# A Box-Type Approximation for General Two-Sample Repeated Measures

## - Technical Report -

Edgar Brunner and Marius Placzek

University of Göttingen, Germany

31. August 2011

# 1 Statistical Model and Hypotheses

Throughout this report we will use the following notation. The $d$-dimensional vector of 1s is denoted by $\mathbf{1}_d = (1, \ldots, 1)'$, $\boldsymbol{I}_d$ denotes the $d \times d$ unit matrix, $\boldsymbol{J}_d = \mathbf{1}_d \mathbf{1}_d'$ the $d \times d$ matrix of 1s, and finally, $\boldsymbol{P}_d = \boldsymbol{I}_d - \frac{1}{d}\boldsymbol{J}_d$ denotes the centering matrix. The Kronecker product of matrices is denoted by $\otimes$, the Kronecker sum by $\oplus$, and the Hadamard product by $\#$.

We consider two groups of independent vectors each containing $d$ repeated measures $\boldsymbol{X}_{ik} = (X_{ik1}, \ldots, X_{ikd})' \sim N(\boldsymbol{\mu}_i, \boldsymbol{V}_i)$, $i = 1, 2;\ k = 1, \ldots, n_i$, where $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{id})' \in \mathbb{R}^d$ and $\boldsymbol{V}_i > 0 \in \mathbb{R}^{d \times d}$. The data are collected in the $(n_i d \times 1)$ data vectors $\boldsymbol{X}_i = (\boldsymbol{X}_{i1}', \ldots, \boldsymbol{X}_{in_i}')'$, $i = 1, 2$, and $\boldsymbol{X} = (\boldsymbol{X}_1', \boldsymbol{X}_2')' \in \mathbb{R}^{Nd \times 1}$, where $N = n_1 + n_2$. This can be regarded as a split-plot repeated measures design where the whole-plot factor $A$ has 2 levels and the sub-plot factor $B$ has $d$ levels. Let $\overline{\mu}_{i\cdot} = \frac{1}{d}\sum_{s=1}^{d} \mu_{is}$. Then the hypotheses typically tested in this design can be expressed as

1. $H_0(A) : \overline{\mu}_{1\cdot} = \overline{\mu}_{2\cdot}$ or in matrix notation $\mathbf{1}_d'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$. This refers to the hypothesis that the averaged effects in both groups are equal.

2. $H_0(B) : \mu_{1k} + \mu_{2k} = \overline{\mu}_{1\cdot} + \overline{\mu}_{2\cdot},\ k = 1, \ldots, d$ or in matrix notation $\boldsymbol{P}_d(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \mathbf{0}$. The interpretation of this hypothesis is that the profile averaged over the two groups is flat.

3. $H_0(AB) : \mu_{1k} - \mu_{2k} = \overline{\mu}_{1\cdot} - \overline{\mu}_{2\cdot},\ k = 1, \ldots, d$ or in matrix notation $\boldsymbol{P}_d(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$. This hypothesis means that the two profiles are parallel.

4. $H_0(A|B) : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$. This means that the two profiles are identical. This hypothesis includes the two hypotheses $H_0(A)$ and $H_0(AB)$.

To generalize these hypotheses to structured repeated measures designs, the centering matrix $\boldsymbol{P}_d$ is replaced by a suitable contrast matrix $\boldsymbol{H}$ and to have a unique equivalent formulation of a hypothesis, we use the projection matrix $\boldsymbol{T} = \boldsymbol{H}'(\boldsymbol{H}\boldsymbol{H}')^{-}\boldsymbol{H}$ and note that for any vector $\boldsymbol{\mu}$ the hypotheses $\boldsymbol{H}\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{T}\boldsymbol{\mu} = \mathbf{0}$ are equivalent. Using this notation, it is obvious that the following general hypotheses in the two-groups split-plot design with structured repeated measures can be tested.

1. $\boldsymbol{T}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) = \mathbf{0}$ as a generalization of $H_0(B)$,

2. $\boldsymbol{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$ as a generalization of $H_0(AB)$.

To derive a statistic for testing these hypotheses we consider the transformed observations $\boldsymbol{Y}_{ik} = \boldsymbol{T}\boldsymbol{X}_{ik} \sim N(\boldsymbol{T}\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$, where $\boldsymbol{\Sigma}_i = \boldsymbol{T}\boldsymbol{V}_i\boldsymbol{T} = Cov(\boldsymbol{Y}_{ik})$. The transformed observations are collected in the $(n_i d \times 1)$ data vectors $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}', \ldots, \boldsymbol{Y}_{in_i}')'$, $i = 1, 2$, and in $\boldsymbol{Y} = (\boldsymbol{Y}_1', \boldsymbol{Y}_2')' \in \mathbb{R}^{(Nd \times 1)}$. Then,

$$\boldsymbol{V} \;=\; Cov(\boldsymbol{Y}) \;=\; \bigoplus_{i=1}^{2} (\boldsymbol{I}_{n_i} \otimes \boldsymbol{\Sigma}_i) \tag{1.1}$$

by independence of the vectors $\boldsymbol{Y}_{ik}$. Finally let $\overline{\boldsymbol{Y}}_{i\cdot} = \frac{1}{n_i}\sum_{k=1}^{n_i} \boldsymbol{Y}_{ik}$, $i = 1, 2$, denote the samples means of the transformed observations.

# 2 Equal Covariance Matrices

## 2.1 The Geisser-Greenhouse Procedure and Box's $\epsilon$

A statistic for testing $H_0(\boldsymbol{T}) : \boldsymbol{T}(\boldsymbol{\mu}_1 \pm \boldsymbol{\mu}_2) = \boldsymbol{0}$ shall be based on the quadratic form

$$Q_N = (\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})' (\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot}). \tag{2.2}$$

Note that under $H_0(\boldsymbol{T})$, $E_{H_0}(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot}) = \boldsymbol{T}(\boldsymbol{\mu}_1 \pm \boldsymbol{\mu}_2) = \boldsymbol{0}$. Then under $H_0$, according to the representation theorem of quadratic forms, $Q_N$ can be written as $Q_N = \sum_{i=1}^d \lambda_i C_i^2$, where the random variables $C_i^2 \sim \chi_1^2$ are independent and the constants $\lambda_i$ are the eigenvalues of $\boldsymbol{\Sigma}_N = Cov(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot}) = \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2$. Under the assumption of equal covariance matrices $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$, it follows that $\boldsymbol{\Sigma}_N = \frac{N}{n_1 n_2}\boldsymbol{\Sigma}$.

Box (1954) approximates the distribution of $\sum_{i=1}^d \lambda_i C_i^2$ by a scaled $\chi^2$-distribution, i.e. $\sum_{i=1}^d \lambda_i C_i^2 \dot{\sim} g \cdot \chi_f^2$ such that he first two moments coincide (the symbol $\dot{\sim}$ means *approximately distributed as*). After some simple algebra one obtains $E(g\chi_f^2) = g \cdot f = \sum_{i=1}^d \lambda_i = \operatorname{tr}(\boldsymbol{\Sigma}_N)$ and $Var(g\chi_f^2) = 2g^2 \cdot f = \sum_{i=1}^d \lambda_i^2 = 2\operatorname{tr}(\boldsymbol{\Sigma}_N^2)$. Solving this system of equations one obtains $g \cdot f = \operatorname{tr}(\boldsymbol{\Sigma}_N)$ and $f = [\operatorname{tr}(\boldsymbol{\Sigma}_N)]^2 / \operatorname{tr}(\boldsymbol{\Sigma}_N^2)$. The quantity $f/(d-1)$ is called *Box's* $\epsilon$.

The ratio $Q_N/(g \cdot f) = Q_N / \operatorname{tr}(\boldsymbol{\Sigma}_N)$ follows, approximately, a $\chi_f^2/f$-distribution. Since $\operatorname{tr}(\boldsymbol{\Sigma}_N)$ is unknown in general, it has to be estimated from the data. To this end let

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i - 1}\sum_{k=1}^{n_i}(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})' \tag{2.3}$$

denote the sample covariance matrix. Finally, let $\widetilde{\boldsymbol{\Sigma}}_N = \frac{N}{n_1 n_2}\widehat{\boldsymbol{\Sigma}}_N$ denote the pooled estimator of $\boldsymbol{\Sigma}_N$, where

$$\widehat{\boldsymbol{\Sigma}}_N = \frac{1}{N-2}\Big((n_1 - 1)\widehat{\boldsymbol{\Sigma}}_1 + (n_2 - 1)\widehat{\boldsymbol{\Sigma}}_2\Big) \tag{2.4}$$

is the pooled estimator of $\boldsymbol{\Sigma}$. Note that we need the trace of the pooled estimator $\widehat{\boldsymbol{\Sigma}}_N$. This trace can be written as

$$\begin{aligned}
\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N) &= \frac{1}{N-2}\Big[\operatorname{tr}((n_1 - 1)\widehat{\boldsymbol{\Sigma}}_1) + \operatorname{tr}((n_2 - 1)\widehat{\boldsymbol{\Sigma}}_2)\Big] \\
&= \frac{1}{N-2}\left(\sum_{i=1}^{2}\sum_{k=1}^{n_i}\operatorname{tr}\left[(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})'\right]\right) \\
&= \frac{1}{N-2}\sum_{i=1}^{2}\sum_{k=1}^{n_i}(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})'(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot}) \qquad (2.5)
\end{aligned}$$

using invariance of the trace under cyclic permutations. The quantity in (2.5) is a quadratic form and applying again a two-moment approximation, Geisser and Greenhouse (1958) showed that $\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N) \dot{\sim} \chi^2_{f_0}/f_0$, where $f_0 = (N-2)f$. Using independence of $(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})$ and $\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)$ by verifying Craig-Sakamoto's condition it follows that

$$F_N = \frac{n_1 n_2}{N} \cdot \frac{(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})'(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})}{\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)} \dot{\sim} F(f, (N-2)f), \qquad (2.6)$$

where

$$f = \frac{[\mathrm{tr}(\mathbf{\Sigma}_N)]^2}{\mathrm{tr}(\mathbf{\Sigma}_N^2)} \qquad (2.7)$$

is unknown and must be estimated from the data. A detailed description of the Geisser-Greenhouse procedure and a comparison with the ANOVA-type statistic is to be found in Bathke et a. (2009).

## 2.2 Degrees of Freedom Estimators

A simple plug-in estimator

$$\widehat{f} = \frac{[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)]^2}{\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2)} \qquad (2.8)$$

of $f$ is obtained by replacing $\mathbf{\Sigma}_N$ in (2.7) by $\widetilde{\mathbf{\Sigma}}_N = \frac{N}{n_1 n_2}\widehat{\mathbf{\Sigma}}_N$, where $\widehat{\mathbf{\Sigma}}_N$ is given in (2.4). This estimator, however, is biased and the bias increases with increasing dimension $d$ as seen by simulations. Therefore, Huynh and Feldt (1976) suggested a modification of $\widehat{f}$ by deriving the expectations of $[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)]^2$ and $\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2)$, namely

$$E\left\{\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)\right]^2\right\} = [\mathrm{tr}(\mathbf{\Sigma})]^2 + \frac{2}{N-2}\mathrm{tr}(\mathbf{\Sigma}^2) \qquad (2.9)$$

$$E\left\{\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2)\right\} = \frac{N-1}{N-2}\mathrm{tr}(\mathbf{\Sigma}^2) + \frac{1}{N-2}[\mathrm{tr}(\mathbf{\Sigma})]^2. \qquad (2.10)$$

Solving this system of equations one obtaines the unbiased estimators

$$E\left\{\frac{(N-1)(N-2)}{N(N-3)}\left[\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)\right]^2 - \frac{2}{N-1}\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2)\right]\right\} = [\mathrm{tr}(\mathbf{\Sigma})]^2 \qquad (2.11)$$

$$E\left\{\frac{(N-2)^2}{N(N-3)}\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2) - \frac{1}{N-2}\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)\right]^2\right]\right\} = \mathrm{tr}(\mathbf{\Sigma}^2). \qquad (2.12)$$

Plugging in (2.11) and (2.12) in (2.7), one obtains

$$\widetilde{f} = \frac{(N-1)\left(\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)\right]^2 - \frac{2}{N-1}\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2)\right)}{(N-2)\left(\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N^2) - \frac{1}{N-2}\left[\mathrm{tr}(\widehat{\mathbf{\Sigma}}_N)\right]^2\right)} = \frac{(N-1)\widehat{f} - 2}{N-2-\widehat{f}}, \qquad (2.13)$$

which is the formula for modifying $\widehat{f}$ given by Huynh and Feldt (1976) in the corrected form as given by Lecoutre (1991).

## 2.3 High-Dimensional Data

The validity of the Huynh-Feldt correction in (2.13) for the high-dimensional case is easily seen from the following considerations. First we rearrange the transformed data vectors $\boldsymbol{Y}_{ik}$ in $(d \times n_i)$ data matrices $\widetilde{\boldsymbol{Y}}_i = (\boldsymbol{Y}_{i1}, \ldots, \boldsymbol{Y}_{in_i}) \in \mathbb{R}^{d \times n_i}$, $i = 1, 2$, and $\widetilde{\boldsymbol{Y}} = (\widetilde{\boldsymbol{Y}}_1 \vdots \widetilde{\boldsymbol{Y}}_2) \in \mathbb{R}^{d \times N}$. Then, the pooled sample covariance matrix $\widehat{\boldsymbol{\Sigma}}_N$ in (2.4) can be re-written as

$$\widehat{\boldsymbol{\Sigma}}_N = \frac{1}{N-2} \widetilde{\boldsymbol{Y}} \, \boldsymbol{P} \, \widetilde{\boldsymbol{Y}}' \in \mathbb{R}^{d \times d},$$

where $\boldsymbol{P} = \boldsymbol{P}_{n_1} \oplus \boldsymbol{P}_{n_2}$ and $\boldsymbol{P}_{n_i} = \boldsymbol{I}_{n_i} - \frac{1}{n_i} \boldsymbol{J}_{n_i}$, $i = 1, 2$, are centering matrices. Note that $\boldsymbol{P}$ and $\boldsymbol{P}_{n_i}$ are projection matrices. Then, by using invariance of the trace under cyclic permutations of matrices one obtaines for $\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N)$

$$\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N) = \frac{1}{N-2} \operatorname{tr}(\boldsymbol{P} \widetilde{\boldsymbol{Y}}' \widetilde{\boldsymbol{Y}} \boldsymbol{P}) = \frac{1}{N-2} \operatorname{tr}(\widehat{\boldsymbol{M}}). \tag{2.14}$$

It is easily seen that $\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N^k) = \operatorname{tr}(\widehat{\boldsymbol{M}}^k)$, $k = 1, 2, 3, \ldots$, since the matrices

$$\begin{aligned} (N-2)\widehat{\boldsymbol{\Sigma}}_N &= \widetilde{\boldsymbol{Y}} \, \boldsymbol{P} \, \widetilde{\boldsymbol{Y}}' = (\widetilde{\boldsymbol{Y}} \, \boldsymbol{P})(\widetilde{\boldsymbol{Y}} \, \boldsymbol{P})' \qquad \text{and} \\ \widehat{\boldsymbol{M}} &= \boldsymbol{P} \widetilde{\boldsymbol{Y}}' \widetilde{\boldsymbol{Y}} \boldsymbol{P} = (\widetilde{\boldsymbol{Y}} \boldsymbol{P})'(\widetilde{\boldsymbol{Y}} \boldsymbol{P}) \end{aligned}$$

have the same non-null eigenvalues. We note that $\widehat{\boldsymbol{\Sigma}}_N \in \mathbb{R}^{d \times d}$ and $\widehat{\boldsymbol{M}} \in \mathbb{R}^{N \times N}$ $\forall d \geq 1$. This property is commonly used when solving the dual eigenvalue problem (see, e.g., Kropf et al., 2005; Läuter, 2004; Läuter et al., 2005). Applying the considerations derived above, one obtains for the estimator $\widehat{f}$ of $f$ in (2.8)

$$\widehat{f} = \frac{[\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N)]^2}{\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N^2)} = \frac{[\operatorname{tr}(\widehat{\boldsymbol{M}})]^2}{\operatorname{tr}(\widehat{\boldsymbol{M}}^2)}. \tag{2.15}$$

Note that always $\widehat{\boldsymbol{M}}, \widehat{\boldsymbol{M}}^2 \in \mathbb{R}^{N \times N}$, independent of the dimension $d$. This justifies the use of the Huynh-Feldt correction also in the high-dimensional case.

When the dimension $d$ is large, the use of $\widehat{\boldsymbol{M}}$ instead of $\widehat{\boldsymbol{\Sigma}}_N$ may save memory space and computation time considerably. Moreover, the use of Hadamard's product may further reduce computation time. This is obvious from the following consideration. Let $\widehat{\boldsymbol{M}} = \boldsymbol{P} \widetilde{\boldsymbol{Y}}' \widetilde{\boldsymbol{Y}} \boldsymbol{P}$, then

$$(N-2)^2 \operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_N^2) = \operatorname{tr}(\widehat{\boldsymbol{M}}^2) = \operatorname{tr}(\widehat{\boldsymbol{M}} \widehat{\boldsymbol{M}}) = \boldsymbol{1}_N' (\widehat{\boldsymbol{M}} \# \widehat{\boldsymbol{M}}) \boldsymbol{1}_N,$$

since $\widehat{\boldsymbol{M}}$ is symmetric. In the last step we used the result from matrix algebra that $\operatorname{tr}(\boldsymbol{A}^2) = \boldsymbol{1}_N' (\boldsymbol{A} \# \boldsymbol{A}') \boldsymbol{1}_N$ if $\boldsymbol{A} \in \mathbb{R}^{N \times N}$.

# 3 Unequal Covariance Matrices

## 3.1 The Box Approximation of $Q_N$

If equal covariance matrices $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$ are not assumed, then $\boldsymbol{\Sigma}_N = Cov(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot}) = \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2$ is estimated by $\widetilde{\boldsymbol{\Sigma}}_N = \frac{1}{n_1}\widehat{\boldsymbol{\Sigma}}_1 + \frac{1}{n_2}\widehat{\boldsymbol{\Sigma}}_2$, where $\widehat{\boldsymbol{\Sigma}}_i$ is given in (2.3). Then, the Box-approximation of $Q_N$ in (2.2) is easily derived in the same way as in Section 2.1. One obtains

$$\begin{aligned}
\frac{Q_N}{\operatorname{tr}(\boldsymbol{\Sigma}_N)} &= \frac{(\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})' \, (\overline{\boldsymbol{Y}}_{1\cdot} \pm \overline{\boldsymbol{Y}}_{2\cdot})}{\operatorname{tr}\left(\frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2\right)} \;\dot{\sim}\; \chi^2_f/f, \\
f &= \frac{[\operatorname{tr}(\boldsymbol{\Sigma}_N)]^2}{\operatorname{tr}(\boldsymbol{\Sigma}_N^2)}. \qquad (3.16)
\end{aligned}$$

As $\operatorname{tr}(\boldsymbol{\Sigma}_N)$ is unknown, it must be estimated from the data. The natural estimator

$$\operatorname{tr}(\widetilde{\boldsymbol{\Sigma}}_N) = \frac{1}{n_1}\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_1) + \frac{1}{n_2}\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}_2)$$

can be written as a quadratic form

$$\operatorname{tr}(\widetilde{\boldsymbol{\Sigma}}_N) = \sum_{i=1}^{2} \frac{1}{n_i(n_i-1)} \sum_{k=1}^{n_i} (\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})'(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot}) = \boldsymbol{Y}'\boldsymbol{W}_N\boldsymbol{Y},$$

where $\boldsymbol{Y} = (\boldsymbol{Y}_1', \boldsymbol{Y}_2')'$ and

$$\boldsymbol{W}_N = \bigoplus_{i=1}^{2} \frac{1}{n_i(n_i-1)} \boldsymbol{P}_{n_i} \otimes \boldsymbol{I}_d \, .$$

The covariance matrix of $\boldsymbol{Y}$ is given in (1.1). Now, under $H_0$, the distribution of $\boldsymbol{Y}'\boldsymbol{W}_N\boldsymbol{Y}$ can be approximated by a scaled $\chi^2$-distribution, i.e., $\boldsymbol{Y}'\boldsymbol{W}_N\boldsymbol{Y} \;\dot{\sim}\; g_0\chi^2_{f_0}$. By a two-moment approximation one obtains under $H_0$

$$\begin{aligned}
E(\boldsymbol{Y}'\boldsymbol{W}_N\boldsymbol{Y}) &= \operatorname{tr}(\boldsymbol{W}_N\boldsymbol{V}) = \operatorname{tr}(\boldsymbol{\Sigma}_N) = g_0 f_0 \\
Var(\boldsymbol{Y}'\boldsymbol{W}_N\boldsymbol{Y}) &= 2\operatorname{tr}[(\boldsymbol{W}_N\boldsymbol{V})^2] = 2\sum_{i=1}^{2} \frac{1}{n_i^2(n_i-1)} \operatorname{tr}(\boldsymbol{\Sigma}_i^2) = 2g_0^2 f_0.
\end{aligned}$$

Solving this system of equations one obtains

$$f_0 = \frac{[\operatorname{tr}(\boldsymbol{\Sigma}_N)]^2}{\operatorname{tr}\left(\frac{1}{n_1^2(n_1-1)}\boldsymbol{\Sigma}_1^2 + \frac{1}{n_2^2(n_2-1)}\boldsymbol{\Sigma}_2^2\right)} \, . \qquad (3.17)$$

The independence of $Q_N$ and $\mathrm{tr}(\widetilde{\mathbf{\Sigma}}_N)$ is easily established by verifying Craig-Sakamoto's condition. This leads to the approximation

$$\frac{Q_N}{\mathrm{tr}(\widetilde{\mathbf{\Sigma}}_N)} = \frac{Q_N}{\mathrm{tr}\left(\frac{1}{n_1}\widehat{\mathbf{\Sigma}}_1 + \frac{1}{n_2}\widehat{\mathbf{\Sigma}}_2\right)} \stackrel{.}{\sim} F(f, f_0), \tag{3.18}$$

where $f$ and $f_0$ are given in (3.16) and (3.17), respectively.

In the special case of $d = 1$, the estimator $f_0$ given in (3.17) reduces to

$$\nu = \frac{\left(\sum_{i=1}^{2} \sigma_i^2/n_i\right)^2}{\sum_{i=1}^{2} (\sigma_i^2/n_i)^2/(n_i - 1)},$$

which is the well-known Satterthwaite-Smith-Welch approximation for the $t$ statistic in the univariate Behrens-Fisher problem (Moser and Stevens, 1992; Moser at al., 1989).

Notice that the degree of freedom, $f$, and Box's epsilon, $\epsilon = f/(d-1)$, respectively, are invariant under the choice of sample sizes in both groups and if the different covariance matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are generated by scaling one covariance matrix $\boldsymbol{V}$ using different scaling factors $\sigma_1$ and $\sigma_2$. Actually this has no impact on $f$ and $\epsilon$, which is easily seen from the following considerations. Let $\mathbf{\Sigma}_1 = \sigma_1 \mathbf{\Sigma} = \boldsymbol{T}\sigma_1\boldsymbol{V}\boldsymbol{T}$ and $\mathbf{\Sigma}_2 = \sigma_2 \mathbf{\Sigma} = \boldsymbol{T}\sigma_2\boldsymbol{V}\boldsymbol{T}$. Then the numerator and denominator of $f$ are given by

$$\begin{aligned}
\left[\mathrm{tr}\left(\frac{1}{n_1}\mathbf{\Sigma}_1 + \frac{1}{n_2}\mathbf{\Sigma}_2\right)\right]^2 &= \left[\mathrm{tr}\left(\left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)\mathbf{\Sigma}\right)\right]^2 \\
&= \left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)^2 \left[\mathrm{tr}\left(\mathbf{\Sigma}\right)\right]^2 \\
\mathrm{tr}\left[\left(\frac{1}{n_1}\mathbf{\Sigma}_1 + \frac{1}{n_2}\mathbf{\Sigma}_2\right)^2\right] &= \mathrm{tr}\left[\left(\left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)\mathbf{\Sigma}\right)^2\right] \\
&= \left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)^2 \mathrm{tr}\left(\mathbf{\Sigma}^2\right).
\end{aligned}$$

Thus,

$$f = \frac{\left[\mathrm{tr}\left(\frac{1}{n_1}\mathbf{\Sigma}_1 + \frac{1}{n_2}\mathbf{\Sigma}_2\right)\right]^2}{\mathrm{tr}\left[\left(\frac{1}{n_1}\mathbf{\Sigma}_1 + \frac{1}{n_2}\mathbf{\Sigma}_2\right)^2\right]} = \frac{\left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)^2 \left[\mathrm{tr}\left(\mathbf{\Sigma}\right)\right]^2}{\left(\frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)^2 \mathrm{tr}\left(\mathbf{\Sigma}^2\right)} = \frac{\left[\mathrm{tr}\left(\mathbf{\Sigma}\right)\right]^2}{\mathrm{tr}\left(\mathbf{\Sigma}^2\right)}. \tag{3.19}$$

This shows that $f$ and in turn Box's $\epsilon$ are invariant under scaling.

Now we investigate the range of $\epsilon$ covered by the covariance structures that are mainly considered in this paper. Table 1 shows the theoretical values of $\epsilon$ for three different covariance structures: (a) compound symmetry, (b) autoregressive structure, and (c) a special Toeplitz structure defined below.

The independent vectors $\boldsymbol{X}_{ik}$, $i = 1, 2$; $k = 1, \ldots, n_i$, and $\boldsymbol{Y}_{ik}$ are obtained by the linear transformations $\boldsymbol{X}_{ik} = \sigma_i \boldsymbol{S}^{1/2} \boldsymbol{Z}_{ik} + c_i B_{ik} \mathbf{1}_d$ and $\boldsymbol{Y}_{ik} = \boldsymbol{T}\boldsymbol{X}_{ik}$, where $\boldsymbol{T} = \boldsymbol{P}_d = \boldsymbol{I}_d - \frac{1}{d}\boldsymbol{J}_d$. The quantities $\boldsymbol{S}, \boldsymbol{Z}_{ik}, B_{ik}, \sigma_i$, and $c_i$ are given by

$$\begin{aligned} \boldsymbol{Z}_{ik} &= (Z_{ik1}, \ldots, Z_{ikd})' \sim N(\mathbf{0}, \boldsymbol{I}_d), \ i = 1, 2; \ k = 1, \ldots, n_i \\ \boldsymbol{B}_i &= (B_{i1}, \ldots, B_{in_i})' \sim N(\mathbf{0}, \boldsymbol{I}_{n_i}), \ i = 1, 2 \\ \sigma_i &\in \{1, 1.2, 1.5, 2\}, \quad c_i \in \{0, 1, 2\}, \ i = 1, 2 \\ \boldsymbol{S} = (s_{k\ell})_{k,\ell=1,\ldots,d} &= \begin{cases} \boldsymbol{I}_d \quad , & \text{for a compound symmetry structure} \\ s_{k\ell} = \frac{1}{1-\rho^2}\, \rho^{|k-\ell|} \ , & 0 < \rho < 1, \text{for an autoregressive structure} \\ s_{k\ell} = d - |k - \ell|, & \text{for a (linearly decreasing) Toeplitz structure.} \end{cases} \end{aligned} \tag{3.20}$$

Then, $\boldsymbol{V}_i = Cov(\boldsymbol{X}_{ik}) = \sigma_i^2 \boldsymbol{S} + c_i^2 \boldsymbol{J}_d$ and $\boldsymbol{\Sigma}_i = Cov(\boldsymbol{Y}_{ik}) = \boldsymbol{T}\boldsymbol{V}_i\boldsymbol{T} = \sigma_i^2 \boldsymbol{T}\boldsymbol{S}\boldsymbol{T} = \sigma_i^2 \boldsymbol{\Sigma}$ since $\boldsymbol{T}\boldsymbol{J}_d = \mathbf{0}$.

We note further that the invariance under scaling of $f$ and $\epsilon$ is no longer true in general, if only $\boldsymbol{S}$ is scaled, i.e. $\boldsymbol{S}_i = \sigma_i \boldsymbol{S}$ and $c_1 \neq 0$ or $c_2 \neq 0$. The degrees of Freedom $f$ and Box's $\epsilon$ remain, however, unchanged by the scaling factors $\sigma_i$ and the samples sizes $n_i$ since the covariance matrices $\boldsymbol{\Sigma}_i = \boldsymbol{T}\boldsymbol{S}_i\boldsymbol{T}$ of the transformed observations can be written as $\Sigma_i = \boldsymbol{T}\sigma_i\boldsymbol{S}\boldsymbol{T} = \sigma_i \boldsymbol{\Sigma}$.

As an example, we present the computations for $\epsilon$ as a function of the dimension $d$ for the one-group setting.

TABLE 1 *Different values of Box's $\epsilon$ related to the original covariance matrix $\boldsymbol{V} = \boldsymbol{S} + \boldsymbol{J}_d = \boldsymbol{\Sigma}$ (version $\epsilon$(or)) and those (version $\epsilon$ $(\boldsymbol{P}_d)$) obtained when multiplying the covariance matrix by the centering matrix $\boldsymbol{P}_d$ for different types of covariance matrices: CS = compound symmetry, AR($\rho$) = autoregressive with correlation $\rho$, and TOEP = Toeplitz-structure with linearly decreasing correlations.*

| | CS | | AR(0.2) | | AR(0.6) | | AR(0.9) | | TOEP | |
|---|---|---|---|---|---|---|---|---|---|---|
| $d$ | $\epsilon$(or) | $\epsilon$ $(\boldsymbol{P}_d)$ | $\epsilon$(or) | $\epsilon$ $(\boldsymbol{P}_d)$ | $\epsilon$(or) | $\epsilon$ $(\boldsymbol{P}_d)$ | $\epsilon$(or) | $\epsilon$ $(\boldsymbol{P}_d)$ | $\epsilon$(or) | $\epsilon$ $(\boldsymbol{P}_d)$ |
| 3 | 0.889 | 1 | 0.790 | 0.985 | 0.545 | 0.900 | 0.380 | 0.824 | 0.568 | 0.800 |
| 5 | 0.800 | 1 | 0.700 | 0.961 | 0.436 | 0.749 | 0.251 | 0.577 | 0.370 | 0.526 |
| 10 | 0.640 | 1 | 0.565 | 0.941 | 0.340 | 0.594 | 0.154 | 0.333 | 0.194 | 0.266 |
| 20 | 0.457 | 1 | 0.415 | 0.931 | 0.266 | 0.519 | 0.105 | 0.204 | 0.099 | 0.130 |
| 30 | 0.356 | 1 | 0.329 | 0.928 | 0.225 | 0.499 | 0.088 | 0.163 | 0.066 | 0.086 |
| 50 | 0.246 | 1 | 0.233 | 0.926 | 0.174 | 0.486 | 0.073 | 0.133 | 0.040 | 0.051 |
| 100 | 0.139 | 1 | 0.135 | 0.925 | 0.112 | 0.478 | 0.056 | 0.115 | 0.020 | 0.025 |
| 200 | 0.074 | 1 | 0.073 | 0.924 | 0.066 | 0.474 | 0.041 | 0.109 | 0.010 | 0.013 |
| 300 | 0.051 | 1 | 0.050 | 0.924 | 0.047 | 0.473 | 0.032 | 0.107 | 0.007 | 0.008 |
| 500 | 0.031 | 1 | 0.031 | 0.923 | 0.030 | 0.472 | 0.023 | 0.106 | 0.004 | 0.005 |
| 1000 | 0.016 | 1 | 0.016 | 0.923 | 0.015 | 0.471 | 0.013 | 0.106 | 0.002 | 0.003 |

One can see that, over all dimensions, $d$, these covariance structures cover the whole range of $\epsilon \in [0, 1]$. This shows that $\epsilon$ is not simply a parameter indicating sphericity: Within the same covariance structure, $\epsilon$ can decrease with increasing dimension $d$.

Now we want to investigate wether the goodness of the approximation proposed in this paper does only depend on Box's $\epsilon$. To this end we simulated type I error rates of the nominal level $\alpha = 0.05$ in 10,000 simulation runs for different covariance structures (CS, AR, TOEP). The sample sizes and scaling factors were chosen in a way that both negatively paired ($n_1 = 20$, $n_2 = 10$, $\sigma_1 = 1$, $\sigma_2 = 3$) and positively paired data ($n_1 = 10$, $n_2 = 20$, $\sigma_1 = 1$, $\sigma_2 = 3$) was simulated for each covariance structure. The results are shown in Table 2.

TABLE 2 *Simulated type I error rates of the nominal level $\alpha = 0.05$ in 10,000 simulation runs for different covariance structures: compound symmetry (CS), autoregressive (AR($\rho$)) and Toeplitz (TOEP) structure. Scaling factors are $\sigma_1 = 1, \sigma_2 = 3$ and sample sizes are $n_1 = 10, n_2 = 20$ for positively paired (pp) and $n_1 = 20,\ n_2 = 10$ for negatively paired (np) data.*

| | CS | | AR(0.2) | | AR(0.6) | | AR(0.9) | | TOEP | |
|---|---|---|---|---|---|---|---|---|---|---|
| $d$ | $\epsilon$(or) | $\epsilon$ ($\boldsymbol{P}_d$) | $\epsilon$(or) | $\epsilon$ ($\boldsymbol{P}_d$) | $\epsilon$(or) | $\epsilon$ ($\boldsymbol{P}_d$) | $\epsilon$(or) | $\epsilon$ ($\boldsymbol{P}_d$) | $\epsilon$(or) | $\epsilon$ ($\boldsymbol{P}_d$) |
| 3 | 0.0480 | 0.0471 | 0.0485 | 0.0468 | 0.0477 | 0.0483 | 0.0488 | 0.0499 | 0.0490 | 0.0495 |
| 5 | 0.0475 | 0.0458 | 0.0473 | 0.0462 | 0.0493 | 0.0503 | 0.0520 | 0.0548 | 0.0504 | 0.0549 |
| 10 | 0.0475 | 0.0455 | 0.0501 | 0.0459 | 0.0521 | 0.0540 | 0.0524 | 0.0570 | 0.0526 | 0.0590 |
| 20 | 0.0468 | 0.0433 | 0.0484 | 0.0453 | 0.0507 | 0.0532 | 0.0540 | 0.0585 | 0.0530 | 0.0596 |
| 30 | 0.0462 | 0.0432 | 0.0482 | 0.0448 | 0.0528 | 0.0536 | 0.0541 | 0.0590 | 0.0536 | 0.0610 |
| 50 | 0.0468 | 0.0417 | 0.0478 | 0.0441 | 0.0511 | 0.0525 | 0.0554 | 0.0582 | 0.0540 | 0.0600 |
| 100 | 0.0473 | 0.0411 | 0.0479 | 0.0439 | 0.0498 | 0.0501 | 0.0535 | 0.0563 | 0.0536 | 0.0597 |

Because of the considerations above the approximations for one covariance structure and a dimension $d$ are based on the same $\epsilon$ for both the negatively and the positively paired data. If the quality of the approximation only depends on Box's $\epsilon$ one would expect that the results for negatively and positively paired data are nearly the same. However, this does not hold for any covariance structure considered in our simulations. In fact the following figures show that for a compound symmetry covariance structure the approximations are a bit more conservative in case of negative pairing. This statement still holds for an autoregressive structure ($\rho = 0.2$). For the autoregressive structure ($\rho = 0.6$) the whole issue goes into reverse: Here the approximations tend to be a bit more liberal in case of negatively pairing. The same result can be seen in case of an autoregressive structure with $\rho = 0.9$ and in case of a Toeplitz structure.

These findings are grahpically displayd in Figure1. The single dots in the graphs refer to the simulated type-I errors for the negative pairing of scaling factors and sample sizes ($n_1 = 20$, $n_2 = 10$, $\sigma_1 = 1$, $\sigma_2 = 3$) matched with the positive pairing $n_1 = 10$, $n_2 = 20$, $\sigma_1 = 1$, $\sigma_2 = 3$) for the dimensions $d = 3, 5, 10, 20, 30, 50, 100$. At large, this leads to the conclusion that the goodness of the approximation does not only depend on Box's $\epsilon$. The underlying covariance structure of the data has an impact on it, too.

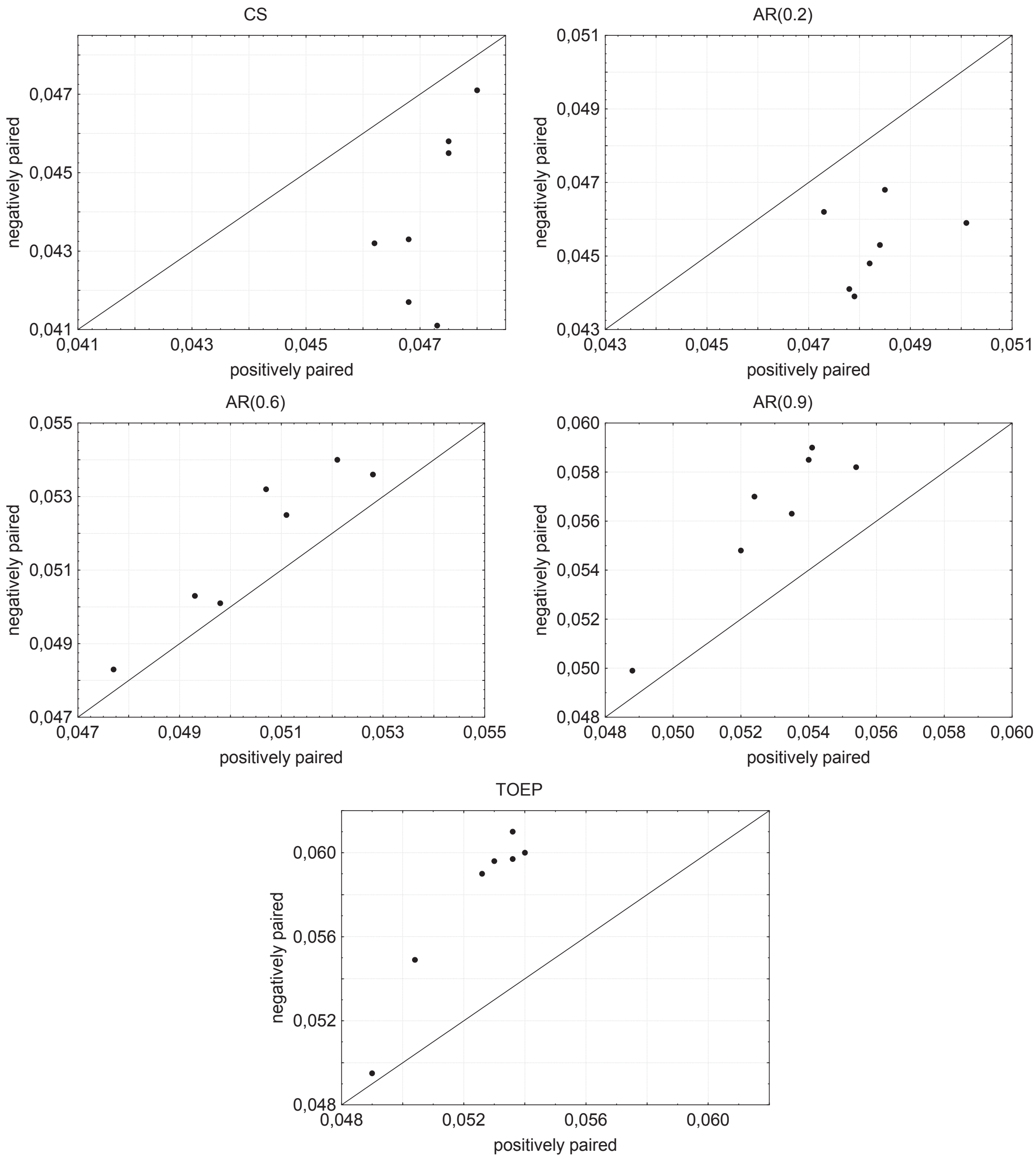


FIGURE 1 *Simulated type-I errors for negatively and positively paired sample sizes and scaling factors in case of different covariance structers but equal values of* $\epsilon$.

## 3.2 Degrees of Freedom Estimators

In the same way as in the case of equal covariance matrices, we will derive unbiased estimators of the numerators and denominators of $f$ and $f_0$. Evaluating these expressions it is easily seen that it suffices to derive unbiased estimators of $[\operatorname{tr}(\boldsymbol{\Sigma}_i)]^2$, $\operatorname{tr}(\boldsymbol{\Sigma}_i^2)$, $\operatorname{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)$, and $\operatorname{tr}(\boldsymbol{\Sigma}_1)\operatorname{tr}(\boldsymbol{\Sigma}_2)$. By independence of $\boldsymbol{Y}_{1k}$ and $\boldsymbol{Y}_{2k'}$, $k = 1, \ldots, n_1$; $k' = 1, \ldots, n_2$, it follows that $E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1\widehat{\boldsymbol{\Sigma}}_2)] = \mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)$ and $E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1)\,\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_2)] = \mathrm{tr}(\boldsymbol{\Sigma}_1)\,\mathrm{tr}(\boldsymbol{\Sigma}_2)$. Thus, it remains to obtain unbiased estimators of $[\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2$ and $\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$, $i = 1, 2$.

To conveniently derive unbiased estimators of the above listed quantities, we return to the vector representation of the transformed observations as considered in Section 1. $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}', \ldots, \boldsymbol{Y}_{in_i}')'$, $i = 1, 2$, $E(\boldsymbol{Y}_i) = \mathbf{1}_{n_i} \otimes \boldsymbol{T}\boldsymbol{\mu}_i$, $\quad Cov(\boldsymbol{Y}_i) = \boldsymbol{I}_{n_i} \otimes \boldsymbol{\Sigma}_i$. Next, $\boldsymbol{Y}_i$ is centered by its mean $\overline{\boldsymbol{Y}}_{i\cdot} = (\frac{1}{n_i}\mathbf{1}_{n_i}' \otimes \boldsymbol{I}_d)\boldsymbol{Y}_i$. The centered vector is denoted by

$$\begin{aligned} \boldsymbol{Z}_i &= \begin{pmatrix} \boldsymbol{Y}_{i1} - \overline{\boldsymbol{Y}}_{i\cdot} \\ \vdots \\ \boldsymbol{Y}_{in_i} - \overline{\boldsymbol{Y}}_{i\cdot} \end{pmatrix} = (\boldsymbol{P}_{n_i} \otimes \boldsymbol{I}_d)\boldsymbol{Y}_i & (3.21) \\ E(\boldsymbol{Z}_i) &= (\boldsymbol{P}_{n_i} \otimes \boldsymbol{I}_d)(\mathbf{1}_{n_i} \otimes \boldsymbol{\mu}_i) = \mathbf{0}, \quad Cov(\boldsymbol{Z}_i) = \boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i. \end{aligned}$$

For the $k$-th and $\ell$-th components $\boldsymbol{Z}_{ik} = \boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot}$ and $\boldsymbol{Z}_{i\ell} = \boldsymbol{Y}_{i\ell} - \overline{\boldsymbol{Y}}_{i\cdot}$ of $\boldsymbol{Z}_i$ it follows that

$$\begin{aligned} Cov(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot}) &= Cov(\boldsymbol{Z}_{ik}) = \frac{n_i - 1}{n_i}\boldsymbol{\Sigma}_i & (3.22) \\ Cov(\boldsymbol{Z}_{ik}, \boldsymbol{Z}_{i\ell}) &= -\frac{1}{n_i}\boldsymbol{\Sigma}_i \quad \text{if } k \neq \ell. & (3.23) \end{aligned}$$

To determine the expectation and variance of $\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)$ we re-write this sample covariance matrix as $\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i - 1}\sum_{k=1}^{n_i} \boldsymbol{S}_{ik}$, where $\boldsymbol{S}_{ik} = \boldsymbol{Z}_{ik}\boldsymbol{Z}_{ik}' = (\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})(\boldsymbol{Y}_{ik} - \overline{\boldsymbol{Y}}_{i\cdot})'$. Let $S_{ik} = \mathrm{tr}(\boldsymbol{S}_{ik}) = \boldsymbol{Z}_{ik}'\boldsymbol{Z}_{ik}$. Then,

$$\begin{aligned} \mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i) &= \frac{1}{n_i - 1}\sum_{k=1}^{n_i} \mathrm{tr}(\boldsymbol{S}_{ik}) = \frac{1}{n_i - 1}\sum_{k=1}^{n_i} \boldsymbol{Z}_{ik}'\boldsymbol{Z}_{ik} \\ &= \frac{1}{n_i - 1}\boldsymbol{Z}_i'\boldsymbol{Z}_i = \frac{1}{n_i - 1}\boldsymbol{Y}_i'(\boldsymbol{P}_{n_i} \otimes \boldsymbol{I}_d)\boldsymbol{Y}_i. & (3.24) \end{aligned}$$

Expectation and variance of $\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)$ are obtained immediately from the above considerations.

$$\begin{aligned} E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)] &= \frac{1}{n_i - 1}\left[\mathrm{tr}(\boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i) + \mathbf{0}'\mathbf{0}\right] = \frac{1}{n_i - 1}(n_i - 1)\,\mathrm{tr}(\boldsymbol{\Sigma}_i) = \mathrm{tr}(\boldsymbol{\Sigma}_i), \\ Var[\widehat{\boldsymbol{\Sigma}}_i] &= \frac{1}{(n_i - 1)^2}\,Var(\boldsymbol{Z}_i'\boldsymbol{Z}_i) = \frac{2}{(n_i - 1)^2}\,\mathrm{tr}(\boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i^2) = \frac{2}{n_i - 1}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2). \end{aligned}$$

The last step follows from the representation theorem of quadratic forms using the normal distribution and $E(\boldsymbol{Z}_i) = \mathbf{0}$. One obtains

$$\boldsymbol{Z}_i'\boldsymbol{Z}_i = \sum_{k=1}^{n_i}\sum_{s=1}^{d} \lambda_{iks}C_{iks}^2,$$

where the random variables $C_{iks}^2 \sim \chi_1^2$ are independent and the constants $\lambda_{iks}$ are the eigenvalues of $Cov(\boldsymbol{Z}_i) = \boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i$. Thus, $Var(\boldsymbol{Z}_i'\boldsymbol{Z}_i) = 2\sum_{k=1}^{n_i}\sum_{s=1}^{d} \lambda_{iks}^2 = 2\,\mathrm{tr}[(\boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i)^2] = 2\,\mathrm{tr}(\boldsymbol{P}_{n_i} \otimes \boldsymbol{\Sigma}_i^2) = 2(n_i-1)\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$. It follows that

$$E\left([\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)]^2\right) = Var[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)] + E^2[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)] = \frac{2}{n_i-1}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + [\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2. \quad (3.25)$$

In the sequel, we will derive the expectation of $\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2)$, which is the plug-in estimator of the denominator of $f$. First, $\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2)$ is represented by means of the matrices $\boldsymbol{S}_{ik} = \boldsymbol{Z}_{ik}\boldsymbol{Z}_{ik}'$.

$$\begin{aligned}
(n_i-1)^2\,\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2) &= \mathrm{tr}\left[\Big(\sum_{k=1}^{n_i}\boldsymbol{S}_{ik}\Big)^2\right] = \mathrm{tr}\left(\sum_{k=1}^{n_i}\sum_{\ell=1}^{n_i}\boldsymbol{S}_{ik}\boldsymbol{S}_{i\ell}\right) \\
&= \mathrm{tr}\left(\sum_{k=1}^{n_i}\sum_{\ell=1}^{n_i}\boldsymbol{Z}_{ik}\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell}\boldsymbol{Z}_{i\ell}'\right) = \sum_{k=1}^{n_i}\sum_{\ell=1}^{n_i}\boldsymbol{Z}_{i\ell}'\boldsymbol{Z}_{ik}\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell} \\
&= \sum_{k=1}^{n_i}\sum_{\ell=1}^{n_i}(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell})^2,
\end{aligned}$$

since $\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell} = \boldsymbol{Z}_{i\ell}'\boldsymbol{Z}_{ik}$. Now consider the expectation

$$E(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell})^2 = Var(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell}) + E^2(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell}).$$

It remains to compute the two quantities $E(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell})$ and $Var(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell})$. To this end, the bilinear form $\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell}$ of the non-independent vectors $\boldsymbol{Z}_{ik}$ and $\boldsymbol{Z}_{i\ell}$ is written as a quardratic form in the combined vector $\boldsymbol{Z}_{ik\ell} = (\boldsymbol{Z}_{ik}', \boldsymbol{Z}_{i\ell}')'$, namely

$$\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell} = (\boldsymbol{Z}_{ik}', \boldsymbol{Z}_{i\ell}')\cdot\left[\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\boldsymbol{I}_d\right]\begin{pmatrix}\boldsymbol{Z}_{ik}\\\boldsymbol{Z}_{i\ell}\end{pmatrix} = \boldsymbol{Z}_{ik\ell}'\left[\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\boldsymbol{I}_d\right]\boldsymbol{Z}_{ik\ell}.$$

The covariance matrix of $\boldsymbol{Z}_{ik\ell}$ is obtained from (3.22) and (3.23)

$$\boldsymbol{G} = Cov(\boldsymbol{Z}_{ik\ell}) = \left(\begin{array}{c|c}\frac{n_i-1}{n_i}\boldsymbol{\Sigma}_i & -\frac{1}{n_i}\boldsymbol{\Sigma}_i\\\hline -\frac{1}{n_i}\boldsymbol{\Sigma}_i & \frac{n_i-1}{n_i}\boldsymbol{\Sigma}_i\end{array}\right) = (\boldsymbol{I}_2 - \tfrac{1}{n_i}\boldsymbol{J}_2)\otimes\boldsymbol{\Sigma}_i.$$

The expectation follows from Lancaster's theorem by noting that $E(\boldsymbol{Z}_{ik}) = \boldsymbol{0}$.

a) if $k \neq \ell$,

$$\begin{aligned}
E(\boldsymbol{Z}_{ik}'\boldsymbol{Z}_{i\ell}) &= E\left(\boldsymbol{Z}_{ik\ell}'\left[\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\boldsymbol{I}_d\right]\boldsymbol{Z}_{ik\ell}\right) = \frac{1}{2}\,\mathrm{tr}\left[\begin{pmatrix}0&1\\1&0\end{pmatrix}\otimes\boldsymbol{I}_d\cdot\boldsymbol{G}\right] \\
&= \frac{1}{2}\,\mathrm{tr}\left(\left[\begin{pmatrix}0&1\\1&0\end{pmatrix} - \tfrac{1}{n_i}\boldsymbol{J}_2\right]\otimes\boldsymbol{\Sigma}_i\right) = -\frac{1}{n_i}\,\mathrm{tr}(\boldsymbol{\Sigma}_i),
\end{aligned}$$

b) if $k = \ell$, $\quad E(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{ik}) = \mathrm{tr}\left(\frac{n_i-1}{n_i}\boldsymbol{\Sigma}_i\right) = \frac{n_i - 1}{n_i}\,\mathrm{tr}(\boldsymbol{\Sigma}_i).$

Note that the normal distribution is assumed and thus one obtains for the variance

a) if $k \neq \ell$,

$$
\begin{aligned}
Var(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{i\ell}) &= Var\left(\boldsymbol{Z}'_{ik\ell}\left[\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \boldsymbol{I}_d\right]\boldsymbol{Z}_{ik\ell}\right) \\
&= 2\,\mathrm{tr}\left(\left[\frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \boldsymbol{I}_d \cdot \boldsymbol{G}\right]^2\right) \\
&= \frac{1}{2}\,\mathrm{tr}\left[\left(\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \tfrac{1}{n_i}\boldsymbol{J}_2\right] \otimes \boldsymbol{\Sigma}_i\right)^2\right] \\
&= \frac{1}{2}\,\mathrm{tr}\left(\boldsymbol{I}_2 - \tfrac{2(n_i-1)}{n_i^2}\boldsymbol{J}_2\right) \cdot \mathrm{tr}(\boldsymbol{\Sigma}_i^2) = \frac{n_i^2 - 2n_i + 2}{n_i^2}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2),
\end{aligned}
$$

b) if $k = \ell$, $\quad Var(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{ik}) = 2\,\mathrm{tr}\left(\frac{(n_i-1)^2}{n_i^2}\boldsymbol{\Sigma}_i^2\right) = \frac{2(n_i - 1)^2}{n_i^2}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2).$

Combining the results derived above it follows that

$$
\begin{aligned}
E\left[(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{i\ell})^2\right] &= Var(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{i\ell}) + E^2(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{i\ell}) \\
&= \begin{cases} \frac{2(n_i-1)^2}{n_i^2}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \frac{(n_i-1)^2}{n_i^2}\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2, & k = \ell \\ \frac{n_i^2-2n_i+2}{n_i^2}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \frac{1}{n_i^2}\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2, & k \neq \ell \end{cases}
\end{aligned}
$$

and finally,

$$
\begin{aligned}
E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2)] &= \frac{1}{(n_i - 1)^2}\sum_{k=1}^{n_i}\sum_{\ell=1}^{n_i} E\left[(\boldsymbol{Z}'_{ik}\boldsymbol{Z}_{i\ell})^2\right] \\
&= \frac{1}{(n_i - 1)^2}\left\{\left[n_i \cdot \frac{2(n_i - 1)^2}{n_i^2} + n_i(n_i - 1) \cdot \frac{n_i^2 - 2n_i + 2}{n_i^2}\right]\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \right. \\
&\qquad \left.\left[n_i \cdot \frac{(n_i - 1)^2}{n_i^2} + n_i(n_i - 1) \cdot \frac{1}{n_i^2}\right]\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2\right\} \\
&= \frac{n_i}{n_i - 1}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \frac{1}{n_i - 1}\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2.
\end{aligned}
$$

Combining this with (3.25) one obtains the system of equations

$$
\begin{aligned}
E\left([\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)]^2\right) &= \frac{2}{n_i - 1}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + [\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2 \\
E[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2)] &= \frac{n_i}{n_i - 1}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \frac{1}{n_i - 1}\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2.
\end{aligned}
$$

Solving this system of equations one obtains unbiased estimators of $[\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2$ and $\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$

$$E\left[\frac{n_i(n_i-1)}{(n_i-2)(n_i+1)}\left\{[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)]^2 - \frac{2}{n_i}\,\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2)\right\}\right] = [\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2 \tag{3.26}$$

$$E\left[\frac{(n_i-1)^2}{(n_i-2)(n_i+1)}\left(\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^2) - \frac{1}{n_i-1}[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i)]^2\right)\right] = \mathrm{tr}(\boldsymbol{\Sigma}_i^2). \tag{3.27}$$

Simulations show that the corrections of $f$ and $f_0$ obtained by (3.26) and (3.27) might be slightly liberal for some covariance structures. This can be turned to a slightly conservativ behavior replacing $(n_i-1)^2$ in (3.27) by $n_i(n_i-1)$.

## 3.3 High-Dimensional Data

As in Section 2.3, we consider the validity of our suggested estimation of Box's $\epsilon$ for high-dimensional data. In particular, we want to avoid handling $d \times d$ matrices if $d > n_i$. All considerations in Section 2.3 can be applied also in the case of unequal covariance matrices.

Let $\boldsymbol{P}_{n_i}$ and $\widetilde{\boldsymbol{Y}}_i$ as defined in Section 2.3. Then,

$$\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i) = \frac{1}{n_i-1}\,\mathrm{tr}(\boldsymbol{P}_{n_i}\widetilde{\boldsymbol{Y}}_i'\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i}) = \frac{1}{n_i-1}\,\mathrm{tr}(\widehat{\boldsymbol{M}}_i).\ i=1,2. \tag{3.28}$$

Then it follows that $\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_i^k) = \mathrm{tr}(\widehat{\boldsymbol{M}}_i^k)$, $k=1,2,3,\ldots$ since the matrices

$$(n_i-1)\widehat{\boldsymbol{\Sigma}}_i = \widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i}\widetilde{\boldsymbol{Y}}_i' = (\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i})(\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i})' \quad \text{and} \tag{3.29}$$

$$\widehat{\boldsymbol{M}}_i = \boldsymbol{P}_{n_i}\widetilde{\boldsymbol{Y}}_i'\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i} = (\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i})'(\widetilde{\boldsymbol{Y}}_i\boldsymbol{P}_{n_i}) \tag{3.30}$$

have the same non-null eigenvalues. We note that $\widehat{\boldsymbol{\Sigma}}_i \in \mathbb{R}^{d\times d}$ and $\widehat{\boldsymbol{M}}_i \in \mathbb{R}^{n_i\times n_i}$, $\forall d \geq 1$. In addition to the considerations in Section 2.3, in the case of unequal covariance matrices it remains to consider handling the mixed product

$$\widehat{\boldsymbol{\Sigma}}_1\widehat{\boldsymbol{\Sigma}}_2 = \frac{1}{(n_1-1)(n_2-1)}\widetilde{\boldsymbol{Y}}_1\boldsymbol{P}_{n_1}\widetilde{\boldsymbol{Y}}_1'\widetilde{\boldsymbol{Y}}_2\boldsymbol{P}_{n_2}\widetilde{\boldsymbol{Y}}_2', \tag{3.31}$$

if $d > n_i$. Let $\widehat{\boldsymbol{M}}_{12} = \boldsymbol{P}_{n_1}\widetilde{\boldsymbol{Y}}_1'\widetilde{\boldsymbol{Y}}_2\boldsymbol{P}_{n_2}$, then it follows by invariance of the trace under cyclic permutations of matrices that

$$(n_1-1)(n_2-1)\,\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1\widehat{\boldsymbol{\Sigma}}_2) = \mathrm{tr}(\widehat{\boldsymbol{M}}_{12}\widehat{\boldsymbol{M}}_{12}') = \mathbf{1}_{n_1}'(\widehat{\boldsymbol{M}}_{12}\#\widehat{\boldsymbol{M}}_{12})\mathbf{1}_{n_2}, \tag{3.32}$$

where we used the result from matrix algebra that $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}') = \mathbf{1}_r'(\boldsymbol{A}\#\boldsymbol{B})\mathbf{1}_s$ if $\boldsymbol{A},\boldsymbol{B} \in \mathbb{R}^{r\times s}$.

# 4 Simulations

## 4.1 Data Generation

Here we report some simulation results which should supplement the simulations presented by Huynh (1978), Algina and Oshima (1994), Keselman et al. (2000), and Keselman et al.(2001) and references cited therein. We will concentrate here on additional simulations involvinng extremely small sample sizes ($n_i = 5, 10$) and / or dimensions $d > 100$.

The observations involving different covariance structures of the repeated measures are generated according to the explanations given in (3.20). We list the simulation results for 100 000 runs for $\alpha = 10\%, 5\%$, and $1\%$ nominal type-I errors and the combinations of sample sizes $(n_1, n_2) = (5, 5), (5, 10), (10, 10)$, and $(10, 20)$ and scaling factors $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$, and $(3, 1)$.

The statistic is given in (3.18) and the degrees of freedom $f$ and $f_0$ are given by

$$
\begin{aligned}
f &= \frac{[\mathrm{tr}(\boldsymbol{\Sigma}_N)]^2}{\mathrm{tr}(\boldsymbol{\Sigma}_N^2)} = \frac{\displaystyle\sum_{i=1}^{2} \frac{1}{n_i^2}[\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2 + \frac{2}{n_1 n_2}\,\mathrm{tr}(\boldsymbol{\Sigma}_1)\,\mathrm{tr}(\boldsymbol{\Sigma}_2)}{\displaystyle\sum_{i=1}^{2} \frac{1}{n_i^2}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2) + \frac{2}{n_1 n_2}\,\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)} \\
f_0 &= \frac{[\mathrm{tr}(\boldsymbol{\Sigma}_N)]^2}{\mathrm{tr}\left(\frac{1}{n_1^2(n_1-1)}\boldsymbol{\Sigma}_1^2 + \frac{1}{n_2^2(n_2-1)}\boldsymbol{\Sigma}_2^2\right)} = \frac{\displaystyle\sum_{i=1}^{2} \frac{1}{n_i^2}[\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2 + \frac{2}{n_1 n_2}\,\mathrm{tr}(\boldsymbol{\Sigma}_1)\,\mathrm{tr}(\boldsymbol{\Sigma}_2)}{\displaystyle\sum_{i=1}^{2} \frac{1}{n_i^2(n_i-1)}\,\mathrm{tr}(\boldsymbol{\Sigma}_i^2)},
\end{aligned}
$$

where an unbiased estimator of $\mathrm{tr}(\boldsymbol{\Sigma}_i)$ is given in (3.24) and in case of $d > n_i$ (3.28) should be used for convenience. An unbiased estimator of $\mathrm{tr}(\boldsymbol{\Sigma}_1)\,\mathrm{tr}(\boldsymbol{\Sigma}_2)$ is the product of the estimators of the traces since the two estimators are independent by assumption. In (3.26) and (3.27), unbiased estimators of $[\mathrm{tr}(\boldsymbol{\Sigma}_i)]^2$ and $\mathrm{tr}(\boldsymbol{\Sigma}_i^2)$ are given while an unbiased estimator of $\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)$ is obtained from (3.31) and in case of $d > n_i$ from (3.32) for convenience.

## 4.2 Simulation Results

TABLE 3 *Simulated type-I error rates in* 100 000 *simulation runs for a compound symmetry structure (CS) and sample sizes* $(n_1, n_2) = (5, 5)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (5, 5)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| CS | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0896 | 0.0425 | 0.0081 | 0.0917 | 0.0454 | 0.0101 |
| 5 | 0.0897 | 0.0420 | 0.0075 | 0.0881 | 0.0432 | 0.0095 |
| 10 | 0.0896 | 0.0428 | 0.0081 | 0.0866 | 0.0416 | 0.0083 |
| 20 | 0.0852 | 0.0408 | 0.0070 | 0.0846 | 0.0406 | 0.0082 |
| 30 | 0.0865 | 0.0406 | 0.0066 | 0.0847 | 0.0403 | 0.0075 |
| 50 | 0.0854 | 0.0396 | 0.0073 | 0.0821 | 0.0393 | 0.0072 |
| 100 | 0.0851 | 0.0390 | 0.0068 | 0.0817 | 0.0382 | 0.0071 |
| 200 | 0.0850 | 0.0399 | 0.0068 | 0.0799 | 0.0379 | 0.0067 |
| 300 | 0.0826 | 0.0390 | 0.0068 | 0.0813 | 0.0371 | 0.0067 |
| 500 | 0.0838 | 0.0391 | 0.0066 | 0.0810 | 0.0371 | 0.0070 |
| 1000 | 0.0839 | 0.0398 | 0.0069 | 0.0799 | 0.0371 | 0.0067 |

TABLE 4 *Simulated type-I error rates in* 100 000 *simulation runs for a compound symmetry structure (CS) and sample sizes* $(n_1, n_2) = (5, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (5, 10)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CS | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0960 | 0.0468 | 0.0097 | 0.0943 | 0.0459 | 0.0080 | 0.0921 | 0.0457 | 0.0113 |
| 5 | 0.0938 | 0.0455 | 0.0096 | 0.0946 | 0.0458 | 0.0092 | 0.0884 | 0.0436 | 0.0096 |
| 10 | 0.0908 | 0.0435 | 0.0088 | 0.0926 | 0.0442 | 0.0084 | 0.0851 | 0.0409 | 0.0079 |
| 20 | 0.0905 | 0.0429 | 0.0088 | 0.0921 | 0.0446 | 0.0080 | 0.0844 | 0.0404 | 0.0078 |
| 30 | 0.0888 | 0.0424 | 0.0081 | 0.0918 | 0.0438 | 0.0085 | 0.0825 | 0.0383 | 0.0076 |
| 50 | 0.0870 | 0.0402 | 0.0078 | 0.0905 | 0.0430 | 0.0079 | 0.0816 | 0.0389 | 0.0075 |
| 100 | 0.0855 | 0.0397 | 0.0073 | 0.0890 | 0.0427 | 0.0077 | 0.0801 | 0.0368 | 0.0066 |
| 200 | 0.0850 | 0.0395 | 0.0073 | 0.0902 | 0.0425 | 0.0078 | 0.0790 | 0.0363 | 0.0066 |
| 300 | 0.0865 | 0.0410 | 0.0072 | 0.0881 | 0.0415 | 0.0075 | 0.0807 | 0.0385 | 0.0073 |
| 500 | 0.0846 | 0.0387 | 0.0067 | 0.0876 | 0.0422 | 0.0076 | 0.0802 | 0.0374 | 0.0064 |
| 1000 | 0.0858 | 0.0401 | 0.0072 | 0.0890 | 0.0427 | 0.0073 | 0.0798 | 0.0367 | 0.0068 |

TABLE 5 *Simulated type-I error rates in* 100 000 *simulation runs for a compound symmetry structure (CS) and sample sizes* $(n_1, n_2) = (10, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (10, 10)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| CS | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0968 | 0.0467 | 0.0085 | 0.0974 | 0.0477 | 0.0091 |
| 5 | 0.0962 | 0.0468 | 0.0088 | 0.0934 | 0.0457 | 0.0087 |
| 10 | 0.0962 | 0.0468 | 0.0084 | 0.0926 | 0.0448 | 0.0082 |
| 20 | 0.0948 | 0.0459 | 0.0082 | 0.0891 | 0.0426 | 0.0081 |
| 30 | 0.0941 | 0.0447 | 0.0084 | 0.0910 | 0.0439 | 0.0080 |
| 50 | 0.0931 | 0.0451 | 0.0085 | 0.0891 | 0.0430 | 0.0077 |
| 100 | 0.0938 | 0.0445 | 0.0087 | 0.0883 | 0.0415 | 0.0073 |
| 200 | 0.0911 | 0.0433 | 0.0079 | 0.0886 | 0.0424 | 0.0079 |
| 300 | 0.0915 | 0.0442 | 0.0086 | 0.0887 | 0.0412 | 0.0073 |
| 500 | 0.0901 | 0.0428 | 0.0079 | 0.0892 | 0.0421 | 0.0079 |
| 1000 | 0.0924 | 0.0444 | 0.0083 | 0.0888 | 0.0417 | 0.0076 |

TABLE 6 *Simulated type-I error rates in* 100 000 *simulation runs for a compound symmetry structure (CS) and sample sizes* $(n_1, n_2) = (10, 20)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (10, 20)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| CS | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0973 | 0.0487 | 0.0094 | 0.0969 | 0.0465 | 0.0088 | 0.0944 | 0.0462 | 0.0087 |
| 5 | 0.0970 | 0.0484 | 0.0094 | 0.0971 | 0.0480 | 0.0094 | 0.0936 | 0.0456 | 0.0090 |
| 10 | 0.0944 | 0.0463 | 0.0089 | 0.0950 | 0.0467 | 0.0088 | 0.0922 | 0.0440 | 0.0087 |
| 20 | 0.0950 | 0.0459 | 0.0083 | 0.0961 | 0.0476 | 0.0090 | 0.0910 | 0.0432 | 0.0079 |
| 30 | 0.0944 | 0.0464 | 0.0087 | 0.0957 | 0.0481 | 0.0088 | 0.0892 | 0.0423 | 0.0079 |
| 50 | 0.0933 | 0.0451 | 0.0084 | 0.0955 | 0.0467 | 0.0088 | 0.0893 | 0.0424 | 0.0076 |
| 100 | 0.0925 | 0.0443 | 0.0080 | 0.0944 | 0.0464 | 0.0087 | 0.0886 | 0.0417 | 0.0074 |
| 200 | 0.0935 | 0.0454 | 0.0084 | 0.0946 | 0.0465 | 0.0082 | 0.0856 | 0.0395 | 0.0072 |
| 300 | 0.0918 | 0.0441 | 0.0081 | 0.0946 | 0.0458 | 0.0085 | 0.0886 | 0.0417 | 0.0074 |
| 500 | 0.0920 | 0.0438 | 0.0080 | 0.0954 | 0.0460 | 0.0087 | 0.0875 | 0.0411 | 0.0071 |
| 1000 | 0.0919 | 0.0436 | 0.0080 | 0.0945 | 0.0464 | 0.0087 | 0.0881 | 0.0417 | 0.0076 |

TABLE 7 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.2) - and sample sizes* $(n_1, n_2) = (5, 5)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (5, 5)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.2) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0908 | 0.0429 | 0.0076 | 0.0938 | 0.0476 | 0.0105 |
| 5 | 0.0900 | 0.0448 | 0.0091 | 0.0921 | 0.0464 | 0.0110 |
| 10 | 0.0883 | 0.0427 | 0.0086 | 0.0900 | 0.0443 | 0.0093 |
| 20 | 0.0895 | 0.0416 | 0.0080 | 0.0876 | 0.0420 | 0.0096 |
| 30 | 0.0908 | 0.0441 | 0.0085 | 0.0871 | 0.0420 | 0.0086 |
| 50 | 0.0893 | 0.0417 | 0.0081 | 0.0835 | 0.0399 | 0.0082 |
| 100 | 0.0888 | 0.0421 | 0.0077 | 0.0839 | 0.0409 | 0.0083 |
| 200 | 0.0866 | 0.0415 | 0.0079 | 0.0837 | 0.0397 | 0.0077 |
| 300 | 0.0865 | 0.0417 | 0.0077 | 0.0839 | 0.0403 | 0.0080 |
| 500 | 0.0864 | 0.0410 | 0.0076 | 0.0838 | 0.0395 | 0.0078 |
| 1000 | 0.0867 | 0.0419 | 0.0076 | 0.0834 | 0.0399 | 0.0074 |

TABLE 8 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.2) - and sample sizes* $(n_1, n_2) = (5, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (5, 10)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.2) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0959 | 0.0485 | 0.0101 | 0.0960 | 0.0474 | 0.0092 | 0.0920 | 0.0468 | 0.0106 |
| 5 | 0.0948 | 0.0474 | 0.0100 | 0.0941 | 0.0458 | 0.0095 | 0.0906 | 0.0450 | 0.0100 |
| 10 | 0.0927 | 0.0456 | 0.0093 | 0.0937 | 0.0456 | 0.0093 | 0.0890 | 0.0442 | 0.0101 |
| 20 | 0.0918 | 0.0456 | 0.0094 | 0.0942 | 0.0454 | 0.0095 | 0.0854 | 0.0420 | 0.0089 |
| 30 | 0.0889 | 0.0431 | 0.0087 | 0.0929 | 0.0451 | 0.0091 | 0.0850 | 0.0411 | 0.0084 |
| 50 | 0.0905 | 0.0441 | 0.0085 | 0.0923 | 0.0452 | 0.0091 | 0.0846 | 0.0408 | 0.0090 |
| 100 | 0.0885 | 0.0432 | 0.0084 | 0.0923 | 0.0447 | 0.0085 | 0.0830 | 0.0398 | 0.0080 |
| 200 | 0.0898 | 0.0435 | 0.0082 | 0.0918 | 0.0442 | 0.0087 | 0.0814 | 0.0387 | 0.0081 |
| 300 | 0.0887 | 0.0431 | 0.0085 | 0.0923 | 0.0459 | 0.0089 | 0.0821 | 0.0397 | 0.0074 |
| 500 | 0.0882 | 0.0417 | 0.0079 | 0.0912 | 0.0445 | 0.0083 | 0.0841 | 0.0399 | 0.0078 |
| 1000 | 0.0862 | 0.0415 | 0.0078 | 0.0916 | 0.0437 | 0.0079 | 0.0821 | 0.0388 | 0.0073 |

TABLE 9 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.2) - and sample sizes* $(n_1, n_2) = (10, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (10, 10)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.2) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0978 | 0.0472 | 0.0091 | 0.0960 | 0.0472 | 0.0096 |
| 5 | 0.0969 | 0.0487 | 0.0097 | 0.0951 | 0.0468 | 0.0092 |
| 10 | 0.0950 | 0.0473 | 0.0095 | 0.0928 | 0.0456 | 0.0091 |
| 20 | 0.0940 | 0.0468 | 0.0097 | 0.0941 | 0.0469 | 0.0089 |
| 30 | 0.0929 | 0.0462 | 0.0093 | 0.0912 | 0.0449 | 0.0092 |
| 50 | 0.0937 | 0.0467 | 0.0095 | 0.0909 | 0.0439 | 0.0085 |
| 100 | 0.0953 | 0.0469 | 0.0089 | 0.0925 | 0.0447 | 0.0089 |
| 200 | 0.0934 | 0.0461 | 0.0091 | 0.0904 | 0.0442 | 0.0084 |
| 300 | 0.0938 | 0.0455 | 0.0089 | 0.0917 | 0.0441 | 0.0082 |
| 500 | 0.0946 | 0.0469 | 0.0092 | 0.0899 | 0.0425 | 0.0078 |
| 1000 | 0.0942 | 0.0469 | 0.0089 | 0.0918 | 0.0446 | 0.0084 |

TABLE 10 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.2) - and sample sizes* $(n_1, n_2) = (10, 20)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (10, 20)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.2) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0968 | 0.0472 | 0.0095 | 0.0987 | 0.0487 | 0.0093 | 0.0939 | 0.0458 | 0.0087 |
| 5 | 0.0955 | 0.0478 | 0.0097 | 0.0976 | 0.0483 | 0.0098 | 0.0947 | 0.0463 | 0.0093 |
| 10 | 0.0969 | 0.0480 | 0.0099 | 0.0962 | 0.0485 | 0.0102 | 0.0937 | 0.0452 | 0.0094 |
| 20 | 0.0970 | 0.0484 | 0.0100 | 0.0957 | 0.0484 | 0.0102 | 0.0921 | 0.0459 | 0.0095 |
| 30 | 0.0930 | 0.0462 | 0.0095 | 0.0962 | 0.0483 | 0.0097 | 0.0917 | 0.0448 | 0.0091 |
| 50 | 0.0951 | 0.0467 | 0.0095 | 0.0966 | 0.0490 | 0.0100 | 0.0908 | 0.0444 | 0.0092 |
| 100 | 0.0944 | 0.0466 | 0.0097 | 0.0965 | 0.0480 | 0.0095 | 0.0894 | 0.0433 | 0.0084 |
| 200 | 0.0942 | 0.0462 | 0.0090 | 0.0963 | 0.0478 | 0.0093 | 0.0901 | 0.0440 | 0.0090 |
| 300 | 0.0949 | 0.0468 | 0.0094 | 0.0980 | 0.0486 | 0.0096 | 0.0909 | 0.0445 | 0.0086 |
| 500 | 0.0943 | 0.0464 | 0.0088 | 0.0968 | 0.0479 | 0.0095 | 0.0886 | 0.0426 | 0.0083 |
| 1000 | 0.0932 | 0.0456 | 0.0090 | 0.0973 | 0.0479 | 0.0094 | 0.0906 | 0.0436 | 0.0085 |

TABLE 11 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.6) - and sample sizes* $(n_1, n_2) = (5, 5)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (5, 5)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.6) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0908 | 0.0439 | 0.0094 | 0.0940 | 0.0487 | 0.0118 |
| 5 | 0.0931 | 0.0483 | 0.0128 | 0.0975 | 0.0529 | 0.0149 |
| 10 | 0.0957 | 0.0523 | 0.0146 | 0.0993 | 0.0555 | 0.0170 |
| 20 | 0.0973 | 0.0533 | 0.0158 | 0.1018 | 0.0582 | 0.0185 |
| 30 | 0.0945 | 0.0503 | 0.0141 | 0.1003 | 0.0561 | 0.0167 |
| 50 | 0.0949 | 0.0501 | 0.0130 | 0.0979 | 0.0543 | 0.0167 |
| 100 | 0.0949 | 0.0495 | 0.0119 | 0.0962 | 0.0529 | 0.0151 |
| 200 | 0.0967 | 0.0492 | 0.0116 | 0.0945 | 0.0511 | 0.0137 |
| 300 | 0.0920 | 0.0468 | 0.0111 | 0.0934 | 0.0488 | 0.0131 |
| 500 | 0.0926 | 0.0460 | 0.0101 | 0.0949 | 0.0497 | 0.0136 |
| 1000 | 0.0925 | 0.0468 | 0.0106 | 0.0935 | 0.0482 | 0.0124 |

TABLE 12 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.6) - and sample sizes* $(n_1, n_2) = (5, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (5, 10)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.6) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0959 | 0.0492 | 0.0112 | 0.0952 | 0.0474 | 0.0103 | 0.0952 | 0.0493 | 0.0126 |
| 5 | 0.0978 | 0.0515 | 0.0139 | 0.0954 | 0.0507 | 0.0131 | 0.0968 | 0.0528 | 0.0157 |
| 10 | 0.0986 | 0.0545 | 0.0161 | 0.0965 | 0.0523 | 0.0141 | 0.1020 | 0.0572 | 0.0186 |
| 20 | 0.0978 | 0.0534 | 0.0156 | 0.0984 | 0.0532 | 0.0149 | 0.1012 | 0.0584 | 0.0182 |
| 30 | 0.0964 | 0.0519 | 0.0150 | 0.0980 | 0.0529 | 0.0135 | 0.1021 | 0.0584 | 0.0187 |
| 50 | 0.0985 | 0.0525 | 0.0146 | 0.0966 | 0.0513 | 0.0132 | 0.0999 | 0.0567 | 0.0174 |
| 100 | 0.0971 | 0.0508 | 0.0126 | 0.0961 | 0.0495 | 0.0120 | 0.0965 | 0.0530 | 0.0154 |
| 200 | 0.0936 | 0.0483 | 0.0118 | 0.0944 | 0.0487 | 0.0110 | 0.0947 | 0.0514 | 0.0143 |
| 300 | 0.0949 | 0.0484 | 0.0119 | 0.0941 | 0.0480 | 0.0112 | 0.0937 | 0.0500 | 0.0139 |
| 500 | 0.0943 | 0.0482 | 0.0110 | 0.0938 | 0.0471 | 0.0100 | 0.0948 | 0.0509 | 0.0145 |
| 1000 | 0.0924 | 0.0466 | 0.0100 | 0.0941 | 0.0472 | 0.0103 | 0.0931 | 0.0495 | 0.0131 |

TABLE 13 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.6) - and sample sizes* $(n_1, n_2) = (10, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (10, 10)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.6) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0966 | 0.0492 | 0.0103 | 0.0954 | 0.0470 | 0.0105 |
| 5 | 0.0951 | 0.0501 | 0.0121 | 0.0953 | 0.0499 | 0.0130 |
| 10 | 0.0968 | 0.0520 | 0.0141 | 0.0974 | 0.0527 | 0.0149 |
| 20 | 0.0983 | 0.0533 | 0.0145 | 0.0991 | 0.0555 | 0.0156 |
| 30 | 0.0978 | 0.0531 | 0.0138 | 0.0992 | 0.0536 | 0.0143 |
| 50 | 0.0976 | 0.0519 | 0.0132 | 0.0965 | 0.0513 | 0.0128 |
| 100 | 0.0975 | 0.0508 | 0.0121 | 0.0960 | 0.0489 | 0.0118 |
| 200 | 0.0987 | 0.0501 | 0.0108 | 0.0959 | 0.0487 | 0.0110 |
| 300 | 0.0963 | 0.0489 | 0.0105 | 0.0955 | 0.0488 | 0.0116 |
| 500 | 0.0959 | 0.0477 | 0.0097 | 0.0951 | 0.0474 | 0.0102 |
| 1000 | 0.0954 | 0.0477 | 0.0092 | 0.0944 | 0.0469 | 0.0099 |

TABLE 14 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.6) - and sample sizes* $(n_1, n_2) = (10, 20)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (10, 20)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.6) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0966 | 0.0497 | 0.0110 | 0.0969 | 0.0488 | 0.0110 | 0.0966 | 0.0499 | 0.0109 |
| 5 | 0.0952 | 0.0508 | 0.0133 | 0.0966 | 0.0510 | 0.0130 | 0.0971 | 0.0516 | 0.0128 |
| 10 | 0.0957 | 0.0521 | 0.0145 | 0.0979 | 0.0528 | 0.0148 | 0.0990 | 0.0543 | 0.0152 |
| 20 | 0.0969 | 0.0519 | 0.0136 | 0.0972 | 0.0532 | 0.0142 | 0.0978 | 0.0531 | 0.0147 |
| 30 | 0.0977 | 0.0529 | 0.0141 | 0.1004 | 0.0534 | 0.0137 | 0.0984 | 0.0533 | 0.0140 |
| 50 | 0.0986 | 0.0520 | 0.0130 | 0.0972 | 0.0521 | 0.0127 | 0.0981 | 0.0519 | 0.0131 |
| 100 | 0.0972 | 0.0505 | 0.0118 | 0.0974 | 0.0500 | 0.0114 | 0.0972 | 0.0505 | 0.0124 |
| 200 | 0.0959 | 0.0481 | 0.0110 | 0.0998 | 0.0505 | 0.0115 | 0.0965 | 0.0491 | 0.0114 |
| 300 | 0.0967 | 0.0495 | 0.0107 | 0.0985 | 0.0503 | 0.0114 | 0.0953 | 0.0485 | 0.0110 |
| 500 | 0.0971 | 0.0487 | 0.0102 | 0.0977 | 0.0484 | 0.0103 | 0.0941 | 0.0478 | 0.0105 |
| 1000 | 0.0966 | 0.0487 | 0.0104 | 0.0979 | 0.0491 | 0.0105 | 0.0939 | 0.0466 | 0.0101 |

TABLE 15 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.9) - and sample sizes* $(n_1, n_2) = (5, 5)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (5, 5)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.9) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0920 | 0.0463 | 0.0106 | 0.0951 | 0.0503 | 0.0132 |
| 5 | 0.0952 | 0.0513 | 0.0154 | 0.1019 | 0.0583 | 0.0196 |
| 10 | 0.0991 | 0.0574 | 0.0194 | 0.1054 | 0.0627 | 0.0233 |
| 20 | 0.1025 | 0.0603 | 0.0202 | 0.1085 | 0.0664 | 0.0248 |
| 30 | 0.1001 | 0.0588 | 0.0203 | 0.1088 | 0.0662 | 0.0245 |
| 50 | 0.0992 | 0.0576 | 0.0196 | 0.1080 | 0.0651 | 0.0243 |
| 100 | 0.0973 | 0.0552 | 0.0173 | 0.1048 | 0.0619 | 0.0220 |
| 200 | 0.0982 | 0.0534 | 0.0146 | 0.1008 | 0.0575 | 0.0184 |
| 300 | 0.0957 | 0.0512 | 0.0138 | 0.1006 | 0.0560 | 0.0178 |
| 500 | 0.0965 | 0.0510 | 0.0129 | 0.1000 | 0.0552 | 0.0167 |
| 1000 | 0.0937 | 0.0489 | 0.0118 | 0.0971 | 0.0524 | 0.0152 |

TABLE 16 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.9) - and sample sizes* $(n_1, n_2) = (5, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (5, 10)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.9) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0956 | 0.0504 | 0.0125 | 0.0952 | 0.0484 | 0.0110 | 0.0976 | 0.0509 | 0.0131 |
| 5 | 0.0990 | 0.0555 | 0.0169 | 0.0946 | 0.0514 | 0.0142 | 0.1031 | 0.0605 | 0.0207 |
| 10 | 0.1017 | 0.0593 | 0.0208 | 0.0960 | 0.0543 | 0.0174 | 0.1104 | 0.0681 | 0.0263 |
| 20 | 0.1024 | 0.0606 | 0.0206 | 0.0997 | 0.0578 | 0.0187 | 0.1098 | 0.0682 | 0.0268 |
| 30 | 0.1022 | 0.0596 | 0.0208 | 0.0986 | 0.0572 | 0.0188 | 0.1102 | 0.0678 | 0.0267 |
| 50 | 0.1020 | 0.0593 | 0.0193 | 0.0988 | 0.0565 | 0.0177 | 0.1082 | 0.0653 | 0.0248 |
| 100 | 0.1005 | 0.0569 | 0.0177 | 0.0989 | 0.0545 | 0.0156 | 0.1044 | 0.0623 | 0.0220 |
| 200 | 0.0995 | 0.0541 | 0.0157 | 0.0987 | 0.0536 | 0.0142 | 0.1027 | 0.0594 | 0.0202 |
| 300 | 0.0973 | 0.0530 | 0.0149 | 0.0975 | 0.0519 | 0.0134 | 0.1021 | 0.0583 | 0.0191 |
| 500 | 0.0969 | 0.0514 | 0.0133 | 0.0970 | 0.0498 | 0.0124 | 0.0996 | 0.0556 | 0.0174 |
| 1000 | 0.0962 | 0.0498 | 0.0128 | 0.0966 | 0.0499 | 0.0125 | 0.0962 | 0.0533 | 0.0163 |

TABLE 17 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.9) - and sample sizes* $(n_1, n_2) = (10, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (10, 10)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| AR(0.9) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0953 | 0.0480 | 0.0110 | 0.0970 | 0.0507 | 0.0122 |
| 5 | 0.0952 | 0.0519 | 0.0140 | 0.0972 | 0.0538 | 0.0158 |
| 10 | 0.0946 | 0.0527 | 0.0158 | 0.0986 | 0.0574 | 0.0189 |
| 20 | 0.0961 | 0.0553 | 0.0172 | 0.0986 | 0.0579 | 0.0189 |
| 30 | 0.0955 | 0.0538 | 0.0166 | 0.1002 | 0.0578 | 0.0185 |
| 50 | 0.0990 | 0.0553 | 0.0174 | 0.0995 | 0.0572 | 0.0184 |
| 100 | 0.0964 | 0.0535 | 0.0149 | 0.1005 | 0.0555 | 0.0167 |
| 200 | 0.0986 | 0.0519 | 0.0137 | 0.0987 | 0.0543 | 0.0148 |
| 300 | 0.0993 | 0.0529 | 0.0132 | 0.0970 | 0.0517 | 0.0135 |
| 500 | 0.0973 | 0.0511 | 0.0124 | 0.0986 | 0.0517 | 0.0129 |
| 1000 | 0.0975 | 0.0500 | 0.0118 | 0.0964 | 0.0503 | 0.0124 |

TABLE 18 *Simulated type-I error rates in* 100 000 *simulation runs for an autoregressive structure - AR(0.9) - and sample sizes* $(n_1, n_2) = (10, 20)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (10, 20)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AR(0.9) | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0953 | 0.0492 | 0.0112 | 0.0953 | 0.0491 | 0.0109 | 0.0957 | 0.0488 | 0.0123 |
| 5 | 0.0964 | 0.0529 | 0.0150 | 0.0941 | 0.0513 | 0.0137 | 0.0969 | 0.0533 | 0.0158 |
| 10 | 0.0953 | 0.0537 | 0.0160 | 0.0935 | 0.0523 | 0.0157 | 0.1016 | 0.0587 | 0.0190 |
| 20 | 0.0969 | 0.0547 | 0.0173 | 0.0955 | 0.0537 | 0.0167 | 0.1004 | 0.0588 | 0.0197 |
| 30 | 0.0970 | 0.0553 | 0.0174 | 0.0961 | 0.0545 | 0.0171 | 0.1012 | 0.0589 | 0.0197 |
| 50 | 0.0977 | 0.0560 | 0.0162 | 0.0981 | 0.0545 | 0.0164 | 0.0999 | 0.0573 | 0.0192 |
| 100 | 0.0992 | 0.0550 | 0.0151 | 0.0986 | 0.0543 | 0.0153 | 0.1002 | 0.0567 | 0.0172 |
| 200 | 0.0996 | 0.0536 | 0.0149 | 0.0981 | 0.0531 | 0.0142 | 0.0993 | 0.0545 | 0.0154 |
| 300 | 0.0987 | 0.0527 | 0.0136 | 0.1000 | 0.0532 | 0.0128 | 0.0994 | 0.0529 | 0.0146 |
| 500 | 0.0990 | 0.0524 | 0.0126 | 0.0997 | 0.0528 | 0.0135 | 0.0977 | 0.0515 | 0.0127 |
| 1000 | 0.0981 | 0.0512 | 0.0114 | 0.0992 | 0.0506 | 0.0115 | 0.0965 | 0.0507 | 0.0120 |

TABLE 19 *Simulated type-I error rates in* 100 000 *simulation runs for a Toeplitz structure with linearly decraesing covariances and sample sizes* $(n_1, n_2) = (5, 5)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (5, 5)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| TOEP | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0941 | 0.0483 | 0.0111 | 0.0972 | 0.0515 | 0.0132 |
| 5 | 0.0977 | 0.0550 | 0.0158 | 0.1033 | 0.0607 | 0.0207 |
| 10 | 0.0995 | 0.0583 | 0.0205 | 0.1085 | 0.0666 | 0.0262 |
| 20 | 0.1000 | 0.0601 | 0.0220 | 0.1096 | 0.0688 | 0.0269 |
| 30 | 0.0999 | 0.0595 | 0.0213 | 0.1105 | 0.0698 | 0.0282 |
| 50 | 0.0999 | 0.0599 | 0.0216 | 0.1099 | 0.0692 | 0.0293 |
| 100 | 0.1013 | 0.0606 | 0.0225 | 0.1110 | 0.0698 | 0.0285 |
| 200 | 0.1012 | 0.0601 | 0.0217 | 0.1108 | 0.0698 | 0.0291 |
| 300 | 0.1031 | 0.0623 | 0.0225 | 0.1112 | 0.0711 | 0.0284 |
| 500 | 0.1009 | 0.0606 | 0.0223 | 0.1106 | 0.0702 | 0.0285 |
| 1000 | 0.1005 | 0.0602 | 0.0217 | 0.1121 | 0.0709 | 0.0291 |

TABLE 20 *Simulated type-I error rates in* 100 000 *simulation runs for a Toeplitz structure with linearly decraesing covariances and sample sizes* $(n_1, n_2) = (5, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (5, 10)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| TOEP | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0957 | 0.0501 | 0.0126 | 0.0946 | 0.0482 | 0.0112 | 0.0975 | 0.0522 | 0.0144 |
| 5 | 0.0990 | 0.0558 | 0.0173 | 0.0965 | 0.0534 | 0.0156 | 0.1041 | 0.0610 | 0.0204 |
| 10 | 0.0994 | 0.0593 | 0.0217 | 0.0989 | 0.0576 | 0.0189 | 0.1099 | 0.0686 | 0.0284 |
| 20 | 0.1022 | 0.0612 | 0.0223 | 0.0973 | 0.0559 | 0.0189 | 0.1126 | 0.0708 | 0.0297 |
| 30 | 0.1035 | 0.0621 | 0.0226 | 0.1004 | 0.0583 | 0.0195 | 0.1144 | 0.0734 | 0.0310 |
| 50 | 0.1014 | 0.0615 | 0.0225 | 0.0994 | 0.0575 | 0.0193 | 0.1111 | 0.0713 | 0.0303 |
| 100 | 0.1037 | 0.0630 | 0.0231 | 0.0976 | 0.0579 | 0.0199 | 0.1128 | 0.0719 | 0.0311 |
| 200 | 0.1037 | 0.0622 | 0.0225 | 0.0979 | 0.0579 | 0.0197 | 0.1119 | 0.0719 | 0.0303 |
| 300 | 0.1020 | 0.0603 | 0.0220 | 0.0984 | 0.0568 | 0.0192 | 0.1143 | 0.0732 | 0.0318 |
| 500 | 0.1023 | 0.0605 | 0.0216 | 0.0975 | 0.0570 | 0.0190 | 0.1141 | 0.0732 | 0.0312 |
| 1000 | 0.1036 | 0.0625 | 0.0233 | 0.0998 | 0.0584 | 0.0198 | 0.1146 | 0.0724 | 0.0314 |

TABLE 21 *Simulated type-I error rates in* 100 000 *simulation runs for a Toeplitz structure with linearly decraesing covariances and sample sizes* $(n_1, n_2) = (10, 10)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3)$.

| $(n_1, n_2) = (10, 10)$ / nsim = 100 000 | | | | | | |
|---|---|---|---|---|---|---|
| TOEP | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | |
| | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0939 | 0.0479 | 0.0113 | 0.0972 | 0.0501 | 0.0125 |
| 5 | 0.0947 | 0.0508 | 0.0139 | 0.0976 | 0.0549 | 0.0164 |
| 10 | 0.0939 | 0.0536 | 0.0172 | 0.0979 | 0.0574 | 0.0190 |
| 20 | 0.0942 | 0.0537 | 0.0169 | 0.0992 | 0.0579 | 0.0207 |
| 30 | 0.0948 | 0.0544 | 0.0169 | 0.0999 | 0.0593 | 0.0207 |
| 50 | 0.0941 | 0.0535 | 0.0171 | 0.0990 | 0.0583 | 0.0203 |
| 100 | 0.0964 | 0.0553 | 0.0175 | 0.1009 | 0.0601 | 0.0204 |
| 200 | 0.0951 | 0.0547 | 0.0176 | 0.0991 | 0.0583 | 0.0201 |
| 300 | 0.0947 | 0.0544 | 0.0163 | 0.0990 | 0.0585 | 0.0200 |
| 500 | 0.0965 | 0.0545 | 0.0171 | 0.1004 | 0.0588 | 0.0208 |
| 1000 | 0.0966 | 0.0556 | 0.0171 | 0.1004 | 0.0582 | 0.0206 |

TABLE 22 *Simulated type-I error rates in* 100 000 *simulation runs for a Toeplitz structure with linearly decraesing covariances and sample sizes* $(n_1, n_2) = (10, 20)$ *for* $(\sigma_1, \sigma_2) = (1, 1), (1, 3), (3, 1)$.

| $(n_1, n_2) = (10, 20)$ / nsim = 100 000 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| TOEP | $(\sigma_1, \sigma_2) = (1, 1)$ | | | $(\sigma_1, \sigma_2) = (1, 3)$ | | | $(\sigma_1, \sigma_2) = (3, 1)$ | | |
| | Nominal level | | | Nominal level | | | Nominal level | | |
| $d$ | 10% | 5% | 1% | 10% | 5% | 1% | 10% | 5% | 1% |
| 3 | 0.0964 | 0.0505 | 0.0117 | 0.0955 | 0.0481 | 0.0112 | 0.0957 | 0.0507 | 0.0124 |
| 5 | 0.0947 | 0.0520 | 0.0146 | 0.0944 | 0.0512 | 0.0139 | 0.0977 | 0.0543 | 0.0170 |
| 10 | 0.0964 | 0.0533 | 0.0164 | 0.0932 | 0.0530 | 0.0154 | 0.1002 | 0.0589 | 0.0207 |
| 20 | 0.0966 | 0.0555 | 0.0174 | 0.0951 | 0.0539 | 0.0166 | 0.1016 | 0.0603 | 0.0212 |
| 30 | 0.0970 | 0.0561 | 0.0179 | 0.0915 | 0.0527 | 0.0160 | 0.1024 | 0.0604 | 0.0214 |
| 50 | 0.0972 | 0.0559 | 0.0181 | 0.0952 | 0.0536 | 0.0161 | 0.1013 | 0.0597 | 0.0213 |
| 100 | 0.0961 | 0.0561 | 0.0181 | 0.0943 | 0.0538 | 0.0160 | 0.0999 | 0.0603 | 0.0219 |
| 200 | 0.0961 | 0.0548 | 0.0176 | 0.0945 | 0.0539 | 0.0162 | 0.1021 | 0.0604 | 0.0210 |
| 300 | 0.0954 | 0.0550 | 0.0175 | 0.0944 | 0.0542 | 0.0165 | 0.1011 | 0.0604 | 0.0212 |
| 500 | 0.0967 | 0.0562 | 0.0179 | 0.0931 | 0.0528 | 0.0165 | 0.0982 | 0.0580 | 0.0205 |
| 1000 | 0.0975 | 0.0562 | 0.0182 | 0.0946 | 0.0531 | 0.0160 | 0.1003 | 0.0591 | 0.0212 |

# Literatur


[1] Bathke, A. C., Schabenberger, O., Tobias, R. D., and Madden, L. V. (2009). Greenhouse-Geisser Adjustment and the ANOVA-Type Statistic: Cousins or Twins?. *The American Statistician* **63**, 239–246.

[2] Box, G. E. P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification. *Annals of Mathematical Statistics* **25**, 484–498.

[3] Geisser, S., and Greenhouse, S. W. (1958). An Extension of Box's Result on the Use of the F Distribution in Multivariate Analysis. *Annals of Mathematical Statistics* **29**, 885–891.

[4] Greenhouse, S. W., and Geisser, S. (1959). On Methods in the Analysis of Profile Data. *Psychometrika* **24**, 2, 95–112.

[5] Huynh, H., and Feldt, L. S. (1976). Estimation of the Box Correction for Degrees of Freedom From Sample Data in Randomized Block and Split-Plot Designs. *Journal of Educational Statistics* **1**, 69–82.

[6] Huynh, H. (1978). Some approximate tests for repeated measures designs. *Psychometrika* **43**, 161–175.

[7] Moser, B. K., Stevens, G. R., and Watts, C. L. (1989), The Two- Sample $t$ Test Versus Satterthwaite's Approximate $F$ Test. *Communications in Statistics-Theory and Methods* **18**, 3963 - 3975.

[8] Moser, B. K., and Stevens, G. R. (1992). Homogeneity of Variance in the Two-Sample Means Test. *The American Statistician* **46**, 19–21.

[9] Lecoutre, B. (1991). A Correction for the $\widetilde{\epsilon}$ Approximative Test in Repeated Measures Designs With Two or More Independent Groups. *Journal of Educational Statistics* **16**, 371–372.

[10] Kropf, S., Läuter, J.Kose, D., von Rosen, D. (2009). Comparison of exact parametric tests for high-dimensional data. *Computational Staistics and Data Analysis* **53**, 776–787.

[11] Läuter, J. (2004). Two new multivariate tests, in particular for a high dimension. *Acta et Commentationis Universitatis Tartuensis de Mathematica* **8**, 179–186.

[12] Läuter, J., Glimm, E., and Eszlinger, M. (2005). Search for relevant sets of variables in a high-dimensional setup keeping the familywise error rate. *Statistics Neerlandica* 298–312.

[13] Quintana, S. M. and Maxwell, S. E. (1994). A Monte Carlo Comparison of Seven $\varepsilon$-Adjustment Procedures in Repeated Measures Designs with Small Sample Sizes. *Journal of Educational Statistics* **19**, 57–71.

[Imhof (1962)]

[14] Imhof, J. P. (1962). Testing the hypothesis of fixed main effects in Scheffe's mixed model. *Annals of Mathematical Statistics* **33**, 1086–1095.

[15] Skene, S. S. and Kenward, M. G. (2010). The analysis of very small samples of repeated measurements II: A modified Box correction. *Statistics in Medicine* **29**, 2838–2856.