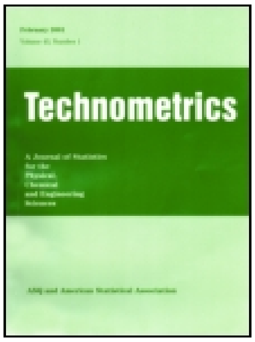

Technometrics

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utch20

Staggered-Level Designs for Experiments With More Than One Hard-to-Change Factor

Heidi Arnouts & Peter Goos

To cite this article: Heidi Arnouts & Peter Goos (2012) Staggered-Level Designs for Experiments With More Than One Hard-to-Change Factor, Technometrics, 54:4, 355-366, DOI: 10.1080/00401706.2012.715834

To link to this article: http://dx.doi.org/10.1080/00401706.2012.715834

Staggered-Level Designs for Experiments With More Than One Hard-to-Change Factor

Heidi Arnouts

Faculty of Applied Economics & StatUa
Center for Statistics
Universiteit Antwerpen
2000 Antwerpen, Belgium
(*heidi.arnouts@ua.ac.be*)

Peter Goos

Faculty of Applied Economics & StatUa
Center for Statistics, Universiteit Antwerpen,
2000 Antwerpen, Belgium
and
Erasmus School of Economics
Erasmus Universiteit Rotterdam
3000 DR Rotterdam, The Netherlands
(*peter.goos@ua.ac.be*)

In many industrial experiments, some of the factors are not independently set for each run. This is due to time and/or cost constraints and to the hard-to-change nature of the levels of these factors. Most of the literature restricts attention to split-plot designs in which all the hard-to-change factors are independently reset at the same points in time. This constraint is to some extent relaxed in split-split-plot designs because these allow the less hard-to-change factors to be reset more often than the most hard-to-change factors. A key feature of the split-split-plot designs, however, is that the less hard-to-change factors are reset whenever the most hard-to-change factors are reset. In this article, we relax this constraint and present a new type of design, which allows the hard-to-change factor levels to be reset at entirely different points in time. We show that the new designs are cost-efficient and that they outperform split-plot and split-split-plot designs in terms of the D- and A-optimality criteria. Because of the fact that the hard-to-change factors are independently reset alternatingly, we name the new designs staggered-level designs. Supplementary materials for this article are available online.

KEY WORDS: A- and D-optimality criterion; Cost; Gauss–Hermite quadrature; OLS and GLS equivalence; Split-plot design; Split-split-plot design.

1. INTRODUCTION

In many industrial experiments, complete randomization with independent settings of all the experimental factors is not feasible. Much research has already been done for situations in which there is only one hard-to-change factor or in which the levels of all hard-to-change factors are reset at the same time. Designs recommended in the literature for these situations are split-plot designs. Bisgaard (2000) recognized that split-plotting is common and much more frequently used than the literature on design of experiments in engineering would suggest. Ganju and Lucas (1999) stated that a split-plot structure should be designed and not, as often happens, be the accidental result of failure to reset factor levels. Ju and Lucas (2002) compared the precision of the estimator of the regression coefficients for various run-order scenarios and showed that classical split-plot designs are superior to random run orders.

Vining, Kowalski, and Montgomery (2005) discussed how to modify the standard central composite and Box–Behnken designs to accommodate a split-plot structure and discovered that some of the constructed designs achieved the equivalence of ordinary least-squares (OLS) and generalized least-squares (GLS) estimation. Goos and Vandebroek (2003) showed that certain two-level factorial and fractional factorial split-plot designs are D-optimal for the estimation of first-order response surface models for a specific number and size of whole plots. Like Jones and Goos (2007), they proposed an algorithm to construct D-optimal split-plot designs. Anbari and Lucas (2008) listed practical split-plot patterns that give higher G-efficiency and higher cost efficiency than completely randomized designs. Macharia and Goos (2010) and Jones and Goos (2012a) showed that many D-optimal and D-efficient split-plot designs also possess the property that the OLS and GLS estimators are equivalent. Jones and Goos (2012b) discussed the construction of I-optimal split-plot designs.

Following the basic idea of split-plot designs and using three strata instead of two leads to split-split-plot designs. This type of design can be useful for situations in which there are two groups of hard-to-change factors, one of which includes factors that are more difficult to reset than the other. The literature on split-split-plot designs is rather limited. Trinca and Gilmour (2001) discussed the design and analysis of multi-stratum experiments, special cases of which are split-plot and split-split-plot designs. They orthogonalized each stratum of the design as much as possible with respect to the higher strata. Schoen (1999) constructed an orthogonal split-split-plot design in a combinatorial way by joining fractional factorial designs to create the desired nesting structure. Jones and Goos (2009) provided a coordinate-exchange algorithm to compute D-optimal split-split-plot designs. Typical for split-split-plot designs is

TECHNOMETRICS, NOVEMBER 2012, VOL. 54, NO. 4
DOI: 10.1080/00401706.2012.715834

Table 1. Wrapper machine example with two hard-to-change factors (spacing and speed) and one easy-to-change factor (temperature) where spacing is set four times, speed is set eight times, and temperature is independently set for each run (see Webb, Lucas, and Borkowski 2004)

Spacing	Speed	Temp
0	1	−1
0	1	1
0	0	0
1	0	1
1	0	−1
1	−1	0
1	1	0
−1	1	0
−1	−1	0
−1	0	−1
−1	0	1
0	0	0
0	−1	−1
0	−1	1
0	0	0

that, when the factors whose levels are most difficult to change are reset, this is also done for the remaining hard-to-change factors.

Real-life examples, however, show that it might also be interesting to reset the various hard-to-change factor levels at different points in time (especially when the levels of some hard-to-change factors are easier to change than others). Such an example is described in Webb, Lucas, and Borkowski (2004). The experiment in the example was conducted at a computer component manufacturing company and aimed at improving the performance of a wrapper machine. Three factors were involved: the spacing of a seal crimper, the speed at which the machine is run, and the temperature of the crimper. The experimenters used a 15-run Box–Behnken design. From the beginning, they considered the levels of the factor spacing as being difficult to change and, therefore, that factor was set only four times. However, when performing the experiment, it became clear that the factor speed was also hard to set, but less so than the factor spacing. As a consequence, the experimenters decided to set the speed only eight times. This resulted in a design, displayed in Table 1, where the factor speed was reset more often than the factor spacing because its level was less hard to change. A key feature of the design is that the levels of the factors speed and spacing were reset at different points in time.

In a second real-life example, researchers did a series of experiments involving two hard-to-change factors, one of which is less difficult to reset than the other. The experiments were concerned with an atmospheric plasma deposition of antibacterial coatings. This is, for example, useful for the lining of a refrigerator and for medical devices. The hard-to-change factors in these experiments are the gap between an electrode and the sample surface, and the frequency of a transformer. The reason why the levels of these factors are hard to change is that a technician is required to modify them. However, the gap is more difficult to change than the frequency, because resetting the gap involves adaptations to the pressure to guarantee favorable conditions for switching on the discharge. To change the frequency, the technician only has to plug in a new transformer. The levels of the other factors, such as power, gas flow rate, and precursor injection, are easy to change and therefore they were reset independently for each run. The experimenters' primary interest was in the main effects and the two-factor interactions of the five factors. The experimenters' preference was to use a 2^5 full factorial design, but they desired a cost-efficient run order for the design.

These two examples show that there is a need in practice for designs that take into account the fact that some of the hard-to-change factors are easier to reset than others and allow the hard-to-change factors to be reset at different points in time. In this article, we introduce a new type of design, called a staggered-level design, that provides the possibility to do this in such a way that the experiment is not only statistically efficient but also cost-efficient. The design also remediates some of the drawbacks of split-plot and split-split-plot designs. It should be noted, however, that the staggered-level designs presented here are useful only when the number of settings of the hard-to-change factor levels is not dictated by the physicalities of the experiment, such as oven sizes or batch sizes.

The remainder of this article focuses on two-level factorial designs and the estimation of models containing main effects and two-factor interaction effects only. In the next section, we describe in general the model used in Webb, Lucas, and Borkowski (2004) for data from experiments involving several hard-to-change factors, some of which are less difficult to reset than others. Next, we discuss the structure and the construction principles of the staggered-level designs. We also compare the staggered-level designs with split-plot and split-split-plot designs, and we investigate the sensitivity of our results to the relative magnitude of the variance components in the model. In Sections 5 and 6, we construct fractional factorial staggered-level designs and staggered-level designs with more than two hard-to-change factors, using a coordinate-exchange algorithm and a Bayesian approach involving a prior distribution about the model's variance components. In Section 7, we discuss the performance of the staggered-level designs in case interval estimates rather than point estimates are desired. Finally, in Section 8, we provide a conclusion and discuss the possible ways in which the staggered-level design can be randomized.

2. MODEL

In the experimental designs considered in this article, the set of f factors is divided in three groups: the f_1 most hard-to-change factors (in the following sections referred to as class-1 hard-to-change factors and denoted by $w_1, \ldots, w_{f_1}$), the f_2 less hard-to-change factors (in the following sections referred to as class-2 hard-to-change factors and denoted by $s_1, \ldots, s_{f_2}$), and the remaining $f_3 = f - f_1 - f_2$ easy-to-change factors, $t_1, \ldots, t_{f_3}$. The experimental runs are partitioned in two ways, one for each class of hard-to-change factors. Such a partitioning is illustrated in Table 1, where the settings of the factor spacing divide the runs in four subsets and the settings of the factor speed divide the runs in eight different subsets. To capture the dependence between runs for which the class-1 hard-to-change factors $w_1, \ldots, w_{f_1}$ are not independently reset, we include random effects δ_i, $i = 1, \ldots, r$, in the model for each of the r independent settings of the class-1 hard-to-change factors.

To capture the dependence between runs for which the class-2 hard-to-change factors $s_1, \ldots, s_{f_2}$ are not independently reset, we also include random effects γ_j, $j = 1, \ldots, g$, in the model for each of the g independent settings of the class-2 hard-to-change factors. The kth response, $k = 1, \ldots, n$, obtained at the ith setting of $w_1, \ldots, w_{f_1}$ and the jth setting of $s_1, \ldots, s_{f_2}$ can then be written as

$$\begin{aligned} Y_{ijk} = \beta_0 &+ \sum_{l=1}^{f_1} \beta_{w_l} w_{li} + \sum_{l=1}^{f_2} \beta_{s_l} s_{lj} + \sum_{l=1}^{f_3} \beta_{t_l} t_{lk} \\ &+ \sum_{l=1}^{f_1-1} \sum_{m=l+1}^{f_1} \beta_{w_l w_m} w_{li} w_{mi} + \sum_{l=1}^{f_2-1} \sum_{m=l+1}^{f_2} \beta_{s_l s_m} s_{lj} s_{mj} \\ &+ \sum_{l=1}^{f_1} \sum_{m=1}^{f_2} \beta_{w_l s_m} w_{li} s_{mj} + \sum_{l=1}^{f_1} \sum_{m=1}^{f_3} \beta_{w_l t_m} w_{li} t_{mk} \\ &+ \sum_{l=1}^{f_2} \sum_{m=1}^{f_3} \beta_{s_l t_m} s_{lj} t_{mk} + \sum_{l=1}^{f_3-1} \sum_{m=l+1}^{f_3} \beta_{t_l t_m} t_{lk} t_{mk} \\ &+ \delta_i + \gamma_j + \varepsilon_k, \end{aligned} \quad (1)$$

where δ_i represents the random effect of the ith setting of the class-1 hard-to-change factors $w_1, \ldots, w_{f_1}$, γ_j is the random effect of the jth setting of the class-2 hard-to-change factors $s_1, \ldots, s_{f_2}$, ε_k is the random error of the response measured at the kth experimental run, and n denotes the number of observations.

In matrix notation, Equation (1) can be written as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_\delta \boldsymbol{\delta} + \mathbf{Z}_\gamma \boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \quad (2)$$

where $\mathbf{Y}$ is the $n \times 1$ vector containing the n responses of the experiment, $\boldsymbol{\beta}$ is a $p \times 1$ vector that contains the $p = 1 + f + f(f-1)/2$ model parameters, $\mathbf{X}$ is the $n \times p$ model matrix (containing the settings of all the factors and their pairwise cross-products), $\mathbf{Z}_\delta$ is the $n \times r$ matrix with (k, i)th entry equal to 1 if the kth run is conducted at the ith setting of the class-1 hard-to-change factors and equal to 0 otherwise, $\mathbf{Z}_\gamma$ is the $n \times g$ matrix with (k, j)th entry equal to 1 if the kth run is conducted at the jth setting of the class-2 hard-to-change factors and equal to 0 otherwise, $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are the $r \times 1$ and $g \times 1$ vectors containing the random effects associated with the independent settings of the class-1 hard-to-change and the class-2 hard-to-change factors, respectively, and $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of random errors.

We assume that $\mathrm{E}(\boldsymbol{\delta}) = \mathbf{0}_r$, $\mathrm{cov}(\boldsymbol{\delta}) = \sigma_\delta^2 \mathbf{I}_r$, $\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_g$, $\mathrm{cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_g$, $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n$, $\mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_n$, $\mathrm{cov}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \mathbf{0}_{r \times g}$, $\mathrm{cov}(\boldsymbol{\delta}, \boldsymbol{\varepsilon}) = \mathbf{0}_{r \times n}$, and $\mathrm{cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{g \times n}$, where $\mathbf{0}_c$ and $\mathbf{I}_c$ represent a c-dimensional zero vector and identity matrix, respectively, and $\mathbf{0}_{c \times d}$ is a zero matrix of dimension $c \times d$. The assumed variance-covariance matrix of the responses, $\mathbf{Y}$, then is

$$\mathbf{V} = \sigma_\varepsilon^2 (\mathbf{I}_n + \eta_\delta \mathbf{Z}_\delta \mathbf{Z}_\delta' + \eta_\gamma \mathbf{Z}_\gamma \mathbf{Z}_\gamma'), \quad (3)$$

where η_δ and η_γ are the variance ratios $\sigma_\delta^2/\sigma_\varepsilon^2$ and $\sigma_\gamma^2/\sigma_\varepsilon^2$ for the class-1 hard-to-change factors and the class-2 hard-to-change factors, respectively. The larger these ratios, the stronger the correlation between runs conducted at the same setting of the class-1 and/or the class-2 hard-to-change factors.

The statistical model in Equation (2) generalizes the split-plot model and the split-split-plot model. For the model to reduce to the split-plot model, it is required that $\mathbf{Z}_\gamma = \mathbf{Z}_\delta$, which means that the class-1 hard-to-change factors and the class-2 hard-to-change factors are reset at the same points in time. In that case, the variance components σ_δ^2 and σ_γ^2 cannot be estimated separately. Only their sum is estimable then. This problem does not occur in split-split-plot designs, where $\mathbf{Z}_\gamma = \mathbf{I}_r \otimes \mathbf{1}_{c_1}$ and $\mathbf{Z}_\delta = \mathbf{I}_g \otimes \mathbf{1}_{c_2}$, with $\otimes$ being the Kronecker product, $\mathbf{1}_{c_i}$ a c_i-dimensional vector of ones, c_1 and c_2 the number of runs in a whole plot and a subplot, respectively, $n = rc_1 = gc_2$, and c_1 being a multiple of c_2.

Under the assumption of normality, the maximum likelihood estimator of the unknown model parameter vector $\boldsymbol{\beta}$ is the GLS estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}. \quad (4)$$

The variance-covariance matrix of this estimator can be expressed as

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}, \quad (5)$$

and the information matrix on the unknown parameter $\boldsymbol{\beta}$ is given by

$$\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}. \quad (6)$$

A criterion to select designs is the D-optimality criterion, which seeks designs that maximize the determinant of the information matrix. Throughout this article, we report the pth root of the determinant of the information matrix, as is customary in the design of experiments literature. To compare two designs with information matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ in terms of the D-optimality criterion, we use the D-efficiency $(|\mathbf{M}_1| / |\mathbf{M}_2|)^{1/p}$. Another optimality criterion we use to compare the various designs presented in this article is the A-optimality criterion. This criterion seeks designs that minimize the sum of the variances of the parameter estimators, given by the trace of the variance-covariance matrix in Equation (5).

Note that optimal designs as well as the relative performance of two designs depend on the relative magnitude of the variance components through $\mathbf{V}$. This is why we perform a sensitivity study in Section 4.2 and use a Bayesian approach to find D-optimal designs in Section 5.1.

3. STRUCTURE OF THE STAGGERED-LEVEL DESIGN

3.1 A 32-Run Example

The key feature of the staggered-level design is the fact that the levels of the class-2 hard-to-change factors are reset at entirely different points in time than the levels of the class-1 hard-to-change factors. To accomplish this, the staggered-level design has a certain structure, which we describe in this section by means of the 32-run five-factor staggered-level design displayed in Table 2 and suitable for a model involving main effects and two-factor interactions. The five factors are one class-1 hard-to-change factor w, one class-2 hard-to-change factor s, and three easy-to-change factors t_1, t_2, and t_3.

The staggered-level design in Table 2 can be easily constructed by hand using the following construction method. In the 32-run staggered-level design, the runs are divided in four

Table 2. Thirty-two-run staggered-level design for a model including main effects and two-factor interactions of two hard-to-change factors, w and s, and three easy-to-change factors t_1, t_2, and t_3

Run	w	s	t_1	t_2	t_3	Run	w	s	t_1	t_2	t_3
1	−1	1	1	1	−1	17	−1	1	1	1	1
2	−1	1	1	−1	1	18	−1	1	1	−1	−1
3	−1	1	−1	1	1	19	−1	1	−1	1	−1
4	−1	1	−1	−1	−1	20	−1	1	−1	−1	1
5	−1	−1	1	1	1	21	−1	−1	1	1	−1
6	−1	−1	1	−1	−1	22	−1	−1	1	−1	1
7	−1	−1	−1	1	−1	23	−1	−1	−1	1	1
8	−1	−1	−1	−1	1	24	−1	−1	−1	−1	−1
9	1	−1	1	1	1	25	1	−1	1	1	−1
10	1	−1	1	−1	−1	26	1	−1	1	−1	1
11	1	−1	−1	1	−1	27	1	−1	−1	1	1
12	1	−1	−1	−1	1	28	1	−1	−1	−1	−1
13	1	1	1	1	−1	29	1	1	1	1	1
14	1	1	1	−1	1	30	1	1	1	−1	−1
15	1	1	−1	1	1	31	1	1	−1	1	−1
16	1	1	−1	−1	−1	32	1	1	−1	−1	1

subsets of size eight created by the settings of the class-1 hard-to-change factor w, and the factor w is at its low level in the first subset. The runs are also divided in subsets through the settings of the class-2 hard-to-change factor s. This division begins and ends with a subset of runs half as large as the subsets defined by the settings of w, that is, with subsets of size four, and with the factor s at its high level. The remaining subsets are of the same size as the subsets determined by w. This results in the structure shown in Table 3, where $\mathbf{1}_d$ represents a d-dimensional vector of ones. The strength of the staggered-level design lies in this specific ordering of the subsets. That ordering ensures that the two levels of the class-2 hard-to-change factor s can be compared with each other within each of the four subsets created by the settings of the class-1 hard-to-change factor. Similarly, the two levels of the class-1 hard-to-change factor can be compared with each other in three subsets formed by the settings of the class-2 hard-to-change factor.

The remaining problem is to determine the levels of the three easy-to-change factors t_1, t_2, and t_3 for each run. It is clear from Table 3 that the design for the easy-to-change factors consists of blocks of size four. Selecting the levels of t_1, t_2, and t_3 therefore comes down to arranging the runs of a quadruplicated 2^3 factorial design in eight blocks of size four. This can be done using block generator $B = t_1t_2t_3$ for each of the four 2^3 factorial designs, and leads to two types of blocks. The first block type involves a 2^{3-1} fraction for which $t_1t_2t_3 = 1$, while the second block type involves a 2^{3-1} fraction for which $t_1t_2t_3 = -1$.

Table 3. Basic structure of the 32-run staggered-level design in Table 2

w	s	t_1, t_2, t_3
$-\mathbf{1}_8$	$\mathbf{1}_4$	Block 1 or 2
	$-\mathbf{1}_8$	Block 1 or 2
$\mathbf{1}_8$		Block 1 or 2
	$\mathbf{1}_8$	Block 1 or 2
$-\mathbf{1}_8$		Block 1 or 2
	$-\mathbf{1}_8$	Block 1 or 2
$\mathbf{1}_8$		Block 1 or 2
	$\mathbf{1}_4$	Block 1 or 2

Each block type is then combined with each possible combination of levels of w and s to ensure that the entire design contains all the points of a 2^5 full factorial design. When these two blocks are used, the columns of $\mathbf{X}$ corresponding to the main effects of t_1, t_2, and t_3 are orthogonal to the subsets of runs determined by the settings of w and s, just like the columns corresponding to all two-factor interactions involving at least one of the three easy-to-change factors. Therefore, the assignment of the blocks to the eight available positions has no impact on the quality of the design if the interest is in the main-effects-plus-two-factor-interactions model.

3.2 General Construction Principles for Two-Level Full Factorial Designs

In general, the construction of a two-level staggered-level design works as follows. The settings of the class-1 hard-to-change factors $w_1, \ldots, w_{f_1}$ divide the experimental runs in groups with a size that is a power of two. The settings of the class-2 hard-to-change factors divide the runs in groups of two different sizes. The first and last group are half as large as the groups determined by $w_1, \ldots, w_{f_1}$, whereas the other groups have the same size. The exact sizes of the groups formed by the settings of both types of hard-to-change factors increase with the number of factors. Since the size of the groups formed by the class-1 hard-to-change factors determines the size of the blocks for the easy-to-change factors, we recommended selecting a size for which there will be no confounding between the two-factor interaction effects of the easy-to-change factors and the subsets of runs dictated by the hard-to-change factors. This is the reason why, for the 32-run staggered-level design in Table 2, groups of runs of size eight and thus four independent settings of the class-1 hard-to-change factor w were chosen. In some cases, using a sufficiently large group size may be impossible, for instance due to the available budget.

Once the number of independent settings of both groups of hard-to-change factors has been determined, a block generator can be used to arrange the various easy-to-change factor-level combinations in blocks of the required size and to spread these blocks over the combinations of the class-1 and class-2 hard-to-change factor levels. For some staggered-level designs, such as the 32-run staggered-level design in Table 2, the assignment of the blocks has no impact on the D- or A-efficiency, while, for other staggered-level designs, it does. The latter is, for example, the case for a 16-run staggered-level design for a model including main effects and two-factor interactions in two hard-to-change factors, w and s, and two easy-to-change factors t_1 and t_2. This design involves four settings of w and five settings of s. As a result, the required block size for t_1 and t_2 is two. Consequently, the two-factor interaction effect involving t_1 and t_2 cannot be made orthogonal to the subsets of runs determined by the independent settings of w and s, and the ordering of the blocks matters. The D-optimal ordering for $\eta_\delta = 3$ and $\eta_\gamma = 2$ is shown in Table 4.

Table 4. Sixteen-run staggered-level design for a model including main effects and two-factor interactions of two hard-to-change factors, w and s, and two easy-to-change factors t_1 and t_2

Run	w	s	t_1	t_2
1	−1	1	1	1
2	−1	1	−1	−1
3	−1	−1	1	−1
4	−1	−1	−1	1
5	1	−1	1	1
6	1	−1	−1	−1
7	1	1	−1	1
8	1	1	1	−1
9	−1	1	−1	1
10	−1	1	1	−1
11	−1	−1	1	1
12	−1	−1	−1	−1
13	1	−1	1	−1
14	1	−1	−1	1
15	1	1	−1	−1
16	1	1	1	1

4. COMPARISON WITH SPLIT-PLOT AND SPLIT-SPLIT-PLOT DESIGNS IN TERMS OF STATISTICAL EFFICIENCY

In this section, we compare the 32-run staggered-level design in Table 2 with a split-plot design and with a split-split-plot design in terms of the D- and A-optimality criteria. A comparison of the designs in terms of cost can be found in the supplementary materials. In the D-optimal 32-run split-plot design, both w and s are set eight times. This design, which is shown in the online supplementary materials, is a 2^5 factorial design arranged in eight whole plots of size four using the contrast columns for w, s, and $t_1t_2t_3$. The split-plot design treats both hard-to-change factors alike by using eight independent settings for each of them, ignoring the fact that the first hard-to-change factor, w, is harder to change than the second, s.

A design that does not treat the two hard-to-change factors alike and uses fewer independent settings of the class-1 hard-to-change factor than of the class-2 hard-to-change factor is a split-split-plot design. The D-optimal 32-run split-split-plot design involving four settings of w and eight settings of s, which is also shown in the supplementary materials, was found using the algorithmic approach of Jones and Goos (2009). It can be constructed combinatorially by using the contrast columns for w and $t_1t_2t_3$ to define the whole plots and the contrast column for s to define the subplots within the whole plots.

4.1 Comparison in Terms of Statistical Efficiency

The staggered-level design is, in many cases, not only more cost-efficient than the split-plot and split-split-plot designs (can be found in the online supplementary materials), but also better in terms of D- and A-efficiency. Detailed computational results for the 32-run designs are given in Table 5. The results displayed were obtained by assuming that $\eta_\delta = 3$, $\eta_\gamma = 2$, and $\sigma_\varepsilon^2 = 1$, but different values for these parameters yield very similar results (see also the sensitivity study in Section 4.2).

Comparing the staggered-level design with the split-plot design in terms of the variance of the parameter estimates in

Table 5. Variances of estimates of fixed model parameters along with D- and A-optimality criterion values for the 32-run staggered-level design in Table 2 and the split-plot and split-split-plot alternatives when $\eta_\delta = 3$, $\eta_\gamma = 2$, and $\sigma_\varepsilon^2 = 1$

Effect	Staggered	Split-plot	Split-split-plot
w	0.823	0.656	1.031
s	0.451	0.656	0.281
t_1	0.031	0.031	0.031
t_2	0.031	0.031	0.031
t_3	0.031	0.031	0.031
ws	0.073	0.656	0.281
Other	0.031	0.031	0.031
D-criterion	16.710	14.948	15.706
A-criterion	2.923	3.000	3.000

Table 5, it is clear that the main effect of the class-2 hard-to-change factor s as well as the interaction effect of the two hard-to-change factors w and s are estimated more precisely from the staggered-level design. The split-plot design allows a more precise estimation of the main effect of w, as a result of the double number of settings of w in comparison with the staggered-level design and the split-split-plot design.

Compared with the split-split-plot design, the staggered-level design leads to more precise estimates of the main effect of w and the interaction effect of w and s. Only the main effect of s is estimated less precisely from the staggered-level design than from the split-split-plot design. Obviously, this is because the number of independent settings of the class-2 hard-to-change factor s is considerably higher in the split-split-plot design. With all three designs, the effects involving the easy-to-change factors are estimated with the smallest possible variance, that is, $1/32 = 0.031$.

The staggered-level design performs best in terms of the D-optimality criterion with a D-criterion value of 16.710. The split-plot design and the split-split-plot design have D-criterion values of 14.948 and 15.706, respectively. The staggered-level design is thus 11.7% better than the split-plot design and 6.3% better than the split-split-plot design in terms of the D-optimality criterion. Looking at the trace of the variance-covariance matrix of the factor effect estimates reveals that the staggered-level design also performs better in terms of the A-optimality criterion. It has an A-criterion value of 2.923 compared with a value of 3 for the split-plot as well as the split-split-plot design. As a result, the design that is cheapest to conduct is, overall, also statistically most efficient.

The correlation matrix of the factor effect estimates from the staggered-level design demonstrates another nice feature of the design. Only 2 out of the 120 pairwise correlations are different from zero. First, there is a negligible correlation of 0.109 between the estimate of the main effect β_s of the class-2 hard-to-change factor, s, and the estimate of the intercept β_0. Second, there is a small negative correlation of −0.163 between the estimate of the main effect β_s and the estimate of the interaction effect β_{ws} involving the two hard-to-change factors w and s. The correlations involve hard-to-change factor effects only. The important thing, however, is that nearly all model parameters can be estimated independently. As a result, the staggered-level design performs almost equally well in this respect as the

Table 6. D-efficiencies of the 32-run staggered-level design relative to the split-plot and split-split-plot designs for various values of η_δ and η_γ

η_δ	η_γ	Split-plot	Split-split-plot
0.1	0.1	1.013	1.006
	1	1.098	1.098
	10	1.384	1.384
1	0.1	1.109	1.004
	1	1.082	1.052
	10	1.235	1.233
10	0.1	1.408	1.004
	1	1.261	1.038
	10	1.137	1.098

split-plot and the split-split-plot design, whose correlation matrices equal the identity matrix.

4.2 Sensitivity Study

The relative performances of the competing designs in terms of D- and A-optimality depend on the two variance ratios η_δ and η_γ. Therefore, it is necessary to investigate the effects of changing the values of these two variance ratios, so that we can confirm that the staggered-level design outperforms the alternatives in all practical instances. We consider nine situations, involving small, average, and large η_δ and η_γ values. We investigate situations where $\eta_\delta > \eta_\gamma$, which is the most realistic scenario given that η_δ corresponds to the factor that is hardest to change, as well as situations where the opposite is true. The results for the sensitivity study for the 32-run designs are shown in Table 6.

For an η_δ value of 0.1, the differences between the staggered-level design and the alternative designs increase with η_γ. The staggered-level design is about 38% better in terms of the D-optimality criterion than both the split-plot and split-split-plot design when η_γ is as large as 10. When η_δ is equal to 1, the staggered-level design performs between 8% and 23% better than the split-plot design in terms of the D-optimality criterion. The D-efficiency of the staggered-level design relative to the split-split-plot design is substantially smaller when η_γ equals only 0.1, but it goes up to 23% when η_γ is equal to 10. Finally, when η_δ is as large as 10, the staggered-level design is considerably better than the split-plot design, for any value of η_γ. The difference between the staggered-level design and the split-split-plot design is negligible when η_γ equals 0.1 and goes up to 10% when η_γ is 10. It is clear that the sensitivity study leads to the conclusion that the staggered-level design is, in all situations considered, statistically the best option when the model to estimate contains all main effects and two-factor interaction effects.

In this article, we focus on models involving main effects and all two-factor interaction effects. For a main effects model, there are conditions for which the 32-run split-plot and split-split-plot designs perform better than the staggered-level design in terms of the D-optimality criterion. For the split-plot design, this is the case when the two variance ratios are about equal. For the split-split-plot design, this is true when η_δ is equal to or larger than η_γ. However, there are still many situations where the staggered-level design remains the most efficient option, for instance when $\eta_\delta = 3$ and $\eta_\gamma = 2$.

5. FRACTIONAL FACTORIAL STAGGERED-LEVEL DESIGNS

Consider an experiment with one class-1 hard-to-change factor w, one class-2 hard-to-change factor s, and four easy-to-change factors t_1, t_2, t_3, and t_4. Given the results from the previous section, a natural design choice would be a staggered-level design using a regular half fraction of a 2^6 factorial design with defining relation $t_4 = wst_1t_2t_3$. It turns out, however, that it is preferable in terms of the D- and A-optimality criteria to use a nonregular half fraction. To find the best possible half fraction, we implemented a coordinate-exchange algorithm that seeks the D-optimal factor levels for any given staggering structure. The algorithm allows us to generate efficient staggered-level designs for a broad range of experimental scenarios.

5.1 A Bayesian D-Optimality Criterion

Since the D-optimality criterion makes use of the information matrix in Equation (6), the D-optimal staggered-level designs depend on the variance-covariance matrix $\mathbf{V}$ of the responses and thus on the variance ratios η_δ and η_γ. In the literature, the standard way to deal with this is to assume some prior value of η_δ and η_γ and find a design that is D-optimal for that value (see, for instance, Goos and Vandebroek 2003 and Jones and Goos 2007, 2009, 2012a). In this article, however, we construct staggered-level designs that are D-optimal over a prior distribution for η_δ and η_γ and apply a Bayesian D-optimality criterion. Denoting the prior distributions for η_δ and η_γ by $\pi_\delta(\eta_\delta)$ and $\pi_\gamma(\eta_\gamma)$, the Bayesian D-optimality criterion we implemented is

$$D_B = \int_0^{+\infty} \int_0^{+\infty} \ln(D(\eta_\delta, \eta_\gamma)) \cdot \pi_\delta(\eta_\delta) \cdot \pi_\gamma(\eta_\gamma) \cdot d\eta_\delta \cdot d\eta_\gamma, \quad (7)$$

where $D(\eta_\delta, \eta_\gamma) = \sigma_\varepsilon^{-2p} |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|$. For the sake of illustration, we assume that the prior distributions $\pi_\delta(\eta_\delta)$ and $\pi_\gamma(\eta_\gamma)$ of the two variance ratios are independent. In practice, this assumption might not be realistic, and a joint prior distribution for η_δ and η_γ might be more appropriate.

5.2 Incorporating Prior Information About η_δ and η_γ

As the variance ratios η_δ and η_γ can only take positive values, a natural choice for their prior distribution is a log-normal distribution. The log-normal distribution has parameters μ and ν, which will often be different for the two variance ratios η_δ and η_γ. The Bayesian D-optimal design criterion then becomes

$$D_B = \int_0^{+\infty} \int_0^{+\infty} \ln(D(\eta_\delta, \eta_\gamma)) \cdot \frac{1}{2\pi \eta_\delta \eta_\gamma \nu_\delta \nu_\gamma} \cdot e^{-\{\ln(\eta_\delta)-\mu_\delta\}^2/2\nu_\delta^2} \cdot e^{-\{ln(\eta_\gamma)-\mu_\gamma\}^2/2\nu_\gamma^2} \cdot d\eta_\delta \cdot d\eta_\gamma, \quad (8)$$

where μ_δ and ν_δ are the parameters of the log-normal prior distribution $\pi_\delta(\eta_\delta)$ of η_δ, and μ_γ and ν_γ are the parameters of the log-normal distribution $\pi_\gamma(\eta_\gamma)$ of η_γ. As there is no analytical expression for this double integral, it has to be approximated numerically.

This can be done using Gauss–Hermite quadrature (see, for example, Bliemer, Rose, and Hess (2009) and Yu, Goos, and Vandebroek (2010) for applications of this technique in transportation research), which means that the Bayesian optimal design criterion is approximated as

$$D_B \approx \frac{1}{\pi} \cdot \sum_{i=1}^{R} \sum_{j=1}^{R} w_i w_j \ln(D(e^{\mu_\delta + a_i \nu_\delta \sqrt{2}}, e^{\mu_\gamma + a_j \nu_\gamma \sqrt{2}})),$$

where the R abscissas a_i and a_j and the corresponding weights w_i and w_j are obtained from the Hermite polynomials. We used eight abscissas to achieve a good approximation of the two-dimensional integral. The eight abscissas and their weights are given in Kythe and Schäferkotter (2005).

When using a log-normal prior distribution for a variance ratio, the parameters μ and ν need to be specified. A key feature of the log-normal distribution is that 99.7% of its probability mass lies in the interval $[m/q^3, mq^3]$, where $m = e^{\mu}$ and $q = e^{\nu}$. If it is highly likely that a variance ratio η is in the interval $[0.1, 10]$, then good choices for μ and ν are 0 and $\ln(\sqrt[3]{10})$. These are also the values we used in our computations for both variance ratios.

5.3 Coordinate-Exchange Algorithm

The input to the algorithm for constructing staggered-level designs consists of

- the desired number of runs n,
- the a priori model,
- the model terms,
- the number of model parameters p,
- the number of experimental factors f,
- the class-1 hard-to-change factors,
- the class-2 hard-to-change factors,
- the number, r, of independent settings of the levels of the class-1 hard-to-change factors (this implies that the levels of the class-2 hard-to-change factors are set $g = r + 1$ times),
- the R abscissas and weights for the Gauss–Hermite quadrature,
- the parameters μ_δ, ν_δ, μ_γ, and ν_γ of the log-normal prior distributions for η_δ and η_γ, and
- the number of random starting designs, T.

To a large extent, our algorithm resembles the original coordinate-exchange algorithm of Meyer and Nachtsheim (1995) and the algorithms of Jones and Goos (2007, 2009, 2012a) and Arnouts, Goos, and Jones (2010) for split-plot, split-split-plot, and strip-plot designs. It first generates a random starting design. For each class-1 hard-to-change factor, r random numbers between -1 and 1 are chosen. Each of these random numbers is used as the level of the class-1 hard-to-change factor for all n/r runs in one of the subsets formed by the settings of the class-1 hard-to-change factor(s). For each class-2 hard-to-change factor, $g = r + 1$ random numbers are chosen. Each of these random numbers is used as the level of the class-2 hard-to-change factor for all the runs in one of the subsets formed by the settings of the class-2 hard-to-change factor(s). For each easy-to-change factor, a random number is chosen for each of the n runs. As a result of this procedure, all runs in a subset formed by the settings of any hard-to-change factor have the same random level for that hard-to-change factor.

In its second step, the coordinate-exchange algorithm improves the random design by considering changes to individual factor levels, which can be viewed as coordinates. For each level of an easy-to-change factor, the Bayesian D-optimality criterion is evaluated over a discrete number of values going from -1 to 1. If the resulting value of the Bayesian D-optimality criterion is better than the current value, the factor's level is replaced. For the level of a hard-to-change factor, the procedure is slightly more complicated. If such a factor level changes, then all the other levels of that factor in the same subset must also change. When the value of the Bayesian D-optimality criterion improves due to the change, the hard-to-change factor's level is changed in all the runs in the same subset.

This procedure for the hard-to-change and easy-to-change factors continues until no changes are made in a whole pass through the entire design. The algorithm repeats the starting and improvement phase T times. Each time, the value of the Bayesian D-optimality criterion found in the current iterate is compared with the best value from previous iterates. If the current value is better, then the current design is stored.

5.4 2^{6-1} Fractional Factorial Staggered-Level Design

The algorithm enables us to construct a 32-run staggered-level design for an experiment with one class-1 hard-to-change factor w, one class-2 hard-to-change factor s, and four easy-to-change factors t_1, t_2, t_3, and t_4. The 32-run design in Table 7 with four settings of the level of w and five settings of the level of s is the best one obtained using our algorithm for estimating all parameters of a main-effects-plus-two-factor-interactions model taking into account the log-normal prior distributions for the variance ratios η_δ and η_γ.

Table 7. D-optimal 32-run staggered-level design for a model including main effects and two-factor interactions of two hard-to-change factors w and s, and four easy-to-change factors t_1, t_2, t_3, and t_4

Run	w	s	t_1	t_2	t_3	t_4	Run	w	s	t_1	t_2	t_3	t_4
1	1	−1	−1	1	−1	−1	17	1	−1	−1	−1	−1	1
2	1	−1	1	1	1	1	18	1	−1	1	1	−1	−1
3	1	−1	1	−1	−1	1	19	1	−1	−1	1	1	1
4	1	−1	−1	−1	1	−1	20	1	−1	1	−1	1	−1
5	1	1	−1	−1	1	1	21	1	1	1	1	−1	1
6	1	1	−1	1	−1	1	22	1	1	−1	−1	−1	−1
7	1	1	1	1	1	−1	23	1	1	1	−1	1	1
8	1	1	1	−1	−1	−1	24	1	1	−1	1	1	−1
9	−1	1	−1	1	−1	−1	25	−1	1	1	1	1	1
10	−1	1	1	−1	−1	1	26	−1	1	1	1	−1	−1
11	−1	1	−1	1	1	1	27	−1	1	−1	−1	1	−1
12	−1	1	1	−1	1	−1	28	−1	1	−1	−1	−1	1
13	−1	−1	−1	−1	−1	−1	29	−1	−1	−1	1	−1	1
14	−1	−1	−1	−1	1	1	30	−1	−1	1	−1	1	1
15	−1	−1	1	1	−1	1	31	−1	−1	1	−1	−1	−1
16	−1	−1	1	1	1	−1	32	−1	−1	−1	1	1	−1

Table 8. Variances of estimates of fixed model parameters along with D- and A-optimality criterion values for the 32-run staggered-level design in Table 7, the split-plot design, the split-split-plot design, and a staggered-level design obtained from a regular 2^{6-1} fractional factorial design when $\eta_\delta = 3$, $\eta_\gamma = 2$, and $\sigma_\varepsilon^2 = 1$

Effect	Staggered	Split-plot	Split-split-plot	Regular fraction
w	0.825	0.656	1.031	0.825
s	0.453	0.656	0.281	0.455
t_1	0.031	0.031	0.031	0.031
t_2	0.031	0.031	0.031	0.031
t_3	0.031	0.031	0.031	0.037
t_4	0.031	0.031	0.031	0.037
ws	0.075	0.656	0.281	0.075
wt_1	0.031	0.031	0.031	0.031
wt_2	0.031	0.031	0.031	0.031
wt_3	0.031	0.031	0.031	0.037
wt_4	0.031	0.031	0.031	0.037
st_1	0.031	0.031	0.031	0.031
st_2	0.031	0.031	0.031	0.031
st_3	0.031	0.031	0.031	0.037
st_4	0.031	0.031	0.031	0.037
t_1t_2	0.055	0.031	0.056	0.045
t_1t_3	0.031	0.041	0.041	0.037
t_1t_4	0.055	0.031	0.041	0.078
t_2t_3	0.031	0.041	0.031	0.071
t_2t_4	0.031	0.031	0.031	0.037
t_3t_4	0.031	0.060	0.031	0.045
D-criterion	18.949	17.428	18.126	17.475
A-criterion	3.165	3.235	3.231	3.279

As benchmarks for the staggered-level design, we use the D-optimal 32-run split-plot design with eight settings of the levels of both hard-to-change factors, and the D-optimal 32-run split-split-plot design involving four settings of the level of w and eight settings of the level of s. These two designs are given in the supplementary materials. To compare the staggered-level design with its competitors, we set the values of η_δ, η_γ, and σ_ε^2 to 3, 2, and 1, as previously. The same conclusions as in Section 4 can be drawn when looking at the variances of the parameter estimates, which are displayed in Table 8. The staggered-level design from Table 7 results in a more precise estimation of the interaction effect of w and s than the two other designs. The main effect of w and the main effect of s are estimated more precisely from the split-plot and split-split-plot design, respectively. For the former design, this is due to a larger number of settings of the class-1 hard-to-change factor's level. For the latter design, this is due to a larger number of settings of the class-2 hard-to-change factor's level. Furthermore, with the staggered-level design, 16 of the 18 effects involving the easy-to-change factors are estimated with the best possible precision for a design involving 32 runs and a σ_ε^2 value of 1, that is, with variance $1/32 = 0.031$. Only the interaction effects between the easy-to-change factors t_1 and t_2, and between t_1 and t_4 are estimated less precisely. For the split-plot and the split-split-plot design, 15 of the effects involving the easy-to-change factors are estimated with the lowest possible variance. In terms of the D-optimality criterion, the staggered-level design is 8.7% and 4.5% better than the split-plot design and the split-split-plot design, respectively. In terms of the A-optimality criterion, the staggered-level design outperforms the split-plot design by 2.2% and the split-split-plot design by 2%.

When taking a closer look at the factor-level combinations of the D-optimal staggered-level design in Table 7, it can be seen that they do not form a regular half fraction of a 2^6 design. Using a regular 2^{6-1} design leads to a staggered-level design that is inferior to the one in Table 7, both in terms of the D-optimality criterion and in terms of the A-optimality criterion. This can be seen by comparing the last column in Table 8 (containing the results for the regular fraction) with the column for the D-optimal staggered-level design in Table 7. The regular fraction was constructed using our coordinate-exchange algorithm under the constraint that $t_4 = wst_1t_2t_3$. The precision of estimation of the main effect of the class-1 hard-to-change factor w as well as the interaction effect between both hard-to-change factors w and s is equally good for the regular and the nonregular design. The estimation of the main effect of the class-2 hard-to-change factor s is slightly more precise for the nonregular design. The main difference is made by the estimates of the interaction effects between the easy-to-change factors: many of these are less precise when the regular 2^{6-1} design is used than when the nonregular design in Table 7 is used.

6. DESIGNS WITH MORE HARD-TO-CHANGE FACTORS

The coordinate-exchange algorithm allows for more extensions of the staggered-level design. It is, for instance, possible to create staggered-level designs for experiments with more than two hard-to-change factors. As an illustration, we consider again a 32-run staggered-level design for an experiment with six factors. This time, however, there are three hard-to-change factors, two class-1 hard-to-change factors w_1 and w_2, and one class-2 hard-to-change factor s. The fact that there are two class-1 hard-to-change factors necessitates the use of more than four settings of the class-1 hard-to-change factors' levels to be able to estimate their main effects, their interaction effect, and the variance component corresponding to these hard-to-change factors, σ_δ^2. The 32-run staggered-level design for this scenario in Table 9 therefore has eight settings of the level of the two class-1 hard-to-change factors and nine settings of the level of the class-2 hard-to-change factor. The design points again form a nonregular 2^{6-1} fraction.

The alternative D-optimal split-split-plot design also involves eight settings of the class-1 hard-to-change factors' levels, but it has 16 settings of the class-2 hard-to-change factor's level. The D-optimal split-plot design for this situation has eight settings for all the hard-to-change factors and therefore is the least expensive design to run. These two benchmark designs are given in the supplementary materials. Comparing the three competing designs in terms of the D- and A-optimality criteria and in terms of the variances of the parameter estimates in Table 10, it can be seen that the split-plot design is not a very interesting option. The split-plot design is about 9% worse than the staggered-level design in terms of the D-optimality criterion and even 53% worse in terms of the A-optimality criterion. This is due to the poor estimation of all the hard-to-change factors' effects and their pairwise interactions in comparison with the staggered-level design.

Comparing the split-split-plot design and the staggered-level design, we can see that, on average, the hard-to-change

Table 9. D-optimal 32-run staggered-level design for estimating a model including main effects and two-factor interactions of two class-1 hard-to-change factors w_1 and w_2, one class-2 hard-to-change factor s, and three easy-to-change factors t_1, t_2, and t_3

Run	w_1	w_2	s	t_1	t_2	t_3	Run	w_1	w_2	s	t_1	t_2	t_3
1	−1	−1	1	1	1	−1	17	−1	−1	1	1	1	1
2	−1	−1	1	−1	−1	1	18	−1	−1	1	−1	−1	−1
3	−1	−1	−1	1	−1	1	19	−1	−1	−1	−1	1	1
4	−1	−1	−1	−1	1	−1	20	−1	−1	−1	1	−1	−1
5	1	−1	−1	−1	−1	1	21	−1	1	−1	1	1	−1
6	1	−1	−1	1	1	−1	22	−1	1	−1	−1	−1	1
7	1	−1	1	−1	1	−1	23	−1	1	1	−1	1	1
8	1	−1	1	1	−1	1	24	−1	1	1	1	−1	−1
9	1	1	1	−1	−1	−1	25	1	1	1	1	1	−1
10	1	1	1	1	1	1	26	1	1	1	−1	−1	1
11	1	1	−1	1	−1	1	27	1	1	−1	1	−1	−1
12	1	1	−1	−1	1	−1	28	1	1	−1	−1	1	1
13	−1	1	−1	1	1	1	29	1	−1	−1	−1	−1	−1
14	−1	1	−1	−1	−1	−1	30	1	−1	−1	1	1	1
15	−1	1	1	1	−1	1	31	1	−1	1	−1	1	1
16	−1	1	1	−1	1	−1	32	1	−1	1	1	−1	−1

factors' effects are estimated a bit more precisely when using the staggered-level design even though the split-plot design is more expensive. Also, using the staggered-level design, 12 of the 15 effects involving the easy-to-change factors are estimated with the largest possible precision, while for the split-split-plot design, this is the case for only eight effects. Additionally, the staggered-level design performs 9.6% better than the split-split-plot design in terms of the D-optimality criterion, and 1.8% better in terms of the A-optimality criterion. For both designs, the correlation matrix of the factor effect estimates is not diagonal. The staggered-level design has 31 nonzero off-diagonal elements, while the split-split-plot design has only three such elements. The maximum correlation for the staggered-level design is 0.640, while for the split-split-plot design it is only 0.333.

Table 10. Variances of estimates of fixed model parameters along with D- and A-optimality criterion values for the 32-run staggered-level design in Table 9 and the corresponding split-plot and split-split-plot designs when $\eta_\delta = 3$, $\eta_\gamma = 2$, and $\sigma_\varepsilon^2 = 1$

Effect	Staggered	Split-plot	Split-split-plot
w_1	0.522	0.656	0.531
w_2	0.454	0.656	0.531
s	0.264	0.656	0.156
t_1	0.031	0.031	0.047
t_2	0.031	0.031	0.031
t_3	0.031	0.031	0.031
w_1w_2	0.418	0.656	0.531
w_1s	0.148	0.656	0.156
w_1t_1	0.031	0.031	0.047
w_1t_2	0.031	0.031	0.031
w_1t_3	0.031	0.031	0.031
w_2s	0.105	0.656	0.156
w_2t_1	0.031	0.031	0.047
w_2t_2	0.031	0.031	0.031
w_2t_3	0.031	0.031	0.031
st_1	0.031	0.031	0.041
st_2	0.031	0.031	0.031
st_3	0.031	0.031	0.031
t_1t_2	0.044	0.031	0.086
t_1t_3	0.205	0.031	0.095
t_2t_3	0.129	0.031	0.156
D-criterion	13.400	12.150	12.220
A-criterion	3.302	5.063	3.362

A variation on the scenario with two class-1 hard-to-change factors is a scenario with two class-2 hard-to-change factors. We discuss this variation in the supplementary materials. In this example, we use 48 runs instead of 32.

7. PERFORMANCE IN THE CASE OF UNKNOWN VARIANCE COMPONENTS

The results we discussed so far assume that the variance components σ_δ^2, σ_γ^2, and σ_ε^2 are known. However, in almost all cases, these variances have to be estimated from the experimental data. In that case, the variance-covariance matrix in Equation (5) is only asymptotically valid, except for those parameters for which the OLS estimator is equivalent to the GLS estimator.

The 32-run benchmark split-plot and split-split-plot designs involving one class-1 hard-to-change factor, one class-2 hard-to-change factor, and three easy-to-change factors, given in the supplementary materials, possess the feature that the OLS and GLS estimators are equivalent for all model parameters. Consequently, knowledge or estimation of the variance components is not required to estimate the main effects and the two-factor interaction effects, and the variance-covariance matrix in Equation (5) is exact for these designs. Comparing the OLS and GLS estimator for the 32-run staggered-level design in Table 2 shows that there is equivalence between OLS and GLS for the easy-to-change factors' main effects and all two-factor interaction effects involving at least one easy-to-change factor. Only for the main effects of both hard-to-change factors and their interaction effect, the OLS and GLS estimators are different. Therefore, for these parameters, we performed 5000 simulations with $\sigma_\delta^2 = 3$, $\sigma_\gamma^2 = 2$, and $\sigma_\varepsilon^2 = 1$ to see how well the staggered-level design in Table 2 performs in practice compared with the split-plot and the split-split-plot designs, with one class-1 hard-to-change factor, one class-2 hard-to-change factor, and three easy-to-change factors in the supplementary materials. We recorded the point estimates for all the factor effects, as well as the 95% confidence limits. For the simulations, we used the mixed procedure in SAS, with REML to estimate the variance components and the Kenward–Roger degrees of freedom option, as recommended by Letsinger, Myers, and Lentner (1996) and Goos, Langhans, and Vandebroek (2006).

The empirical variances of the main effect estimates of both hard-to-change factors and the interaction effect estimate involving these factors are shown in Table 11. From the results, it is clear that the overall conclusions from Section 4, assuming known σ_δ^2, σ_γ^2, and σ_ε^2, still hold when these variances need to be estimated. The main effect of the class-1 hard-to-change factor w is estimated most precisely from the split-plot design, followed by the staggered-level design. The main effect of the class-2 hard-to-change factor s is estimated most precisely from the split-split-plot design, followed by the staggered-level design. Finally, the estimate of the interaction effect between the two hard-to-change factors has the smallest variance when the staggered-level design is used. In terms of the empirical D-optimality criterion (i.e., the D-criterion based on simulated parameter estimates) and

Table 11. Empirical variances of parameter estimates along with empirical D- and A-optimality criterion values for the 32-run staggered-level design in Table 2 and the corresponding split-plot and split-split-plot designs obtained by a simulation study

Effect	Staggered	Split-plot	Split-split-plot
w	0.842	0.631	1.045
s	0.440	0.648	0.284
ws	0.090	0.624	0.280
D-criterion	16.637	15.082	15.806
A-criterion	2.903	2.933	2.988

the empirical A-optimality criterion, the staggered-level design outperforms the two other designs.

A comparison of the OLS and GLS estimator for the 32-run staggered-level designs in Tables 7 and 9 shows, again, that for nearly all the easy-to-change factors' main effects and all two-factor interaction effects involving at least one easy-to-change factor, the OLS and GLS estimators are equivalent. Only when a two-factor interaction effect is not estimated with the minimum variance of $1/32$, as is the case for the interaction effect between t_1 and t_2 in Table 7, the equivalence does not hold. In the case of the 48-run staggered-level design briefly mentioned in Section 6 and shown in the supplementary materials, the variance of the estimates for most effects involving the easy-to-change factors is slightly higher than the minimum variance of $1/48 = 0.021$. For these effects, as well as for similar effects in the 32-run design cases in Tables 7 and 9, the GLS estimates depend on the responses' variance-covariance matrix $\mathbf{V}$ only to a very small extent. So, the OLS and GLS estimates are only marginally different for these effects. The results for unknown variance components are therefore not substantially different from those for known variance components.

The simulation results in Table 11 are completely in line with the earlier calculations and show that the staggered-level design performs better than the split-plot and split-split-plot designs in terms of the precision of point estimates. However, as pointed out by Gilmour and Trinca (2012), many researchers are also interested in significance tests and/or the construction of confidence intervals for the factor effects. It is, of course, well known from the literature on split-plot designs that inference concerning the main effects of hard-to-change factors and concerning their interactions is problematic. Given the low numbers of settings of the hard-to-change factors' levels in the staggered-level designs, we cannot expect powerful inference concerning the hard-to-change factors' effects from these designs either. This is confirmed by the average 95% confidence intervals for the staggered-level design in Table 2 and the corresponding split-plot and split-split-plot designs, which we display in Table 12 and which we obtained by simulating data from a null model.

The confidence intervals for the main effects of the two hard-to-change factors, w and s, are the widest for the staggered-level design. So, when it comes to detecting nonzero main effects of the hard-to-change factors, the staggered-level design performs worse than the split-plot and split-split-plot designs. This should not come as a surprise given that the staggered-level design is the least expensive design to run, with the smallest number of independent settings of the hard-to-change factors' levels. One consequence of this is that the precision of the estimates for σ_δ^2 and σ_γ^2 from the staggered-level design is smaller than from the split-plot and split-split-plot designs, with, occasionally, very large estimates (in our simulations σ_δ^2 was once overestimated by a factor of about 10.7 for the staggered-level design and 10.2 for the split-split-plot design; for the split-plot design, $\sigma_\delta^2 + \sigma_\gamma^2$ was only overestimated by a factor of about 4.1). This results in inflated estimated standard errors for the hard-to-change factors' main effects, and, hence, in wider confidence intervals. The impact of the poorer precision in the estimates for σ_δ^2 and σ_γ^2 on the other effects is almost nonexistent because the estimated standard errors for these effects largely depend on the estimate for σ_ε^2, which is as precise for the staggered-level design as for the other designs.

A feature that distinguishes the staggered-level design from the split-plot and split-split-plot designs is its narrow average confidence interval for the interaction effect between the hard-to-change factors. The confidence interval is so narrow due to the small standard error associated with the interaction effect and to its large degrees of freedom (10.02 vs. 3.63 and 3.48 for the main effects of the hard-to-change factors).

For the main effects of the easy-to-change factors and the interaction effects involving them, there is no substantial difference between the average confidence intervals for the staggered-level design, the split-plot design, and the split-split-plot design.

From this simulation study, it should be clear that the practitioner has to decide in advance what the primary goal of the experiment is: point estimation or statistical inference. If the main interest lies in detecting the significant factor effects, the split-plot design is a better choice than the staggered-level

Table 12. Average degrees of freedom and 95% confidence limits for the factor effects in a model containing main effects and two-factor interactions for the 32-run designs in Section 4 when a GLS analysis is used

	Staggered-level			Split-plot			Split-split-plot		
Effect	D.F.	Lower	Upper	D.F.	Lower	Upper	D.F.	Lower	Upper
w	3.63	−4.20	4.25	4.04	−2.10	2.08	2.46	−3.48	3.70
s	3.48	−2.71	2.70	4.04	−2.08	2.10	3.72	−1.77	1.71
t_i	12.36	−0.37	0.38	12.01	−0.37	0.37	12.23	−0.39	0.35
ws	10.02	−0.63	0.62	4.04	−2.07	2.10	3.72	−1.72	1.76
wt_i	12.36	−0.39	0.36	12.01	−0.37	0.37	12.23	−0.37	0.38
st_i	12.36	−0.37	0.38	12.01	−0.37	0.38	12.23	−0.39	0.36
t_it_j	12.36	−0.38	0.37	12.01	−0.37	0.37	12.23	−0.39	0.36

design in this example. The split-plot design is the most costly option of the three designs but it has the best inferential ability for the factors with few settings. When the primary focus is point estimation, the staggered-level design is the best option. As illustrated in this article, this design option is not only the most efficient one for point estimation but it is also the least expensive design option.

In some practical experiments, the exact run order and the time points at which the factors are independently reset is not faithfully recorded. In those cases, it is impossible to perform the correct GLS analysis, and an OLS approach is undoubtedly the most commonly used back-up alternative. The results of a simulation study for the OLS analysis are described in the supplementary materials.

8. DISCUSSION

In this article, we have presented a new type of design that is useful for experiments involving several hard-to-change factors. The novel feature of the designs is that the hard-to-change factors' levels are set at different points in time. The resulting designs are cost-efficient as well as statistically efficient. The use of a coordinate-exchange algorithm to select D-optimal factor levels for any given staggering structure allows for the construction of a wide variety of staggered-level designs other than those shown in this article. In most of the designs described here, the sizes of the subsets and the numbers of runs were powers of two. This is often too limiting for practical use. However, as we briefly mentioned in Section 6 and as has been shown in the supplementary materials, the algorithm also allows the construction of staggered-level designs when the number of runs is a multiple of four, for any number of hard-to-change and easy-to-change factors.

When constructing the staggered-level designs, we accounted for the uncertainty about the relative magnitudes of the variance components in the model using a log-normal prior distribution and Gauss–Hermite quadrature. In the instances reported in this article, this did not lead to a different design than when assuming that all the variance components are known. Hence, for the staggered-level structure used in this article, the optimal staggered-level designs are not sensitive to the prior information on the variance components. This is in line with earlier research on split-plot and split-split-plot designs.

All the designs discussed in this article suffer from some form of restricted randomization, due to the hard-to-change nature of some experimental factors. This does, however, not mean that nothing can be randomized. A key feature of the split-plot design options is that the order of the whole plots can be randomized as well as the order of the runs within the whole plots, without losing any statistical efficiency. Similarly, randomizing the order of the whole plots, the subplots within whole plots, and the runs within subplots in a split-split-plot design has no impact on the design's efficiency. For the staggered-level design, the groups of runs defined by the settings of the class-1 and class-2 hard-to-change factors cannot be re-ordered randomly. For the design's efficiency, it is crucial that the groups possess the staggering pattern as well as alternating levels for the hard-to-change factors. This is clearly shown in Table 13, where we displayed the six possible orders for the class-1 hard-to-change factor level in the staggered-level design in Table 2, along with the corresponding D-efficiencies. The table shows that only the two alternating patterns yield a D-efficiency of 100%. The worst ordering has only one change of level and results in an efficiency loss of 9%. What can be randomized in staggered-level designs without incurring an efficiency loss is the order of the runs within the groups of size $n/2r$ at each combination of hard-to-change factor levels. Additionally, the class-1 hard-to-change factors can be randomly assigned to the columns for $w_1, \ldots, w_{f_1}$, the class-2 hard-to-change factors can be randomly assigned to the columns for $s_1, \ldots, s_{f_2}$, and the easy-to-change factors can be randomly assigned to the columns for $t_1, \ldots, t_{f_3}$. Also, the factor levels can be randomly assigned. Finally, the experimenter can randomly determine whether the design is run from top to bottom or from bottom to top.

Table 13. Six possible orders of the levels of the class-1 hard-to-change factor, w, in the 32-run staggered-level design in Table 2 and the corresponding D-efficiencies

Setting	Order 1	Order 2	Order 3	Order 4	Order 5	Order 6
1	$-\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$
2	$\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$
3	$-\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$
4	$\mathbf{1}_8$	$\mathbf{1}_8$	$-\mathbf{1}_8$	$-\mathbf{1}_8$	$-\mathbf{1}_8$	$\mathbf{1}_8$
D-efficiency	1.000	0.910	0.910	1.000	0.933	0.933

Since the newly proposed staggered-level designs involve only a limited number of settings of the level of the hard-to-change factors, just like split-plot designs and split-split-plot designs, they share the same kind of weakness: powerful statistical inference regarding the main effects of the hard-to-change factors and regarding the interaction effects involving only hard-to-change factors is impossible. Of course, statistically speaking, it would be much better to set the levels of the hard-to-change factors more often. However, in many practical settings, this is impossible and many experimenters would forgo experimentation if they have to carry out more independent settings of the hard-to-change factors' levels. The nice thing about split(-split)-plot designs and the staggered-level designs is that the inference about all other factors effects is very powerful, and that, through their interaction effects with easy-to-change factors, it is possible to learn much about the hard-to-change factors.

In this article, we restricted our attention to designs involving two classes of hard-to-change factors only. We believe that, when there are more than two hard-to-change factors, it might be better to work with more than two classes. Also, in this article, when constructing the designs, we ignored the fact that, along with the factor effects, variance components need to be estimated. Ideally, the D- and A-optimality criteria should take into account the estimation of the variance components as well as the degrees of freedom associated with them, in the spirit of Gilmour and Trinca (2012). These are two interesting topics for future research.

SUPPLEMENTARY MATERIALS

The following supplementary materials are contained in a single archive and can be obtained via a single download from the *Technometrics* website.

Readme file: Table of content of the supplementary materials and explanation.

D-optimal 32-run split-plot and split-split-plot design I ($f_1 = f_2 = 1$, $f_3 = 3$): The D-optimal 32-run split-plot and split-split-plot design for a problem involving one class-1 hard-to-change factor, w, one class-2 hard-to-change factor, s, and three easy-to-change factors, t_1, t_2, and t_3. These design options are discussed in Section 4 as alternatives to the 32-run staggered-level design in Table 2.

Cost comparison of the 32-run staggered-level design and its two competitors: In Section 4, the 32-run staggered-level design involving five factors is compared with the D-optimal 32-run split-plot design and split-split-plot design in terms of statistical efficiency. It turns out that the staggered-level design is more D- and A-efficient than its competitors. In this file, a similar comparison is made in terms of cost. The results of the comparison are also in favor of the staggered-level design.

D-optimal 32-run split-plot and split-split-plot design II ($f_1 = f_2 = 1$, $f_3 = 4$)**:** The D-optimal 32-run split-plot and split-split-plot design for a problem involving one class-1 hard-to-change factor, w, one class-2 hard-to-change factor, s, and four easy-to-change factors, t_1, t_2, t_3, and t_4. These design options are discussed in Section 5 as alternatives to the 32-run staggered-level design in Table 7.

D-optimal 32-run split-plot and split-split-plot design III ($f_1 = 2$, $f_2 = 1$, $f_3 = 3$)**:** The D-optimal 32-run split-plot and split-split-plot design for a design problem involving two class-1 hard-to-change factors, w_1 and w_2, one class-2 hard-to-change factor, s, and three easy-to-change factors, t_1, t_2, and t_3. These design options are discussed in Section 6 as alternatives to the 32-run staggered-level design in Table 9.

A 48-run design problem involving two class-2 hard-to-change factors: In this file, the D-optimal 48-run staggered-level design for a design problem involving one class-1 hard-to-change factor, w, two class-2 hard-to-change factors, s_1 and s_2, and three easy-to-change factors, t_1, t_2, and t_3, is discussed. This design option is compared to three alternatives, that is, the D-optimal 48-run split-plot design involving 12 whole plots, the D-optimal 48-run split-split-plot design involving 4 whole plots and 8 subplots, and the D-optimal 48-run split-split-plot design involving 6 whole plots and 12 subplots. The staggered-level design outperforms the three alternative designs in terms of the D- and A-optimality criteria.

OLS analysis for the 32-run designs in Section 4: Impact of performing the OLS analysis instead of the proper GLS analysis on the average degrees of freedom and the 95% confidence limits for the factor effects.

ACKNOWLEDGMENTS

The authors thank Sabine Paulussen, Sanaa Sarghini, and Dirk Vangeneugden from the Vlaamse Instelling voor Technologische Ontwikkeling (VITO) for granting permission to use the anti-bacterial coating experiment as an example in this article, and the three referees, the associate editor, and the editor for their constructive suggestions.

[Received December 2009. Revised May 2012.]

REFERENCES

Anbari, F. T., and Lucas, J. M. (2008), "Designing and Running Super-Efficient Experiments: Optimum Blocking With One Hard-to-Change Factor," *Journal of Quality Technology*, 40, 31–45. [355]

Arnouts, H., Goos, P., and Jones, B. (2010), "Design and Analysis of Industrial Strip-Plot Experiments," *Quality and Reliability Engineering International*, 26, 127–136. [361]

Bisgaard, S. (2000), "The Design and Analysis of $2^{k-p} \times 2^{q-r}$ Split-Plot Experiments," *Journal of Quality Technology*, 32, 39–56. [355]

Bliemer, M. C. J., Rose, J. M., and Hess, S. (2009), "Approximation of Bayesian Efficiency in Experimental Choice Designs," *Journal of Choice Modelling*, 1, 98–127. [361]

Ganju, J., and Lucas, J. M. (1999), "Detecting Randomization Restrictions Caused by Factors," *Journal of Statistical Planning and Inference*, 81, 129–140. [355]

Gilmour, S. G., and Trinca, L. A. (2012), "Optimum Design of Experiments for Statistical Inference," *Applied Statistics*, 61, 345–401. [364,365]

Goos, P., Langhans, I., and Vandebroek, M. (2006), "Practical Inference From Industrial Split-Plot Designs," *Journal of Quality Technology*, 38, 162–179. [363]

Goos, P., and Vandebroek, M. (2003), "D-Optimal Split-Plot Designs With Given Numbers and Sizes of Whole Plots," *Technometrics*, 45, 235–245. [355,360]

Jones, B., and Goos, P. (2007), "A Candidate-Set-Free Algorithm for Generating D-Optimal Split-Plot Designs," *Applied Statistics*, 56, 347–364. [355,360,361]

——— (2009), "D-Optimal Design of Split-Split-Plot Experiments," *Biometrika*, 96, 67–82. [355,359,360,361]

——— (2012a), "An Algorithm for Finding D-Efficient Equivalent-Estimation Second-Order Split-Plot Designs," *Journal of Quality Technology*, 44, 363–374. [355]

——— (2012b), "I-Optimal versus D-Optimal Split-Plot Response Surface Designs," *Journal of Quality Technology*, 44, 85–101. [355]

Ju, H. T., and Lucas, J. M. (2002), "L^k Factorial Experiments With Hard-to-Change and Easy-to-Change Factors," *Journal of Quality Technology*, 34, 411–421. [355]

Kythe, P. K., and Schäferkotter, M. R. (2005), *Handbook of Computational Methods for Integration*, London: Chapman and Hall/CRC. [361]

Letsinger, J. D., Myers, R. H., and Lentner, M. (1996), "Response Surface Methods for Bi-Randomization Structures," *Journal of Quality Technology*, 28, 381–397. [363]

Macharia, H., and Goos, P. (2010), "D-Optimal and D-Efficient Equivalent-Estimation Second-Order Split-Plot Designs," *Journal of Quality Technology*, 42, 358–372. [355]

Meyer, R. K., and Nachtsheim, C. J. (1995), "The Coordinate-Exchange Algorithm for Constructing Exact Optimal Experimental Designs," *Technometrics*, 37, 60–69. [361]

Schoen, E. D. (1999), "Designing Fractional Two-Level Experiments With Nested Error Structures," *Journal of Applied Statistics*, 26, 495–508. [355]

Trinca, L. A., and Gilmour, S. G. (2001), "Multi-Stratum Response Surface Designs," *Technometrics*, 43, 25–33. [355]

Vining, G. G., Kowalski, S. M., and Montgomery, D. C. (2005), "Response Surface Designs Within a Split-Plot Structure," *Journal of Quality Technology*, 37, 115–129. [355]

Webb, D. F., Lucas, J. M., and Borkowski, J. J. (2004), "Factorial Experiments When Factor Levels are Not Necessarily Reset," *Journal of Quality Technology*, 36, 1–11. [356]

Yu, J., Goos, P., and Vandebroek, M. (2010), "Comparing Different Sampling Schemes for Approximating the Integrals Involved in the Efficient Design of Stated Choice Experiments," *Transportation Research Part B: Methodological*, 44, 1268–1289. [361]