6).$$

Proof. There would exist some $\varepsilon_{J_l} (> 0)$ such that $\sum_{i=1}^k b_i^2 \sigma_{ij_l}^2 / \tau_{ij_l} > \sum_{i=1}^k b_i^2 \sigma_{ij'}^2 / \tau_{ij'l} + \varepsilon_{J_l}$ for all $j' (\neq j_l)$. Define $J_{j_l} = \{|\sum_{i=1}^k b_i^2 \sigma_{ij}^2 / N_i - \sum_{i=1}^k b_i^2 \sigma_{ij}^2 / C_i| < \varepsilon_{J_l}(d^2/2a)\}$, $j = 1, \dots, \ell$, and $J_l = \cap_{j=1}^{\ell} J_{j_l}$. Note that

$$j_{l_N} = j_l \text{ when } \left(\sum_{i=1}^k b_i^2 \frac{\sigma_{i1}^2}{N_i}, \dots, \sum_{i=1}^k b_i^2 \frac{\sigma_{i\ell}^2}{N_i} \right) \in J_l.$$

Now, from (A.8), we evaluate that

$$\begin{aligned} & E_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)} Y_{ijl}} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right|^t \right\} \\ & \leq E_{\boldsymbol{\theta}} \left\{ \left(\frac{1}{u_{l(j_l)}} \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{Y_{ijl}} \left| \frac{u_{l(j_l)} Y_{ijl}}{a \tau_{ijl}} - 1 \right| \right)^t \right\} = O(\nu^{-t/2}) \quad (t \geq 2). \end{aligned} \quad (\text{A.15})$$

Next, define $D_{il} = \{1 + [u_l d^{-2} \tilde{Y}_{il}] > m\}$, $i = 1, \dots, k$, and $D_l = \cap_{i=1}^k D_{il}$. Then, from Lemmas 1-2, we evaluate that

$$\begin{aligned} & P_{\boldsymbol{\theta}}(J_l^c) \\ & \leq \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}}(J_{j_l}^c) = \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}} \left(\left| \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{N_i} - \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{C_i} \right| > \varepsilon_{J_l}(d^2/2a) \right) \\ & \leq \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}} \left(\left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{N_i} - \frac{b_i^2 \sigma_{ij}^2}{C_i} \right| > \varepsilon_{J_l}(d^2/2a) \right) \cap A_l \cap D_l \right) + P_{\boldsymbol{\theta}}(A_l^c) + P_{\boldsymbol{\theta}}(D_l^c) \\ & = \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}} \left(\left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{[u_l \tilde{Y}_{il}/d^2] + 1} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}/d^2} \right| > \varepsilon_{J_l}(d^2/2a) \right) \cap A_l \right) \\ & \quad + P_{\boldsymbol{\theta}}(A_l^c) + P_{\boldsymbol{\theta}}(D_l^c) \\ & = \sum_{j=1}^{\ell} P_{\boldsymbol{\theta}} \left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)} Y_{ijl} + d^2 U_i} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right| > \varepsilon_{J_l}/2a \right) + O(d^6), \end{aligned} \quad (\text{A.16})$$

where $U_i = 1 + [u_{l(j_l)}Y_{ijl}/d^2] - u_{l(j_l)}Y_{ijl}/d^2$, $i = 1, \dots, k$. Here, using the Taylor expansion and (A.15), we evaluate in (A.16) that

$$\begin{aligned}
& P_{\boldsymbol{\theta}} \left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)}Y_{ijl} + d^2 U_i} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right| > \varepsilon_{J_l}/2a \right) \\
& \leq P_{\boldsymbol{\theta}} \left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)}Y_{ijl}} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right| + d^2 \sum_{i=1}^k U_i \frac{b_i^2 \sigma_{ij}^2}{(u_{l(j_l)}Y_{ijl} + d^2 U_i')^2} > \varepsilon_{J_l}/2a \right) \\
& \leq P_{\boldsymbol{\theta}} \left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)}Y_{ijl}} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right| + d^2 \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{(u_{l(j_l)}Y_{ijl})^2} > \varepsilon_{J_l}/2a \right) \\
& \leq \left(\frac{2a}{\varepsilon_{J_l}} \right)^6 E_{\boldsymbol{\theta}} \left\{ \left(\sum_{i=1}^k \left| \frac{b_i^2 \sigma_{ij}^2}{u_{l(j_l)}Y_{ijl}} - \frac{b_i^2 \sigma_{ij}^2}{a \tau_{ijl}} \right| + d^2 \sum_{i=1}^k \frac{b_i^2 \sigma_{ij}^2}{(u_{l(j_l)}Y_{ijl})^2} \right)^6 \right\} \\
& = O(d^6), \tag{A.17}
\end{aligned}$$

where $U_i' \in (0, U_i)$, $i = 1, \dots, k$. From (A.16) and (A.17), we conclude the result. $\square$

The following result can be obtained similarly to Lemma 6 given in Aoshima and Yata (2007).

Lemma 7. *For the two-stage procedure (7)–(9) with (11), one has from (A.13)–(A.14) as $d \rightarrow 0$ that $E_{\boldsymbol{\theta}}(\mathfrak{R}) = o(d^2)$ in (15).*

Lemma 8 *For the two-stage procedure (20)–(21) with (22), let $\tilde{u}\tilde{Y}_i$, $i = 1, \dots, k$, satisfy that*

$$\begin{aligned}
& \sum_{i=1}^k \max \left\{ m, \left\lceil \frac{\tilde{u}\tilde{Y}_i}{d^2} \right\rceil + 1 \right\} \\
& = \min \left(\sum_{i=1}^k \max \left\{ m, \left\lceil \frac{u_1 \tilde{Y}_{i1}}{d^2} \right\rceil + 1 \right\}, \dots, \sum_{i=1}^k \max \left\{ m, \left\lceil \frac{u_{\ell} \tilde{Y}_{i\ell}}{d^2} \right\rceil + 1 \right\} \right),
\end{aligned}$$

where $m = [ad^{-2}\tau_{\sigma^*}] + 1$, and u_j is the design constant with axis number j ($= 1, \dots, \ell$). Then, we have as $d \rightarrow 0$ that

$$E_{\boldsymbol{\theta}}(\tilde{Y}_i) = E_{\boldsymbol{\theta}}(Y_{ij_{o_l}}) + O(d^3).$$

Proof. Let $\hat{u}\hat{Y}_i$, $i = 1, \dots, k$, satisfy that

$$\sum_{i=1}^k \left(1 + \left\lceil \frac{\hat{u}\hat{Y}_i}{d^2} \right\rceil \right) = \min \left(\sum_{i=1}^k \left(1 + \left\lceil \frac{u_{1(j_1)}Y_{ij_1}}{d^2} \right\rceil \right), \dots, \sum_{i=1}^k \left(1 + \left\lceil \frac{u_{\ell(j_\ell)}Y_{ij_\ell}}{d^2} \right\rceil \right) \right).$$

From (A.2), we note that $u_{l(j_l)}Y_{ijl} = u_l \tilde{Y}_{il}$ when $(Y_{1l}, \dots, Y_{\ell l}) \in A_l$, $i = 1, \dots, k$; $l = 1, \dots, \ell$. We note that $\max(m, [u_l d^{-2} \tilde{Y}_{il}] + 1) = [u_l d^{-2} \tilde{Y}_{il}] + 1$ when

$u_l \tilde{Y}_{il} \in D_{il}$, $i = 1, \dots, k$; $l = 1, \dots, \ell$. Let $A = \cap_{l=1}^{\ell} A_l$ and $D = \cap_{l=1}^k \cap_{l=1}^{\ell} D_{il}$. Let $\mathbf{Y}_l = (Y_{1l}, \dots, Y_{\ell l})$, $l = 1, \dots, \ell$, and $\tilde{\mathbf{u}}\tilde{\mathbf{Y}} = (\sum_{i=1}^k u_1 \tilde{Y}_{i1}, \dots, \sum_{i=1}^k u_{\ell} \tilde{Y}_{i\ell})$. Note that

$$\tilde{Y}_i = \hat{Y}_i \text{ when } (\mathbf{Y}_1, \dots, \mathbf{Y}_{\ell}) \in A \text{ and } \tilde{\mathbf{u}}\tilde{\mathbf{Y}} \in D \quad (\text{A.18})$$

for $i = 1, \dots, k$.

Next, let us write $\sum_{i=1}^k \tau_{ij_o l_o} = \min(\sum_{i=1}^k \tau_{ij_1 1}, \dots, \sum_{i=1}^k \tau_{ij_{\ell} \ell})$. Recall the definition of (l_o, j_o) , which was given right after Theorem 3.2, to claim that

$$\sum_{i=1}^k \tau_{ij_o l_o} < \sum_{i=1}^k \tau_{ij_{l'} l'} \text{ for all } l' (\neq l_o).$$

Then, there would exist some $\varepsilon_B (> 0)$ such that $\sum_{i=1}^k \tau_{ij_o l_o} + \varepsilon_B < \sum_{i=1}^k \tau_{ij_{l'} l'}$ for all $l' (\neq l_o)$. Define $B_l = \{|\sum_{i=1}^k (1 + [u_{l(j_l)} d^{-2} Y_{ij_l l}]) - a d^{-2} \sum_{i=1}^k \tau_{ij_l l}|\} < \varepsilon_B(a/2d^2)\}$, $l = 1, \dots, \ell$, and $B = \cap_{l=1}^{\ell} B_l$. Let $\mathbf{u}\mathbf{Y} = (\sum_{i=1}^k u_{1(j_1)} d^{-2} Y_{ij_1 1}, \dots, \sum_{i=1}^k u_{\ell(j_{\ell})} d^{-2} Y_{ij_{\ell} \ell})$. Note that

$$\hat{Y}_i = Y_{ij_o l_o} \text{ when } \mathbf{u}\mathbf{Y} \in B. \quad (\text{A.19})$$

From (A.8), we have that

$$\begin{aligned} E_{\boldsymbol{\theta}} & \left\{ \left| \sum_{i=1}^k d^2 \left(1 + \left[\frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right] - \frac{a}{d^2} \tau_{ij_l l} \right) \right|^t \right\} \\ &= E_{\boldsymbol{\theta}} \left\{ \left| \sum_{i=1}^k (u_{l(j_l)} Y_{ij_l l} - a \tau_{ij_l l}) + d^2 \sum_{i=1}^k \left(1 + \left[\frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right] - \frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right) \right|^t \right\} \\ &\leq E_{\boldsymbol{\theta}} \left\{ \left(\left| \sum_{i=1}^k u_{l(j_l)} Y_{ij_l l} - a \tau_{ij_l l} \right| + d^2 \sum_{i=1}^k \left(1 + \left[\frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right] - \frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right) \right)^t \right\} \\ &\leq E_{\boldsymbol{\theta}} \left\{ \left(\sum_{i=1}^k a \tau_{ij_l l} \left| \frac{u_{l(j_l)}}{a} \frac{Y_{ij_l l}}{\tau_{ij_l l}} - 1 \right| + d^2 k \right)^t \right\} \\ &= O(d^t) \quad (t \geq 2). \end{aligned}$$

Then, we evaluate that

$$\begin{aligned} P_{\boldsymbol{\theta}}(B^c) &\leq \sum_{l=1}^{\ell} P_{\boldsymbol{\theta}}(B_l^c) = \sum_{l=1}^{\ell} P_{\boldsymbol{\theta}} \left\{ \left| \sum_{i=1}^k \left(1 + \left[\frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right] - \frac{a}{d^2} \tau_{ij_l l} \right) \right| > \varepsilon_B(a/2d^2) \right\} \\ &\leq \left(\frac{2}{a\varepsilon_B} \right)^6 \sum_{l=1}^{\ell} E_{\boldsymbol{\theta}} \left\{ \left| \sum_{i=1}^k d^2 \left(1 + \left[\frac{u_{l(j_l)}}{d^2} Y_{ij_l l} \right] - \frac{a}{d^2} \tau_{ij_l l} \right) \right|^6 \right\} \leq O(d^6). \end{aligned}$$

Let I_B be the indicator function. Note that $|E_{\boldsymbol{\theta}}(Y_{ij_l} I_{B^c})| \leq \sqrt{E_{\boldsymbol{\theta}}(Y_{ij_l}^2) P_{\boldsymbol{\theta}}(B^c)} = O(d^3)$.

From (A.3), we evaluate that

$$P_{\boldsymbol{\theta}}(A^c) \leq \sum_{j=1}^{\ell} \sum_{l=1}^{\ell} P_{\boldsymbol{\theta}}(A_{jl}^c) = O(\nu^{-3}),$$

From (A.7), we evaluate that

$$P_{\boldsymbol{\theta}}(D^c) \leq \sum_{i=1}^k \sum_{l=1}^{\ell} P_{\boldsymbol{\theta}}(D_{il}^c) = O(d^6).$$

Note that $|E_{\boldsymbol{\theta}}(Y_{ijl}I_{A^c})| = O(\nu^{-3/2})$ and $|E_{\boldsymbol{\theta}}(Y_{ijl}I_{D^c})| = O(\nu^{-3/2})$. Let $I_{(A \cap B \cap D)}$ be the indicator function. We have that

$$E_{\boldsymbol{\theta}}(\tilde{Y}_i I_{(A \cap B \cap D)^c}) \leq E_{\boldsymbol{\theta}}(\tilde{Y}_i (I_{A^c} + I_{B^c} + I_{D^c})) = O(d^3). \quad (\text{A.20})$$

From (A.18), (A.19) and (A.20), we obtain that

$$\begin{aligned} E_{\boldsymbol{\theta}}(\tilde{Y}_i) &= E_{\boldsymbol{\theta}}(\tilde{Y}_i I_{(A \cap B \cap D)}) + E_{\boldsymbol{\theta}}(\tilde{Y}_i I_{(A \cap B \cap D)^c}), \\ &= E_{\boldsymbol{\theta}}(Y_{ijolo}) + O(d^3). \end{aligned}$$

The proof is completed. □

References

- Aoshima, M., 2005. Statistical inference in two-stage sampling. Am. Math. Soc. Trans. 215, 125–145.
- Aoshima, M. and Takada, Y., 2006. Second-order efficiency for two-stage estimation of a linear function of normal mean vectors when covariance matrices have some structures. Seq. Anal. 25, 327–345.
- Aoshima, M. and Yata, K., 2007. Asymptotic second-order consistency for two-stage estimation methodologies and its applications, revised.
- Ghosh, M., Mukhopadhyay, N., and Sen, P. K., 1997. Sequential Estimation. Wiley, New York.
- Hall, P., 1981. Asymptotic theory of triple sampling for sequential estimation of a mean. Ann. Statist. 9, 1229–1238.
- Mukhopadhyay, N. and Duggan, W. T., 1999. On a two-stage procedure having second-order properties with applications. Ann. Inst. Statist. Math. 51, 621–636.
- Stein, C., 1945. A two-sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist. 16, 243–258.
- Takada, Y., 2004. Asymptotic second-order efficiency of a two-stage procedure for estimating a linear function of normal means. Seq. Anal. 23, 103–120.

Fig.2. When the sample is taken up to size $n_i = (a/d^2)|b_i|\sigma_{i1} \sum_{i'=1}^k |b_{i'}|\sigma_{i'1}$ for each π_i

Fig.4. When the sample is taken up to size $n_i = (a/d^2) \max_{1 \leq j \leq \ell} |b_i|\sigma_{ij} \sum_{i'=1}^k |b_{i'}|\sigma_{i'j}$ for each π_i

Fig.3. When the sample is additionally taken until the total sample for each π_i becomes of size $n_i = (a\delta/d^2)|b_i|\sigma_{i1} \sum_{i'=1}^k |b_{i'}|\sigma_{i'1}$

Table 1. Simulated results

	u	$\bar{n}$	$s(\bar{n})$	$\bar{p}$	$s(\bar{p})$	$\bar{n} - C$	$E(N - C)$	
$k = 2, (\rho_1, \rho_2) = (1/10, 1/5), \tau_{1\star} = 1.920$ $m = 30$ ($d = 0.707$)								
Remark 2 with (24)								
C	90	8.211	95.21	0.175	0.9712	0.00167	5.21	5.12
C_1	42.67		45.18	0.091			2.51	2.43
C_2	47.33		50.03	0.104			2.70	2.68
Remark 2 with (25)								
C	90	8.261	95.75	0.174	0.9720	0.00165	5.75	5.66
C_1	42.67		45.44	0.091			2.77	2.69
C_2	47.33		50.31	0.104			2.98	2.97
$k = 3, (\rho_1, \rho_2, \rho_3) = (1/10, 3/20, 1/5), \tau_{1\star} = 2.881$ $m = 30$ ($d = 0.866$)								
Remark 2 with (24)								
C	135	8.127	141.50	0.211	0.9707	0.00169	6.50	5.87
C_1	42.65		44.67	0.080			2.02	1.87
C_2	45.04		47.24	0.087			2.20	1.96
C_3	47.31		49.60	0.092			2.29	2.04
Remark 2 with (25)								
C	135	8.396	146.06	0.217	0.9746	0.00157	11.06	10.38
C_1	42.65		46.10	0.083			3.45	3.29
C_2	45.04		48.76	0.089			3.72	3.46
C_3	47.31		51.20	0.095			3.88	3.62
$k = 4, (\rho_1, \rho_2, \rho_3, \rho_4) = (1/10, 2/15, 1/6, 1/5), \tau_{1\star} = 3.841$ $m = 30$ ($d = 1.001$)								
Remark 2 with (24)								
C	180	8.085	186.56	0.239	0.9701	0.00170	6.56	6.62
C_1	42.64		44.23	0.074			1.59	1.59
C_2	44.25		45.85	0.078			1.60	1.63
C_3	45.80		47.43	0.081			1.63	1.68
C_4	47.31		49.05	0.085			1.74	1.72
Remark 2 with (25)								
C	180	8.530	196.62	0.256	0.9742	0.00159	16.62	16.60
C_1	42.64		46.61	0.079			3.97	3.95
C_2	44.25		48.32	0.084			4.07	4.09
C_3	45.80		50.00	0.086			4.19	4.22
C_4	47.31		51.70	0.090			4.39	4.35

Table 2. Simulated results

		u	$\bar{n}$	$s(\bar{n})$	$\bar{p}$	$s(\bar{p})$	$\bar{n} - C$	$E(N - C)$
$k = 2, d = 0.458$								
$(\rho_1, \rho_2) = (-0.2, 0.1), (\tau_{1\star}, \tau_{2\star}) = (0.785, 1.357)$								
(20)–(21) with (22): $m = 30$								
C	155	7.948	163.22	0.176	0.9153	0.00278	8.22	3.02
C_1	83.06		84.34	0.131			1.28	1.60
C_2	71.94		78.87	0.149			6.94	1.42
(7)–(9) with (11) having $l = 2$: $m = 51$								
C	155	7.867	160.30	0.138	0.9171	0.00276	5.30	1.75
C_1	83.06		85.92	0.084			2.86	0.91
C_2	71.94		74.38	0.068			2.44	0.84
Aoshima and Takada (2006): $m = 30$								
C	158.04	8.261	174.23	0.182	0.9368	0.00243	16.20	10.07
C_1	83.06		88.46	0.106			5.39	5.41
C_2	74.97		85.78	0.129			10.80	4.67
$k = 4, d = 0.648$								
$(\rho_1, \rho_2, \rho_3, \rho_4) = (-0.2, -0.2, 0.1, 0.1), (\tau_{1\star}, \tau_{2\star}) = (1.570, 3.135)$								
(20)–(21) with (22): $m = 30$								
C	310	7.873	318.27	0.246	0.9091	0.00287	8.27	3.91
C_1	83.06		84.18	0.099			1.12	1.02
C_2	83.06		84.20	0.099			1.13	1.02
C_3	71.94		74.98	0.103			3.04	0.93
C_4	71.94		74.91	0.100			2.98	0.93
(7)–(9) with (11) having $l = 2$: $m = 59$								
C	310	7.801	313.18	0.169	0.9100	0.00286	3.17	1.96
C_1	83.06		83.87	0.061			0.80	0.49
C_2	83.06		83.86	0.061			0.80	0.49
C_3	71.94		72.72	0.051			0.78	0.49
C_4	71.94		72.72	0.051			0.79	0.49
Aoshima and Takada (2006): $m = 30$								
C	316.08	8.530	355.89	0.266	0.9426	0.00233	39.82	29.92
C_1	83.06		90.92	0.093			7.86	8.13
C_2	83.06		90.93	0.093			7.87	8.13
C_3	74.97		87.02	0.112			12.04	6.83
C_4	74.97		87.02	0.111			12.05	6.83

Table 3. Simulated results

	u		$\bar{n}$	$s(\bar{n})$	$\bar{p}$	$s(\bar{p})$	$\bar{n} - C$	$E(N - C)$
$k = 2, d = 0.567$								
$(\rho_1, \rho_2) = (-0.1, 0.2), (\tau_{1\star}, \tau_{2\star}) = (1.231, 1.217)$								
(20)–(21) with (22): $m = 30$								
C	107	8.084	116.41	0.188	0.9275	0.00259	9.41	6.33
C_1	42.60		50.12	0.087			7.52	2.53
C_2	64.40		66.29	0.167			1.88	3.80
(7)–(9) with (11) having $l = 2$: $m = 30$								
C	115.64	8.134	124.24	0.241	0.9429	0.00232	8.60	7.15
C_1	62.41		67.09	0.140			4.67	3.89
C_2	53.23		57.15	0.110			3.93	3.25
Aoshima and Takada (2006): $m = 30$								
C	113.78	8.261	122.98	0.171	0.9428	0.00232	9.20	7.37
C_1	49.38		54.42	0.064			5.05	3.38
C_2	64.40		68.55	0.143			4.15	3.99
$k = 4, d = 0.801$								
$(\rho_1, \rho_2, \rho_3, \rho_4) = (-0.1, -0.1, 0.2, 0.2), (\tau_{1\star}, \tau_{2\star}) = (2.462, 2.853)$								
(20)–(21) with (22): $m = 30$								
C	214	7.983	225.60	0.271	0.9233	0.00266	11.59	8.05
C_1	42.60		47.16	0.076			4.56	1.66
C_2	42.60		47.12	0.077			4.52	1.66
C_3	64.40		65.66	0.137			1.26	2.37
C_4	64.40		65.66	0.137			1.26	2.37
(7)–(9) with (11) having $l = 2$: $m = 35$								
C	231.28	7.909	239.04	0.169	0.942	0.00234	7.76	6.60
C_1	62.41		64.49	0.061			2.08	1.77
C_2	62.41		64.50	0.061			2.10	1.77
C_3	53.23		55.01	0.051			1.78	1.53
C_4	53.23		55.03	0.051			1.80	1.53
Aoshima and Takada (2006): $m = 30$								
C	227.56	8.530	251.82	0.250	0.952	0.00214	24.27	21.43
C_1	49.38		55.50	0.054			6.13	4.93
C_2	49.38		55.51	0.054			6.13	4.93
C_3	64.40		70.43	0.125			6.03	5.79
C_4	64.40		70.38	0.125			5.98	5.79