NISS

A Class of Experimental Designs for Estimating a Response Surface and Variance Components

Bruce E. Ankenman, Hui Liu, Alan F. Karr,
and Jeffrey D. Picka

Technical Report Number 78
March, 1998 (Revised November, 2000)

National Institute of Statistical Sciences
19 T. W. Alexander Drive
PO Box 14006
Research Triangle Park, NC 27709-4006
www.niss.org

A Class of Experimental Designs for Estimating a Response Surface and Variance Components

November 22, 2000

Bruce E. Ankenman and Hui Liu
Dept. of Industrial Engineering
and Management Sciences
Northwestern University
Evanston, IL 60208

Alan F. Karr
National Institute of Statistical Sciences
Research Triangle Park, NC 27709

Jeffrey D. Picka
Department of Mathematics
University of Maryland
College Park, MD 20742

This article introduces a new class of experimental designs, called *split factorials*, which allow for the estimation of both response surface effects (fixed effects of crossed factors) and variance components arising from nested random effects. With an economical run size, split factorials provide flexibility in dividing the degrees of freedom among the different estimations. For a split factorial design, it is shown that the OLS estimators for the fixed effects are BLUE and the variance component estimators from the mean squared errors on the ANOVA table are also minimum variance among unbiased quadratic estimators. An application involving concrete mixing demonstrates the use of a split factorial experiment.

KEY WORDS: Staggered nested factorial, Mixed effects model, Nested factors, REML, Blocking schemes, Split factorial, Fractional factorials.

In many experimental settings, the measured response is affected not only by the fixed effects of crossed factors, but also by the random effects (usually nested) of sampling and measurement procedures. For example, in an experiment to study certain critical dimensions on a molded part, machine settings such as mold zone temperatures or screw speed could be the crossed factors of interest while shift-to-shift variation, part-to-part variation, and measurement-to-measurement variation might be the random effects of interest. The fixed effect estimates can be used to optimize the process, and knowing which variation source is largest could help to focus quality improvement efforts.

The fixed effects of crossed factors are often studied with 2^{k-p} experiments, where k is the number of crossed factors, p is the degree of fractionation, and 2^{k-p} is the number of design points. The variances of nested random effects are called variance components (see Searle, Casella, and McCulloch, 1992), and are typically estimated by means of hierarchical or nested designs (see Figure 1). If the *i*th nested random factor in a q-stage hierarchical design has the same number of levels, m_i, at each level of the (i-1)st factor, then the design is balanced. If $m_i = m$ for all i, then the design will have m^q observations. Figure 1 shows a balanced hierarchical design for two random factors: batches and samples nested within batches.

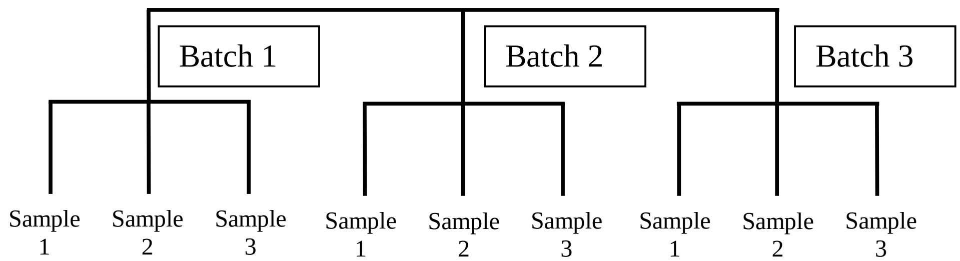

Figure 1. A Balanced Nested Design for m_1=3 Batches and m_2=3 Samples.

Both crossed factor effects and variance components could be estimated by performing an m^q nested design at each design point in a 2^{k-p} design. However, this would require $m^q \times 2^{k-p}$ observations, which often is not feasible or economical.

In this article, we construct a new class of experimental designs, called *split factorial designs*. A split factorial is a subset of an $m^q \times 2^{k-p}$ experiment that preserves the ability to estimate both the crossed factor effects (with a specified resolution) and the q variance components. Although other subsets could be used for these situations, the split factorial is chosen here because it is easy to design, run and analyze. These desirable properties result because the split factorial retains many of the characteristics of balanced designs including equal number of observations at each of the 2^{k-p} design points, the use of simple methods for parameter estimation, and an easily understood structure that can facilitate implementation of the experiment.

In the next section, a design methodology for split factorial experiments is introduced. Section 2 discusses analysis of split factorial designs and compares split factorials with existing designs for the few practical cases where they are comparable. In Section 3, an experiment involving concrete mixing, with three crossed factors and two variance components, is used to motivate and demonstrate the use of split factorial experiments. A discussion section concludes the article.

1. DESIGN METHODOLOGY

A methodology is now introduced for designing a split factorial experiment that has k crossed factors, each at two levels, and q=2^d (where d is an integer) variance components associated with nested random effects. The methodology takes a 2^{k-p} design with n observations at each of the 2^{k-p} design points. The design points are split into q sub-experiments by d blocking (splitting) generators. The experiment is then called a $2^{(k+d)-(d+p)}\times n$ split factorial. Each of the sub-experiments gathers information on only one of the q variance components. The design steps for a $2^{(k+d)-(d+p)}\times n$ split factorial are as follows:

1) Select n and p such that 2^{k-p} degrees of freedom (df.) are enough for estimating the fixed effects and $(n-1)2^{k-d-p}$ df. are sufficient for each variance component.
2) Choose a $2^{(k+d)-(d+p)}$ blocked factorial using blocking generators from a reference such as Bisgaard (1994); Sun, Wu, and Chen (1997); or Sitter, Chen, and Feder (1997). The q blocks (here called sub-experiments) will each have 2^{k-d-p} design points.

3) Let the variance components (1 to q) be such that the random effects of the (i+1)st variance component are nested under the effects of the ith variance component.

4) In the i^{th} sub-experiment (for i from 1 to q), a nested design that branches only at the i^{th} level (into n branches) will be run at each of the 2^{k-d-p} design points.

Example 1: This split factorial is too small for actual use, but is useful for demonstration of the design procedures. Let a 2^3 experiment (k=3, p=0) be split into 4 sub-experiments for 4 variance components (q=4, d=2) with n=3 observations at each design point. The blocking generators $\mathbf{B}_1$=AB and $\mathbf{B}_2$=AC can be used to split the experiment into sub-experiments, each with 2 design points (see Tables 1 & 2). Figure 2 shows the nesting structure at each design point. As described in design step 4, the nesting structures only branch at one level and the branching level is different for each sub-experiment. For example, the nesting structures at the two design points in sub-experiment 3 are circled in Figure 2 and branch only at the third level of nesting.

Table 1. Coding for Converting 2 Columns, $\mathbf{B}_1$ and $\mathbf{B}_2$, from a Two-Level Factorial into a Single Column Designating Sub-experiment or Block.

$\mathbf{B}_1$	$\mathbf{B}_2$		Sub-experiment, Level, or Block
-1	-1	→	1
1	-1	→	2
-1	1	→	3
1	1	→	4

Table 2. The $2^{(3+2)-(2+0)}\times 2$ Split Factorial using $\mathbf{B}_1$=AB and $\mathbf{B}_2$=AC for Splitting.

Design Point	A	B	C	$\mathbf{B}_1$=AB	$\mathbf{B}_2$=AC	Sub-exp.
1	-1	-1	-1	1	1	4
2	1	-1	-1	-1	-1	1
3	-1	1	-1	-1	1	3
4	1	1	-1	1	-1	2
5	-1	-1	1	1	-1	2
6	1	-1	1	-1	1	3
7	-1	1	1	-1	-1	1
8	1	1	1	1	1	4

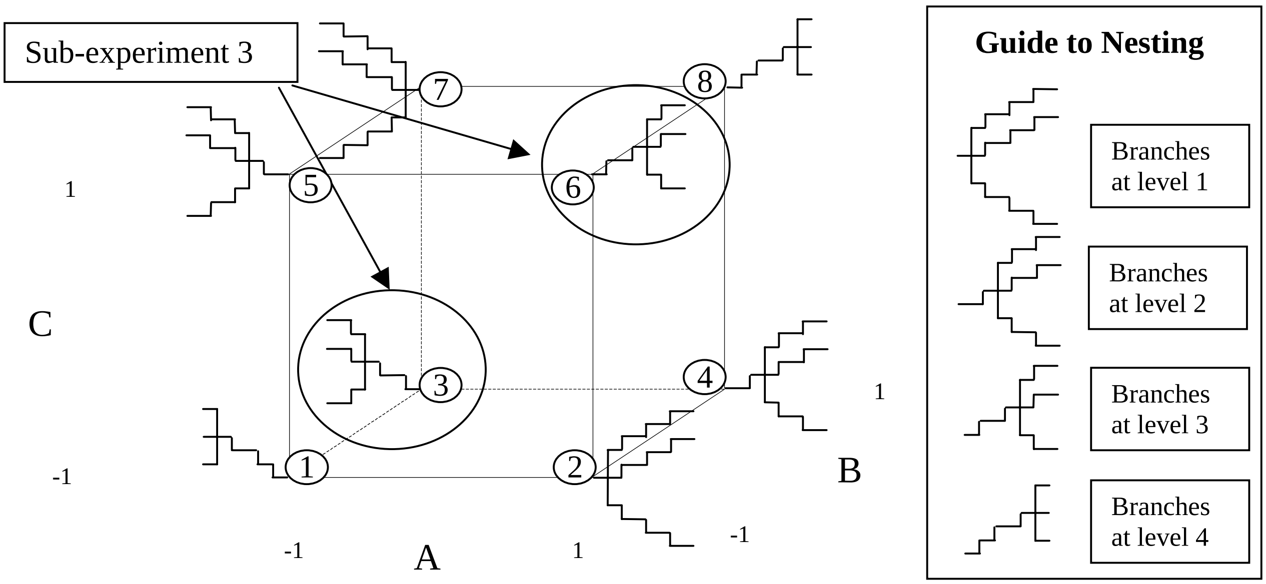

Figure 2. The Nesting Structure for k=3, p=0, q=4, n=3.

2. ANALYSIS OF SPLIT FACTORIAL DESIGNS

In this section, a linear mixed effects model is presented for the analysis of a split factorial. Using this model, estimation and tests of the variance components are discussed, including the difficulty of avoiding negative variance estimates. Correlation between the fixed effects estimators is discussed, and then the estimates of the variance components are used to perform approximate tests for the fixed effects. Finally, split factorials are compared with an alternative experimental design methodology.

Model and Variance Structure: A response model for a $2^{(k+d)-(d+p)}\times n$ split factorial is:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \sum_{i=1}^{q} \mathbf{Z}_i \mathbf{u}_i \ , \tag{1}$$

where $r = 2^{k-p}$, $\mathbf{X}$ is an $nr\times r$ matrix of estimable response surface contrasts including a constant column, and $\mathbf{b}$ is a vector of r unknown coefficient parameters. The matrix $\mathbf{Z}_i$ is an $\left(nr \times \frac{inr+(q-i)r}{q}\right)$ indicator matrix associated with the ith variance component, and $\mathbf{u}_i$ is a vector of length $\frac{inr+(q-i)r}{q}$ consisting of normally distributed independent random effect parameters associated with the ith variance component such that $\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{I}\sigma_i^2)$. Each random effect in $\mathbf{u}_i$ is nested under the treatment combinations and the (i-1) random effects above it. The usual random error term is $\mathbf{u}_q$. The quantity $\frac{inr+(q-i)r}{q}$ is derived by observing that there are $\frac{nr}{q}$ levels of the ith random factor in the first i sub-experiments and $\frac{r}{q}$ levels in the remaining (q-i) sub-experiments.

Assuming that the variance components do not depend on the crossed factors, then

$$\mathbf{V} = Var(\mathbf{y}) = \sum_{i=1}^{q} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i'. \qquad (2)$$

Given k factors, 2^{k-p}design points, q variance components, and n observations per design point, then expressions for $\mathbf{X}$ and $\mathbf{Z}_i$ can be derived for a split factorial design. Let $\mathbf{X}_1$ be the full rank $r \times r$ design matrix (including the constant) for a single replicate of the 2^{k-p} design, then $\mathbf{X}_1\mathbf{X}_1' = \mathbf{X}_1'\mathbf{X}_1 = r\mathbf{I}_r$, where $\mathbf{I}_r$ is the $r\times r$ identity matrix. The observations are ordered such that $\mathbf{X} = \mathbf{X}_1 \otimes \mathbf{1}_n$, where $\mathbf{1}_n$ is an n-length vector of ones and $\otimes$ represents the Kronecker product. Let $\mathbf{x}_{s,t}'$ be the row in $\mathbf{X}_1$, which is the tth observation in the sth sub-experiment. Now sort the rows of $\mathbf{X}_1$ in ascending order first by t and then by s such that

$$\mathbf{X}_1 = \begin{bmatrix} \mathbf{x}'_{1,1} \\ \mathbf{x}'_{2,1} \\ \vdots \\ \mathbf{x}'_{q,1} \\ \mathbf{x}'_{1,2} \\ \mathbf{x}'_{2,2} \\ \vdots \\ \mathbf{x}'_{q,r/q} \end{bmatrix} \text{ and thus } \mathbf{X} = \begin{bmatrix} \mathbf{x}'_{1,1} \otimes \mathbf{1}_n \\ \mathbf{x}'_{2,1} \otimes \mathbf{1}_n \\ \vdots \\ \mathbf{x}'_{q,1} \otimes \mathbf{1}_n \\ \mathbf{x}'_{1,2} \otimes \mathbf{1}_n \\ \mathbf{x}'_{2,2} \otimes \mathbf{1}_n \\ \vdots \\ \mathbf{x}'_{q,r/q} \otimes \mathbf{1}_n \end{bmatrix}. \tag{3}$$

The Kronecker sum is defined such that $\mathbf{P} \oplus \mathbf{Q} = \begin{bmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix}$ for any matrices, **P** and **Q**. If this notation is extended to a Kronecker summation, then the ordering in (3) leads to

$$\mathbf{Z}_i = \mathbf{I}_{r/q} \otimes \left(\left(\bigoplus_{s=1}^{i} \mathbf{I}_n \right) \oplus \left(\bigoplus_{s=i+1}^{q} \mathbf{1}_n \right) \right). \tag{4}$$

Estimation and Testing of Variance Components: The variance component estimators can be derived by the method of moments from the expected mean squares in Table 3. In the table, the sum of squares for the *i*th level of nesting is given by

$$SS_1 = \sum_{g=1}^{r} \sum_{m} (\bar{y}_{gm**\cdots*} - \bar{y}_{g**\cdots*})^2 \text{ , and}$$

$$SS_i = \sum_{g=1}^{r} \cdots \sum_{l} \sum_{m} \cdots \sum_{h} (\bar{y}_{g\cdots lm**\cdots*} - \bar{y}_{g\cdots l**\cdots*})^2 \text{ , for } i=2 \ldots q,$$

where g is the subscript related to the design points (treatment combinations) of the crossed design and l, m and h are the subscripts related to the $(i\text{-}1)^{\text{st}}$, ith, and qth level of nesting, respectively. The star subscript indicates averaging over that level of nesting.

Due to the simplicity of the expected mean squares for a split factorial, the method of moments estimator for σ_i^2 is $\hat{\sigma}_i^2 = \text{MS}_i - \text{MS}_{i+1}$ for i = 1 to q-1. Under normality, these ANOVA estimators are not only unbiased, but also are the Uniformly Minimum Variance Unbiased translation-Invariant Quadratic (UMVUIQ) estimators (see Appendix 2).

Table 3. ANOVA table for a Split Factorial.

Source	df	SS	MS	Expected MS	F ratio
Fixed Effects (not corrected for the mean)	2^{k-p}	SS_{FE}	MS_{FE}	$\mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b}/2^{k-p} + \sum_{i=1}^{q}\left(n - \frac{i(n-1)}{q}\right)\sigma_i^2$	
Variance Component 1	$(n-1)2^{k-d-p}$	SS_1	MS_1	$\sum_{j=0}^{q-1}\sigma_{q-j}^2$	$F_1 = MS_1/MS_2$
⋮	⋮	⋮	⋮	⋮	
Variance Component *i*	$(n-1)2^{k-d-p}$	SS_i	MS_i	$\sum_{j=0}^{q-i}\sigma_{q-j}^2$	$F_i = MS_i/MS_{i+1}$
⋮	⋮	⋮	⋮	⋮	
Error (Var. Comp. *q*)	$(n-1)2^{k-d-p}$	SS_q	MS_q	σ_q^2	
Total	$n2^{k-p}$	SS_T			

Tests for the variance components are also simple under normality. It can be shown that all terms in SS_i are zero except those deriving from observations in sub-experiment *i*. Since each sub-experiment is balanced when treated alone, each of the sums of squares for variance components in Table 3 when divided by its expected mean square has a chi-squared distribution with $(n-1)2^{k-d-p}$ degrees of freedom. Thus, standard F-tests as shown in Table 3 can be used to test if any variance component is zero.

Unfortunately, ANOVA variance estimates can be negative. Searle, Casella, and McCulloch (1992, p.130) discuss various methods of coping with this possibility. A common strategy is to assume that those variance components with negative estimates are zero or at least negligible. Alternatively, maximum likelihood methods like those implemented in many software packages always produce non-negative estimates. Negative estimates will tend to occur unless the variance components with lower subscripts are substantially larger than those with higher subscripts. In split factorials, increasing either n or r will reduce the problem of negative estimates. Under normality, Searle, Casella, and McCulloch (1992, p.137) provide an expression which, when applied to the split factorial ANOVA table in Table 3, shows that

$$\Pr\left\{\hat{\sigma}_i^2 < 0\right\} = \Pr\left\{F_{f,f} < \frac{\sum_{j=0}^{q-(i+1)} \sigma_{q-j}^2}{\sum_{j=0}^{q-i} \sigma_{q-j}^2}\right\},$$

where $f = (n-1)2^{k-d-p}$ and $F_{f,f}$ has an F-distribution with f and f degrees of freedom. Clearly, one needs some knowledge of the relative size of the variance components in order to determine the probability of negative variance estimates.

When all the ANOVA estimates are positive and normality is assumed, they are equivalent to restricted maximum likelihood (REML) estimates of the variance components. This is due to the fact that each sub-experiment, treated alone, contains balanced data (see Anderson, et al., 1984). REML estimators are consistent and have an approximately normal distribution in large samples (Searle, Casella, McCulloch, 1992). This equivalence provides a closed-form expression for the REML estimates.

Estimation, Correlation, and Tests of Fixed Effects: The condition under which the OLS estimators of the parameters in **b** are the best linear unbiased estimators (BLUE) is that an invertible matrix **A** exists such that $\mathbf{V}^{-1}\mathbf{X}=\mathbf{X}\mathbf{A}$ (Seber, 1977, p. 63). Appendix 1 shows that this condition is satisfied for data from a split factorial. Thus, estimating the coefficients of the response surface can be done by simple OLS regression techniques

$$\hat{\mathbf{b}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{nr}\mathbf{X}'\mathbf{y}\,. \tag{5}$$

In addition to the usual confounding caused by the fractionation in the 2^{k-p} design, there will also be non-zero covariance between certain fixed effect estimators due to the covariance between certain observations in the split factorial. The effect estimators that are correlated can be determined by creating a new defining relation for the experiment, called a *correlation relation*, that uses all the generators including the d generators that are used to split the factorial into sub-experiments. However, the splitting generators use a different operator, "~" which means "correlated with". Suppose, for example, the defining relation of a factorial design is I=ABCF and the splitting generators are: $\mathbf{B}_1$~ABE; $\mathbf{B}_2$~BCDE. Since there are no expected block effects, these effects are eliminated. The words ABE and BCDE are then used to extend the defining relation (see Box, Hunter and Hunter, 1978, page 409) to a correlation relation as follows:

I=ABCF~ABE~CEF~ BCDE~ADEF~ACD~BDF.

Multiplying any effect by this correlation relation shows the confounding and correlation pattern. Concepts similar to resolution and aberration (see Fries and Hunter, 1980) can now be used to select splitting generators for split factorials.

The sign and amount of correlation will be evident in the variance-covariance matrix for the coefficient estimators, which is

$$\mathbf{H} = Var\left(\hat{\mathbf{b}}\right) = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}. \tag{6}$$

Commonly, the estimates of the variance components derived by the method of moments above are used in (2) to obtain a $\hat{\mathbf{V}}$ that can be substituted for $\mathbf{V}$ in (6). Let b_t be the tth element of $\mathbf{b}$. Under the null hypothesis that $\mathbf{b} = \mathbf{0}$, the expected mean square associated with b_t can be shown to be $\sum_{i=1}^{q}\left(n - \frac{i(n-1)}{q}\right)\sigma_i^2$. However, if any of the effects in the correlation string for b_t are non-null, they will bias the mean square. Assuming all the effects in the correlation string are null, an approximate F-test that b_t=0 can be performed using the test statistic $F_t = \hat{b}_t^2 / \hat{h}_{tt}$, where $\hat{h}_{tt}$ is the tth diagonal element of $\hat{\mathbf{H}} = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}$. Equivalently,

$$F_t = MS_{FEt} \Big/ \sum_{i=1}^{q}\left(n - \frac{i(n-1)}{q}\right)MS_i \ , \tag{7}$$

where MS_{FEt} is the mean square due only to the tth factor. It can be shown that the denominator in (7) is equal to $\mathbf{a}'\mathbf{m}$, where the ith element of $\mathbf{m}$ and $\mathbf{a}$ are $m_i = \mathrm{MS}_i$ and

$$a_i = \begin{cases} n - {}^{(n-1)}\!/_{q} & \text{for } i = 1 \\ \quad - {}^{(n-1)}\!/_{q} & \text{else} \end{cases}$$

respectively. Satterthwaite's approximation now can be used to determine appropriate degrees of freedom for this approximate F-test. The numerator has one degree of freedom and the denominator degrees of freedom are approximated by

$$df_{\text{denominator}} = \frac{(n-1)2^{k-d-p}\left(\mathbf{a}'\mathbf{m}\right)^2}{\sum_{i=1}^{q}\left(a_i m_i\right)^2}.$$

Comparison with existing design methodology: The only class of designs in the literature for this type of experimentation is the staggered nested factorial proposed by Smith and Beverly (1981). Staggered nested designs were first introduced by Bainbridge (1965) and are unbalanced hierarchical nested designs with a single branch at each level of nesting. These designs split the degrees of freedom equally among the variance components. The staggered nested factorial places a staggered nested design at each point of a crossed factor design. Staggered nested factorials exist only if n, the number of observations at each design point, is q+1, where q is the number of variance components.

For two designs to be comparable, they must have the same crossed design and equal degrees for each of the variance components. Thus, for any split factorial with q=2 and n=3, there is a comparable staggered nested factorial. Comparable designs for a 2^3 crossed design are shown in Figure 3. Table 4 shows that the staggered nested design produces lower variance estimators than the split factorial. Unlike the split factorial, the staggered nested factorial is orthogonal if the crossed factor design is orthogonal. However, due to the imbalance of the staggered nested design, the OLS estimators for the fixed effects will not be BLUE, as they are for the split factorial, and there is no guarantee that the ANOVA estimators of the variance components are UMVUIQ.

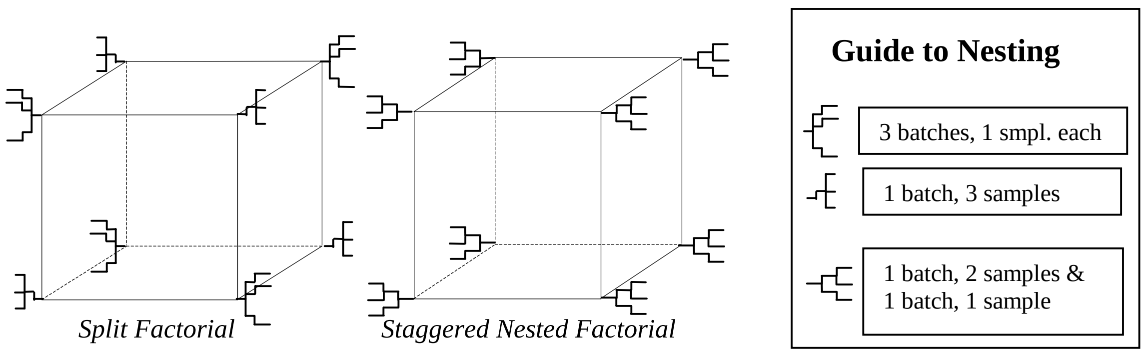

Figure 3. Comparable designs for k=3, p=0, q=2, n=3.

For Table 4, the variances of the variance component estimators can be found from the formulas provided in Searle, Casella and McCulloch, (1992, Appendix F.1, part **c**). Since the variance components are nested under the treatment combinations, we can ignore the fixed effects for this calculation. Using their model and notation,

$$y_{ij} = \mu + \alpha_i + e_{ij},$$
$$i = 1,2,\ldots,a \text{ and } j = 1,2,\ldots,n_i,$$
$$\sum n_i = N, \quad \sum n_i^2 = S_2 \quad \sum n_i^3 = S_3,$$

where a is the total number of batches in the experiment. Both the staggered nested factorial and the split factorial will have a=2r, and N=3r, where r is the number of design points in the crossed factor design. For the split factorial, $n_i = 1$ for 3*r*/2 of the batches and $n_i = 3$ for the remaining r/2 batches, thus $S_2 = 6r$ and $S_3 = 15r$. For the staggered nested factorial, $n_i = 1$ for *r* of the batches and $n_i = 2$ for the remaining r/2 batches, thus $S_2 = 5r$ and $S_3 = 9r$. Let $\tau = \sigma_1^2 / \sigma_2^2$. Substituting these values into the formulas provided by Searle, Casella and McCulloch (1992) and simplifying gives:

$$Var_{Split}(\hat{\sigma}_1^2) = \frac{2(1-5r+6r^2-4r\tau+6r^2\tau-6r\tau^2+6r^2\tau^2)}{r(3r-2)^2},$$

$$Var_{Stag}(\hat{\sigma}_1^2) = \frac{2(9-45r+54r^2-30r\tau+54r^2\tau-29r\tau^2+45r^2\tau^2)}{r(9r-5)^2},$$

and the difference, $D_{v1} = Var_{Split}(\hat{\sigma}_1^2) - Var_{Stag}(\hat{\sigma}_1^2)$, is

$$D_{v1} = \frac{2(-11+73r-156r^2+108r^3+20r\tau-66r^2\tau+54r^3\tau-34r\tau^2+162r^2\tau^2-225r^3\tau^2+81r^4\tau^2)}{r(3r-2)^2(9r-5)^2}.$$

For r>2, it can be shown that D_{v1} is a parabola in τ with a minimum point that is greater than zero. Since D_{v1} is positive for all σ_1^2 and σ_2^2, then any staggered nested factorial with *q*=2 and *n*=3 will have a smaller variance for $\hat{\sigma}_1^2$, than the competing split factorial. Similar results were found by simulation for the case of *q*=4 and *n*=5.

Table 4. Comparison of Split Factorial and Staggered Nested Factorial.

	Split Factorial*	Stag. Nested Factorial*	Difference (Split-Stag)
$Var(\hat{b}_s)\quad \forall s$ (*q*=2, *n*=3)	$\frac{2\sigma_1^2+\sigma_2^2}{3r}$	$\frac{2\sigma_1^2+2\sigma_1^2\sigma_2^2+\sigma_2^2}{r(4\sigma_1^2+3\sigma_2^2)}$	$\frac{\sigma_1^2(2\sigma_1^2+\sigma_2^2)}{3r(4\sigma_1^2+3\sigma_2^2)} > 0 \quad \forall \sigma_1^2, \sigma_2^2$
$Var(\hat{\sigma}_1^2)$ (*q*=2, *n*=3)	$Var_{Split}(\hat{\sigma}_1^2)$	$Var_{Stag}(\hat{\sigma}_1^2)$	$D_{v1} = Var_{Split}(\hat{\sigma}_1^2) - Var_{Stag}(\hat{\sigma}_1^2) > 0 \quad \forall \sigma_1^2, \sigma_2^2$
$Var(\hat{\sigma}_2^2)$ (*q*=2, *n*=3)	$\frac{2\sigma_2^4}{r}$	$\frac{2\sigma_2^4}{r}$	$0, \quad \forall \sigma_1^2, \sigma_2^2$

**r is the number of design points in the crossed factor design.*

For these cases, and for the case of three variance components (q=3 and n=4) where there is no comparable split factorial, the staggered nested factorials are preferable designs unless a simple analysis is very important. There are, however, many other cases where $n \neq q+1$ and staggered nested designs are not available. In particular, for the practical case of q=2 and n=2, there is no comparable design for the split factorial. The existence and optimality of designs when $n \neq q+1$ and n>2 is left for future research.

3. CONCRETE PERMEABILITY APPLICATION

The split factorial designs were motivated by an experiment run at the NSF Center for Advanced Cement-Based Materials at Northwestern University as a part of a joint project with the National Institute of Statistical Sciences. For exposition purposes, this application has been simplified. The full data set is in Appendix 3.5 of Jaiswal (1998).

The response is the electric charge (in Coulombs) passing through a sample in the rapid chloride permeability test (RCPT), see ASTM (1991). Lower charge implies lower permeability of concrete to chloride ions and thus better performance.

Concrete is made by combining water, cement, and aggregate (rocks, sand) of various sizes. The experiment included 2 levels of water-to-cement ratio (W/C), 4 aggregate grades, and 2 maximum aggregate sizes. The goal of the experiment was to relate these variables to the chloride permeability and to estimate the *batch-to-batch* and *sample-to-sample* variance components. Since the RCPT is destructive, no repeated measurements can be made, and thus measurement error is confounded with the sample-to-sample variance component. For simplicity, we take measurement error to be negligible.

In this application, there were two primary reasons for estimating the variance components: (1) to understand the variation of permeability in concrete structures where multiple batches of concrete are poured together and (2) to gain intuition on whether the mixing or casting process might produce larger variation.

A design with 4 observations at each of the 16 design points is presented in Table 5 and shown graphically in Figure 4. Table 1 is again used to convert two columns from the 2^4 full factorial, A and B, into a single column, X, for the four-level factor. Many authors have described this procedure, including Ankenman (1999) and Montgomery (1997, p.364). Although this full design was desirable, resource constraints required a design with only 32 observations. If the design were reduced by simply eliminating one half of the recipes from the full design, many interaction terms would not be estimable.

The split factorial, shown in Table 6 and Figure 5, also reduces the design to 32 runs. The design has k=4 two-level factors, two of which are converted to a four-level factor. It has n=2 observations at each design point, and there are two sub-experiments so d=1 and q=2. Since it is a full factorial, p=0, thus there is no defining relation and $r=2^{k-p}=16$. The correlation relation is I~ACD. For the split factorial, all response surface effects can be estimated, though some of these estimators are correlated. There are 8 degrees of freedom for estimating each variance component.

Figure 6 shows the measured charge (in Coulombs) for the observations in the split factorial. Lower aggregate grade levels and larger maximum aggregate size reduce the charge, suggesting that including larger aggregate improves performance.

The variance components can be estimated from the mean square in Table 7 as $\hat{\sigma}_1^2 = 272{,}891 - 135{,}560 = 137{,}331$ and $\hat{\sigma}_2^2 = 135{,}560$. In this case, the variance components are roughly the same size. The F-test in Table 7 has a p-value of 0.17

suggesting the batch-to-batch variance may be zero. However, this probably just means that the batch-to-batch variance is not much larger than the sample-to-sample variance.

Table 8 shows the approximate tests for the fixed effects. As in Section 2, Satterthwaite's method was used. For this example, $\mathbf{a}' = (3/2 \quad -1/2)$ and $\mathbf{m}' = (272891.18 \quad 135559.75)$, and thus the denominator for the F-test in (7) is $\mathbf{a}'\mathbf{m} = 341557.18$ and $df_{\text{denominator}} = 5.42$. The tests suggest that Factor D, the max. aggregate size, and factor X, the aggregate grade, have significant effects, confirming the observations from the cube plot in Figure 6. Both the quadratic and cubic terms for aggregate grade were found to be insignificant, thus only the linear contrast (X in Table 6) and max. aggregate size (D in Table 6) are included in the response surface model,

$$\text{Permeability} = 3987 + 654\ X - 826\ D.$$

This model allows for predictions of the permeability in the experimental region. More description of the results can be found in Jaiswal, S. S., Picka, J. D., et al. (2000).

Table 7. The ANOVA Table for the Concrete Permeability Example.

Source	DF	Type I SS	Type I MS	EMS	F	p
X	3	17387343.84	5795781.28			
C	1	94721.28	94721.28			
D	1	21801455.28	21801455.28			
X*C	3	2957371.09	985790.36			
X*D	3	4013875.59	1337958.53			
C*D	1	1168538.28	1168538.28			
X*C*D	3	1102912.09	367637.36			
BATCH(X*C*D)	8	2183129.50	272891.18	$\sigma_1^2 + \sigma_2^2$	2.01	0.17
Sample(BATCH)	8	1084478.00	135559.75	σ_2^2		
Corrected Total	31	51793824.96				

Table 8. Tests for Fixed Effects.

Source	NDF	DDF	Type III F	Pr > F
X	3	5.42	16.97	0.0036
C	1	5.42	0.28	0.6193
D	1	5.42	63.83	0.0003
X*C	3	5.42	2.89	0.1339
X*D	3	5.42	3.92	0.0809
C*D	1	5.42	3.42	0.1191
X*C*D	3	5.42	1.08	0.4332

4. DISCUSSION AND EXTENSIONS

Split factorial designs have attractive characteristics for estimating both response surface effects and variance components. 1) The experimenter can divide the degrees of freedom between the response surface effects and the variance components. 2) Each variance component is estimated with equal degrees of freedom. 3) The ANOVA and OLS estimates are often adequate, resulting in simple analysis. 4) The symmetry of the split factorial facilitates the implementation and analysis of the experiment.

To illustrate 4) above, note that due to the symmetry of the split factorial, the estimate of the sample-to-sample variance is just the pooled variance of the pairs of observations in the shaded circles in Figure 6. Similarly the pooled variance from the unshaded circles is the estimate of the sum of the two variance components. Also, although missing observations change the correlation structure of the fixed effect estimators, they do not affect the property that the OLS estimators for the fixed effects are BLUE or that the ANOVA estimators are UMVUIQ, since they only change the size of the identity matrices and length of the vectors of ones in equations (3) and (4). However, adding observations, such as an additional sample to any batch in sub-experiment 1 on Table 6, can destroy these properties. With such an addition, generalized least squares and REML estimates would be needed for the fixed effects and variance components, respectively.

More flexibility can be introduced into the split factorial designs by allowing each sub-experiment to have a different number, n_i, of observations at each design point. This would result in the sum of squares for the ith variance component having $(n_i - 1)2^{k-d-p}$ degrees of freedom and the total number of observations being $\sum_{i=1}^{q} n_i 2^{k-d-p}$.

Table 5. The Full 64-Run Design for the Concrete Permeability Experiment.

	X	A	B	C	D	Batch 1		Batch 2	
Recipe #	Grade	Code 1	Code 2	W/C Ratio	Max. Ag.	Cyl. 1	Cyl. 2	Cyl. 1	Cyl. 2
1	1	-1	-1	-1	-1	×	×	×	×
2	2	1	-1	-1	-1	×	×	×	×
3	3	-1	1	-1	-1	×	×	×	×
4	4	1	1	-1	-1	×	×	×	×
5	1	-1	-1	1	-1	×	×	×	×
6	2	1	-1	1	-1	×	×	×	×
7	3	-1	1	1	-1	×	×	×	×
8	4	1	1	1	-1	×	×	×	×
9	1	-1	-1	-1	1	×	×	×	×
10	2	1	-1	-1	1	×	×	×	×
11	3	-1	1	-1	1	×	×	×	×
12	4	1	1	-1	1	×	×	×	×
13	1	-1	-1	1	1	×	×	×	×
14	2	1	-1	1	1	×	×	×	×
15	3	-1	1	1	1	×	×	×	×
16	4	1	1	1	1	×	×	×	×

× Indicates an Observation

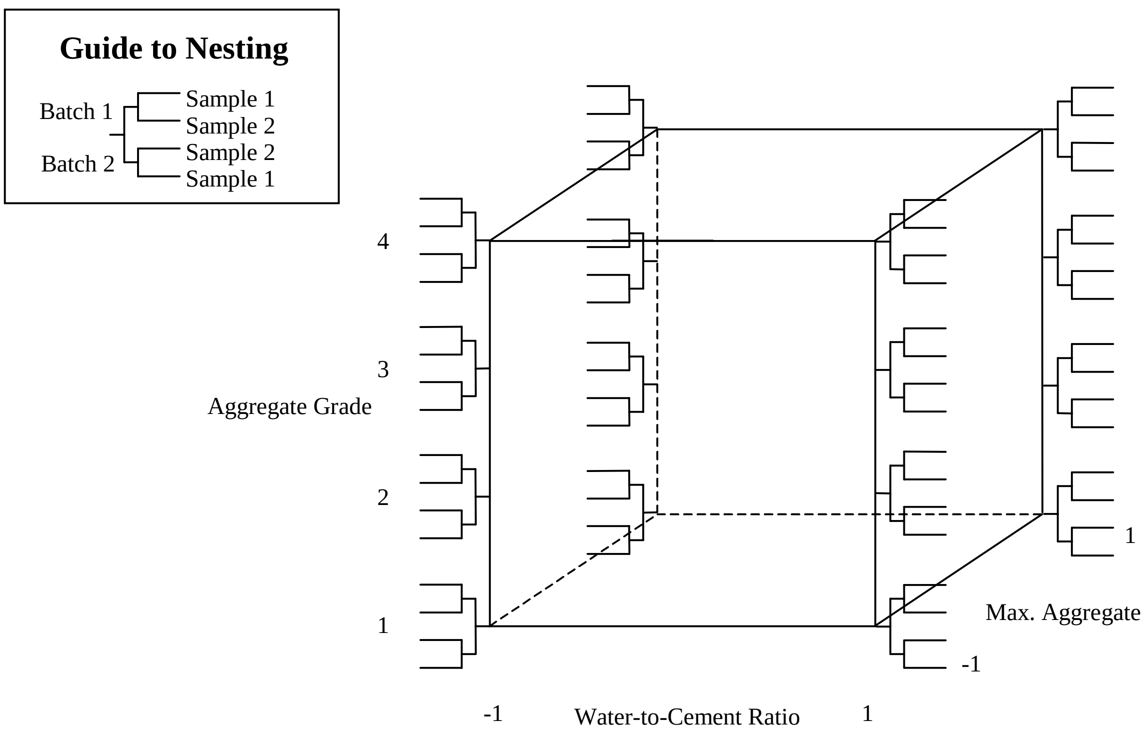

Figure 4. The Full Design.

Table 6. The Split Factorial Design for the Concrete Permeability Experiment.

	X	A	B	C	D	**B**$_1$~ ACD		Batch 1		Batch 2	
Recipe #	Grade	Code 1	Code 2	W/C Ratio	Max. Ag.		Sub-Exp.	Cyl. 1	Cyl. 2	Cyl. 1	Cyl. 2
1	1	-1	-1	-1	-1	-1	1	×		×	
10	2	1	-1	-1	1	-1	1	×		×	
3	3	-1	1	-1	-1	-1	1	×		×	
12	4	1	1	-1	1	-1	1	×		×	
13	1	-1	-1	1	1	-1	1	×		×	
6	2	1	-1	1	-1	-1	1	×		×	
15	3	-1	1	1	1	-1	1	×		×	
8	4	1	1	1	-1	-1	1	×		×	
9	1	-1	-1	-1	1	1	2	×	×		
2	2	1	-1	-1	-1	1	2	×	×		
11	3	-1	1	-1	1	1	2	×	×		
4	4	1	1	-1	-1	1	2	×	×		
5	1	-1	-1	1	-1	1	2	×	×		
14	2	1	-1	1	1	1	2	×	×		
7	3	-1	1	1	-1	1	2	×	×		
16	4	1	1	1	1	1	2	×	×		

× Indicates an Observation

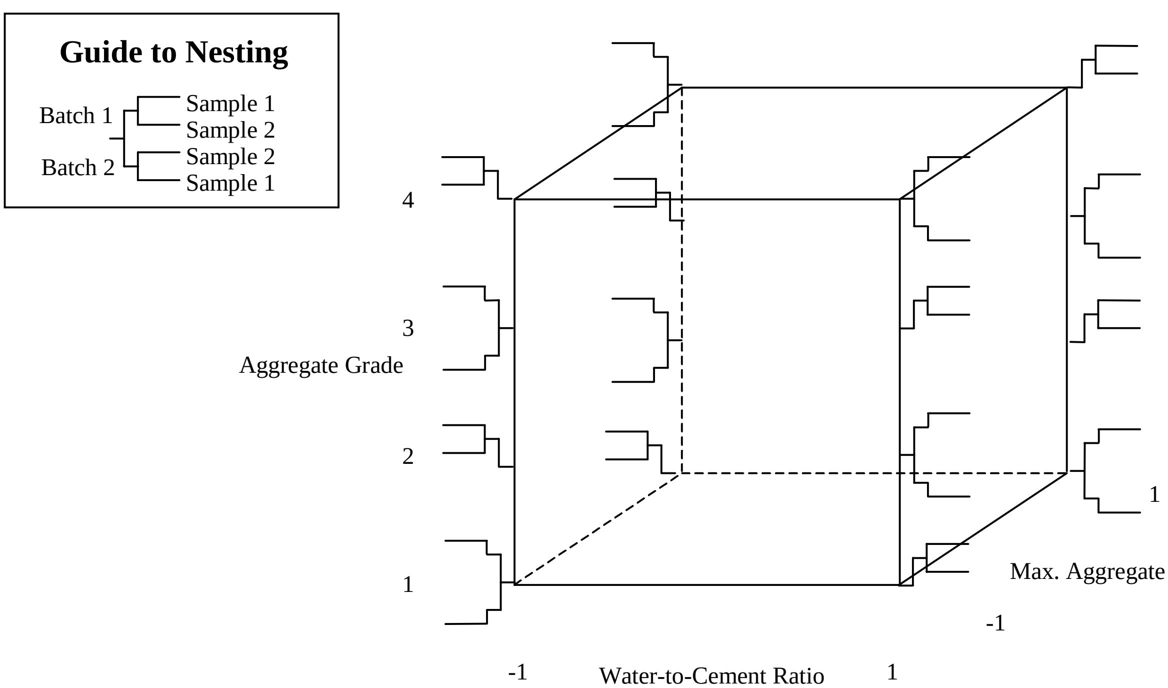

Figure 5. Split Factorial Design for the Concrete Permeability Experiment.

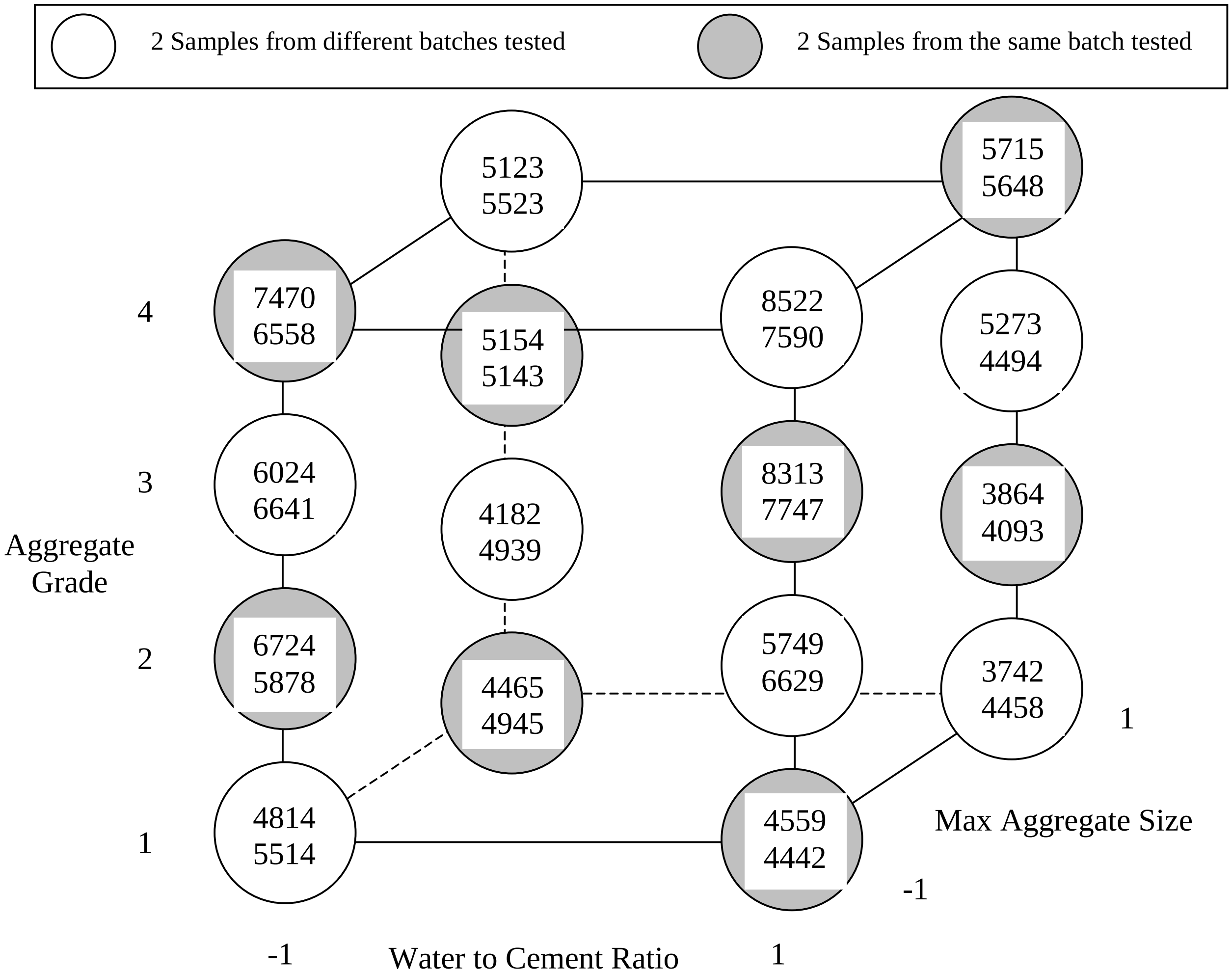

Figure 6. A Cube Plot of the Data from the Concrete Experiment.

APPENDIX 1:

Proof that OLS Estimators are BLUE for the Fixed Effects in a Split Factorial.

The well-known condition, under which the OLS estimators of fixed effects from a design, $\mathbf{X}$, are BLUE, is that there exists an invertible matrix $\mathbf{A}$ such that $\mathbf{V}^{-1}\mathbf{X}=\mathbf{XA}$, where $\mathbf{V}=Var(\mathbf{y})$. See Seber (1977, p. 63).

For a split factorial, we will show that $\mathbf{V}^{-1}\mathbf{X}=\mathbf{XA}$ for

$$\mathbf{A}=\frac{1}{r}\mathbf{X}_1'\Delta_1\mathbf{X}_1,$$

where $\mathbf{X}_1$ is the full $r \times r$ contrast matrix (including the constant column) for a single replicate of a 2^{k-p} design and Δ_1 is an $r \times r$ diagonal matrix.

Given the ordering of **X** in (3) for a split factorial and the resulting **Z** in (4), then

$$\sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i' = \bigoplus_{t=1}^{r/q} \left(\left(\bigoplus_{s=1}^{i} \sigma_i^2 \mathbf{I}_n \right) \oplus \left(\bigoplus_{s=i+1}^{q} \sigma_i^2 \mathbf{J}_n \right) \right). \tag{A.1.1}$$

From (2) and (A.1.1), it can be seen that $\mathbf{V} = \bigoplus_{j=1}^{r} \left(\alpha_j \mathbf{I}_n + \beta_j \mathbf{J}_n \right)$, where

$\alpha_j = \sum_{i=1}^{q} \sigma_i^2$ and $\beta_j = \sum_{i=1}^{j} \sigma_i^2$ and $r = 2^{k-p}$. Assume that $\sigma_q^2 > 0$. Since $\sigma_i^2 \geq 0 \quad \forall i$, it follows that $\alpha_j > 0, \beta_j \geq 0 \quad \forall j$. The inverse of **V** is then

$$\mathbf{V}^{-1} = \bigoplus_{j=1}^{r} \left(\alpha_j^* \mathbf{I}_n + \beta_j^* \mathbf{J}_n \right), \tag{A.1.2}$$

where $\alpha_j^* = 1 / \alpha_j$ and $\beta_j^* = \dfrac{-\beta_j}{\alpha_j \left(\alpha_j + n\beta_j \right)}$.

Let Δ_1 be an $r \times r$ diagonal matrix such that $\delta_{jj} = \alpha_j^* + n\beta_j^*$ is the jth diagonal element of Δ_1, then using (A.1.2),

$$\mathbf{V}^{-1}\mathbf{X} = \Delta\mathbf{X}, \tag{A.1.3}$$

where $\Delta = \Delta_1 \otimes \mathbf{I}_n$. Since $\delta_{jj} = \dfrac{1}{\alpha_j + n\beta_j}$, then $\delta_{jj} > 0$, thus both Δ_1 and Δ are invertible. Using (A.1.3) and $\mathbf{X}_1\mathbf{X}_1' = \mathbf{X}_1'\mathbf{X}_1 = r\mathbf{I}_r$, then

$$\mathbf{XA} = \left(\mathbf{X}_1 \otimes \mathbf{1}_n \right) \left(\frac{1}{r} \mathbf{X}_1' \Delta_1 \mathbf{X}_1 \otimes 1 \right) = \Delta_1 \mathbf{X}_1 \otimes \mathbf{I}_n \mathbf{1}_n = \left(\Delta_1 \otimes \mathbf{I}_n \right) \left(\mathbf{X}_1 \otimes \mathbf{1}_n \right) = \Delta\mathbf{X} = \mathbf{V}^{-1}\mathbf{X}.$$

Since $\mathbf{A}^{-1} = \frac{1}{r} \mathbf{X}_1' \Delta_1^{-1} \mathbf{X}_1$, **A** is invertible and therefore, the OLS estimators of the coefficients **b** in (1) are BLUE, if **X** is the contrast matrix of a split factorial.

APPENDIX 2:

Proof that ANOVA Estimators of the Variance Components for Split Factorials are UMVUIQ.

The proof of this result follows the argument given in Searle, Casella, and McCulloch (1992, pp. 417-421). It is necessary to assume that there is no kurtosis associated with the random effects.

The argument has two steps, both of which involve constructing a linearized version of the quadratic ANOVA estimators of the variance components. The proof in Appendix 1 implies that the argument on pp. 420-421 of Searle *et al.* (1992) is true for the ANOVA estimators for split factorial variance components, and hence that they are the best quadratic unbiased estimators of these variance components. By construction they are invariant, and to show that they have uniformly minimum variance among all such estimators, a condition specified by Seely (1971) must be satisfied. The remainder of this appendix shows that this condition is satisfied, and that hence the ANOVA estimators are UMVUIQ estimators.

For this proof, we use the model in (1). Using $\mathbf{X}_1$ as in (3),

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \frac{1}{nr}\mathbf{X}\mathbf{X}' = \frac{1}{nr}(\mathbf{X}_1 \otimes \mathbf{1}_n)(\mathbf{X}_1' \otimes \mathbf{1}_n') = \frac{1}{n}(\mathbf{I}_r \otimes \mathbf{J}_n).$$

Let us define a matrix $\mathbf{M} = (\mathbf{I}_{nr} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')$. Some manipulation shows that:

$$\mathbf{M} = \mathbf{I}_{r/q} \otimes \left(\bigoplus_{s=1}^{q} \mathbf{I}_n - \tfrac{1}{n}\mathbf{J}_n \right).$$

Seely's condition concerns the set,

$$\Omega = \left\{ \sum_{i=1}^{q} c_i \mathbf{M}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{M} \middle| c_i \in \Re \right\},$$

of all linear combinations of the $\mathbf{MZ}_i\mathbf{Z}_i'\mathbf{M}$ matrices. Searle, et al. refer to Theorem 6 of Kleffe and Pincus (1974) to state that if Ω is a quadratic subspace of symmetric matrices, then UMVUIQ estimators exist. A quadratic subspace is a set of matrices is such that if any matrix $\mathbf{B}$ is in the set, then $\mathbf{B}^2$ is also in the set. Using Lemma 1 condition (c) of Seely (1971), Ω is a quadratic subspace if

$$(\mathbf{MZ}_v\mathbf{Z}_v'\mathbf{M})(\mathbf{MZ}_w\mathbf{Z}_w'\mathbf{M}) = \sum_{s=1}^{q} c_{v,w,s}\mathbf{MZ}_s\mathbf{Z}_s'\mathbf{M} \quad \forall v, w \quad , \tag{A.2.1}$$

where $\{c_{v,w,s}\}$ is a $q \times q \times q$ tensor of constants. The condition, simply stated, is that the product of any pair of $\mathbf{MZZ'M}$ matrices is some linear combination of the set of original $\mathbf{MZZ'M}$ matrices.

We will now show that any split factorial experiment will satisfy this condition. Using the expression for $\mathbf{Z}_i$ in (4),

$$\mathbf{Z}_i\mathbf{Z}_i' = \mathbf{I}_{r/q} \otimes \left(\left(\bigoplus_{s=1}^{i} \mathbf{I}_n \right) \oplus \left(\bigoplus_{s=i+1}^{q} \mathbf{J}_n \right) \right).$$

Since $\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\mathbf{J}_n = \mathbf{0}_n$, where $\mathbf{0}_n$ is an $n \times n$ matrix of zeros, then

$$\mathbf{M'Z}_i\mathbf{Z}_i'\mathbf{M} = \mathbf{I}_{r/q} \otimes \left(\left(\bigoplus_{s=1}^{i} \mathbf{I}_n - \tfrac{1}{n}\mathbf{J}_n \right) \oplus \left(\bigoplus_{s=i+1}^{q} \mathbf{0}_n \right) \right).$$

Note that $\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right) = \left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)$. Thus if $v \le w$, then

$$(\mathbf{MZ}_v\mathbf{Z}_v'\mathbf{M})(\mathbf{MZ}_w\mathbf{Z}_w'\mathbf{M}) = (\mathbf{MZ}_w\mathbf{Z}_w'\mathbf{M})(\mathbf{MZ}_v\mathbf{Z}_v'\mathbf{M}) = \mathbf{MZ}_v\mathbf{Z}_v'\mathbf{M},$$

and the condition is satisfied.

ACKNOWLEDGEMENTS

This research was supported in part by the National Institute of Statistical Sciences and the NSF Center for Science and Technology for Advanced Cement-Based Materials through NSF grant numbers DMS-9313013 and DMS-9208758. The authors would like to thank Jeff Wu, Max Morris, Ivelisse Avilés, the editors, and referees for their helpful comments on an earlier draft of this paper.

REFERENCES

Anderson, R. D., Henderson, H. V., Pukelshiem, F., and Searle, S. R. (1984), "Best Estimation of Variance Components from Balanced Data with Arbitrary Kurtosis," *Math. Operationsforsh. Stat., Ser. Stat.*, 15, 163-176.

Ankenman, B. E., (1999), "Design of Experiments with Two-Level and Four-Level Factors," *Journal of Quality Technology*, 31, 363-375.

ASTM (1991), "Standard Test Method for Electrical Indication of Concrete's Ability to Resist Chloride Ion Penetration," *American Standards for Testing and Measurement*, pp. C1202-91.

Bainbridge, T. R. (1965), "Staggered, Nested Designs for Estimating Variance Components," *Industrial Quality Control*, 22, 12-20.

Bisgaard, S. (1994), "Blocking Generators for Small 2^{k-p} Designs," *Journal of Quality Technology*, 26, 288-296.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters*, New York: John Wiley.

Fries, A. and Hunter, W. G. (1980), "Minimum Aberration 2^{k-p} Designs," *Technometrics*, 22, 601-608.

Jaiswal, S. S., Picka, J. D., Igusa, T., Karr, A. F. Shah, S. P., Ankenman, B. E., Styer, P. (2000). "Experimental Studies of the Chloride Permeability of Concrete," *Concrete Science and Engineering*. (to appear)

Jaiswal, S. S. (1998), *Effect of Aggregates on Chloride Conductivity of Concrete: An Experimental and Modeling Approach*, Department of Civil Engineering, Northwestern University, unpublished Ph.D. dissertation.

Montgomery, D. C. (1997), *Design and Analysis of Experiments*, Fourth Edition, New York: John Wiley.

Montgomery, D. C. and Myers, R. H., (1995), *Response Surface Methodology*, New York: John Wiley.

Searle, S. R., Casella, G., and McCulloch, C. E. (1992), *Variance Components*, New York: John Wiley.

Seber, G. A. F. (1977), *Linear Regression Analysis*, New York: John Wiley.

Sitter, R. R., Chen, J., and Feder, M. (1997), "Fractional Resolution and Minimum Aberration in 2^{n-k} Designs," *Technometrics*, 39, 382-390.

Smith, J. R. And Beverly, J. M. (1981), "The Use and Analysis of Staggered Nested Factorial Designs," *Journal of Quality Technology*, 13, 166-173.

Sun, D. X., Wu, C. F. J., and Chen, Y. (1997), "Optimal Blocking Schemes for 2^n and 2^{k-p} Designs," *Technometrics*, 39, 298-307.

Wolfinger, R. D., and Tobias, R. D. (1998), "Joint Estimation of Location, Dispersion, and Random Effects in Robust Design," *Technometrics*, 40, 62-71.