KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0144

ESTIMATING THE INTERCEPT IN AN ORTHOGONALLY BLOCKED EXPERIMENT WHEN THE BLOCK EFFECTS ARE RANDOM

by
P. GOOS
M. VANDEBROEK

D/2001/2376/44

Estimating the Intercept in an Orthogonally Blocked Experiment When the Block Effects Are Random

Peter Goos
Martina Vandebroek

Katholieke Universiteit Leuven, Belgium

Abstract

For an orthogonally blocked experiment, Khuri (1992) has shown that the ordinary least squares estimator and the generalized least squares estimator of the factor effects in a response surface model with random block effects coincide. However, the equivalence does not hold for the estimation of the intercept when the block sizes are heterogeneous. When the block sizes are homogeneous, ordinary and generalized least squares provide an identical estimate for the intercept.

Keywords: orthogonal blocking, random block effects, combined intra- and inter-block estimator, equivalence of OLS and GLS

1 Introduction

In many experimental situations, response surface designs are divided into blocks in order to control for an extraneous source of variation. Khuri (1992) and Gilmour and Trinca (2000) pointed out that many experimental situations exist in which the blocks are randomly selected from a population of blocks, such that the block effects should be treated as random. The statistical model corresponding to a response surface experiment with n observations and b random block effects is given by

$$\mathbf{y} = \beta_0 \mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \qquad (1)$$

where $\mathbf{y}$ is a vector of n observations on a certain response, β_0 is the intercept, $\mathbf{1}_n$ is an n-dimensional vector of ones, $\mathbf{X}$ is the $n \times p$-dimensional design matrix, $\boldsymbol{\beta} = [\ \beta_1\ \beta_2\ \ldots\ \beta_p\ ]'$ is the p-dimensional vector of factor effects, $\mathbf{Z}$ is an $n \times b$ matrix of zeroes and ones assigning the n observations to the b blocks, $\boldsymbol{\gamma} = [\ \gamma_1\ \gamma_2\ \ldots\ \gamma_b\ ]'$ is the vector containing the b block effects and $\boldsymbol{\varepsilon}$ is a random error vector. Further, it is assumed that both $\boldsymbol{\gamma}$ and $\boldsymbol{\varepsilon}$ are normally distributed, and that $\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_b$,

$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n$, $\mathrm{cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_{b\times b}$, $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_{n\times n}$, and $\mathrm{cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b\times n}$. The variance-covariance matrix of the observations $\mathrm{cov}(\mathbf{y})$ can then be written as

$$\boldsymbol{\Sigma} = \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\gamma^2 \mathbf{Z}\mathbf{Z}' = \sigma_\varepsilon^2 \mathbf{A}. \tag{2}$$

Suppose the entries of $\mathbf{y}$ are grouped per block, then

$$\mathbf{A} = \mathrm{diag}[\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_b], \tag{3}$$

where

$$\mathbf{A}_i = \mathbf{I}_{k_i} + \eta \mathbf{1}_{k_i} \mathbf{1}'_{k_i}, \tag{4}$$

k_i is the size of block i, and

$$\eta = \sigma_\gamma^2 / \sigma_\varepsilon^2. \tag{5}$$

If the variance components are known, the best linear unbiased estimator (BLUE) of the unknown $\boldsymbol{\tau} = [\ \beta_0\ , \boldsymbol{\beta}'\ ]'$ is given by the generalized least squares estimator

$$\hat{\boldsymbol{\tau}}_{\mathrm{GLS}} = (\mathbf{W}'\mathbf{A}^{-1}\mathbf{W})^{-1}\mathbf{W}'\mathbf{A}^{-1}\mathbf{y}, \tag{6}$$

with $\mathbf{W} = [\ \mathbf{1}_n : \mathbf{X}\ ]$. The variance of this estimator is

$$\mathrm{var}(\hat{\boldsymbol{\tau}}_{\mathrm{GLS}}) = \sigma_\varepsilon^2 (\mathbf{W}'\mathbf{A}^{-1}\mathbf{W})^{-1}. \tag{7}$$

The generalized least squares estimator is also referred to as the combined intra- and inter-block estimator. Unlike the intra-block estimator, which is obtained by treating the block effects as fixed, the generalized least squares estimator does not suppress the inter-block information. However, Khuri (1992) has shown that both estimators produce the same estimate for $\boldsymbol{\beta}$ when the experiment is orthogonally blocked, that is when

$$\frac{1}{k_i}\mathbf{X}_i'\mathbf{1}_{k_i} = \frac{1}{n}\mathbf{X}'\mathbf{1}_n, \qquad (i = 1, \ldots, b), \tag{8}$$

where $\mathbf{X}_i$ represents the part of $\mathbf{X}$ corresponding to the ith block. In addition, he shows that the generalized least squares estimator for $\boldsymbol{\beta}$ is then also equivalent to that obtained from ordinary least squares regression, that is

$$\hat{\boldsymbol{\beta}}_{\mathrm{OLS}} = \hat{\boldsymbol{\beta}}_{\mathrm{GLS}}. \tag{9}$$

The obvious advantage thereof is that the estimates for the factor effects $\boldsymbol{\beta}$ do not depend on η. We denote the ordinary least squares estimator for $\boldsymbol{\tau}$ by

$$\hat{\boldsymbol{\tau}}_{\mathrm{OLS}} = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{y}, \tag{10}$$

and its variance is given by

$$\mathrm{var}(\hat{\boldsymbol{\tau}}_{\mathrm{OLS}}) = \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{A}\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}. \tag{11}$$

From this expression, it can be seen that, even in an orthogonally blocked experiment, knowledge of the variance components remains indispensible for statistical inference. In case the variance components are unknown, they can be estimated using restricted maximum likelihood (for a detailed discussion, see Gilmour and Trinca 2000).

Although ordinary and generalized least squares estimation produce the same estimates for the factor effects in an orthogonally blocked experiment, they produce different estimates for the intercept β_0 when the block sizes are heterogeneous. Also the variances of both estimators are different. When predicting the response is one of the experimenter's goals, for example to find those settings of the experimental factors to achieve a target value for the response, these difference should not be ignored. The analytical results in this paper are illustrated by means of an orthogonally blocked experiment conducted at the research center of a food additive producer. It is also shown that the ordinary and generalized least squares estimator for β_0 are equivalent when the block sizes are homogeneous, and that, in this case, the variances (7) and (11) are equal. In the remainder of the paper, we will denote the ordinary and the generalized least squares estimators for β_0 by $\hat{\beta}_{0,\mathrm{OLS}}$ and $\hat{\beta}_{0,\mathrm{GLS}}$ respectively. Finally, it should be stressed that all designs considered in this paper are orthogonally blocked and that only the case where $\mathbf{W}$ is full rank is considered.

2 Heterogeneous block sizes

When the block sizes are heterogeneous, $\hat{\beta}_0$ depends on the estimation method used. Heterogeneous block sizes frequently occur when an orthogonally blocked second order standard design is used. For example, Box and Hunter (1957) propose an orthogonally blocked central composite design for three variables with 2 factorial blocks of size 6 and one axial block of size 8. Another example of an orthogonally blocked central composite design is given in Table 1. This 54-run experiment was carried out in the laboratory of a multinational producing ingredients for food and beverage applications to investigate the effect of adding salt in a starch extraction process on its yield (expressed in %). Five blocks corresponding to batches of raw material originating from five of the countries in which the company was active were used in the experiment. Two of the blocks contain 18 runs, whereas the others contains only 6 runs. Five equally spaced salt levels were used in the experiment and the middle level was replicated twice as much as the other levels. As the yield was expected to increase at a slackening pace, a quadratic model was fitted using both ordinary and generalized least squares. Ordinary least squares estimation gave

$$\mathrm{E}(y) = 49.44 + 5.51x - 2.58x^2,$$

Table 1: Orthogonally blocked starch extraction experiment.

x	Block 1			Block 2			Block 3	Block 4	Block 5
-1	40.5	39.4	42.3	40.8	39.4	38.6	49.3	48.5	38.1
-0.5	43.2	41.7	41.3	40.1	43.5	44.4	51.0	52.8	44.7
0	50.4	43.6	46.6	45.7	42.8	48.4	55.8	62.2	50.6
0	47.4	45.0	48.5	47.6	48.9	45.8	59.4	55.5	52.1
0.5	49.1	53.1	51.3	46.9	47.6	50.8	56.4	60.5	51.3
1	50.0	48.3	50.1	49.1	48.3	50.2	60.0	60.1	52.6

whereas generalized least squares produced

$$\mathrm{E}(y) = 51.41 + 5.51x - 2.58x^2.$$

The latter model was obtained using the ratio of the restricted maximum likelihood estimates $\hat{\sigma}_\gamma^2 = 26.68$ and $\sigma_\varepsilon^2 = 4.21$ as an estimate for η. It is clear that, apart from the intercept, both models are identical. The difference between both intercepts amounts to 1.97, which is quite large when compared to the standard deviation of the predicted value at the center point (2.35). In cases where the interest is in estimating the location of a stationary point on a response surface or in estimating the difference between the responses for two combinations of the factor levels, the value of the intercept does not matter. However, the goals of an experiment often include the prediction of the response for certain combinations of the factor levels. This is important when a target value for the response has to be achieved. Of course, any deviation from the target is undesirable, such that the estimation method in this case really matters.

In order to assess the difference between the ordinary and generalized least squares estimator over a range of examples, consider the analytical expressions for $\hat{\beta}_{0,\mathrm{OLS}}$ and $\hat{\beta}_{0,\mathrm{GLS}}$ derived in Appendix A:

$$\hat{\beta}_{0,\mathrm{OLS}} = \frac{1}{n}\mathbf{1}_n'\mathbf{y} - \frac{1}{n}\mathbf{1}_n'\mathbf{X}\hat{\boldsymbol{\beta}}, \tag{12}$$

$$\hat{\beta}_{0,\mathrm{GLS}} = \frac{\sum_{i=1}^{b} \frac{\mathbf{1}_{k_i}'\mathbf{y}_i}{1+k_i\eta}}{\sum_{i=1}^{b} \frac{k_i}{1+k_i\eta}} - \frac{1}{n}\mathbf{1}_n'\mathbf{X}\hat{\boldsymbol{\beta}}, \tag{13}$$

where $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_{\mathrm{OLS}} = \hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$. The difference between both estimators equals

$$\hat{\beta}_{0,\mathrm{GLS}} - \hat{\beta}_{0,\mathrm{OLS}} = \frac{\sum_{i=1}^{b} \frac{k_i\bar{y}_i}{1+k_i\eta}}{\sum_{i=1}^{b} \frac{k_i}{1+k_i\eta}} - \bar{y}, \tag{14}$$

where $\bar{y}$ is the average response of the experiment, and $\bar{y}_i$ represents the average response in the ith block. Remarkably, this difference only depends on the block

Table 2: Differences between $\text{var}(\hat{\beta}_{0,\text{OLS}})$ and $\text{var}(\hat{\beta}_{0,\text{GLS}})$ for several block sizes and several values of η.

n	k_1	k_2	k_3	$\eta = 1$	$\eta = 2$	$\eta = 5$
10	4	6		0.0828	0.1208	0.1600
10	3	7		0.3231	0.4766	0.6364
10	3	3	4	0.0261	0.0390	0.0526
10	2	4	4	0.0941	0.1471	0.2051
20	9	11		0.0227	0.0317	0.0408
20	8	12		0.0906	0.1267	0.1633
20	6	6	8	0.0293	0.0416	0.0541
20	6	7	7	0.0072	0.0103	0.0135
32	10	11	11	0.0030	0.0041	0.0053
32	8	12	12	0.0469	0.0658	0.0849

effects δ_i, the random errors ε_i, and k_i, but not on the design matrix $\mathbf{X}$ or $\boldsymbol{\tau}$. This is proven in Appendix B. Since both the ordinary and generalized least squares estimators are unbiased, the expected value of (14) is zero and the choice between both estimators has to be based on their variances.

In Appendix C, it is shown that the variances of $\hat{\beta}_{0,\text{OLS}}$ and $\hat{\beta}_{0,\text{GLS}}$ are different, unlike the variances of $\hat{\boldsymbol{\beta}}_{\text{OLS}}$ and $\hat{\boldsymbol{\beta}}_{\text{GLS}}$. The difference does not depend on $\mathbf{X}$ or $\boldsymbol{\tau}$, and equals

$$\text{var}(\hat{\beta}_{0,\text{GLS}}) - \text{var}(\hat{\beta}_{0,\text{OLS}}) = \sigma_\varepsilon^2\left(\sum_{i=1}^{b} \frac{k_i^2\eta}{n^2} - \frac{\sum_{i=1}^{b} \frac{k_i^2\eta}{n^2(1+k_i\eta)}}{1 - \sum_{i=1}^{b} \frac{k_i^2\eta}{n(1+k_i\eta)}}\right), \tag{15}$$

and is always positive (for a formal proof, see Appendix D). From the formula, it can be seen that the difference between the ordinary and generalized least squares estimator increases with the total variance in the responses. In order to illustrate how the difference depends on η and the block sizes k_i, we have computed values of (15) for 10 instances with heterogeneous block sizes, holding $\sigma_\varepsilon^2 + \sigma_\gamma^2 = 10$. The results are displayed in Table 2. It can be seen from the table that the extent to which the variance associated with the generalized least squares estimator is smaller than that associated with the ordinary least squares estimator increases with η and with block size heterogeneity.

In order to obtain some sense for the practical consequences of the difference between both estimators, consider again the starch extraction experiment. It is known that the yield of the extraction process increases with the amount of salt added. However, the amount of salt added could not be increased infinitely for economical reasons and a target yield of 45% was determined. Confidence intervals for $\text{E}(y)$ were then used to determine the amount of salt needed to achieve the target value. Of course, the 90% confidence interval for E(y) is given by

$$\mathrm{E}(y) \pm 1.645\sqrt{\mathrm{var}(y)},$$

and it depends on the estimation method used. Using the ordinary least squares estimate for the intercept and the corresponding variance, it turns out that $x = -0.0115$ is needed to achieve the 45% target yield in 90% of the cases. It can be verified that the corresponding 90% confidence interval is

$$[45.00, 53.75].$$

Using the generalized least squares estimate for he intercept and the corresponding variance, a value of $x = -0.3923$ was obtained instead. The corresponding confidence interval is

$$[45.00, 52.71],$$

which is substantially smaller than the interval obtained by using ordinary least squares. From this example, it is clear that ordinary and generalized least squares may provide entirely different solutions to a practical problem. The solution given by the generalized least squares method is much cheaper than that suggested by the ordinary least squares approach. In view of the smaller variance associated with the generalized least squares approach, the cheaper option was most inspiring.

3 Homogeneous block sizes

When the block sizes are homogeneous, ordinary and generalized least squares produce identical estimators of β_0 and the variances of both estimators are equal. As an illustration, consider the data in Table 3 from a small reactor study introduced by Box and Draper (1987) and revisited by Khuri (1994). In the experiment, the effect of 3 factors (flow rate, concentration of a catalyst, and temperature) on the concentration of a product was investigated. The 24 runs of the experiment were performed sequentially in 4 blocks of size 6. A full quadratic model was fitted to the data. It can be verified that is equal to

$$\begin{aligned} y = 51.79 &+ 0.74x_1 + 4.81x_2 + 8.01x_3 + 0.38x_1x_2 \\ &+ 10.35x_1x_3 - 2.83x_2x_3 - 3.83x_1^2 + 1.22x_2^2 - 6.26x_3^2, \end{aligned}$$

no matter whether ordinary of generalized least squares is used. A formal proof of the equivalence between the ordinary and generalized least squares estimators for the intercept is obtained by noting that the right hand side of (14) is zero when $k = k_1 = k_2 = \cdots = k_b$. The difference (15) between the variances of $\hat{\beta}_{0,\mathrm{OLS}}$ and $\hat{\beta}_{0,\mathrm{GLS}}$ is then also equal to zero. As a consequence, the variance-covariance matrices of the ordinary and generalized least squares estimator are identical when an orthogonally blocked experiment with homogeneous block sizes is used.

Table 3: Orthogonally blocked central composite design with four blocks of size six.

Block 1				Block 2				Block 3				Block 4			
x_1	x_2	x_3	y	x_1	x_2	x_3	y	x_1	x_2	x_3	y	x_1	x_2	x_3	y
-1	-1	1	40.0	-1	-1	-1	39.5	$-\alpha$	0	0	43.0	$-\alpha$	0	0	39.2
1	-1	-1	18.6	1	-1	1	59.7	α	0	0	43.9	α	0	0	46.3
-1	1	-1	53.8	-1	1	1	42.2	0	$-\alpha$	0	47.0	0	$-\alpha$	0	44.9
1	1	1	64.2	1	1	-1	33.6	0	α	0	62.8	0	α	0	58.1
0	0	0	53.5	0	0	0	54.1	0	0	$-\alpha$	25.6	0	0	$-\alpha$	27.0
0	0	0	52.7	0	0	0	51.0	0	0	α	49.7	0	0	α	50.7

$\alpha = \sqrt{2}$

4 Conclusion

In this paper, it is shown that ordinary and generalized least squares produce a different estimate for the intercept in a response surface model with random block effects when an orthogonally blocked experiment with heterogeneous block sizes is used. The difference between both estimates increases with the total variance in the response, the extent to which the responses are correlated and the heterogeneity of the block sizes. This result is surprising in view of Khuri's (1992) proof that both methods yield identical estimates of the factor effects. Another interesting point is that the variance-covariance matrices of the ordinary and generalized least squares estimator differ in only one element, namely the variance of the intercept. When the block sizes are homogeneous, both methods yield the same estimate for the intercept. In this case, the ordinary and generalized least squares estimators have identical variance-covariance matrices. The estimation of the intercept is important when the goal of the experiment is to predict the response for combinations of the experimental factors and when a target value for the response has to be achieved. This was illustrated in Section 2.

The results described in this paper emphasize the importance of using an orthogonally blocked experiment with homogeneous block sizes. In doing so, the estimation of all regression parameters, including the intercept, is insensitive to the estimation method used (ordinary or generalized least squares) nor to the estimate of the variance components. In addition, inference procedures will also be independent from the estimator used. When heterogeneous block sizes are inevitable, the generalized least squares estimator is recommended for the intercept because it has a smaller variance than the ordinary least squares estimator. For all other regression parameters, the estimation method does not matter.

In the light of these results, it is interesting to point out that orthogonal blocking is an optimal strategy to assign the experimental runs of a given design $\mathbf{X}$ to the blocks (Goos and Vandebroek 2001). Therefore, algorithms for computing optimal blocked experiments (Atkinson and Donev 1989, Goos and Vandebroek 2001) will generate orthogonally blocked designs whenever possible. This is also true for the algorithms of Trinca and Gilmour (2000, 2001).

Acknowledgements

The research for this paper was carried out while the first author was a Postdoctoral Researcher of the Fund for Scientific Research - Flanders (Belgium). The authors are grateful to Willy Gochet for his help in developing the proof in Appendix D.

Appendix A. The difference between $\hat{\beta}_{0,\text{GLS}}$ and $\hat{\beta}_{0,\text{OLS}}$ when the block sizes are heterogeneous.

Assume that the block sizes are heterogeneous and equal to k_i. The generalized least squares estimator (6) becomes

$$\begin{bmatrix}\hat{\beta}_{0,\text{GLS}}\\ \hat{\boldsymbol{\beta}}_{\text{GLS}}\end{bmatrix} = \begin{bmatrix}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{1}_n & \mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}\\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{1}_n & \mathbf{X}'\mathbf{A}^{-1}\mathbf{X}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y}\\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{y}\end{bmatrix}. \tag{16}$$

We have from Harville's (1997) Theorem 18.2.8 that

$$\mathbf{A}_i^{-1} = \mathbf{I}_{k_i\times k_i} - \frac{\eta}{1+k_i\eta}\mathbf{1}_{k_i}\mathbf{1}_{k_i}'.$$

Hence

$$\begin{aligned}
\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{1}_n &= \sum_{i=1}^{b}\mathbf{1}_k'\mathbf{A}_i^{-1}\mathbf{1}_k,\\
&= \sum_{i=1}^{b}\mathbf{1}_{k_i}'(\mathbf{I}_{k_i\times k_i} - \frac{\eta}{1+k_i\eta}\mathbf{1}_{k_i}\mathbf{1}_{k_i}')\mathbf{1}_{k_i},\\
&= \sum_{i=1}^{b}(k_i - \frac{\eta}{1+k_i\eta}k_i^2),\\
&= \sum_{i=1}^{b}\frac{k_i}{1+k_i\eta},\\
&= \sum_{i=1}^{b}k_ic_i,\\
&= d,
\end{aligned}$$

and by using Theorem 8.5.11 of Harville (1997) the generalized least squares estimator becomes

$$\begin{bmatrix}\hat{\beta}_{0,\text{GLS}}\\ \hat{\boldsymbol{\beta}}_{\text{GLS}}\end{bmatrix} = \begin{bmatrix} d & \mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X} \\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{1}_n & \mathbf{X}'\mathbf{A}^{-1}\mathbf{X}\end{bmatrix}^{-1} \begin{bmatrix}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y}\\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{y}\end{bmatrix},$$

$$= \begin{bmatrix} d^{-1} + d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}\mathbf{Q}_1 & -d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}\mathbf{Q}_2 \\ -\mathbf{Q}_1 & \mathbf{Q}_2\end{bmatrix} \begin{bmatrix}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y}\\ \mathbf{X}'\mathbf{A}^{-1}\mathbf{y}\end{bmatrix},$$

$$= \begin{bmatrix} d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y} + d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}(\mathbf{Q}_1\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y} - \mathbf{Q}_2\mathbf{X}'\mathbf{A}^{-1}\mathbf{y}) \\ -\mathbf{Q}_1\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y} + \mathbf{Q}_2\mathbf{X}'\mathbf{A}^{-1}\mathbf{y}\end{bmatrix},$$

where

$$\mathbf{Q}_1 = \mathbf{Q}_2\mathbf{X}'\mathbf{A}^{-1}\mathbf{1}_n d^{-1},$$

and

$$\mathbf{Q}_2 = \{\mathbf{X}'\mathbf{A}^{-1}\mathbf{X} - \mathbf{X}'\mathbf{A}^{-1}\mathbf{1}_n d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}\}^{-1}.$$

As a result,

$$\hat{\beta}_{0,\text{GLS}} = d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y} - d^{-1}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X}\hat{\boldsymbol{\beta}}_{\text{GLS}}.$$

Since

$$\begin{aligned}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{y} &= \sum_{i=1}^{b}\mathbf{1}_{k_i}'\mathbf{A}_i^{-1}\mathbf{y}_i,\\ &= \sum_{i=1}^{b}\mathbf{1}_{k_i}'(\mathbf{I}_{k_i\times k_i} - \frac{\eta}{1+k_i\eta}\mathbf{1}_{k_i}\mathbf{1}_{k_i}')\mathbf{y}_i,\\ &= \sum_{i=1}^{b}(1 - \frac{k_i\eta}{1+k_i\eta})\mathbf{1}_{k_i}'\mathbf{y}_i,\\ &= \sum_{i=1}^{b}c_i\mathbf{1}_{k_i}'\mathbf{y}_i,\end{aligned}$$

where $\mathbf{y}_i$ is that part of $\mathbf{y}$ corresponding to the ith block, and

$$\begin{aligned}\mathbf{1}_n'\mathbf{A}^{-1}\mathbf{X} &= \sum_{i=1}^{b}\mathbf{1}_{k_i}'\mathbf{A}_i^{-1}\mathbf{X}_i,\\ &= \sum_{i=1}^{b}\mathbf{1}_{k_i}'(\mathbf{I}_{k_i\times k_i} - \frac{\eta}{1+k_i\eta}\mathbf{1}_{k_i}\mathbf{1}_{k_i}')\mathbf{X}_i,\\ &= \sum_{i=1}^{b}(1 - \frac{k_i\eta}{1+k_i\eta})\mathbf{1}_{k_i}'\mathbf{X}_i,\\ &= n^{-1}(\sum_{i=1}^{b}k_ic_i)\mathbf{1}_n'\mathbf{X},\\ &= dn^{-1}\mathbf{1}_n'\mathbf{X},\end{aligned}$$

this estimator becomes

$$\hat{\beta}_{0,\text{GLS}} = d^{-1}(\sum_{i=1}^{b} c_i \mathbf{1}'_{k_i}\mathbf{y}_i) - n^{-1}\mathbf{1}'_n\mathbf{X}\hat{\boldsymbol{\beta}}_{\text{GLS}}. \tag{17}$$

The OLS estimator for the intercept can be obtained by substituting $\mathbf{A} = \mathbf{I}_{n\times n}$ in (16) and is equal to

$$\hat{\beta}_{0,\text{OLS}} = n^{-1}\mathbf{1}'_n\mathbf{y} - n^{-1}\mathbf{1}'_n\mathbf{X}\hat{\boldsymbol{\beta}}_{\text{OLS}}. \tag{18}$$

Using the fact that $\hat{\boldsymbol{\beta}}_{\text{OLS}} = \hat{\boldsymbol{\beta}}_{\text{GLS}}$, it can easily be derived from (17) and (18) that $\hat{\beta}_{0,\text{GLS}}$ and $\hat{\beta}_{0,\text{OLS}}$ are equivalent when the block sizes are equal.

Appendix B. The independence of $\hat{\beta}_{0,\text{GLS}} - \hat{\beta}_{0,\text{OLS}}$ from X and $\boldsymbol{\beta}$.

Substituting (1) in (12) and (13) and subtracting both equations, we obtain

$$\begin{aligned}
\hat{\beta}_{0,\text{GLS}} - \hat{\beta}_{0,\text{OLS}} &= (\sum_{i=1}^{b} k_i c_i)^{-1} \sum_{i=1}^{b} c_i \mathbf{1}'_{k_i}(\beta_0 \mathbf{1}_{k_i} + \mathbf{X}_i\boldsymbol{\beta} + \gamma_i \mathbf{1}_{k_i} + \boldsymbol{\varepsilon}_i) \\
&\quad - n^{-1}\mathbf{1}'_n(\beta_0\mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}), \\
&= (\sum_{i=1}^{b} k_i c_i)^{-1} \sum_{i=1}^{b} (k_i c_i \beta_0 + c_i \mathbf{1}'_{k_i}\mathbf{X}_i\boldsymbol{\beta} + k_i c_i \gamma_i + c_i \mathbf{1}'_{k_i}\boldsymbol{\varepsilon}_i) \\
&\quad - n^{-1}(n\beta_0 + \mathbf{1}'_n\mathbf{X}\boldsymbol{\beta} + \mathbf{1}'_n\mathbf{Z}\boldsymbol{\gamma} + \mathbf{1}'_n\boldsymbol{\varepsilon}), \\
&= (\sum_{i=1}^{b} k_i c_i)^{-1} \sum_{i=1}^{b} (c_i \mathbf{1}'_{k_i}\mathbf{X}_i\boldsymbol{\beta} + k_i c_i \gamma_i + c_i \mathbf{1}'_{k_i}\boldsymbol{\varepsilon}_i) \\
&\quad - n^{-1}(\mathbf{1}'_n\mathbf{X}\boldsymbol{\beta} + \mathbf{1}'_n\mathbf{Z}\boldsymbol{\gamma} + \mathbf{1}'_n\boldsymbol{\varepsilon}), \\
&= (\sum_{i=1}^{b} k_i c_i)^{-1} \sum_{i=1}^{b} (k_i c_i n^{-1}\mathbf{1}'_n\mathbf{X}\boldsymbol{\beta} + k_i c_i \gamma_i + c_i \mathbf{1}'_{k_i}\boldsymbol{\varepsilon}_i) \\
&\quad - n^{-1}(\mathbf{1}'_n\mathbf{X}\boldsymbol{\beta} + \mathbf{1}'_n\mathbf{Z}\boldsymbol{\gamma} + \mathbf{1}'_n\boldsymbol{\varepsilon}), \\
&= (\sum_{i=1}^{b} k_i c_i)^{-1} \sum_{i=1}^{b} (k_i c_i \gamma_i + c_i \mathbf{1}'_{k_i}\boldsymbol{\varepsilon}_i) - n^{-1}(\mathbf{1}'_n\mathbf{Z}\boldsymbol{\gamma} + \mathbf{1}'_n\boldsymbol{\varepsilon}),
\end{aligned}$$

From this result, it is clear that the difference between $\hat{\beta}_{0,\text{GLS}}$ and $\hat{\beta}_{0,\text{OLS}}$ only depends on the random block effects δ_i and the random errors ε_i, but not on the design matrix $\mathbf{X}$ and the parameters β_0 and $\boldsymbol{\beta}$.

Appendix C. The variances of $\hat{\boldsymbol{\beta}}_{\text{GLS}}$ and $\hat{\boldsymbol{\beta}}_{\text{OLS}}$.

The variance-covariance matrix of the GLS estimator for $\boldsymbol{\tau}$ is given by

$$\begin{aligned}
\text{var}(\hat{\boldsymbol{\tau}}_{\text{GLS}}) &= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{A}^{-1}\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2(\sum_{i=1}^{b}\mathbf{W}_i'\mathbf{A}_i^{-1}\mathbf{W}_i)^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W} - \sum_{i=1}^{b}\frac{\eta}{1+k_i\eta}(\mathbf{W}_i'\mathbf{1}_{k_i})(\mathbf{1}_{k_i}'\mathbf{W}_i))^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W} - \sum_{i=1}^{b}\frac{\eta}{1+k_i\eta}(\frac{k_i}{n}\mathbf{W}'\mathbf{1}_n)(\frac{k_i}{n}\mathbf{1}_n'\mathbf{W}_i))^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W} - \sum_{i=1}^{b}\frac{k_i^2\eta}{n^2(1+k_i\eta)}(\mathbf{W}'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{W}_i))^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W} - c_1(\mathbf{W}'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{W}))^{-1}, \\
&= \sigma_\varepsilon^2((\mathbf{W}'\mathbf{W})^{-1} + c_1c_2^{-1}(\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{W})(\mathbf{W}'\mathbf{W})^{-1}), \\
&= \sigma_\varepsilon^2((\mathbf{W}'\mathbf{W})^{-1} + c_1c_2^{-1}\mathbf{u}_1\mathbf{u}_1'),
\end{aligned}$$

where $\mathbf{u}_1'$ is the p-dimensional unit vector $[\,1\,,\,0\,,\,\ldots\,,\,0\,]$,

$$c_1 = \sum_{i=1}^{b}\frac{k_i^2\eta}{n^2(1+k_i\eta)},$$

and

$$\begin{aligned}
c_2 &= 1 - c_1(\mathbf{1}_n'\mathbf{W}_i)(\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\mathbf{1}_n), \\
&= 1 - c_1\mathbf{u}_1'(\mathbf{W}'\mathbf{1}_n), \\
&= 1 - c_1 n.
\end{aligned}$$

The variance-covariance matrix of the OLS estimator for $\boldsymbol{\tau}$ is given by

$$\begin{aligned}
\text{var}(\hat{\boldsymbol{\tau}}_{\text{OLS}}) &= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{A}\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1} + \sigma_\gamma^2(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{Z}\mathbf{Z}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1} + \sigma_\gamma^2(\mathbf{W}'\mathbf{W})^{-1}(\sum_{i=1}^{b}(\mathbf{W}_i'\mathbf{1}_{k_i})(\mathbf{1}_{k_i}'\mathbf{W}_i))(\mathbf{W}'\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1} + \sigma_\gamma^2(\mathbf{W}'\mathbf{W})^{-1}(\sum_{i=1}^{b}(\frac{k_i}{n}\mathbf{W}'\mathbf{1}_n)(\frac{k_i}{n}\mathbf{1}_n'\mathbf{W}))(\mathbf{W}'\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1} + \sigma_\gamma^2(\sum_{i=1}^{b}\frac{k_i^2}{n^2})(\mathbf{W}'\mathbf{W})^{-1}(\mathbf{W}'\mathbf{1}_n)(\mathbf{1}_n'\mathbf{W})(\mathbf{W}'\mathbf{W})^{-1}, \\
&= \sigma_\varepsilon^2((\mathbf{W}'\mathbf{W})^{-1} + (\sum_{i=1}^{b}\frac{k_i^2\eta}{n^2})\mathbf{u}_1\mathbf{u}_1'),
\end{aligned}$$

It is clear from these expressions that $\mathrm{var}(\hat{\boldsymbol{\tau}}_{\text{OLS}})$ and $\mathrm{var}(\hat{\boldsymbol{\tau}}_{\text{GLS}})$ only differ in one element, namely the variance of the intercept estimator in the upper left hand corner. When the block sizes are all equal to k, both $\mathrm{var}(\hat{\boldsymbol{\tau}}_{\text{OLS}})$ and $\mathrm{var}(\hat{\boldsymbol{\tau}}_{\text{GLS}})$ reduce to

$$\sigma_\varepsilon^2(\mathbf{W}'\mathbf{W})^{-1} + \frac{\sigma_\gamma^2}{b}\mathbf{u}_1\mathbf{u}_1'.$$

Appendix D. Proof that var $(\hat{\beta}_{0,\text{GLS}}) < \mathrm{var}(\hat{\beta}_{0,\text{OLS}})$.

The variance associated with the generalized least squares estimator is smaller than that associated with the ordinary least squares estimator if (15) is positive. Since $\sigma_\varepsilon^2 \geq 0$, $\eta \geq 0$, and

$$1 - \sum_{i=1}^{b} \frac{k_i^2\eta}{n(1+k_i\eta)} \geq 0,$$

the difference (15) is positive if

$$(\sum_{i=1}^{b} \frac{k_i^2}{n^2})(1 - \sum_{i=1}^{b} \frac{k_i^2\eta}{n(1+k_i\eta)}) - \sum_{i=1}^{b} \frac{k_i^2}{n^2(1+k_i\eta)} \geq 0.$$

Substituting k_i/n by r_i and $n\eta$ by λ, we find that

$$\begin{aligned}
&(\sum_{i=1}^{b} r_i^2)(1 - \lambda\sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i}) - \sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i} \\
&= \sum_{i=1}^{b} r_i^2 - \sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i}(1 + \lambda\sum_{i=1}^{b} r_i^2), \\
&= \sum_{i=1}^{b} \frac{r_i^2(1+\lambda r_i)}{1+\lambda r_i} - \sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i}(1 + \lambda\sum_{i=1}^{b} r_i^2), \\
&= \lambda(\sum_{i=1}^{b} \frac{r_i^3}{1+\lambda r_i} - (\sum_{i=1}^{b} r_i^2)\sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i}).
\end{aligned}$$

Omitting the positive constant λ, we have

$$\begin{aligned}
&\sum_{i=1}^{b} \frac{r_i^3}{1+\lambda r_i} - (\sum_{i=1}^{b} r_i^2)\sum_{i=1}^{b} \frac{r_i^2}{1+\lambda r_i} \\
&= \frac{r_1^2}{1+\lambda r_1}(r_1 - \sum_{i=1}^{b} r_i^2) + \frac{r_2^2}{1+\lambda r_2}(r_2 - \sum_{i=1}^{b} r_i^2) + \cdots + \frac{r_b^2}{1+\lambda r_b}(r_b - \sum_{i=1}^{b} r_i^2), \\
&= \frac{r_1^2(r_1(1-r_1) - r_2^2 - \cdots - r_b^2)(1+\lambda r_2)\ldots(1+\lambda r_b)}{\prod_{i=1}^{b}(1+\lambda r_i)} \\
&\quad + \cdots + \frac{r_b^2(r_b(1-r_b) - r_1^2 - \cdots - r_{b-1}^2)(1+\lambda r_1)\ldots(1+\lambda r_{b-1})}{\prod_{i=1}^{b}(1+\lambda r_i)}.
\end{aligned}$$

Using the fact that $\sum_{i=1}^{b} r_i = 1$, and omitting the denominators from this expression, we obtain

$$
\begin{aligned}
& r_1^2(r_1(1-r_1) - r_2^2 - \cdots - r_b^2)(1+\lambda r_2)\ldots(1+\lambda r_b) \\
& \quad + \cdots + r_b^2(r_b(1-r_b) - r_1^2 - \cdots - r_{b-1}^2)(1+\lambda r_2)\ldots(1+\lambda r_b) \\
& = r_1^2(r_1(r_2 + \cdots + r_b) - r_2^2 - \cdots - r_b^2)(1+\lambda r_2)\ldots(1+\lambda r_b) \\
& \quad + \cdots + r_b^2(r_b(r_1 + \cdots + r_{b-1}) - r_1^2 - \cdots - r_{b-1}^2)(1+\lambda r_2)\ldots(1+\lambda r_b), \\
& = (r_1 - r_2)^2 r_1 r_2 (1+\lambda r_3)\ldots(1+\lambda r_b) \\
& \quad + \cdots + (r_{b-1} - r_b)^2 r_{b-1} r_b (1+\lambda r_1)\ldots(1+\lambda r_{b-2}), \\
& = \sum_{i=1}^{b} \sum_{j=i+1}^{b} (r_i - r_j)^2 r_i r_j \Big(\prod_{\substack{l=1 \\ l\neq i,j}}^{b} (1+\lambda r_l)\Big),
\end{aligned}
$$

which is clearly positive, so that we can conclude that (15) is positive. This expression becomes 0 if the block sizes are homogeneous because all r_i are then equal.

References

Atkinson, A. and Donev, A. (1989). The construction of exact D-optimum experimental designs with application to blocking response surface designs, *Biometrika* **76**: 515–526.

Box, G. and Draper, N. (1987). *Empirical Model-Building and Response Surfaces*, New York: Wiley.

Box, G. and Hunter, J. (1957). Multi-factor experimental designs for exploring response surfaces, *Annals of Mathematical Statistics* **28**: 195–241.

Gilmour, S. and Trinca, L. (2000). Some practical advice on polynomial regression analysis from blocked response surface designs, *Communications in Statistics: Theory and Methods* **29**: 2157–2180.

Goos, P. and Vandebroek, M. (2001). D-optimal response surface designs in the presence of random block effects, *Computational Statistics and Data Analysis* **37**: 433–453.

Harville, D. (1997). *Matrix Algebra from a Statistician's Perspective*, New York: Springer.

Khuri, A. (1992). Response surface models with random block effects, *Technometrics* **34**: 26–37.

Khuri, A. (1994). Effect of blocking on the estimation of a response surface, *Journal of Applied Statistics* **21**: 305–316.

Trinca, L. and Gilmour, S. (2000). An algorithm for arranging response surface designs in small blocks, *Computational Statistics and Data Analysis* **33**: 25–43.

Trinca, L. and Gilmour, S. (2001). Multi-stratum response surface designs, *Technometrics* **43**: 25–33.