Optimal Design of Blocked Experiments in the Presence of Supplementary Information About the Blocks

2 authors:

Bradley Jones
SAS Institute
105 PUBLICATIONS   **3,243** CITATIONS

Peter Goos
KU Leuven
247 PUBLICATIONS   **4,675** CITATIONS

Some of the authors of this publication are also working on these related projects:

SRN / WOG Empirical and methodological challenges in choice experiments View project

Design of experiments for variance component estimation View project

Optimal Design of Blocked Experiments in the Presence of Supplementary Information About the Blocks[*]

BRADLEY JONES

SAS Institute, Cary, NC 27513, USA

PETER GOOS

Faculty of Bioscience Engineering & Leuven Statistics Research Centre,
Katholieke Universiteit Leuven, Belgium
Faculty of Applied Economics & StatUa Center for Statistics, Universiteit Antwerpen, Belgium

In some designed experiments, measurements of characteristics of the experimental units may be available prior to performing the runs. If the investigators believe that these measured characteristics may have some effect on the response of interest, then it seems natural to include these characteristics as factors in the experiment even though they are not under direct control. It may also be possible to apply multiple runs to a given experimental unit by subdividing it into multiple pieces, each having the same characteristics. A similar scenario involves using a person or an animal as an experimental unit multiple times but with different treatments. Here, the measured information about the subjects may not change over the experiment. In either of these cases, the fact that several runs employ the same experimental unit means that the responses for those runs are correlated. This correlation, in addition to the natural variability of the measured characteristics over the sample of available experimental units, requires new methodology for creating optimal designs. Specifically, the methodology must choose a subset of the experimental units and determine the number of treatments applied to each experimental unit in addition to choosing the level combinations of the controllable factors for each run. In this article, we provide two methods for generating optimal designs in the presence of additional information about the experimental units. The first method fixes the number of runs performed on each experimental unit. The second method allows for varying numbers of runs applied to each experimental unit, subject to a constraint on the total number of runs. We discuss several illustrative examples using each method as well as a real experiment using previously fabricated batches of polypropylene as experimental units in a study on the effects of a subsequent plasma treatment.

Key Words: Blocked Experiment; Concomitant Variable; Covariate; D-Optimality; I-Optimality; Split-Plot Experiment; V-Optimality.

1. Introduction

PRIOR to running an experiment, an investigator may possess information about certain characteristics of the experimental units that could have a substantial effect on the response. It is natural to take these characteristics into account when designing the experiment and when analyzing the resulting data. We call the factors that measure the characteristics of the experimental units covariates or concomitant variables. The covariates may be continuous or categorical. They are different from the ex-

perimental factors in that they cannot be controlled directly. Designing an experiment around the levels of the covariates reduces the standard errors of the factor-effect estimates compared with an experiment that ignores the covariate information and uses complete randomization of the experimental units. In this way, covariates are similar to blocking factors.

Because instruments often experience drift over time as the output of a sequence of experimental runs is measured, the design of experiments literature contains many references dealing with time as a covariate. Joiner and Campbell (1976) gave four interesting examples of experiments involving a time trend or drift, while Hill (1960) summarized the early literature on optimal designs in the presence of a time trend. These optimal designs are generally called trend-robust designs because they lead to estimates of the experimental factors' effects that are insensitive to the time trend. Daniel and Wilcoxon (1966) proposed trend-robust arrangements of both full factorial and fractional factorial experiments. Follow-up work, in which a trend-robust arrangement was sought for a given set of factor level combinations, was done by Cheng and Jacroux (1988), Cheng (1990), John (1990), Cheng and Steinberg (1991), and Mee and Romanova (2010). Atkinson and Donev (1996) proposed an algorithm for generating general trend-robust designs.

The general problem of covariate information in experimental design, involving covariates other than time, received much less attention in the literature than the search for good experimental designs in the presence of time-trend effects. Harville (1974), Cook and Thibodeau (1980), Nachtsheim (1989), and Goos and Jones (2011) discussed the design of experiments in the presence of general covariates.

A key feature of the published work on design of experiments in the presence of covariates is that it assumes that every experimental unit is used exactly once and that all the responses are independent. In this paper, we study situations where each experimental unit (about which supplementary information is available, in the form of one or more covariates) is divisible into subunits that can undergo different treatments. Hence, we face a design-of-experiments problem in which multiple experimental runs involve the same covariate values. We distinguish two scenarios. In the first scenario, we consider experiments where it is desirable to use an equal number of treatments for every experimental unit. In the second sce-

nario, we allow the number of treatments to differ from unit to unit.

The fact that we consider experiments involving more than one treatment per experimental unit makes our designs similar to blocked experiments. The design of blocked experiments has received much attention in the design-of-experiments literature (see, for instance, Wu and Hamada (2000) for an overview of the work on blocking factorial and fractional factorial designs and Atkinson and Donev (1989), Cook and Nachtsheim (1989), Goos and Vandebroek (2001), and Goos and Donev (2006) for algorithms on the optimal design of general blocked experiments involving fixed and random block effects). However, covariates corresponding to the blocks are not taken into account in any of this work. Moreover, these authors always assume that there are no interactions between the blocks and the treatment factors. In this paper, we have covariate information for each block, so we can model interactions involving the covariates and the treatments. The model in our paper therefore shows some similarity with the work of Khuri (1996), who allowed for interactions between the blocks and the experimental factors using a random-effects model. A difference between Khuri's work and ours is that we use fixed covariates to explain the exact nature of the interactions instead of treating the interactions as random effects.

The model we consider in this paper is very similar to that used for data from split-plot experiments in Letsinger et al. (1996), Vining et al. (2005), and Goos et al. (2006). This is because our model involves a random intercept and two types of explanatory variables, some of which (the covariates) are related to the larger experimental units (the blocks) and some of which (the controlled experimental factors) are applied to the individual runs. In a split-plot model, there is also a random intercept, and one set of explanatory variables (the whole-plot factors) is applied to the larger experimental units, whereas another set (the subplot factors) is applied to the individual runs. A major difference between split-plot experiments and the experiments considered in this paper is that, unlike the settings of the whole-plot factors in a split-plot experiment, the values of the covariates are not under the investigator's control. However, the investigator is in the position to choose which of the available supply of experimental units to use and to determine how many subunits of the selected experimental units are used.

This paper is organized as follows. In Section 2, we describe the two examples that motivated us to study the problem of designing experiments involving covariates. In Section 3, we discuss the model we use to analyze data from blocked experiments in the presence of covariate information about the blocks. In Section 4, we define the D- and the I-optimality criteria to select designs. Next, in Section 5, we focus on small artificial examples to clarify the difference between the two scenarios that are central to the paper. Finally, we return to the two motivating examples in Section 6, and end the paper with a discussion in Section 7.

2. Motivating Examples

2.1. Scenario 1: Equal Numbers of Runs Within Every Block

Many experiments involve subjects each of whom tests or evaluates several products. Typical examples include shampoos, cars, credit-card offerings, or wines. Our example involves the testing of running shoes. An important quality characteristic of a running shoe is its wear. The wear is clearly a function of the design of the running shoe. Key design factors that drive the wear of a shoe are the sole's thickness, the shoe brand, and whether or not the shoe has extra cushioning or shock absorption. There are, however, also other important factors that have an impact on the wear. First, the weekly mileage of the runner plays a major role, as well as the runner's weight. Also, whether the runner's strike point is the heel or the midfoot is an influential factor.

The usual experimental practice is to recruit a pool of runners to test prototype shoes. Decisions that affect the efficiency of the experimental design are the selection of participants from the runners in the pool and the assignment of the prototypes to each of the selected runners. The selection of any particular runner automatically implies a value for each of the covariates. Hence, technically speaking, the experimenter needs to choose the covariate values from a pool of potential covariate values and to assign several combinations of treatment-factor levels to each chosen covariate value.

To control cost, we require each participant in the study to test several factor-level combinations, i.e., several types of shoes. This inevitably leads to sets of correlated responses because the wear patterns for any given runner will be more homogeneous than the wear patterns from different runners. For logistical

reasons, we preferred to have each participant test an equal number of shoes. The number of runners in the study was determined by the budget and fixed in advance.

Designing the running-shoe experiment requires an experimental design methodology that chooses the runners (and thus their covariate values), while simultaneously optimizing the settings for the shoe-design factors over the set of experimental runs.

2.2. Scenario 2: Unequal Numbers of Runs Within Every Block

Jones and Goos (2007) and Goos and Gilmour (2012) discuss a large polypropylene experiment involving several stages. In the first stage, batches of polypropylene plates were made. Every batch was made according to different formulations involving the ingredients polypropylene, ethylene, ethylene propylene diene monomer (EPDM), ethylene vinyl acetate (EVA), lubricant, UV stabilizer, talcum, and mica. The second stage of the experiment involved a gas plasma treatment in an oven. The gas plasma treatment was defined by the oven pressure, the power utilized, the reaction time, and the type of gas used.

The polypropylene production in the first stage of the experiment yielded large batches of a few hundred polypropylene plates (each with size 4 cm × 8 cm × 3 mm). For the initial experiment, only a few dozen plates from each batch were required. The remainder of the plates were individually packed so as to ensure there was no loss of quality. As a result, the investigators ended up with a stock of several batches of polypropylene plates with different, but known, compositions.

The stock of plates makes it possible to conduct follow-up experiments to test, for instance, other gas plasma treatments and other oven types on a wide variety of plates. The goal of all the experiments was to tailor the gas plasma treatment to possible polypropylene formulations that an existing or new customer might have in mind for his/her product. The customers included producers of mobile phones, car manufacturers, and producers of medical devices, all of whom use polypropylene, but for entirely different products. Obviously, different end uses of the polypropylene impose different requirements on the formulation. For the success of the gas plasma treatment company, it is of crucial importance to be able to react flexibly to the changing needs of the customers and to accumulate knowledge

about the interactions between the gas plasma treatments and the polypropylene formulations. The large historical stock of polypropylene plates from different batches allows them to conduct many gas plasma treatment experiments inexpensively, i.e., without having to go through the trouble of ordering new sets of polypropylene plates. This raised the question of how to design the follow-up experiments in such a way as to take into account the available information on each batch. Here, the available information about each batch is the polypropylene formulation, so the ingredients of the polypropylene comprise the matrix of covariate observations.

When discussing the optimal design of the follow-up experiments, we observed the following:

- Even though the gas plasma treatment factors are independently reset for every experimental run, the runs that involve plates from the same batch are dependent. Hence, their responses are correlated, as are the runs in the same block of a blocked experiment and the runs in the same whole plot of a split-plot experiment.

- There is no limitation on the number of batches used in each follow-up experiment because the available batches are numerous. Also, it is as easy to select five plates from five different batches as it is to select five plates from a single batch. This is because every batch is just a box of small plates sitting on a shelf in a warehouse. This makes the current type of experiment different from a split-plot experiment because the cost of a split-plot experiment is primarily driven by the number of whole plots, whereas, in the polypropylene experiment, the experimental cost is driven by the total number of plates used.

- There is also no realistic upper bound on the number of plates selected from any given batch because every batch consists of hundreds of plates and every follow-up experiment requires at most a few dozen plates usually from different batches. This is a second major difference between the current type of experiment and a split-plot experiment. In a split-plot experiment, the number of runs within a whole plot is often dictated by the logistics of the experiment (for instance, the size of an oven, the number of runs that can be done in a day, or on a plot of land).

As a result, for follow-up polypropylene experiments, we need an optimal experimental de-sign methodology that simultaneously identifies the batches (and thus the covariate values) to use, determines the number of plates to use from each chosen batch, and assigns the settings for the gas plasma treatment factors to each plate. We desire the design produced by this methodology to be the most efficient possible for estimating the model effects, given that the choice of covariate values is limited.

In experiments other than the polypropylene experiment, it may be desirable to limit the number of batches. For instance, the available batches may be spread over different geographical locations, in which case using many batches might require a major logistical cost. In such instances, it makes sense to define an upper bound for the number of batches. Even in that scenario, it is useful to optimize the number of batches used because the optimal number of batches depends to some extent on the values of the covariates and is therefore not necessarily the maximum number the budget allows.

3. Statistical Model and Analysis

In the presence of covariate information on the blocks, the model we use for analyzing data from experiments with b blocks, the ith of which involves k_i runs, is

$$Y_{ij} = \mathbf{f}'(\mathbf{c}_i, \mathbf{x}_{ij})\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}, \qquad (1)$$

where Y_{ij} is the response measured at the jth run in the ith block, $\mathbf{c}_i$ is the vector containing the values of the covariates for block i, $\mathbf{x}_{ij}$ is the vector that contains the levels of the experimental factors at the jth run in the ith block, $\mathbf{f}(\mathbf{c}_i, \mathbf{x}_{ij})$ is the model expansion of the two types of explanatory variables (the covariates and the experimental factors), and $\boldsymbol{\beta}$ contains the intercept, the covariates' effects, the factor effects and the interaction effects that are in the model. The term γ_i represents the random effect of the ith block and ε_{ij} is the random error associated with the jth run in block i. We denote the dimension of $\mathbf{f}(\mathbf{c}_i, \mathbf{x}_{ij})$ and $\boldsymbol{\beta}$ by p.

For an experiment with sample size n and b blocks, the model can be written in matrix notation as

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

where $\boldsymbol{Y}$ is the vector of responses, $\mathbf{X}$ represents the $n \times p$ model matrix containing the values of the covariates $\mathbf{c}$ and the experimental factors $\mathbf{x}$ and their model expansions, $\boldsymbol{\beta}$ is again the p-dimensional vector containing the p fixed effects of the covariates and the experimental factors in the model, $\mathbf{Z}$ is an $n \times b$

matrix of zeroes and ones assigning the n runs to the b blocks, $\boldsymbol{\gamma}$ is the b-dimensional vector containing the random effects of the b blocks, and $\boldsymbol{\varepsilon}$ is the n-dimensional vector containing the random errors. We assume that

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n \text{ and } \mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_n, \qquad (3)$$
$$\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_b \text{ and } \mathrm{cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_b, \qquad (4)$$

and

$$\mathrm{cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b \times n}.$$

Under these assumptions, the covariance matrix of the responses, $\mathrm{Var}(\boldsymbol{Y})$, is

$$\mathbf{V} = \sigma_\varepsilon^2 \mathbf{I}_n + \sigma_\gamma^2 \mathbf{Z}\mathbf{Z}'. \qquad (5)$$

When the entries of $\boldsymbol{Y}$ are grouped by block, then

$$\mathbf{V} = \mathrm{diag}(\mathbf{V}_1, \ldots, \mathbf{V}_b), \qquad (6)$$

where

$$\mathbf{V}_i = \sigma_\varepsilon^2 \mathbf{I}_{k_i} + \sigma_\gamma^2 \mathbf{1}_{k_i}\mathbf{1}'_{k_i} = \sigma_\varepsilon^2(\mathbf{I}_{k_i} + \eta \mathbf{1}_{k_i}\mathbf{1}'_{k_i}),$$

k_i is the number of runs in block i, and the variance ratio $\eta = \sigma_\gamma^2/\sigma_\varepsilon^2$ is a measure for the extent to which responses from runs within the same block are correlated. The larger η, the more the responses within one block are correlated.

When the random error terms as well as the block effects are normally distributed, the maximum-likelihood estimator of the unknown model parameter vector $\boldsymbol{\beta}$ is the generalized least squares (GLS) estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{Y}, \qquad (8)$$

with covariance matrix

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \qquad (9)$$

The information matrix for the parameter vector $\boldsymbol{\beta}$ is given by

$$\mathbf{M} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}. \qquad (10)$$

4. Criteria for Selecting Designs

The most commonly used criterion to select experimental designs is the D-optimality criterion, which seeks designs that maximize the determinant of the information matrix,

$$|\mathbf{M}| = |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|.$$

We use the D-efficiency to compare the quality of two designs with information matrices $\mathbf{M}_1$ and $\mathbf{M}_2$. The D-efficiency of a design with information matrix $\mathbf{M}_1$

relative to a design with information matrix $\mathbf{M}_2$ is defined as

$$\text{D-efficiency} = \left(\frac{|\mathbf{M}_1|}{|\mathbf{M}_2|}\right)^{1/p}.$$

A D-efficiency larger than one indicates that design 1 is better than design 2 in terms of the D-optimality criterion. In general, the D-optimal design depends on the variance ratio η through the covariance matrix $\mathbf{V}$ of the responses, as does the D-efficiency of one design relative to another.

An I-optimal design minimizes the average prediction variance

average variance
$$= \frac{\int_\chi \mathbf{f}'(\mathbf{c}, \mathbf{x})(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{f}(\mathbf{c}, \mathbf{x})d(\mathbf{c}', \mathbf{x}')'}{\int_\chi d(\mathbf{c}', \mathbf{x}')'}$$
$$(11)$$

over the experimental region χ. Jones and Goos (2012) show that the average prediction variance can be calculated as

average variance
$$= \frac{1}{\int_\chi d(\mathbf{c}', \mathbf{x}')'}\mathrm{tr}\big[(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}\big], \quad (12)$$

where

$$\mathbf{B} = \int_\chi \mathbf{f}(\mathbf{c}, \mathbf{x})\mathbf{f}'(\mathbf{c}, \mathbf{x})d(\mathbf{c}', \mathbf{x}')' \qquad (13)$$

is the moments matrix. Expressions for the moments matrix for quadratic models are given in Hardin and Sloane (1991a, b), for cuboidal and spherical experimental regions.

If P_1 is the average prediction variance of one design and P_2 is the average prediction variance of a second design, then the I-efficiency of the former design compared with the latter is computed as

$$\text{I-efficiency} = P_2/P_1.$$

An I-efficiency larger than one indicates that design 1 is better than design 2 in terms of the average prediction variance. The I-optimal design and the I-efficiency of one design relative to another generally depend on the variance ratio η through the covariance matrix $\mathbf{V}$.

A technical difficulty with finding a D- or I-optimal design in this paper is that the matrix $\mathbf{V}$ and therefore also the D- and I-optimality criteria depend on the unknown variances σ_γ^2 and σ_ε^2. Fortunately, the optimal split-plot designs do not depend on the absolute magnitude of these two variances, but

only on their relative magnitude. Therefore, generating optimal split-plot designs only requires input of the relative magnitude of σ_γ^2 and σ_ε^2.

For the purpose of generating an optimal design, an educated guess of the variance ratio $\eta = \sigma_\gamma^2/\sigma_\varepsilon^2$ is usually good enough, because a design that is optimal for one variance ratio is generally also optimal for a broad range of variance ratios around the specified one. Moreover, whenever different variance ratios lead to different designs, the quality of these designs is almost identical. Goos (2002) recommends using a variance ratio of one for finding optimal split-plot designs in the absence of detailed *a priori* information about it. We therefore use an η value of one to generate D- and I-optimal designs in this article. A computationally more demanding approach to deal with the dependence of optimal designs on the variance ratio η, which properly accounts for uncertainty about the exact η value, is described by Mylona et al. (2014).

We computed D- and I-optimal designs in the presence of covariates using a new algorithm. The key feature of the algorithm is that it combines a point-exchange procedure with a coordinate-exchange procedure. The algorithm uses a point-exchange procedure to optimize the values of the covariates $\mathbf{c}$ that have to be chosen from a predefined list of possible covariate values. This list of covariate values acts as a set of candidate points and necessitates a point-exchange algorithm (see, for instance, Fedorov (1972), Cook and Nachtsheim (1980), and Goos and Vandebroek (2003)). To optimize the settings of the experimental factors $\mathbf{x}$, we use a coordinate-exchange algorithm (see, for instance, Meyer and Nachtsheim (1995) and Jones and Goos (2007)).

For Scenario 1 in Section 2, our algorithm generates designs for a fixed number of blocks b and a fixed number of runs $k = k_1 = \cdots = k_b$ within the blocks and, given b and k, it optimizes the treatment factor level combinations within every block. The algorithm is therefore similar to the algorithms of Goos and Vandebroek (2001, 2003), which seek optimal blocked/split-plot designs with given numbers of blocks/whole plots and runs per block/whole plot. For Scenario 2 in Section 2, our algorithm also optimizes the number of blocks and the number of runs within a block. Therefore, our algorithm for that scenario resembles the algorithms of Goos and Vandebroek (2004) and Kessels et al. (2008), which deal with finding optimal numbers of whole plots/blocks as well as optimal whole-plot sizes/block sizes for

split-plot and blocked experiments with given numbers of runs.

5. Proof-of-Concept Examples

In this section, we discuss several small, artificial examples to demonstrate the key concepts. For each scenario, the initial focus is on D-optimal designs for main-effects models and models including two-factor interactions. Later, we turn our attention toward models involving quadratic effects, for which we compute D- and I-optimal designs.

5.1. Scenario 1

5.1.1. A Categorical Covariate

Consider an experimental study involving two continuous experimental factors, x_1 and x_2, and a single two-level categorical covariate, say gender. Assume that a large pool of potential participants is available, but the budget allows for eight participants and two experimental runs per participant. If the interest is in a main-effects model

$$Y_{ij} = \beta_0 + \beta_1 c_i + \beta_2 x_{1ij} + \beta_3 x_{2ij} + \gamma_i + \varepsilon_{ij}, \quad (14)$$

where c is the coded covariate (-1 for males, 1 for females), then the D-optimal design uses four males and four females. Each participant either tests the factor level combinations $(x_1, x_2) = (-1, 1)$ and $(x_1, x_2) = (1, -1)$, or the factor level combinations $(x_1, x_2) = (-1, -1)$ and $(x_1, x_2) = (1, 1)$. The resulting design, which is shown in Figure 1, is orthogonal. Assuming that $\sigma_\varepsilon^2 = \sigma_\gamma^2 = 1$, the design has a diagonal information matrix equal to

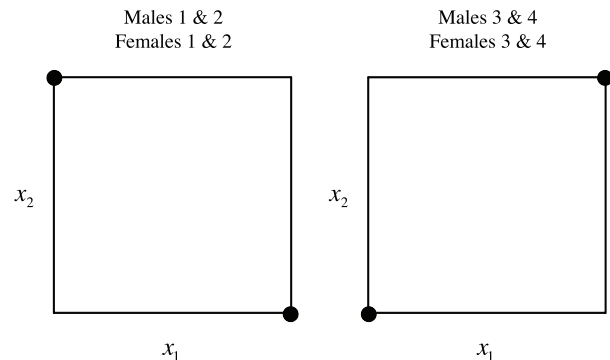$$\mathbf{M} = \mathrm{diag}(16/3, 16/3, 16, 16),$$

FIGURE 1. D-Optimal Design Involving Blocks of Size Two and Two Experimental Factors x_1 and x_2 for a Main-Effects Model with One Two-Level Categorical Covariate, Gender.
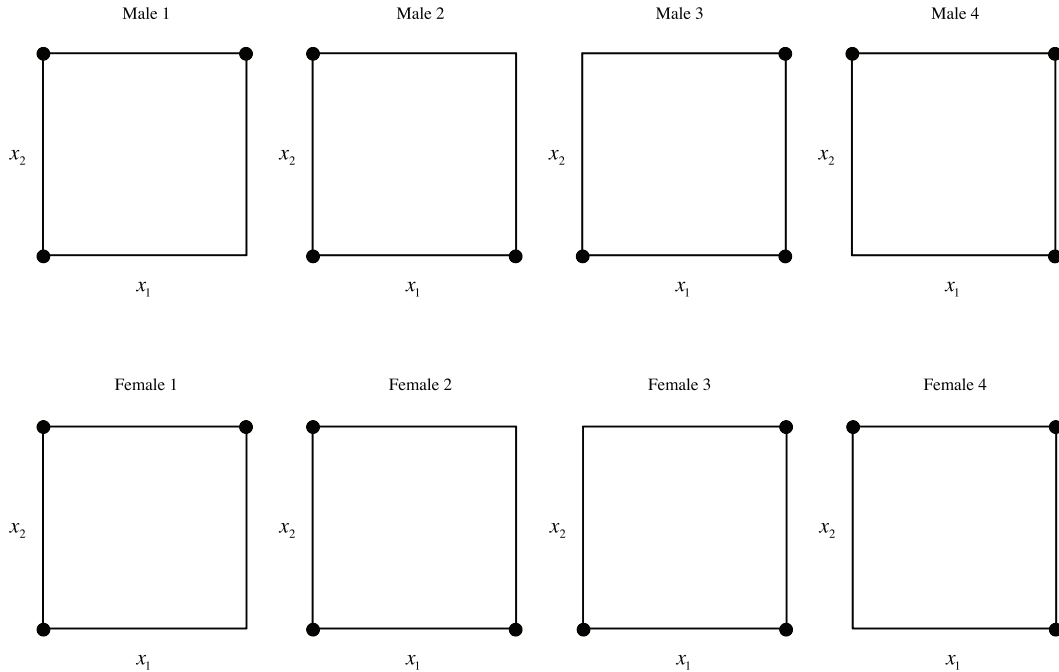
FIGURE 2. D-Optimal Design Involving Blocks of Size Three and Two Experimental Factors x_1 and x_2 for a Main-Effects Model with One Two-Level Categorical Covariate, Gender.

where the diagonal elements correspond to the intercept and the main effects of the covariate and the two experimental factors. The fact that the last two diagonal elements equal 16, the total number of runs, shows that no information is lost concerning the main effects of x_1 and x_2 due to the blocking.

Now, consider a situation where three experimental runs are possible for each of the eight participants, yielding a total of 24 experimental runs. A D-optimal design for that situation is shown in Figure 2. Every block of that design involves the 2^2 factorial design in x_1 and x_2 minus one point. The overall projection of the design, obtained by ignoring the blocks, involves six replicates of the 2^2 factorial design. The overall projection for each gender results in three replicates of the 2^2 design. The design results in a diagonal information matrix,

$$\mathbf{M} = \mathrm{diag}(6, 6, 22, 22),$$

which implies that all model effects can be estimated independently. This is remarkable given the blocks of size three, which makes it impossible to have an orthogonal two-level design. The nonorthogonality of the design can be seen from the fact that the last two diagonal elements of the information matrix are smaller than the number of runs, 24.

5.1.2. Two Quantitative Covariates

Consider an experimental study involving two continuous experimental factors, x_1 and x_2, and two quantitative covariates, c_1 and c_2. The model of interest involves the main effects and second-order interactions involving x_1, x_2, c_1, and c_2,

$$\begin{aligned} Y_{ij} = {}& \beta_0 + \beta_1 c_{1i} + \beta_2 c_{2i} + \beta_3 x_{1ij} + \beta_4 x_{2ij} \\ & + \beta_5 c_{1i} c_{2i} + \beta_6 c_{1i} x_{1ij} + \beta_7 c_{1i} x_{2ij} + \beta_8 c_{2i} x_{1ij} \\ & + \beta_9 c_{2i} x_{2ij} + \beta_{10} x_{1ij} x_{2ij} + \gamma_i + \varepsilon_{ij}. \end{aligned} \quad (15)$$

The set of covariate values involves 16 draws from a bivariate uniform distribution on the $[-1, 1] \times [-1, 1]$ square. This set is shown in Figure 3.

In this section, we assume that eight blocks of three runs each are desired and therefore compute a D-optimal design involving eight blocks of size three. The design we obtained is shown as design 1 in the left pane of Table 1. The covariate combinations shown by means of a square in Figure 3 are the ones selected by the D-optimal design, while the filled circles indicate those not selected. A key result is that the combinations of covariate values selected in the D-optimal design are located close to the vertices of the square in Figure 3. So our algorithm for finding a D-optimal design picks the covariate combinations
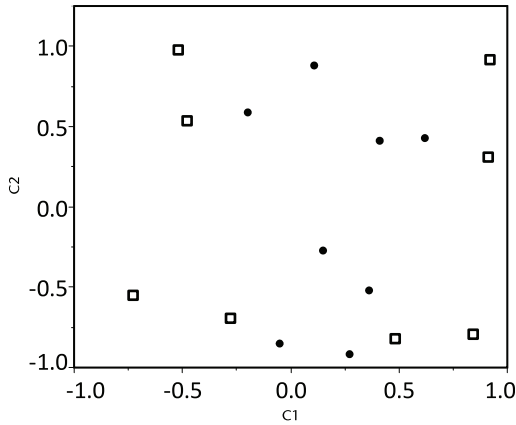
FIGURE 3. Scatter Plot of the Values for the Two Quantitative Covariates c_1 and c_2 in Section 5.1.2. Squares indicate the covariate combinations selected by the D-optimal 24-run design with eight blocks of size 3 (labeled design 1 in Table 1), while filled circles correspond to covariate combinations that are not used in the D-optimal design.

that are closest to the factor level combinations of a 2^2 design in the covariates.

5.2. Scenario 2

5.2.1. Model with Main Effects and Interactions

Consider again an experimental study involving 24 runs, two continuous experimental factors, x_1 and x_2, and two quantitative covariates, c_1 and c_2, and assume the model of interest is given by Equation (15). As the set of covariate values, we utilize the set of 16 draws from a bivariate uniform distribution on the square $[-1, 1] \times [-1, 1]$ shown in Figure 3. Now, however, we neither impose a fixed number of blocks nor a fixed number of runs per block. Instead, we seek the optimal number of blocks and the optimal numbers of runs in those blocks. Because we start from 16 combinations of covariate values, the maximum number of blocks in this study is 16. The

TABLE 1. D-Optimal Designs for a Model with Main Effects and Second-Order Interactions Involving Two Quantitative Covariates, c_1 and c_2, and Two Quantitative Experimental Factors, x_1 and x_2

|	Design 1					Design 2					Design 3			
Block	c_1	c_2	x_1	x_2	Block	c_1	c_2	x_1	x_2	Block	c_1	c_2	x_1	x_2
1	−0.73	−0.55	−1	−1	1	−0.73	−0.55	−1	−1	1	−0.73	−0.55	−1	−1
1	−0.73	−0.55	−1	1	1	−0.73	−0.55	−1	1	1	−0.73	−0.55	−1	1
1	−0.73	−0.55	1	−1	1	−0.73	−0.55	1	−1	1	−0.73	−0.55	1	−1
2	−0.52	0.98	−1	−1	1	−0.73	−0.55	1	1	1	−0.73	−0.55	1	1
2	−0.52	0.98	1	−1	2	−0.52	0.98	−1	−1	2	−0.52	0.98	−1	−1
2	−0.52	0.98	1	1	2	−0.52	0.98	−1	1	2	−0.52	0.98	−1	1
3	−0.48	0.54	−1	1	2	−0.52	0.98	1	−1	2	−0.52	0.98	1	−1
3	−0.48	0.54	1	−1	2	−0.52	0.98	1	1	2	−0.52	0.98	1	1
3	−0.48	0.54	1	1	3	−0.48	0.54	−1	−1	3	−0.28	−0.69	−1	−1
4	−0.28	−0.69	−1	−1	3	−0.48	0.54	1	1	3	−0.28	−0.69	1	1
4	−0.28	−0.69	−1	1	4	−0.28	−0.69	−1	1	4	0.48	−0.82	−1	1
4	−0.28	−0.69	1	1	4	−0.28	−0.69	1	−1	4	0.48	−0.82	1	−1
5	0.48	−0.82	−1	−1	5	0.48	−0.82	−1	−1	5	0.84	−0.79	−1	−1
5	0.48	−0.82	1	−1	5	0.48	−0.82	1	1	5	0.84	−0.79	−1	1
5	0.48	−0.82	1	1	6	0.84	−0.79	−1	−1	5	0.84	−0.79	1	−1
6	0.84	−0.79	−1	1	6	0.84	−0.79	−1	1	5	0.84	−0.79	1	1
6	0.84	−0.79	1	−1	6	0.84	−0.79	1	−1	6	0.91	0.31	−1	−1
6	0.84	−0.79	1	1	6	0.84	−0.79	1	1	6	0.91	0.31	1	1
7	0.91	0.31	−1	−1	7	0.91	0.31	−1	1	7	0.92	0.92	−1	−1
7	0.91	0.31	−1	1	7	0.91	0.31	1	−1	7	0.92	0.92	−1	1
7	0.91	0.31	1	1	8	0.92	0.92	−1	−1	7	0.92	0.92	−1	1
8	0.92	0.92	−1	−1	8	0.92	0.92	−1	1	7	0.92	0.92	1	−1
8	0.92	0.92	−1	1	8	0.92	0.92	1	−1	7	0.92	0.92	1	−1
8	0.92	0.92	1	−1	8	0.92	0.92	1	1	7	0.92	0.92	1	1
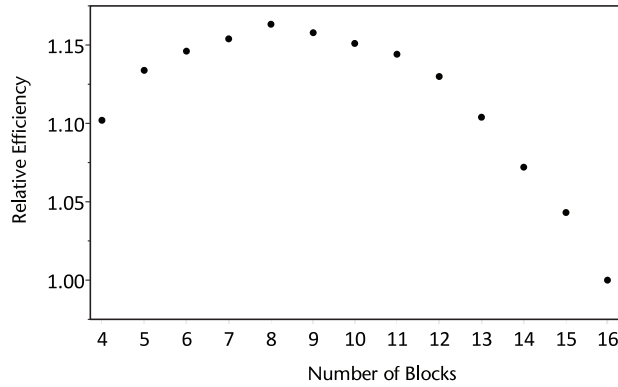
FIGURE 4. D-Efficiencies of the D-Optimal Designs with Numbers of Blocks Ranging from 4 to 16, Relative to the Design with 16 Blocks.
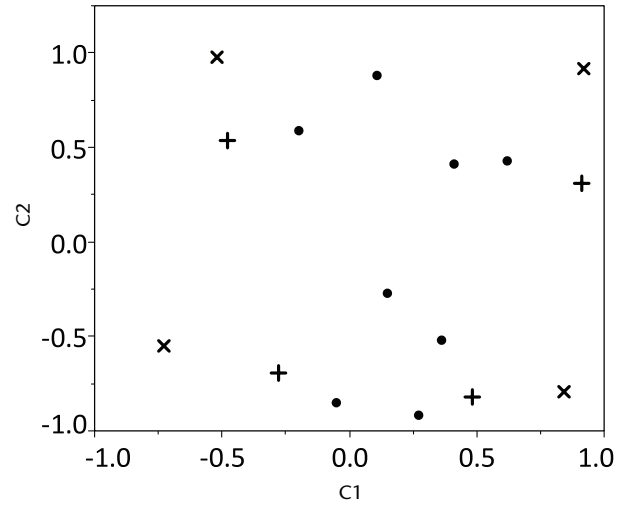

FIGURE 5. Scatter Plot of the Values for the Two Quantitative Covariates c_1 and c_2 in Section 5.2.1. Crosses indicate the covariate combinations corresponding to blocks of four runs in the D-optimal 24-run design with eight unequally sized blocks (labeled design 2 in Table 1), while plus signs correspond to blocks of two runs and filled circles correspond to covariate combinations that are not used in the D-optimal design.

minimum number is 4, to ensure estimability of the main effects of the two covariates and the interaction involving them.

We generated D-optimal 24-run designs with numbers of blocks ranging from 4 to 16 and display the D-efficiencies of the resulting designs in Figure 4 (where the D-efficiencies are relative to the 16-block design). The figure shows that the D-optimal design with eight blocks outperforms all others, and that it is 16.3% more efficient than the design with 16 blocks. Remarkably, the D-optimal design with eight blocks does not involve equal numbers of runs in each block, despite the fact that the number of blocks, 8, is a divisor of the number of runs, 24. Instead, the D-optimal 8-block design involves four blocks of four runs each and four blocks of two runs each. The complete design is shown as design 2 in the middle pane of Table 1. One feature of the design is that it has even block sizes, allowing it to be orthogonally blocked for the main effects. The selected covariate combinations, along with the corresponding block sizes, are shown in Figure 5. The block sizes are symbolized using crosses and plus signs: the crosses indicate covariate combinations that correspond to blocks of four runs and the plus signs indicate covariate combinations corresponding to blocks of two runs.

We compared the D-optimal design with four blocks of size 4 and four blocks of size 2 to the D-optimal design with 8 blocks of size 3, shown as design 1 in the left pane of Table 1. The 8-block design with equal block sizes is about 7% less D-efficient than the 8-block design with unequal block sizes. A resemblance between both 8-block designs is that

they use the same set of eight covariate combinations. This can be seen by comparing Figures 3 and 5. One reason why the design with equal block sizes is less efficient is that it is not orthogonally blocked for the main effects of x_1 and x_2, due to its odd block size.

The optimal block sizes in the D-optimal design are only two units apart. In some of the D-optimal designs we obtained, the block sizes were even more heterogeneous than that. The D-optimal design with seven blocks, shown as design 3 in the right pane of Table 1, has one block involving six runs, three blocks involving four runs, and three blocks involving two runs. As shown in Figure 4, this design is only 1% less D-efficient than the design with eight unequally sized blocks. It uses the same covariate combinations as the two 8-block designs, except for $(c_1, c_2) = (-0.48, 0.54)$.

5.2.2. Full Quadratic Model

Consider a 20-run experiment involving three quantitative experimental factors, x_1, x_2, and x_3, and one quantitative covariate, c. Assume that the covariate's values are -1, 0, and 1 and that each of these three values occurs four times in the pool of covariates. The model of interest is the full quadratic

TABLE 2. I-Optimal Design for the Full Quadratic Model Involving One Quantitative Covariate, c, and Three Quantitative Experimental Factors, x_1, x_2, and x_3, in Section 5.2.2

Block	c	x_1	x_2	x_3
1	0	1	1	−1
1	0	0	0	1
1	0	−1	−1	−1
2	−1	−1	−1	1
2	−1	1	1	1
2	−1	1	−1	−1
2	−1	−1	1	0
3	1	−1	0	−1
3	1	1	0	1
3	1	0	−1	0
4	1	0	1	0
4	1	1	−1	−1
4	1	−1	−1	1
5	−1	0	1	−1
5	−1	1	0	0
6	0	0	0	0
7	0	0	0	0
7	0	1	−1	1
7	0	−1	1	1
8	0	0	0	0

model in x_1, x_2, x_3, and c,

$$
\begin{aligned}
Y_{ij} = {} & \beta_0 + \beta_1 c_i + \beta_2 x_{1ij} + \beta_3 x_{2ij} + \beta_4 x_{3ij} + \beta_5 c_i^2 \\
& + \beta_6 x_{1ij}^2 + \beta_7 x_{2ij}^2 + \beta_8 x_{3ij}^2 + \beta_9 c_i x_{1ij} \\
& + \beta_{10} c_i x_{2ij} + \beta_{11} c_i x_{3ij} + \beta_{12} x_{1ij} x_{2ij} \\
& + \beta_{13} x_{1ij} x_{3ij} + \beta_{14} x_{2ij} x_{3ij} + \gamma_i + \varepsilon_{ij}.
\end{aligned}
$$

Given that the pool of covariates consists of 12 possible blocks and that the model is full quadratic, the minimum and maximum numbers of blocks are 3 and 12. We therefore computed I-optimal designs with any numbers of blocks ranging from 3 to 12. It turns out that the I-optimal number of blocks is eight. The corresponding design has an average variance of prediction of 0.733. In terms of I-efficiency, the 8-block design is 2% better than the I-optimal design with the maximum number of blocks, 12. That design has an average variance of prediction of 0.748. The I-optimal 8-block design is displayed in Table 2. The design involves one block of four runs, four blocks of three runs, one block of two runs, and two blocks with just one run. The two blocks with just one run involve the center point $(x_1, x_2, x_3) = (0, 0, 0)$ at the middle level of the covariate.

5.2.3. A Mixture-Process Variable Model

In some cases, a stock of several mixtures is available for experimentation. In that case, the covariate values are the proportions of the ingredients composing the mixture and sum to one. Finding a D-optimal design for a model containing the covariates as well as some experimental factors then requires the use of a mixture-process variable model (Cornell (2002), Smith (2005)). We computed 21-run I- and D-optimal designs involving three ingredient proportions, c_1, c_2, and c_3 (the covariates), and two quantitative factors, x_1 and x_2, for the following mixture-process variable model:

$$
\begin{aligned}
Y_{ij} = {} & \beta_1 c_{1i} + \beta_2 c_{2i} + \beta_3 c_{3i} + \beta_4 c_{1i} c_{2i} + \beta_5 c_{1i} c_{3i} \\
& + \beta_6 c_{2i} c_{3i} + \beta_7 c_{1i} x_{1ij} + \beta_8 c_{2i} x_{1ij} + \beta_9 c_{3i} x_{1ij} \\
& + \beta_{10} c_{1i} x_{2ij} + \beta_{11} c_{2i} x_{2ij} + \beta_{12} c_{3i} x_{2ij} \\
& + \beta_{13} x_{1ij} x_{2ij} + \beta_{14} x_{1ij}^2 + \beta_{15} x_{2ij}^2 + \gamma_i + \varepsilon_{ij}.
\end{aligned}
\tag{16}
$$

As the set of covariate values for this example, we used a duplicated simplex-centroid design on a constrained mixture region defined by the inequality

$$0.2 \le x_i \le 0.4,$$

for $i = 1, 2, 3$. Hence, the set of covariates involves seven different combinations of covariate values twice. These seven combinations are shown in Figure 6. The duplication of the seven mixtures corresponds to a scenario in which 14 different batches are available, some of which have an identical formulation but have been prepared at different points in time or in different production plants. We assume that multiple samples can be taken from each batch for experimentation.

The 21-run I- and D-optimal designs we obtained are shown side by side in Table 3. The I-optimal design involves 11 blocks, whereas the D-optimal design has 12 blocks. Both designs have multiple blocks of size 1. The I-optimal design has four such blocks and the D-optimal design has six such blocks. So, only one sample is taken from several of the available batches. In addition to the four blocks of size 1, the I-optimal design has four blocks of size 2 and three blocks of size 3. In addition to the six blocks of size 1, the D-optimal design has four blocks of size 2, one block of size 3, and one block of size 4. So, both designs in-
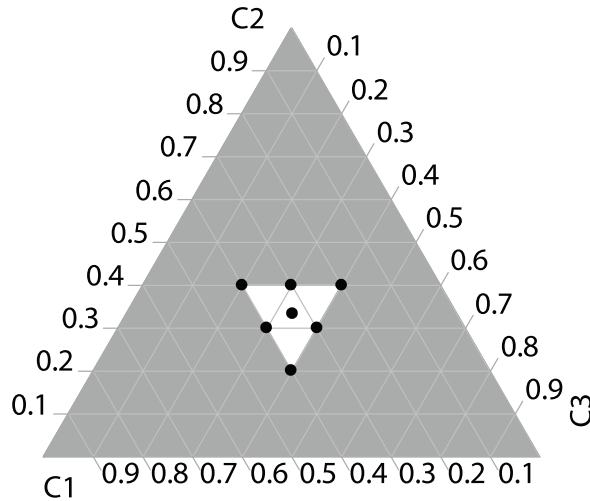
FIGURE 6. Covariate Values Used for the Optimal Designs for the Mixture-Process Variable Model in Section 5.2.3.

volve heterogeneous block sizes and neither of them utilizes the 14 available batches. A major difference between the D-optimal design and the I-optimal design is that the former does not use the covariate combination $(c_1, c_2, c_3) = (0.333, 0.333, 0.333)$. The I-optimal design in Table 3 results in an average variance of prediction of 0.692, while the D-optimal design gives an average variance of prediction of 0.850. Hence, the D-optimal design has an I-efficiency of 81.4%. The D-efficiency of the I-optimal design is 84.1%.

During our search for the I- and D-optimal designs in Table 3, we also computed D-optimal designs with other numbers of blocks. It turns out that the D-optimal 10-block, 11-block, 13-block, and 14-block designs are less than 1% less D-efficient than the 12-block design. Similarly, the I-optimal 12-block, 13-block, and 14-block designs perform nearly as well as the I-optimal 11-block design. The practical implica-

TABLE 3. I- and D-Optimal Designs for the Mixture-Process Variable Model Involving Three Quantitative Covariates c_1, c_2, and c_3 (Composing the Mixture) and Two Quantitative Experimental Factors x_1 and x_2 (the Process Variables) in Section 5.2.3

	I-optimal design						D-optimal design				
Block	c_1	c_2	c_3	x_1	x_2	Block	c_1	c_2	c_3	x_1	x_2
1	0.200	0.400	0.400	1	-1	1	0.200	0.400	0.400	1	-1
1	0.200	0.400	0.400	0	1	1	0.200	0.400	0.400	-1	1
1	0.200	0.400	0.400	-1	0	2	0.200	0.400	0.400	1	1
2	0.300	0.300	0.400	0	-1	2	0.200	0.400	0.400	-1	-1
2	0.300	0.300	0.400	1	1	3	0.300	0.300	0.400	0	1
3	0.300	0.300	0.400	-1	0	4	0.300	0.300	0.400	1	0
4	0.300	0.400	0.300	0	0	5	0.300	0.400	0.300	-1	1
4	0.300	0.400	0.300	-1	1	6	0.300	0.400	0.300	0	0
5	0.300	0.400	0.300	1	0	7	0.400	0.200	0.400	1	-1
5	0.300	0.400	0.300	-1	-1	7	0.400	0.200	0.400	-1	-1
6	0.333	0.333	0.333	0	0	7	0.400	0.200	0.400	-1	1
7	0.333	0.333	0.333	0	0	7	0.400	0.200	0.400	0	0
8	0.400	0.200	0.400	0	1	8	0.400	0.200	0.400	1	1
8	0.400	0.200	0.400	-1	-1	8	0.400	0.200	0.400	0	-1
8	0.400	0.200	0.400	1	-1	8	0.400	0.200	0.400	-1	0
9	0.400	0.300	0.300	-1	1	9	0.400	0.300	0.300	1	0
9	0.400	0.300	0.300	1	0	10	0.400	0.300	0.300	0	1
10	0.400	0.300	0.300	0	-1	11	0.400	0.400	0.200	1	-1
11	0.400	0.400	0.200	-1	0	11	0.400	0.400	0.200	-1	1
11	0.400	0.400	0.200	1	1	12	0.400	0.400	0.200	-1	-1
11	0.400	0.400	0.200	1	-1	12	0.400	0.400	0.200	1	1

TABLE 4. D-Optimal Design for the Running-Shoe Experiment Involving Eight Runners Testing Three Shoes Each

Runner	Mileage	Weight	Strike point	Thickness	Brand	Cushioning
1	16	146	Midfoot	-1	1	-1
1	16	146	Midfoot	1	-1	1
1	16	146	Midfoot	-1	1	1
2	33	146	Midfoot	-1	1	1
2	33	146	Midfoot	-1	-1	-1
2	33	146	Midfoot	1	1	-1
3	16	146	Heel	-1	1	-1
3	16	146	Heel	1	-1	1
3	16	146	Heel	-1	-1	-1
4	37	141	Heel	-1	1	1
4	37	141	Heel	1	-1	-1
4	37	141	Heel	-1	-1	1
5	28	138	Midfoot	1	1	-1
5	28	138	Midfoot	-1	-1	-1
5	28	138	Midfoot	1	-1	1
6	15	150	Heel	1	1	-1
6	15	150	Heel	1	1	1
6	15	150	Heel	-1	-1	-1
7	32	139	Heel	-1	1	1
7	32	139	Heel	1	-1	1
7	32	139	Heel	1	1	-1
8	23	152	Midfoot	-1	-1	1
8	23	152	Midfoot	1	1	1
8	23	152	Midfoot	1	-1	-1

tion of this result is that the investigator may choose the number of blocks for logistical convenience.

6. The Motivating Examples Revisited

6.1. Running-Shoe Example

For the running-shoe experiment, we need to create a D-optimal design involving 24 runs and eight runners for estimating a main-effects model in three covariates and three experimental factors. The eight runners have to be selected from a pool of 100. The available covariate information consists of the runners' weekly mileage, their weight, and their strike point. The strike point is a categorical covariate with

two levels, 'heel' and 'midfoot'. The two quantitative covariates, mileage and weight, have a substantial negative correlation of -0.6736. The weekly mileage ranges from 15 to 37 and the weight ranges from 138 to 152 pounds. We provide the covariate values for each of the 100 available runners in the supplementary materials available at http://www.asq.org/pub/jqt/. The three experimental factors in the running-shoe experiment are thickness, brand, and cushioning. The latter two experimental factors are two-level categorical factors.

The D-optimal design is shown in Table 4. The D-optimal selection of eight runners out of the pool of 100 available runners is graphically shown in Figure
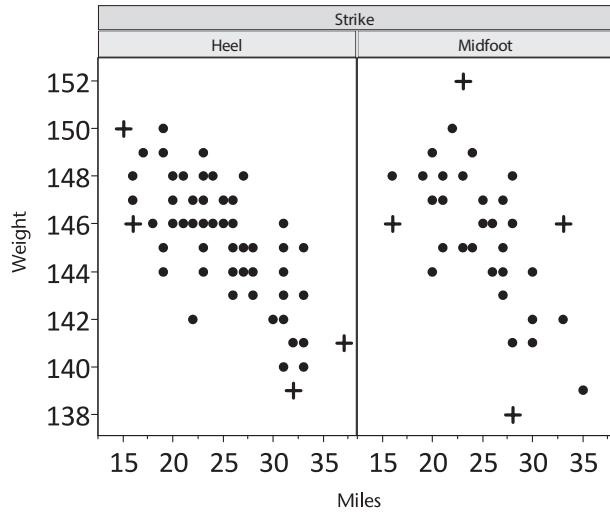
FIGURE 7. Scatter Plots Showing the Covariate Values in the Running-Shoe Example. Plus signs indicate the eight runners appearing in the D-optimal design.

7. In the figure, the plus signs indicate the selected runners, whereas the filled circles indicate those not selected. From the figure, it can be seen that half of the selected runners land on their heels and that the other half of them land on their midfoot. Also, the selected runners have extreme values for the two quantitative covariates, mileage and weight. For the subset of eight selected runners, the correlation between mileage and weight is still -0.6020.

Due to the strong negative correlation between the mileage and the weight of the selected runners, the information matrix corresponding to the D-optimal design is not diagonal. Assuming that $\sigma_\varepsilon^2 = \sigma_\gamma^2 = 1$, Matrix A is the information matrix and its determinant is 2.096×10^6. Matrix B, the variance–covariance matrix of the estimates shows that, with the D-optimal design, almost all pairs of parameter estimates are either uncorrelated or nearly uncorrelated. Note that the first row and column of the information matrix and the variance–covariance matrix correspond to the intercept and that the next ones correspond to the weekly mileage, weight, strike point, thickness, brand, and cushioning. The levels of strike point, brand, and cushioning were coded as -1 and 1. The meaning of the diagonal lower-right 4×4 submatrix is that the estimates of the main effects of the experimental factors thickness, brand, and cushioning are uncorrelated and that these factors are orthogonal to the categorical covariate strike point.
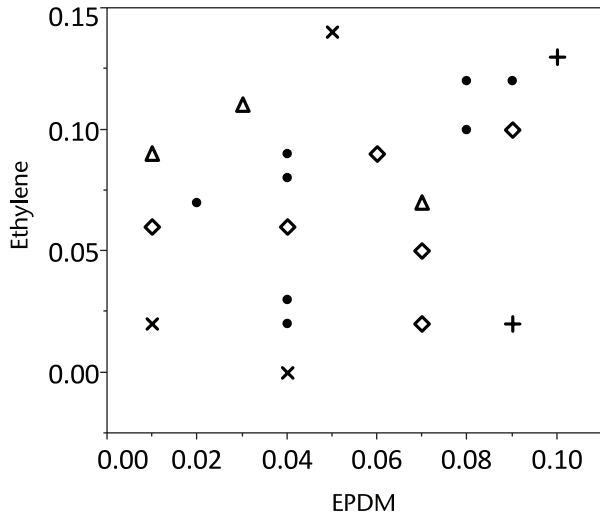
6.2. Polypropylene Example

Consider an experiment focusing on the 23 polypropylene formulations without EVA listed in Goos and Jones (2011), where the interest is in three quantitative experimental factors (reaction time, gas flow rate, and power), in addition to the two quantitative covariates (the proportion of EPDM and the proportion of ethylene in the formulation). The

MATRIX A

$$
\begin{bmatrix}
6 & -0.545 & -0.214 & 0 & 0 & 0 & 0 \\
-0.545 & 3.347 & -1.782 & 0 & -0.091 & -0.182 & 0.364 \\
-0.214 & -1.782 & 2.724 & -0.643 & 0 & 0.143 & -0.071 \\
0 & 0 & -0.643 & 6 & 0 & 0 & 0 \\
0 & -0.091 & 0 & 0 & 22 & 0 & 0 \\
0 & -0.182 & 0.143 & 0 & 0 & 22 & 0 \\
0 & 0.364 & -0.071 & 0 & 0 & 0 & 22
\end{bmatrix}
$$

MATRIX B

$$
\begin{bmatrix}
0.174 & 0.056 & 0.051 & 0.006 & 0 & 0 & -0.001 \\
0.056 & 0.484 & 0.329 & 0.035 & 0.002 & 0.002 & -0.007 \\
0.051 & 0.329 & 0.602 & 0.064 & 0.001 & -0.001 & -0.003 \\
0.006 & 0.035 & 0.064 & 0.174 & 0 & 0 & 0 \\
0 & 0.002 & 0.001 & 0 & 0.045 & 0 & 0 \\
0 & 0.002 & -0.001 & 0 & 0 & 0.045 & 0 \\
-0.001 & -0.007 & -0.003 & 0 & 0 & 0 & 0.046
\end{bmatrix}
$$

(a) I-optimal design
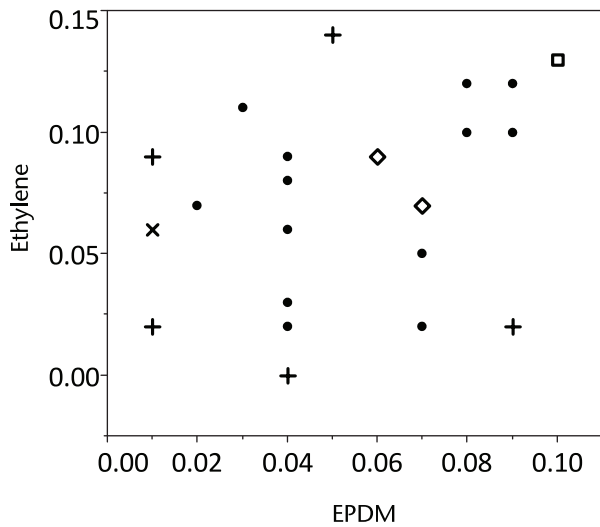
(b) D-optimal design

FIGURE 8. Selected and Nonselected Combinations of Covariate Values in the Optimal Designs for the Polypropylene Experiment with Two Quantitative Covariates and a Full Quadratic Model. A square indicates a selected combination with block size 5. A plus indicates a selected combination with block size 4. A cross indicates a selected combination with block size 3. A triangle indicates a selected combination with block size 2 and a diamond indicates a selected combination with block size 1. Filled circles indicate nonselected combinations.

model of interest is the full quadratic model in the five variables, involving 21 parameters.

We generated I- and D-optimal 30-run designs for this model, with the number of blocks ranging from 7 (the minimum required for the covariates' effects and the variance of the block effects to be estimable) to 23 (the number of boxes with polypropylene formulations available). The I-optimal design option has 15 blocks, whereas the D-optimal design only has 9 blocks. In both designs, the blocks have unequal sizes. The I-optimal design has two blocks of size 4, three blocks of size 3, three blocks of size 2, and seven blocks of size 1. The D-optimal design has one block of size 5, five blocks of size 4, one block of size 3, no blocks of size 2, and two blocks of size 1.

Figure 8 shows which covariate combinations correspond to large blocks, which combinations correspond to small blocks, and which combinations are not selected in the I- and D-optimal designs. In the figure, filled circles represent combinations of covariate values that were not selected in the optimal designs. The I-optimal design clearly involves a larger number of covariate combinations in the center of the experimental region than the D-optimal design. These centrally located combinations are all indicated using a diamond or a triangle, meaning that the corresponding block size is only 1 or 2. This result is in line with the results of Jones and Goos (2012), who compared D- and I-optimal split-plot designs and observed that I-optimal split-plot designs tend to involve larger numbers of center runs for the full quadratic model.

The largest number of selected combinations of covariates values in the I- and D-optimal designs for the polypropylene example are located near the edges of the experimental region. The block sizes corresponding to these combinations are 2, 3, 4, and 5. These block sizes are indicated using a triangle, a cross, a plus, and a square, respectively.

While Figure 8 focuses on the covariates, Table 5 presents the complete I- and D-optimal designs for the polypropylene experiment and the full quadratic model. The I-optimal design involves three center runs in the factors time, flow, and power, whereas the D-optimal design involves only center run. The I-optimal design visits the middle level of the factors time, flow, and power 10, 9, and 9 times, respectively. For the D-optimal design, the frequency with which the middle level of the factors time, flow, and power appears is 5, 6, and 6, respectively.

Table 6 lists the I- and D-efficiencies of all designs we constructed in our search for the overall I- and D-optimal designs, along with the average variance of

TABLE 5. I- and D-Optimal Designs for a Polypropylene Experiment Involving Two Quantitative Covariates (EPDM and Ethylene), Three Quantitative Experimental Factors (Time, Flow, and Power) and a Full Quadratic Model

I-optimal design						D-optimal design					
Block	EPDM	Ethyl.	Time	Flow	Power	Block	EPDM	Ethyl.	Time	Flow	Power
1	0.01	0.02	−1	1	−1	1	0.01	0.02	1	−1	0
1	0.01	0.02	0	1	0	1	0.01	0.02	−1	1	−1
1	0.01	0.02	1	−1	−1	1	0.01	0.02	−1	−1	1
2	0.01	0.06	0	0	1	1	0.01	0.02	1	0	1
3	0.01	0.09	−1	−1	0	2	0.01	0.06	−1	−1	0
3	0.01	0.09	1	1	−1	2	0.01	0.06	0	1	1
4	0.03	0.11	0	0	0	2	0.01	0.06	1	0	−1
4	0.03	0.11	−1	1	1	3	0.01	0.09	1	1	0
5	0.04	0.00	0	0	0	3	0.01	0.09	0	−1	−1
5	0.04	0.00	−1	−1	1	3	0.01	0.09	1	−1	1
5	0.04	0.00	1	1	1	3	0.01	0.09	−1	0	1
6	0.04	0.06	1	0	0	4	0.04	0.00	1	−1	−1
7	0.04	0.06	0	1	−1	4	0.04	0.00	0	−1	1
8	0.05	0.14	0	1	0	4	0.04	0.00	1	1	1
8	0.05	0.14	1	−1	1	4	0.04	0.00	−1	0	0
8	0.05	0.14	−1	0	−1	5	0.05	0.14	1	1	−1
9	0.06	0.09	0	−1	−1	5	0.05	0.14	−1	−1	−1
10	0.07	0.02	0	1	−1	5	0.05	0.14	−1	1	1
11	0.07	0.05	−1	0	0	5	0.05	0.14	1	−1	1
12	0.07	0.07	0	0	1	6	0.06	0.09	−1	1	0
12	0.07	0.07	1	1	0	7	0.07	0.07	0	0	−1
13	0.09	0.02	1	0	−1	8	0.09	0.02	−1	−1	−1
13	0.09	0.02	1	−1	1	8	0.09	0.02	−1	1	1
13	0.09	0.02	−1	−1	−1	8	0.09	0.02	1	−1	1
13	0.09	0.02	−1	1	1	8	0.09	0.02	1	1	−1
14	0.09	0.10	0	0	0	9	0.10	0.13	1	1	1
15	0.10	0.13	−1	−1	1	9	0.10	0.13	−1	−1	1
15	0.10	0.13	1	−1	−1	9	0.10	0.13	1	−1	−1
15	0.10	0.13	1	1	1	9	0.10	0.13	−1	1	−1
15	0.10	0.13	−1	1	−1	9	0.10	0.13	0	0	0

prediction for the I-optimal designs and the determinant of the information matrix for the D-optimal designs. The efficiencies show that restricting the number of blocks to 15 in case the I-optimality criterion is preferred and to 9 in case the D-optimality criterion is preferred results in substantially higher efficiencies than when all 23 available combinations of the covariate values would be used. The gain in I-efficiency is 11.2%, while the gain in D-efficiency is 35.3%.

7. Discussion

When setting up experiments, it is desirable to use all available information. In this article, we show

how to take advantage of covariate information that is available about potential participants in a study or about batches of materials or mixtures that might be used. One main challenge then is to select the participants or batches to be used in the experiment. Motivated by two practical examples, we distinguish between a scenario in which it is convenient to use each participant or batch equally often and a scenario in which there is no such constraint.

In general terms, we deal with the design of blocked experiments, where certain properties of the blocks have been measured using one or more categorical or quantitative variables, called covariates.

TABLE 6. I- and D-Efficiencies for the I- and D-Optimal Designs for the Polypropylene Experiment

Number of blocks	Average variance of prediction	I-efficiency	Determinant of information matrix	D-efficiency
7	1.1439	0.852	9.9635E+17	1.335
8	1.0239	0.952	1.1470E+18	1.343
9	0.9577	1.018	1.3325E+18	1.353
10	0.9309	1.047	1.1475E+18	1.344
11	0.9211	1.058	9.3051E+17	1.330
12	0.9064	1.075	8.3749E+17	1.324
13	0.8809	1.106	6.2721E+17	1.305
14	0.8884	1.097	5.2843E+17	1.295
15	0.8762	1.112	3.3504E+17	1.267
16	0.8870	1.099	2.6868E+17	1.254
17	0.8924	1.092	1.4911E+17	1.219
18	0.8931	1.091	8.2315E+16	1.185
19	0.9092	1.072	4.2793E+16	1.149
20	0.9253	1.053	3.2954E+16	1.135
21	0.9436	1.033	1.1977E+16	1.081
22	0.9727	1.002	6.1154E+15	1.047
23	0.9745	1.000	2.3260E+15	1.000

The approach we propose allows the experimenter to quantify the impact of the covariates on the response and also to estimate possible interactions between the covariates and the experimental factors. Therefore, our approach relaxes the traditional assumption that there are no interactions between the blocking factor of an experiment and the experimental factors. This is possible because we treat the block effects as random effects. Therefore, our model is similar to that for data from split-plot experiments.

Using various proof-of-concept examples and an actual polypropylene experiment, we show that, in many situations, it is better to use blocks with unequal numbers of runs than equally sized blocks. In the polypropylene example, carrying out an experiment with unequally sized blocks is as convenient as carrying out an experiment with equally sized blocks. Moreover, the cost of the polypropylene experiment is independent of the number of blocks. In some practical situations, however, it may be desirable to limit the number of blocks for logistical or cost reasons. In that case, it still makes sense to use our methodology because it may be better to use fewer blocks than the budget allows.

The computation of I-optimal designs in the presence of covariate information requires some serious consideration of the region of interest. If the avail-

able covariate values cover the region of interest well, then it makes sense to compute the average variance of prediction of any given design over the region covered by the covariates. If, however, the region of interest is different from the area covered by the covariate values, the region over which to compute the average prediction variance has to be redefined. If, for instance, an I-optimal design were desired for the running-shoe example and predictions need to be made for heavy-weight runners with a large weekly mileage and low-weight runners with a low weekly mileage, then the region over which to compute the average prediction variance should be larger than if these types of runners would not be considered.

We supply JMP Scripting Language (JSL) scripts for each algorithm in the supplementary materials. In general, the algorithm allowing for unequally sized blocks generates designs that are more statistically efficient. Use of the algorithm for equally sized blocks should be limited to cases where logistical considerations demand a fixed number of experimental units combined with a fixed number of runs for each experimental unit.

In cases similar to the running-shoe example, it could be that the correlations of observations within a subject might not be constant as it is in our models. One possibility is that an AR(1) structure for the

correlation might be desirable in some cases. This is an area under current investigation.

References

ATKINSON, A. C. and DONEV, A. N. (1989). "The Construction of Exact D-Optimum Experimental Designs with Application to Blocking Response Surface Designs". *Biometrika* 76, pp. 515–526.

ATKINSON, A. C. and DONEV, A. N. (1996). "Experimental Designs Optimally Balanced for Trend". *Technometrics* 38, pp. 333–341.

CHENG, C.-S. (1990). "Construction of Run Orders of Factorial Designs". In *Statistical Design and Analysis of Industrial Experiments*, Ghosh, S., ed., pp. 423–439. New York, NY: Marcel Dekker.

CHENG, C.-S. and JACROUX, M. (1988). "The Construction of Trend-Free Run Orders of Two-Level Factorial Designs". *Journal of the American Statistical Association* 83, pp. 1152–1158.

CHENG, C.-S. and STEINBERG, D. M. (1991). "Trend Robust Two-Level Factorial Designs". *Biometrika* 78, pp. 325–336.

COOK, R. D. and NACHTSHEIM, C. J. (1980). "A Comparison of Algorithms for Constructing Exact D-Optimal Designs". *Technometrics* 22, pp. 315–324.

COOK, R. D. and NACHTSHEIM, C. J. (1989). "Computer-Aided Blocking of Factorial and Response-Surface designs". *Technometrics* 31, pp. 339–346.

COOK, R. D. and THIBODEAU, L. A. (1980). "Marginally Restricted D-Optimal Designs". *Journal of the American Statistical Association* 75, pp. 366–371.

CORNELL, J. A. (2002). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*. New York, NY: Wiley.

DANIEL, C. and WILCOXON, F. (1966). "Factorial 2^{p-q} Plans Robust Against Linear and Quadratic Trends". *Technometrics* 8, pp. 259–278.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. New York, NY: Academic Press.

GOOS, P. (2002). *The Optimal Design of Blocked and Split-plot Experiments*. New York, NY: Springer.

GOOS, P. and DONEV, A. N. (2006). "Blocking Response Surface Designs". *Computational Statistics and Data Analysis* 51, pp. 1075–1088.

GOOS, P. and GILMOUR, S. (2012). "Analyzing Categorical Data from Split-Plot and Other Multi-Stratum Experiments." *Technometrics* 54, pp. 340–354.

GOOS, P. and JONES, B. (2011). *Design of Experiments: A Case-Study Approach*. New York, NY: Wiley.

GOOS, P.; LANGHANS, I.; and VANDEBROEK, M. (2006). "Practical Inference from Industrial Split-Plot Designs". *Journal of Quality Technology* 38, pp. 162–179.

GOOS, P. and VANDEBROEK, M. (2001). "D-Optimal Response Surface Designs in the Presence of Random Block Effects". *Computational Statistics and Data Analysis* 37, pp. 433–453.

GOOS, P. and VANDEBROEK, M. (2003). "D-Optimal Split-Plot Designs with Given Numbers and Sizes of Whole Plots". *Technometrics* 45, pp. 235–245.

GOOS, P. and VANDEBROEK, M. (2004). "Outperforming Completely Randomized Designs". *Journal of Quality Technology* 36, pp. 12–26.

HARDIN, R. H. and SLOANE, N. J. A. (1991a). "Computer-Generated Minimal (and Larger) Response-Surface Designs. I: The Sphere". http://www2.research.att.com/njas/doc/doeh.pdf.

HARDIN, R. H. and SLOANE, N. J. A. (1991b). "Computer-Generated Minimal (and Larger) Response-Surface Designs. II: The Cube". http://www2.research.att.com/njas/doc/meatball.pdf.

HARVILLE, D. A. (1974). "Nearly Optimal Allocation of Experimental Units Using Observed Covariate Values". *Technometrics* 16, pp. 589–599.

HILL, H. M. (1960). "Experimental Designs to Adjust for Time Trends". *Technometrics* 2, pp. 67–82.

JOHN, P. W. M. (1990). "Time Trends and Factorial Experiments". *Technometrics* 32, pp. 275–282.

JOINER, B. L. and CAMPBELL, C. (1976). "Designing Experiments When Run Order Is Important". *Technometrics* 18, pp. 249–259.

JONES, B. and GOOS, P. (2007). "A Candidate-Set-Free Algorithm for Generating D-Optimal Split-Plot Designs". *Journal of the Royal Statistical Society, Series C: Applied Statistics* 56, pp. 347–364.

JONES, B. and GOOS, P. (2012). "I-Optimal Versus D-Optimal Split-Plot Response Surface Designs". *Journal of Quality Technology* 44, pp. 85–101.

KESSELS, R.; GOOS, P.; and VANDEBROEK, M. (2008). "Optimal Designs for Rating-Based Conjoint Experiments". *Computational Statistics and Data Analysis* 52, pp. 2369–2387.

KHURI, A. I. (1996). "Response Surface Models with Mixed Effects". *Journal of Quality Technology* 28, pp. 177–186.

LETSINGER, J. D.; MYERS, R. H.; and LENTNER, M. (1996). "Response Surface Methods for Bi-Randomization Structures". *Journal of Quality Technology* 28, pp. 381–397.

MEE, R. W. and ROMANOVA, A. V. (2010). "Constructing and Analyzing Two-Level Trend-Robust Designs". *Quality Engineering* 22, pp. 306–316 and "Corrigendum". *Quality Engineering* 23, 112.

MEYER, R. K. AND Nachtsheim, C. J. (1995). "The Coordinate-Exchange Algorithm for Constructing Exact Optimal Experimental Designs". *Technometrics* 37, pp. 60–69.

MYLONA, K.; GOOS, P.; AND Jones, B. (2014). "Optimal Design of Blocked and Split-Plot Experiments for Fixed Effects and Variance Component Estimation". *Technometrics* 56, pp. 132–144.

NACHTSHEIM, C. J. (1989). "On the Design of Experiments in the Presence of Fixed Covariates". *Journal of Statistical Planning and Inference* 22, pp. 203–212.

SMITH, W. F. (2005). *Experimental Design for Formulation*. Philadelphia, PA: SIAM.

VINING, G. G.; KOWALSKI, S. M.; and MONTGOMERY, D. C. (2005). "Response Surface Designs Within a Split-Plot Structure". *Journal of Quality Technology* 37, pp. 115–128.

WU, C. F. J. and HAMADA, M. (2000). *Experiments: Planning, Analysis, and Parameter Design Optimization*. New York, NY: Wiley.

∼