Journal of Statistical Software

November 2009, Volume 32, Code Snippet 1. *http://www.jstatsoft.org/*

Blocking Mixed-Level Factorials with SAS

John Lawson
Brigham Young University

G. Bruce Schaalje
Brigham Young University

Bruce Jay Collings
Brigham Young University

Abstract

Orthogonal array screening designs for mixed-level or asymmetric factorials have recently become popular. Tables of designs and software for creating these designs are readily available to practitioners. However, confounded block designs for mixed-level factorials are not as popular partly due to the fact that software for creating these designs has not been well publicized. Classical methods for creating confounded-block mixed-level factorials normally described in textbooks utilize modular arithmetic or finite fields. In the recent literature optimal design theory has also been proposed as method for creating these designs. Although no examples are shown in the online documentation, both classical and optimal confounded-block mixed-level factorials can be easily created using SAS data step programming in conjunction with `proc plan`, `proc factex` or `proc optex`. In this article we show examples of creating these designs in SAS, and we compare the properties of designs created by classical methods and optimal design theory.

Keywords: complete confounding, partial confounding, asymmetric design, SAS.

1. Introduction

Orthogonal array screening experiments for mixed-level or asymmetric factorials are used frequently in engineering and marketing experiments (Kuhfeld and Tobias 2005). This is partly due to the fact that tables and software for creating these designs are readily available (Addelman 1962; Taguchi 1987; Wu and Hamada 2000). Once the important factors have been identified in a screening experiment, a follow-up full-factorial with the important factors is the next logical step. The purpose for using a full-factorial experiment is to estimate interactions among the factors. However, the number of treatment combinations in a full-factorial is often too large to test in a homogenous block of experimental units. Incomplete block factorial experiments can be used to reduce the variability of experimental units within a block and therefore increase the precision for estimating and testing treatment effects. For this reason, experimenters conducting full factorial experiments should always consider blocking their experiments by batch of raw materials, day of experimentation, or any other convenient way of grouping homogeneous experimental units. However, with factorial experiments, the advantages of blocking can be negated when important interactions are confounded with block effects.

When the block size, or number of experimental units in a block, is smaller than the number of treatment combinations in a factorial design, some type of incomplete block design is required. In a completely confounded-block design only one repetition or replicate of the factorial design is run, and the ability to estimate some interactions will be lost because they must be confounded or aliased with blocks. If there are several factors in a design, and many higher order interactions, a completely confounded design is useful if the higher order interactions can be confounded with blocks leaving the lower order interactions estimable. When there are only two or three factors in a factorial design it would be undesirable to completely confound any interaction with blocks. In this case, if more than one replication of the experiment can be run, different interactions can be confounded in different replicates, allowing all interactions to be estimated. This is called a partially confounded design.

For symmetric 2^p factorials, the methods for finding the best possible confounding for a completely confounded design and the methods for creating a partially confounded design are well known and are available in many menu-driven software packages that are routinely used by practitioners. Use of such statistical software for creating experimental designs has become commonplace in research. However, methods for creating confounded asymmetric or mixed-level factorials using statistical software are less well known. Few standard textbooks on experimental design even describe how to create confounded mixed-level factorials manually, and these methods are not available in the popular menu driven software used for designing experiments such as Minitab (Minitab Inc. 2008), SAS **ADX** (SAS Institute Inc. 2008), or **Design-Ease** (**Stat-Ease**, Inc. 2008). Even though blocked mixed-level factorials are very useful, few examples of confounded mixed-level factorial designs can be found in the statistical applications literature. One reason for so few applications of blocked mixed-level factorials may be a lack of knowledge on the part of practitioners regarding how to create these designs easily with standard statistical software. Thus, the upside-down situation occurs in which research problems are modified to fit designs that are available rather than creating a design that suits the research problem. In this article we will illustrate how SAS, a package that is widely used in industry for analysis of experimental data, can be used to easily create a wide variety of blocked mixed-level factorial designs with desirable properties.

2. Completely confounded mixed-level factorials

A mixed-level or asymmetric factorial can be represented by $s_1^{m_1} \times s_2^{m_2} \times \cdots s_\gamma^{m_\gamma}$, involving $n = \sum_{i=1}^{\gamma} m_i$ factors where m_i factors each have s_i levels. If the experimental units are heterogeneous, the precision for detecting main effects and interactions can be greatly enhanced if the experimental units can be grouped into blocks of size k. For industrial experiments, blocks might represent raw material from different batches or experiments performed on different days in a plant or laboratory. If the practical block size k is less than the total number of combinations of treatment factor levels n, some type of incomplete block design is required. If the number of experiments is restricted to one full replicate of the factorial, some interactions must be confounded with block contrasts and will be inestimable. This is called a completely confounded-block design. The challenge is to find a design that does not confound block contrasts with the lower order interactions and interactions of interest to the experimenter.

The classical method of confounding in symmetric factorials where each factor has the same (prime) number of levels was described in a paper by Bose (1947). In symmetrical designs, Franklin (1985) proposed an algorithm to find a block-defining contrast that will allow estimation of the effects and interactions specified by the experimenter. There have been many generalizations of the classical method of confounding to asymmetrical or mixed-level factorials. Voss (1986) explains that many of these methods are equivalent to applying the classical method separately to symmetric sub-experiments that each involve only prime-leveled factors. He shows that using this classical method separately for each symmetric sub-experiment can result in an overall design with a predetermined confounding pattern.

The model for a blocked factorial experiment can be written in matrix notation as

$$y = X\tau + Z\beta + \epsilon \tag{1}$$

where y is the $n \times 1$ vector of responses, τ is the vector of estimable treatment effects and interactions, and β is the vector of block effects. One optimality criterion that has been proposed for blocked designs is the D_s criteria, see Atkinson, Donev, and Tobias (2007). A D_s optimal design is one that minimizes the covariance matrix of the least squares estimator for τ or equivalently maximizes the determinant of $X^\top QX$ where

$$Q = I - Z(Z^\top Z^{-1} Z^\top). \tag{2}$$

Designs where blocks are orthogonal to treatment effects (i.e., $X^\top Z = 0$ or $X^\top QX = X^\top X$) are 100% D_s efficient.

Applying the classical method separately to symmetric sub-experiments results in designs that have known confounding patterns and are 100% D_s efficient for estimating the effects and interactions, τ, that are not confounded with the block contrasts (since they will be orthogonal to blocks). However, in practical situations, use of the classical approach does not provide much flexibility in the choice of block size or in the choice of interactions to be confounded with block differences. Since the sub-experiments are often defined in terms of pseudo-factors, interactions of interest often become confounded with blocks.

Cook and Nachtsheim (1989) describe a more general computer algorithm for creating blocked designs by beginning with a non-singular starting design, then sequentially exchanging treatment combinations assigned to one block with those assigned to other blocks in order to maximize $|X^\top QX|$. The designs resulting from this algorithm may not be 100% D_s efficient for estimating τ, but greater choices of block sizes and estimable interactions are possible.

Although no examples are shown in the SAS online documentation, SAS `proc factex` is capable of creating completely confounded mixed-level factorial designs using the classical method for each symmetric sub-experiment. In addition, SAS `proc optex` is capable of creating blocked designs using the algorithm of Cook and Nachtsheim (1989). In the next section we will illustrate the use of these two SAS procedures and present an example that illustrates the difference between the designs that can be obtained from them.

2.1. An example

An example from Collings (1984) involves blocking a $3 \times 4 \times 6$ factorial into blocks of size 12. Factor A has three levels. Factor B has four levels that can be represented as all combinations of two two-level pseudo-factors b_1 and b_2. Factor C has six levels that can be represented as all possible combinations of a two-level pseudo-factor c_1 and a three-level pseudo-factor c_2. Thus the treatment combinations of the $3 \times 4 \times 6$ factorial in factors A, B and C can be represented by the treatment combinations in a $3^2 \times 2^3$ factorial. The symmetric 3^2 sub-experiment involves factors A and c_2, and the symmetric 2^3 sub-experiment involves factors b_1, b_2, and c_1. In this experiment there are $(3-1)+(4-1)+(6-1) = 10$ degrees of freedom for main effects, $2 \times 3 + 2 \times 5 + 3 \times 5 = 31$ degrees of freedom for two-factor interactions, $2 \times 3 \times 5 = 30$ degrees of freedom for the three-factor interaction, and $(3 \times 4 \times 6)/12 - 1 = 5$ degrees of freedom for blocks.

Using the classical method separately for each sub-experiment, the 3^2 sub-experiment can be split into 3 blocks by confounding two degrees of freedom from the interaction between the three-level factor A and three-level pseudo-factor c_2. The 2^3 sub-experiment can be split into two blocks of four by confounding the $b_1 \times b_2 \times c_1$ interaction. Care must be taken to not confound an interaction involving only pseudo-factors of the same main effect, such as $b_1 \times b_2$, in a sub-experiment, because that will confound part of the main effect B. After confounding in the sub-experiments, each block from the first sub-experiment is combined with each block of the second sub-experiment to result in $3 \times 2 = 6$ blocks involving all the factors. The $5 = 6 - 1$ degrees of freedom for blocks will be confounded with two degrees of freedom from the AC interaction that resulted from confounding the interaction between A and pseudo-factor c_2 in the first sub-experiment, one degree of freedom from the BC interaction that resulted from confounding $b_1 \times b_2 \times c_1$ in the second sub-experiment, and two degrees of freedom from the ABC interaction that resulted from the generalized interaction of $A \times c_2$ and $b_1 \times b_2 \times c_1$ when the two sub-experiments were combined. The main effects and the AB interaction are not confounded with blocks and can be estimated with 100% D_s efficiency from this design. Each of the six blocks will contain 12 experimental units.

In this case the classical method allows creation of a design where the estimable effects are known, but there is no flexibility. Any other choice of confounded effects in the sub-experiments would result in confounding a main effect, and if the convenient block size was smaller than 12, there would be no way to proceed since the block size is determined by the combination of blocks from each sub-experiment. However, since there are 30 degrees of freedom for the three-factor interaction ABC, using Cook and Nachtsheim's algorithm enables the user to find a design that does not confound any two-factor interaction. In general, when using the classical method to confound within the sub-experiments, the block size will be determined from the combination of blocks from sub-experiments, and the interactions that are confounded are determined by those confounded within the sub-experiments and their generalized interactions. Since there are fewer factors in the sub-experiments than the whole factorial, there is less flexibility in choice of interactions to confound.

2.2. Description of SAS code

The SAS code shown in Table 1 illustrates the use of `proc factex` to create a design using the classical method within each sub-experiment.

The first use of `proc factex` does the confounding in the 3^2 sub-experiment. The `blocks`

```
proc factex;
   factors A c2/nlev=3;
   blocks nblocks=3;
   model estimate=(A c2);
   examine aliasing;
   output out=d1 ;

proc factex;
   factors b1 b2 c1;
   blocks nblocks=2;
   model estimate=(b1|b2|c1@2);
   examine aliasing;
   output out=d2 pointrep=d1;

proc print; run;
```

Table 1: This code illustrates the classical method of confounding in sub-experiments; the first use of `proc factex` confounds the $A \times c_2$ interaction in the 3^2 sub-experiment, and the second use of `proc factex` confounds the $b_1 \times b_2 \times c_1$ interaction in the 2^3 sub-experiment.

statement specifies three blocks. The `model` statement asks `proc factex` to search for a block defining contrast that will allow estimation of the two main effects A and c_2. This means the two-factor interaction will be confounded with blocks. The `examine aliasing` statement will print a table showing what block defining contrast was used. The `output` statement directs the resulting design to a SAS file `d1`.

The second use of `proc factex` does the confounding in the 2^3 sub-experiment. Since there is no `nlev` option on the factors statement, the default (two-levels) is used. This time the `model` statement asks `proc factex` to search for a block defining contrast that will allow estimation of all main effects and two-factor interactions. This means the three-factor interaction will be confounded with blocks. The `pointrep=d1` option on the `output` statement tells `proc factex` to combine each block of the resulting design with each block of the design from the first sub-experiment that was stored in `d1`.

A look at the resulting output shows that `proc factex` used the label `BLOCK` for the block indicators from the 2^3 sub-experiment, and the label `BLOCK2` for the block indicators from the 3^2 sub-experiment. The levels of the two-level factors are `-1` and `1`, and the levels of the three-level factors are `-1`, `0`, and `1`. The next SAS data step shown in Table 2 combines levels of the pseudo-factors into the levels of the four-level factor B and the six-level factor C and combines the block labels.

Table 3 shows how to use `proc optex` to create a Cook-Nachtsheim blocked design that allows estimation of all two-factor interactions.

In the first block of code, `proc plan` is used to create the full $3 \times 4 \times 6$ factorial and stores it in the file (`cdesign`). In the second block of code, the `model` statement in `proc optex` specifies that all main effects and two-factor interactions be estimable (or not completely confounded

```
data df; set d2;
   A=A+1;
   B=(b2+1)+((b1+1)/2);
   C=(c2+1)+(3*(c1+1)/2);
   block=3*(block-1)+block2;
   keep block A B C;

proc sort; by block;
proc print; run;
```

Table 2: This code combines psuedo-factor levels and block indicators.

```
proc plan;
  factors A=3 B=4 C=6;
  output out=cdesign A nvals=(0 1 2)
                     B nvals=(0 1 2 3)
                     C nvals=(0 1 2 3 4 5);

proc optex data=cdesign coding=orthcan seed=73565;
   class A B C;
   model A B C A*B A*C B*C;
   blocks structure=(6)12 init=chain noexchange;
   generate initdesign=cdesign method=sequential;
   output out=bdesign blockname=blk; run;

proc print data=bdesign; run;
```

Table 3: This code creates a D_s optimal block design. `proc plan` creates the full factorial and `proc optex` optimally blocks the design into 6 blocks of size 12.

with blocks). The `blocks` and the `generate` statements specify that a design with all 72 treatment combinations be created that is blocked into 6 blocks of size 12. The resulting design is stored in the file `bdesign`. When this code is run, `proc optex` finds a design that is 87.35% D_s efficient for estimating the main effects and two-factor interactions. These estimable effects are not orthogonal to the blocks, and the analysis of data from this model must be made with a general linear model program like SAS `proc glm`, but no part of any two-factor interaction is completely confounded with blocks, unlike the case with the design created using the classical approach. Furthermore, using Cook and Nachtsheim's algorithm there is much more flexibility in choice of block size. By changing the `blocks` statement to `structure=(9)8;` a design with 9 blocks of size 8, and D_s efficiency of 83.50% is found. Changing to `blocks structure=(12)6;` results in a design in 12 blocks of size 6 with 78.94% D_s efficiency for estimating the main effects and two-factor interactions, etc.

3. Partially confounded mixed-level factorials

When a factorial experiment includes only two or three factors, normally one of the main purposes of running the experiment is to estimate all of the interaction effects. If interactions are confounded with blocks in order to reduce the block size and increase the precision, the ability to estimate all the interactions will be lost. However, by running more than one replicate of a factorial design, certain interactions can be confounded with blocks in one replicate of the design while different interactions are confounded with blocks in other replicates. By combining the replicates, the experimenter can recover the estimability of the confounded interactions, although they will be estimated with less precision than the main effects and interactions that were not confounded in any replicate. Designs set up in this way are called partially confounded factorial designs. Fisher (1942) described how partially confounded designs can be created with symmetric factorials.

For asymmetric or mixed-level factorials, partially confounded designs can also be created. Different approaches have been discussed in the literature for creating partially confounded mixed-level factorials. One approach is to create a balanced confounded design so that (1) the information recovered for each degree of freedom for any partially confounded interaction is the same, and (2) any contrast of a partially confounded interaction is estimable independently of any contrast of another partially confounded interaction. The information recovered for the ith degree of freedom in the model (1)is calculated as c_{ii}/c'_{ii}. c_{ii} is the diagonal of $X^\top X^{-1}$ matrix corresponding to particular single degree of freedom, and $\sigma^2 c_{ii}$ is the variance of $\hat{\tau}_i$ in a design where the treatment effects are orthogonal to blocks. c'_{ii} is the diagonal $X^\top Q X^{-1}$, and $\sigma'^2 c'_{ii}$ is the variance of $\hat{\tau}_i$ in the partially confounded design where Q is defined in Equation 2. In partially confounded designs $c'_{ii} > c_{ii}$, but σ'^2 should be much smaller than σ^2 due to the the fact that the experimental units are more homogeneous within the blocks of reduced size.

Das (1960) developed a general method for creating balanced-confounded designs by linking asymmetric designs to fractions of suitable symmetrical factorials. Using his method, the ratio of the total number of treatment combinations, n, to the block size, k, must be a prime or prime power. With the proper restrictions, the D_s optimality criteria can also be used to create a balanced confounded mixed-level design using Cook and Nachtsheim's algorithm. This method allows for additional choices for block sizes and number of replicates. For combinations of block size and number of replicates where no balanced confounded design exists, D_s optimality criteria can also be used to find a nearly balanced design where all main effects and interactions are estimable.

In this section we will illustrate the use of SAS to create a balanced partially confounded design using Das's method, and illustrate the use SAS `proc optex` to create partially confounded designs using Cook and Nachtsheim's algorithm.

3.1. Examples of creating balanced confounded designs

Using Das's method the ratio of the total number of treatment combinations to the block size, $r = n/k$, must be a prime number or a prime power, i.e., $r = p^m$, where m is some positive integer greater than or equal to 1. The factors in the design that have p levels are called real factors, and the factors that do not have p levels are called factors of asymmetry. The factors of asymmetry are represented by p-level pseudo factors. To illustrate Das's method, consider creating a partially confounded 3×2^2 factorial in blocks of size 4. One replicate of this design would consist of 3 blocks of size 4. The ratio of the total number of treatment combinations to the block size, $(3 \times 2^2)/4 = 3$, is a prime number. Thus, the three-level-factor A is a real factor and the two two-level factors B and C are factors of asymmetry. Three-level pseudo factors b and c are used to represent factors B and C, and the corresponding symmetrical factorial is a 3^3. One replicate of the design is constructed by confounding one interaction in the symmetrical factorial and then obtaining a fraction of the symmetrical design to represent the asymmetrical design. For example, confounding the Abc interaction splits the symmetric 3^3 design into three blocks of size 9. Next, a 4/9 fraction of the symmetric design is created by eliminating treatment combinations that have level 2 of factors b and c.

The defining relation for the 4/9 fraction of the symmetric design is

$$I = b = b^2 = c = c^2 = bc = b^2c = bc^2 = b^2c^2 \qquad (3)$$

which is the same as the defining relation for a 1/9 fraction. However, in the 1/9 fraction every effect that can be estimated is completely confounded with eight other non-estimable effects, while in the 4/9 fraction they are only partially confounded.

In the symmetric 3^3 design Abc interaction was confounded with blocks by choice, and if all factors actually had three levels this interaction would be completely confounded with blocks and lost. However, due the the 4/9 fraction Abc is only partially confounded with blocks and other effects become partially confounded by the fractionation as well. Multiplying through the defining relation we see that the following effects (Ab^2c, Ac, Abc^2, Ab, Ab^2c^2, A, Ac^2, and Ab^2) also become partially confounded by the fractionation. Four of the confounded interactions in terms of the pseudo factors all represent part of the ABC interaction in the asymmetric design. These are the Abc interaction that was originally chosen to confound with the blocks, and the interactions Ab^2c, Abc^2, and Ab^2c^2 that were confounded due to the fractionization. To create the balanced confounded design, using Das's method, each of these interactions, that represent different parts of the interaction originally chosen to be confounded, must be confounded in an additional replication of the design. Therefore three additional replicates must be added. In one additional replicate Ab^2c is confounded; in a second Abc^2 is confounded, and in the third Ab^2c^2 is confounded. This design can be easily created using `proc plan` and the SAS data step as shown in Table 4.

In the first block of code, `proc plan` was used to create the symmetric 3^3 factorial in factors A, and pseudo factors b and c. Next the data step is used to create four replicates of the design confounding Abc in the first replicate, Ab^2c in the second replicate, Abc^2 in the third replicate and Ab^2c^2 in the fourth replicate. Finally the last section of code combines the replicates and drops level 2 of factors b and c to create the 4/9 replicate.

Due to the fractionation four parts of the ABC interaction (ABc, Ab^2c, Abc^2, and Ab^2c^2) become partially confounded with blocks. In addition, two parts of the AC interaction (Ac and Ac^2), two parts of the AB interaction (Ab and Ab^2) and the main effect A all become partially confounded with blocks. On the other hand, main effects B, C and their interaction BC are not confounded with blocks. This information, along with the fact that four replicates each containing three blocks of four experimental units, are required to create a balanced confounded design using Das's method, leads to an alternative way of constructing an equivalent balanced confounded design.

Using the `prior=` option in `proc optex` with orthogonal coding, the user can control essentially how many prior observations worth of data he has for various terms in the model. When this option is used in addition to the `block` statement, `proc optex` uses the Cook-

```
proc plan;
   factors A=3 b=3 c=3;
   output out=fact A nvals=(0 1 2)
                   b nvals=(0 1 2)
                   c nvals=(0 1 2);

data rep1; set fact; rep=1; block=mod(A+b+c,3)+1;
data rep2; set fact; rep=2; block=mod(A+2*b+c,3)+1;
data rep3; set fact; rep=3; block=mod(A+b+2*c,3)+1;
data rep4; set fact; rep=4; block=mod(A+2*b+2*c,3)+1;

data bdesign; set rep1 rep2 rep3 rep4;
if b<2 and c<2;

proc sort data=bdesign; by rep block;
proc print data=bdesign; run;
```

Table 4: This code creates a balanced 3×2^2 in blocks of 4.

```
proc plan;
  factors A=3 B=2 C=2;
  output out=can A nvals=(0 1 2)
                 B nvals=(0 1)
                 C nvals=(0 1);

proc optex data=can coding=orth;
  class A B C;
  model B C B*C, A, A*C A*B, A*B*C/prior=0, 12, 24, 48;
  block structure=(12)4 init=chain noexchange niter=1000 keep=10;
  generate initdesign=can method=sequential;
  output out=bdesign;

proc print; run;
```

Table 5: This code creates a balanced confounded design using the `prior=` option. `proc plan` creates the full factorial and `proc optex` optimally blocks the design into 12 blocks of size 4.

Nachtsheim algorithm to maximize $|X^\top QX + P|$, where Q is given in Equation 2 and P is a diagonal matrix with the number of prior observations worth of data available for each effect specified on the diagonal. To create a design, equivalent to the design created with Das's method above, main effects B, C and their interaction BC should not be confounded with

	DBC[(12)4]	DSO[(12)4]
No. Reps:	4	4
$A(l)$	0.9375	0.9375
$A(q)$	0.9375	0.9375
B	1.0000	1.0000
C	1.0000	1.0000
BC	1.0000	1.0000
$A(l)B$	0.8125	0.8125
$A(q)B$	0.8125	0.8125
$A(l)C$	0.8125	0.8125
$A(q)C$	0.8125	0.8125
$A(l)BC$	0.4375	0.4375
$A(q)BC$	0.4375	0.4375

Table 6: The information recovered for 3×2^2 design in blocks of size 4.

blocks, therefore the user should specify that zero observations worth of prior information is available for these model terms, and `proc optex` will attempt to create a design that allows estimation of these terms with maximum precision. Main effect A is confounded once, in four reps of 12 experimental units, so the user should specify that he has 12 observations worth of prior information regarding this term which will alow it to be partially confounded in one replicate of the design. Interactions AC and AB are each confounded twice, so the user should specify that he has 24 observations worth of prior information regarding these two model terms. Finally the three way interaction ABC was confounded in all four reps so, the user should specify that he has 48 observations worth of prior information regarding this model term. Specifying these priors along with the the fact that 12 blocks of 4 experimental units per block are required (as indicated in the `block structure=(12)4;`) will allow `proc optex` to create a design equivalent to the one created by Das's method above (given sufficient iterations; generally 1000 seems to be enough). The code to do this is shown in Table 5.

Table 6 shows the information recovered for each single degree of freedom for the 3×2^2 design. It can be seen that the information recovered for each single degree of freedom in the design created by Das's method (labled DBC[(12)4] in the second column of the design to represent 12 blocks of 4) is the same as the information recovered for each single degree of freedom for the design created using the Cook-Nachtsheim algorithm (labeled DSO[(12)4] in the third column of the table.

This shows that the two designs are equivalent. It can also be seen that no information was lost for the two-level factors B, C and their interaction BC. Finally, it can be seen that that the information recovered for each degree of freedom for the three-level main effect A is the same (labeled $A(l)$ and $A(q)$ to represent the linear and quadratic contrasts), and the information recovered for each single degree of freedom of the AB interaction, and the ABC interactions are the same. This shows that the design has balanced confounding.

As another illustration of creating a balanced confounded design, consider blocking the 3×2^2 factorial in blocks of size 3. One replicate of this design would consist of 4 blocks of size 3. In this case the ratio of the total number of treatment combinations to the block size, $(3 \times 2^2)/3 = 4 = 2^2$, is a prime power. Thus, the three-level factor A is a factor of asymmetry and the two two-level factors B and C are real factors. Two two-level pseudo factors, a_1 and a_2 must be used to represent the three-level factor of asymmetry A, and the corresponding symmetric factorial will be a 2^4. One replicate of the design is constructed by confounding two interactions and their generalized interaction in the symmetric factorial and then obtaining a 3/4 fraction of the symmetrical factorial to represent the asymmetric design. In this case the 3/4 fraction is created by eliminating experiments that have the high level for both pseudo factors a_1 and a_2. The defining relation for the 3/4 fraction is

$$I = a_1 = a_2 = a_1a_2 \tag{4}$$

which is the same as the defining relation for a 1/4 fraction.

When choosing two interactions to confound with blocks in the symmetric factorial, an interaction involving only pseudo factors such as a_1a_2 should not be chosen because that will result in confounding the main effect A. Also, an interaction involving only real factors should not be chosen, to avoid completely confounding that interaction. When two interactions in the symmetric factorial such as a_1BC and a_2B are chosen to confound, it will also confound their generalized interaction $a_1BC \times a_2B = a_1a_2C$. To find out what additional interactions become partially confounded by the fractionation, multiply each of the interactions chosen to be confounded, and their generalized interactions, through the defining relation to obtain:

$$\begin{aligned} a_1BC &= BC = a_1a_2BC = a_2BC \\ a_2B &= a_1a_2B = B = a_1B \\ a_1a_2C &= a_2C = a_1C = C. \end{aligned}$$

By this, it can be seen that two additional parts of the ABC interaction (a_1a_2BC, and a_2BC), two additional parts of the AB interaction (a_2B, a_1a_2B, and a_1B), and two additional parts of the AC (a_1a_2C, a_2C and a_1C), that were chosen to be confounded, all become partially confounded due to the fractionation. In order to construct a balanced confounded design in this case three reps of four blocks must be constructed. To do this confound a_1BC and a_2B and their generalized interaction a_1a_2C in one rep, confound a_1a_2BC, a_1B and their generalized interaction a_2C in a second rep, and finally confound a_2BC, a_1a_2B and their generalized interaction a_1a_2C in a third rep. This can be easily done using `proc plan` and the SAS data step as shown in Table 7.

An equivalent design can also be created using the `prior=` option in `proc optex`. The main effect A for the factor of asymmetry was never confounded. Main effects B, C and their interaction BC were confounded only once, while the interactions AB, AC and ABC were confounded three times in three reps of 12 experimental units. Thus the code in Table 8 produces a balanced confounded design equivalent to the design produced by Das's Method.

Table 9 shows the information recovered for the design created by Das's method and the Cook-Nachtsheim algorithm. Again it can be seen that both methods produce equivalent balanced confounded designs.

Since use of Das's method requires manually determining the defining relationship for the asymmetric fraction of a symmetric factorial, determining the number of replicates required, and choosing the interactions to be confounded, it would be desirable to have a program or macro to do this automatically. The next section describes a macro to do this.

```
proc plan;
  factors a1=2 a2=2 B=2 C=2;
  output out=fact a1 nvals=(0 1)
                  a2 nvals=(0 1)
                  B nvals=(0 1)
                  C nvals=(0 1);

data rep1;
   set fact;
          rep=1;
          blk1=mod(a1+B+C,2);
          blk2=mod(a2+B,2);

data rep2;
   set fact;
          rep=2;
          blk1=mod(a1+a2+B+C,2);
          blk2=mod(a1+B,2);

data rep3;
   set fact;
          rep=3;
          blk1=mod(a2+B+C,2);
          blk2=mod(a1+a2+B,2);

data bdesign;
   set rep1 rep2 rep3;
          if a1=1 and a2=1 then delete;
          A=a1+2*a2;
          Block = blk1+2*blk2;
          keep rep Block A B C ;

proc sort data=bdesign; by rep Block;
proc print data=bdesign; var rep Block A B C; run;
```

Table 7: This code creates a 3×2^2 in blocks of 3. `proc plan` creates the full factorial. The next three data steps create the reps confounding a_1BC, a_2B and their generalized interaction in the first rep; a_1a_2BC, a_1B and their generalized interaction in the second rep; and a_2BC, a_1a_2B and their generalized interaction in the third rep. The final data step and creates levels for factor A from the pseudo factor levels and the block numbers from the block defining contrasts.

```
proc plan;
factors A=3 B=2 C=2;
output out=can A nvals=(0 1 2)
               B nvals=(0 1)
               C nvals=(0 1);

proc optex data=can coding=orth;
  class A B C;
  model A, B C B*C, A*B A*C A*B*C/prior=0, 12, 36;
  block structure=(12)3 init=chain noexchange niter=1000 keep=10;
  generate initdesign=can method=sequential;
  output out=besign;

proc print; run;
```

Table 8: This code creates a balanced confounded design using the `prior=` option. `proc plan` creates the full factorial and `proc optex` optimally blocks the design into 12 blocks of size 3.

	DBC[(12)3]	DSO[(12)3]
No. Reps:	4	4
$A(l)$	1.0000	1.0000
$A(q)$	1.0000	1.0000
B	0.8889	0.8889
C	0.8889	0.8889
BC	0.8889	0.8889
$A(l)B$	0.5556	0.5556
$A(q)B$	0.5556	0.5556
$A(l)C$	0.5556	0.5556
$A(q)C$	0.5556	0.5556
$A(l)BC$	0.5556	0.5556
$A(q)BC$	0.5556	0.5556

Table 9: The information recovered for 3×2^2 design in blocks of size 3.

3.2. SAS macro for creating Das's balanced confounded designs

Because Das's method of creating balanced confounded designs requires multiple replications, it is useful when the mixed-level factorial does not have too many treatment combinations. Otherwise, the total number of experiments would be excessive. The macro `%BalConf` produces balanced confounded designs using Das's method for asymmetric of mixed-level factorials with two or three factors and 32 or fewer treatment combinations. The macro is invoked by the statement `%BalConf`, and since the user must supply input such as the number and levels of the factors and the block size, the macro is interactive. When invoked the macro

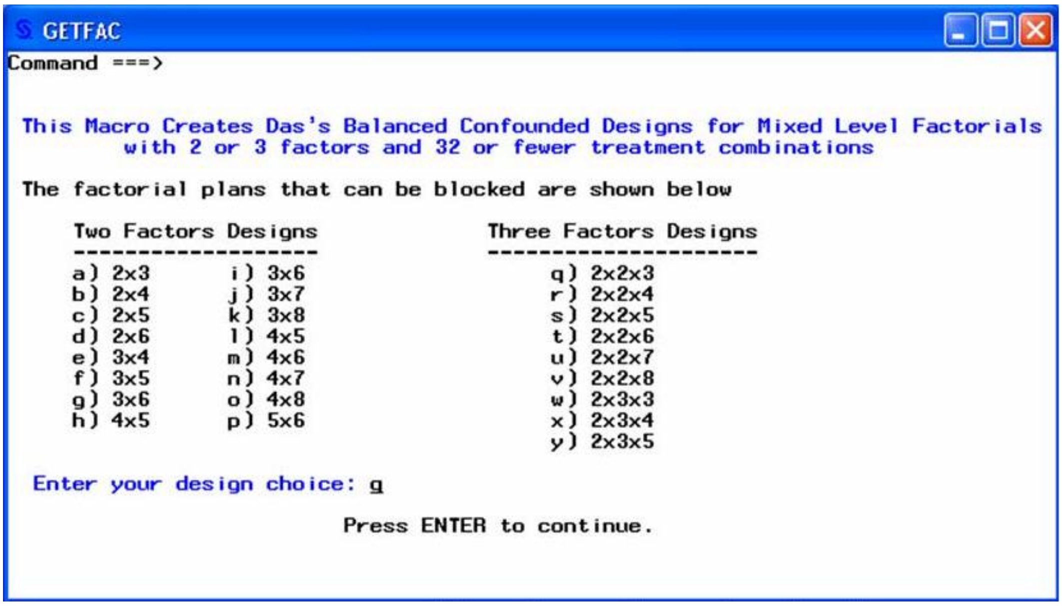

Figure 1: Design choice template for the `%BalConf` macro.

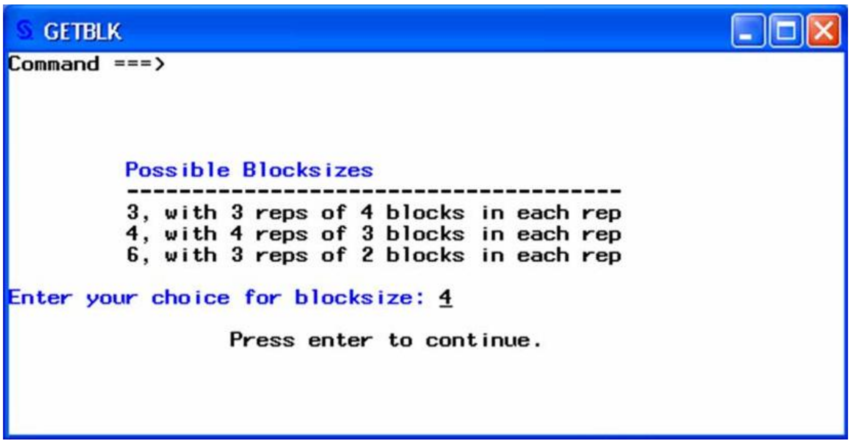

Figure 2: Block size choice template for the `%BalConf` macro.

displays the design choice template shown in Figure 1.

The user must type the letter of the design desired followed by the enter key, and the macro next displays the block size choice template shown in Figure 2. Only a limited number of block sizes are available since the total number of treatment combinations divided by the block size must be a prime number or prime power.

After choosing the block size desired followed by the enter key, the macro prints the design to the output window and stores it in the SAS file `work.bdesign`.

3.3. Comparison of partially confounded designs

A balanced incomplete block design with t treatments and blocks of size $k < t$ can always be found by choosing all possible subsets of size k of the t treatment levels. Similarly, a balanced confounded factorial design can usually be found using Das's method, but the block size and number of replicates required for balanced confounded designs created by Das's method are limited. In Section 3.1 it was shown that balanced confounded designs equivalent to the ones created by Das's method can also be created using optimal design theory. There it was shown that by specifying the number of observations worth of prior information available for various terms in the model that correspond to the terms confounded using Das's method, a balanced confounded design with the same amount of information recovered for each single degree of freedom in the model can be created using the optimal design approach.

Balanced incomplete block designs with fewer blocks than are required by the design composed of all subsets of k taken from t can usually be found for most combinations of t and k. Likewise, balanced confounded factorial designs can usually be found with block sizes that would not be possible with Das's method, and with fewer replicates than would be required by Das's method. In this section we show that these balanced confounded designs that are different than the ones created by Das's method can be created by using the optimal design approach. For example, to create a balanced-confounded design for a $2 \times 2 \times 5$ factorial in blocks of 4 using Das's method requires 16 replicates or 80 blocks. However, a balanced confounded design in 8 replicates can be obtained using optimal design theory. For combinations of block sizes and number of replicates where a balanced confounded design does not exist, a nearly balanced design that still allows estimation of all main effects and interactions in the model can also be constructed using the optimal design approach.

To illustrate the creation of alternate balanced confounded designs using the optimal design approach, consider some alternative blocked designs for the 3×2^2 factorial. By changing the model statement in Table 8 from `model A, B C B*C, A*B A*C A*B*C/prior=0, 12, 36;` to `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 36;`, a balanced confounded design that is different from the one shown in Table 9 is created. It is shown in the first column of Table 10. The information recovered for each degree of freedom from confounded terms are the same, showing the design has balanced confounding, but the information recovered for each degree of freedom is not the same as in Table 9. It can be seen that main effect A is not orthogonal to blocks in this design and that more information is recovered for two factor interactions at the cost of less information recovered for the three factor interaction when compared to the designs in Table 9.

If the priors are removed from the model statement in Table 8 so that it appears as `model A|B|C;`, a partially confounded design that is nearly balanced is created. This is shown in the second column of Table 10. Here it can be seen that all model terms, both main effects and interactions, have nearly the same information recovered.

If the model statement in Table 8 is changed to `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 36;` and the number of blocks is reduced to eight (i.e., two replicates of the design), `proc optex` is able to produce a partially confounded design where less information is recovered for the three way interaction than the main effects and two-way interactions, but the design does not have balanced confounding. The information recovered for each degree of freedom from this design is shown in the third column of Table 10. It would not be possible to create a partially confounded design with two replicates and blocks of size three using Das's method.

	DSO[(12)3]	DSO[(12)3]	DSO[(8)3]	DSO[(8)3]
No. Reps:	3	3	2	2
$A(l)$	0.8333	0.7222	0.7083	0.5833
$A(q)$	0.8333	0.6667	0.6250	0.7500
B	0.8889	0.7407	0.8889	0.6667
C	0.8889	0.5925	0.8889	0.7778
BC	0.8889	0.7407	0.8889	0.8889
$A(l)B$	0.7222	0.7778	0.8750	0.7500
$A(q)B$	0.7222	0.7592	0.9028	0.7500
$A(l)C$	0.7222	0.7778	0.8750	0.7500
$A(q)C$	0.7222	0.7963	0.9028	0.8056
$A(l)BC$	0.3889	0.6111	0.2083	0.5833
$A(q)BC$	0.3889	0.8148	0.2361	0.6944

Table 10: The information recovered for alternate 3×2^2 design in blocks of size 3. The first column was created using the statement `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 36;` with 12 blocks of 3; the second column was created without using priors by the statement `model A|B|C;` with 12 blocks of 3; the third column was created using the statement `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 36;` with 8 blocks of 3; and the fourth column was created without using priors by the statement `model A|B|C;` with 8 blocks of 3.

The final column in Table 10 shows the information recovered for a partially confounded design with two replicates that was created by removing the `prior=` option from the model statement. In this design the information recovered is more uniform across all model terms when compared to the design represented in the third column of Table 10.

Table 11 shows some confounded-block designs that can be created with `proc optex` having blocks of size 2. Designs with blocks of size 2 cannot be created using Das's method. The prior option for each design is shown in the caption for the table. It can be seen that with three replicates of the design, a balanced confounded design can be produced using the prior options shown for the first column in Table 11. The second column of the table shows a nearly balanced design, with the same number of replicates, that can be created by removing the prior option from the model statement. When the number of replicates is reduced to 2, `proc optex` cannot find a balanced confounded design, but by using the same prior options that were used for the design shown in the first column, a design that concentrates the loss of information in the three factor interaction can be created. This design is shown in the third column of Table 11. Finally a design with two replicates and more uniform information lost can be obtained by removing the prior option. This design is shown in the last column of Table 11.

Table 12 shows some partially confounded designs designs that can be created with blocks of size 4. The prior option for each design is shown in the caption for the table. With blocks of size 4, balanced confounded designs can be found for 2, 3, or 4 replicates using the optimal design approach. The information recovered for designs with 2 and 4 replicates are shown in the Table. The information recovered shown in the first column in the table are for the balanced confounded design created with Das's method in Section 3.1. A two replicate design with the identical information recovered for each degree of freedom is shown in the second

No. Reps:	DSO[(18)2] 3	DSO[(18)2] 3	DSO[(12)2] 2	DSO[(12)2] 2
$A(l)$	0.7500	0.6250	0.5000	0.6250
$A(q)$	0.7500	0.3750	0.7500	0.7500
B	0.6667	0.4444	0.6667	0.6667
C	0.6667	0.5556	0.5000	0.5000
BC	0.6667	0.5556	0.8333	0.5000
$A(l)B$	0.4167	0.6250	0.5000	0.5000
$A(q)B$	0.4167	0.4306	0.5833	0.4583
$A(l)C$	0.4167	0.6250	0.7500	0.5000
$A(q)C$	0.4167	0.6528	0.5000	0.3750
$A(l)BC$	0.4167	0.4583	0.2500	0.6250
$A(q)BC$	0.4167	0.6528	0.1667	0.5000

Table 11: The information recovered for 3×2^2 design in blocks of size 2. The first column was created using the statement `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 36;` with 18 blocks of 2; the second column was created without using priors by the statement `model A|B|C;` with 18 blocks of 2; the third column was created using the statement `model A|B|C@2, A|B|C@3/prior=0, 24;` with 12 blocks of 2; and the fourth column was created without using priors by the statement `model A|B|C;` with 12 blocks of 2.

No. Reps:	DBC[(12)4] 4	DSO[(6)4] 2	DSO[(6)4] 2	DSO[(6)4] 2
$A(l)$	0.9375	0.9375	0.7500	0.7500
$A(q)$	0.9375	0.9375	0.7500	0.9375
B	1.0000	1.0000	1.0000	0.9167
C	1.0000	1.0000	1.0000	0.9167
BC	1.0000	1.0000	1.0000	0.7500
$A(l)B$	0.8125	0.8125	1.0000	0.8750
$A(q)B$	0.8125	0.8125	1.0000	0.6458
$A(l)C$	0.8125	0.8125	1.0000	0.8750
$A(q)C$	0.8125	0.8125	1.0000	0.7708
$A(l)BC$	0.4375	0.4375	0.2500	0.7500
$A(q)BC$	0.4375	0.4375	0.2500	0.8125

Table 12: The information recovered for 3×2^2 design in blocks of size 4. The first column is Das's balanced confounded design. The second column was created with `proc optex` using the statement `model A|B|C@1, A|B|C@2, A|B|C@3/prior=0, 12, 24;` with 6 blocks of 4; the third column was created using the statement `model A|B|C@2, A|B|C@3/prior=0, 24;` with 6 blocks of 4; and the fourth column was created with no priors and 6 blocks of 4.

column of the table. It would not be possible to create a design with two replicates using Das's method. The third column in the table shows the information for another balanced confounded design that can be created in two replicates by changing the prior option. The

Total number of runs	Number of replicates	Block size	Method of construction
36	3	3	Das
48	4	4	Das
36	3	6	Das
48	4	2	Cook-Nachtsheim
36	3	3	Cook-Nachtsheim
48	4	4	Cook-Nachtsheim
30	$2\frac{1}{2}$	5	Cook-Nachtsheim
40	$3\frac{1}{2}$	5	Cook-Nachtsheim
60	5	5	Cook-Nachtsheim
36	3	6	Cook-Nachtsheim

Table 13: Comparison of block size and number of replicates for partially confounded designs for 3×2^2 factorials.

last column in the design shows the information recovered for a nearly balanced design in two replicates that was created by removing the prior option on the model statement.

In general, using Cook and Nachtsheim's algorithm in `proc optex` cannot always produce a balanced confounded design as Das's method will, but in some cases it can create a different balanced confounded than can produced by Das's method. It can also produce balanced confounded designs with block sizes not possible to obtain from Das's method and designs with fewer replicates than required by Das's method. For combinations of block size and replicates where no balanced confounded design can be found, Cook and Nachtsheim's algorithm may still be able to produce a design that is nearly balanced with all main effects and interactions estimable. Therefore, when the number of treatment combinations is so large that the number of replicates required to create a balanced confounded design using Das's method is excessive, or if the block size choices are too limited, the D_s optimality criteria and Cook and Nachtsheim's algorithm may be a better way to obtain a design. Table 13 illustrates the flexibility in choices of block size and number of replicates possible using the optimal design approach compared to Das's method, again using the 3×2^2 factorial.

4. Discussion

Although completely confounded asymmetric factorials can be easily created with SAS, this has not been illustrated in the standard SAS online documentation. We illustrate how the classical approach (taught in some experimental design classes) of confounding in separate symmetric sub-experiments can be accomplished using the `pointrep` option on the `output` statement in `proc factex`. We then compare the resulting designs to those that can be produced using the Cook-Nachtsheim algorithm that is available in SAS `proc optex`. Designs produced using the Cook-Nachtsheim algorithm, although not 100% D_s optimal, are more flexible in the sense that they allow more options for block size and less confounding of lower order interactions (at the cost of more confounding of some main effects).

For partially confounded asymmetric factorials, Das's method of creating balanced-confounded designs can be easily implemented using the SAS data step in combination with `proc plan`. Use of the macro in this paper automates the procedure. SAS `proc optex` can also be used to create balanced and nearly balanced confounded designs for asymmetric factorials using the Cook-Nachtsheim algorithm. Again, there is a wider choices of block size and number of replicates when creating designs with the Cook-Nachtsheim algorithm than with Das's method. However, since this method is not guaranteed to produce a balanced confounded design like Das's method, the information recovered must be calculated for each single degree of freedom to verify whether the design has balanced confounding or not.

By illustrating how these useful blocked designs can be easily created with SAS, we hope that they will be utilized more frequently by practitioners.

References

Addelman S (1962). "Orthogonal Main-Effect Plans for Asymptotic Factorial Experiments." *Technometrics*, **4**, 21–26.

Atkinson AC, Donev AN, Tobias RD (2007). *Optimum Experimental Designs with SAS*. Oxford University Press, Oxford.

Bose RC (1947). "Mathematical Theory of the Symmetrical Factorial Design." *Sankhya*, **8**, 107–166.

Collings BJ (1984). "Generating the Intrablock and Interblock Subgroups for Confounding in General Factorial Experiments." *The Annals of Statistics*, **12**, 1500–1509.

Cook RD, Nachtsheim CJ (1989). "Computer-Aided Blocking of Factorial and Response Surface Designs." *Technometrics*, **31**, 339–346.

Das MN (1960). "Fractional Replicates as Asymmetrical Factorial Designs." *Journal of the Indian Society of Agricultural Statistics*, **12**, 159–174.

Fisher RA (1942). "The Theory of Confounding in Fractional Experiments in Relation to the Theory of Groups." *Annals of Eugenics*, **11**, 341.

Franklin MF (1985). "Selecting Defining Contrasts and Confounded Effects in p^{n-m} Factorial Experiments." *Technometrics*, **27**, 165–172.

Kuhfeld WF, Tobias RD (2005). "Large Factorial Designs for Product Engineering and Marketing Research Applications." *Technometrics*, **47**, 132–141.

Minitab Inc (2008). *Minitab 15*. State College, PA. URL http://www.minitab.com/.

SAS Institute Inc (2008). *SAS* ***ADX*** *Interface for Design of Experiments, Version 9*. Cary, NC. URL http://www.sas.com/.

Stat-Ease, Inc (2008). ***Design-Ease****, Version 7.1*. Minneapolis, MN. URL http://www.statease.com/.

Taguchi G (1987). *System of Experimental Design, Volume 2*. Unipub Kraus International Publications, White Plains NY.

Voss DT (1986). "On Generalizations of the Classical Method of Confounding to Asymmetrical Factorial Experiments." *Communications in Statistics – Theory and Methods*, **15**, 1299–1314.

Wu CFJ, Hamada M (2000). *Experiments – Planning Analysis and Parameter Design Optimization*. John Wiley & Sons, NY.

Affiliation:

John Lawson
Department of Statistics
Brigham Young University
Provo, UT, United States of America
E-mail: lawson@byu.edu

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 32, Code Snippet 1 *Submitted:* 2008-06-07
November 2009 *Accepted:* 2009-06-22